UNIVERSITY OF LEEDS

THE EVALUATION OF NEW LOCAL RAIL STATIONS IN WEST YORKSHIRE

J.M. PRESTON

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

School of Economic Studies

(Institute for Transport Studies)

JUNE 1987

ABSTRACT

Between 1982 and 1984 six new stations were opened on passenger railways in West Yorkshire whilst additional sites were being considered. The aim of this thesis is to assess and evaluate demand at the six existing and up to 28 potential new stations in West Yorkshire. This involved three inter-related strands of research. Firstly, market research was carried out at the six new stations opened and was particularly useful in determining the proportions of travellers generated and abstracted. Secondly, statistical models, based on aggregate simultaneous and disaggregate mode split structures, were developed and their forecasting abilities assessed. In this part of our work a subsidiary aim emerged; namely to assess the trade-off between complexity and accuracy in modelling new station demand. Thirdly, an evaluation framework, using one set of demand forecasts, was developed, taking into account the costs and benefits to Public Transport operators, new station users and society as a whole.

It was found that the six new stations opened in West Yorkshire may be judged a success in both financial and social terms, whilst up to 10 sites were identified as representing good social investments. In terms of our subsidiary aim it was found that, given limited resources, simple modelling approaches, such as provided by an aggregate simultaneous model, may be preferable to more complex approaches when evaluating small-scale new station programmes. Our findings are shown to have implications beyond West Yorkshire as simple guide lines for new station site identification and evaluation have been determined.

ACKNOWLEDGEMENTS

This work was financed by the Science and Engineering Research Council and the British Railways Board as a CASE (Collaborative Award in Science and Engineering) project.

I would like to thank David Wharton-Street and John Ferro of British Rail, Bob Tebb and Roger Cobbe of the West Yorkshire Passenger Transport Executive and Phil Beaumont and Gary Billington of West Yorkshire Councy Council who, along with many others in their respective organisations, gave their full co-operation to this project. In addition, I would like to thank the many members of the Institute who have helped to produce an atmosphere that was conducive to research. In particular I must record my appreciation of the support given by my two supervisors, Peter Bonsall and Dr. Chris Nash, without which this work could not have been carried out. Also I must thank Peter Hopkinson, Keith Tolley, Mark Wardman and Jim Whelan for helping with the surveys, Ferri Ghahri-Saremi for his assistance with computing, Dr. Neil Douglas for being my first year "mentor", Frank Montgomery for providing the data used to estimate the effects on road congestion and Dr. Tony Fowkes and Dr. Huw Williams for making clearer some of the vagaries of logit modelling. Mark Wardman should also be thanked for proof reading several chapters. All mistakes are, of course, purely my own.

Lastly, I would like to thank Sue Nemes for her excellent job in typing the final manuscript and Teela for putting up with the anti-social working practices (and behaviour) that producing this thesis has involved.

TABLE OF CONTENTS

Page

Page

Page

Page

LIST OF FIGURES

LIST OF TABLES

LIST OF WIDELY USED ABBREVIATIONS

AD	=	Absolute Deviation measure (see Table 6.12 for definition)
ASC	=	Alternative Specific Constant
ASM	=	Aggregate Simultaneous Model
BA	=	Best Alternative
BCR	=	Benefit Cost Ratio
BR	=	British Rail
CASE	=	Collaborative Award in Science and Engineering
CBA	=	Cost Benefit Analysis
CO	=	Car Owning
CR	=	Chance Recovery
DA	=	Disaggregate Approach
EIL	=	Extended Incremental Logit
EL	=	Extended Logit
FPR	=	First Preference Recovery
FYRR	=	First Year Rate of Return
HL	=	Hierarchical Logit
HRT	=	Heavy Rapid Transit
IAA	=	Independence from Irrelevant Alternatives
IID	=	Identically and Independently Distributed
IL	=	Incremental Logit
IRR	=	Internal Rate of Return
IVT	=	In Vehicle Time
JTW	=	Journey to Work
LOS	=	Level of Service
LRT	=	Light Rapid Transit
LUTS	=	Land Use and Transportation Study

MNL	=	Multinomial Logit
MNP	=	Multinomial Probit
MSR	=	Market Share Recovery
NBC	=	National Bus Company
NCO	=	Non Car Owning
NPAAS	=	National Passenger Accountancy and Analysis System
NPV	=	Net Present Value
NWM	=	Non Work Model
OD	=	Origin/Destination
OPCS	=	Office of Population Censuses and Surveys
OVT	=	Out of Vehicle Time
PTA	=	Passenger Transport Authority
PTE	=	Passenger Transport Executive
PTS	=	Passenger Train Survey
RMSE	=	Root Mean Square Error (See Table 6.12 for definition)
RP	=	Revealed Preference
RUM	=	Random Utility Model
SAS	=	Statistical Analysis System
SASPAC	=	Small Area Statistics Package
SE	=	Socio Economic
SERC	=	Science and Engineering Research Council
SP	=	Stated Preference
SPSS	=	Statistical Package for the Social Sciences
SWS	=	Special Workplace Statistics
TP	=	Transfer Price
TRM	=	Trip Rate Model
VOT	=	Value of Time
WK	=	Walk time
WMCC	=	West Midlands County Council

WT	=	Wait time
WYMCC	=	West Yorkshire Metropolitan Councy Council
WYPTE	=	West Yorkshire Passenger Transport Executive
WYTS	=	West Yorkshire Transportation Study
YM	=	Yesterday's Mode

PART ONE RESEARCH FRAMEWORK

CHAPTER ONE INTRODUCTION - DEFINING THE RESEARCH PROBLEM

1.1 BACKGROUND TO THE RESEARCH

This thesis describes research carried out, in the main, between October 1982 and September 1985. The starting point for this study was the programme of opening new stations on the West Yorkshire local rail network, initiated by the Passenger Transport Executive (PTE). This was seen as, at that time, providing a near unique opportunity to monitor and evaluate the effects of new local rail stations in Britain. Initially the programme was envisaged as involving the opening of nine stations in the period 1982-85*, with up to a further 15 stations thereafter.** In fact, due to unforeseen engineering and political problems, only six new stations were opened during the period 1982-5. These were:

(i) Fitzwilliam, on the Leeds to Doncaster line, opened 1st March, 1982

(ii) Deighton, on the Huddersfield to Leeds/Wakefield lines, opened 26th April 1982.

(iii) Crossflatts, on the Keighley to Leeds/Bradford lines, opened 17th May 1982.

(iv) Slaithwaite, on the Huddersfield to Manchester line, opened 13th December 1982.

(v) Bramley (Swinnow Road), on the Leeds to Bradford line, opened 12th September, 1983

(vi) Saltaire, on the Keighley to Leeds/Bradford lines, opened 9th April 1984.

In Chapter Two we shall go on to put this new station programme into context, in terms of time, place and scale, and thus identify the objectives, as far as the PTE is concerned, of such a programme.

---

* Fitzwilliam, Crossflatts, Deighton, Slaithwaite, Bramley, Lightcliffe, East Garforth and Hawksworth (Sully, 1981)

** East Ardsley, Hemsworth, Low Moor, Luddendenfoot, Milnsbridge, Osmondthorpe, Salter Hebble, Thornhill, Wrenthorpe, Outwood, Cottingley, Gamble Hill, Frizinghall, Methley and Stanningley (WYPTE, 1984A)

1.2 RESEARCH AIMS

The goal of this research is to develop a quantitative, scientific approach to the measurements of the costs and benefits of six existing new stations and a number (as yet unspecified) of potential sites. Our study's objectives were initially taken as given, as this research was financed by a CASE award, sponsored by British Rail (BR) and supported by WYPTE, with the co-operation of West Yorkshire Metropolitan County Council (WYMCC). These objectives were as follows:

(i) To assess the volume and characteristics of the patronage at new stations and to examine its dependence on factors such as the location and characteristics of the population served, access to the station by alternative modes and the nature of the rail service provided in comparison with other modes of transport.

(ii) To establish the extent to which this traffic is achieved at the expense of other modes, and the extent to which it is newly generated.

(iii) To measure the costs and benefits of the new station to users, rail and bus operators and the public at large.

(iv) To assess the effects on accessibility to jobs, shopping and other facilities for the inhabitants concerned using both objective and attitudinal measures.

In carrying out these objectives we shall be adding to the work done by WYPTE (1984B).

As we shall demonstrate in 2.3.2, it is likely that the secondary effects (on activity levels) and the tertiary effects (on land use) of new stations will be limited and hence the bulk of this research will concentrate on the primary effects on trip patterns (i.e. objectives (i), (ii) and (iii)).

1.3 RESEARCH METHODS

As will be shown in Chapter Three, in order to achieve the above objectives, this study's methodology will consist of three strands:

(i) Market Research. Due to resource constraints this will be limited to interviewing boarding passengers at new stations, with the main surveys, carried out at six sites between 1982 and 85, covering 921 respondents. The survey strategy and results will be discussed in Chapter Five.

(ii) Statistical Modelling. In order to evaluate potential station sites, approaches that forecast the likely level and nature of demand need to be developed. In 3.2 we shall show that a number of statistical models have been developed to forecast rail demand, with simple regression and trip rate models most commonly used to assess the types of new stations we are interested in. By comparing the theoretical advantages/disadvantages of different model types (discussed in 3.3) and the practical advantages/disadvantages, particularly with reference to data requirements (discussed in Chapter Four) two approaches were identified as being worthy of further investigation:

(a) An aggregate approach. As Chapter Six shows, this was based on the development of multiple regression equations.

(b) A disaggregate approach. As Chapter Seven shows, this was based on logit models of mode choice for work journeys.

(iii) Evaluation. The information provided by the market research and statistical modelling stages will be used as input to the evaluation stage. As discussed in 3.4 this will include both a financial appraisal and social cost-benefit analysis, the results of which will be presented in Chapter Ten.

Figure 1.1 shows how the aims, methods and results are inter-related by our research framework. An important aspect of this framework is the feedback between the study's aims and results. In particular we wish to determine to what extent this feedback is positive (thus leading to a change in the "equilibrium" state i.e. new station(s) opened) and what extent this feedback is negative (no change in the "equilibrium" state). In determining this, key regulators will be the search procedure (discussed in 9.1.1) and decision rule(s) chosen (discussed in Chapter Ten).

As a result of the methodology to be used in this study a number of subsidiary objectives may be developed. For example the aggregate and disaggregate approaches may be compared, both in terms of practical (and in particular, as will be shown in Chapter Nine, predictive) and theoretical capabilities. In addition we shall be able to test the practical applicability of

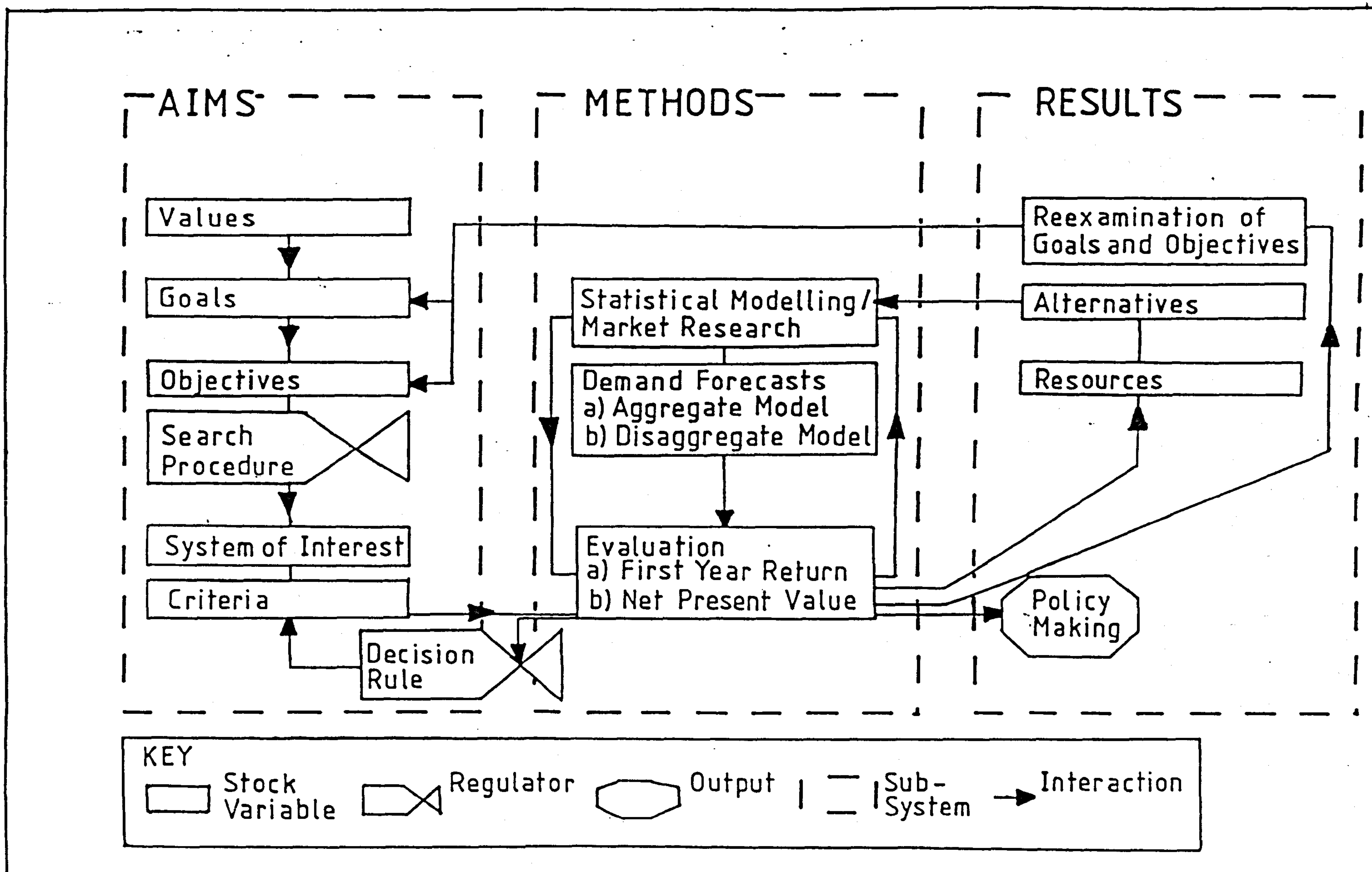


FIGURE 1.1 RESEARCH FRAMEWORK

the forecasting methods developed and in particular that of our disaggregate approach, as will be shown in Chapter Eight. Furthermore, we will be able to illustrate the effects of extending the evaluation approaches most commonly used to assess rail investment in the U.K. This will be done in Chapter Ten.

1.4 LEVELS OF RESOLUTION

Before progressing with this study the level of resolution, in particular in terms of spatial, temporal and theoretical dimensions, need to be determined. The spatial level of resolution should already be apparent. We are interested in existing and potential new stations in West Yorkshire, although as a by-product of our work, studies were carried out elsewhere including Derbyshire, Leicestershire, North Yorkshire and Tyne and Wear. The temporal and theoretical levels of resolution are not so easily identifiable and are thus discussed below.

1.4.1 Temporal Level of Resolution

At least three levels of temporal resolution may be identified:

(i) Cross sectional. This will provide the basis for most of our modelling work.

(ii) Time-series (or longitudinal). Such an approach is limited by data availability but will be used to assess changes in new station usage over time (section 9.4).

(iii) Before and after studies. The scope for such an approach was limited by the fact that three of the six new stations were opened before this research began, but this problem was reduced by including recall questions in our market research.

It can thus be seen that this research adopts a flexible approach in that it attempts to incorporate elements of all three temporal horizons . In addition it should be noted that our study period (October 1982 to September 1985) was one in which there were numerous fare changes, in particular the introduction of off-peak schemes and the extension of Metrocard to rail (see Table 2.7), whilst new rolling stock was also introduced (WYPTE and BR, 1982). Moreover, not long after our study period finished two important changes occured. Firstly, the 1985 Local Government Act led to the abolition of WYMCC in April 1986. Secondly, the 1985 Transport Act led to the deregulation of stage carriage bus services, outside Greater London, in

October 1986.

1.4.2 Theoretical Level of Resolution

Alonso (1968) demonstrated that there are important trade-offs in modelling between specification and input error (see Figure 1.2). Cast more generally this implies that, for any study, the accuracy of the results is dependent on the quality of the input and hence there are likely to be trade-offs between theoretical and practical levels of resolution (or, expressed another way, between complexity and accuracy). In our study input quality is limited by:

(i) Data constraints. As Chapter Four demonstrates this study is dependent on existing data sources and as these data sources were not specifically collected with the objectives of this research in mind there are numerous application problems.

(ii) Time and cost constraints. This research was envisaged as involving approximately three man years of labour and a very limited monetary budget. This has restricted the scale of market research that can be undertaken and has also meant that a number of interesting side issues could not be investigated fully.

As Figure 1.1 has shown, our research framework has been designed so as to produce practical results, particularly to determine the effects of a West Yorkshire new station policy. Given this aim and the constraints listed above it is evident that we will face trade-offs between theory and practice. This is examined further in 3.3 and is especially evident in Chapters Seven and Eight. Where such trade-offs occur we have concentrated on producing practical results in order to achieve our research objectives whilst attempting, given our resource constraints, to minimise theoretical problems as far as possible. Thus, to an extent, this research may be seen as an attempt to determine the right balance between complexity (particularly in terms of resource requirements) and accuracy in order to successfully evaluate new local rail stations.

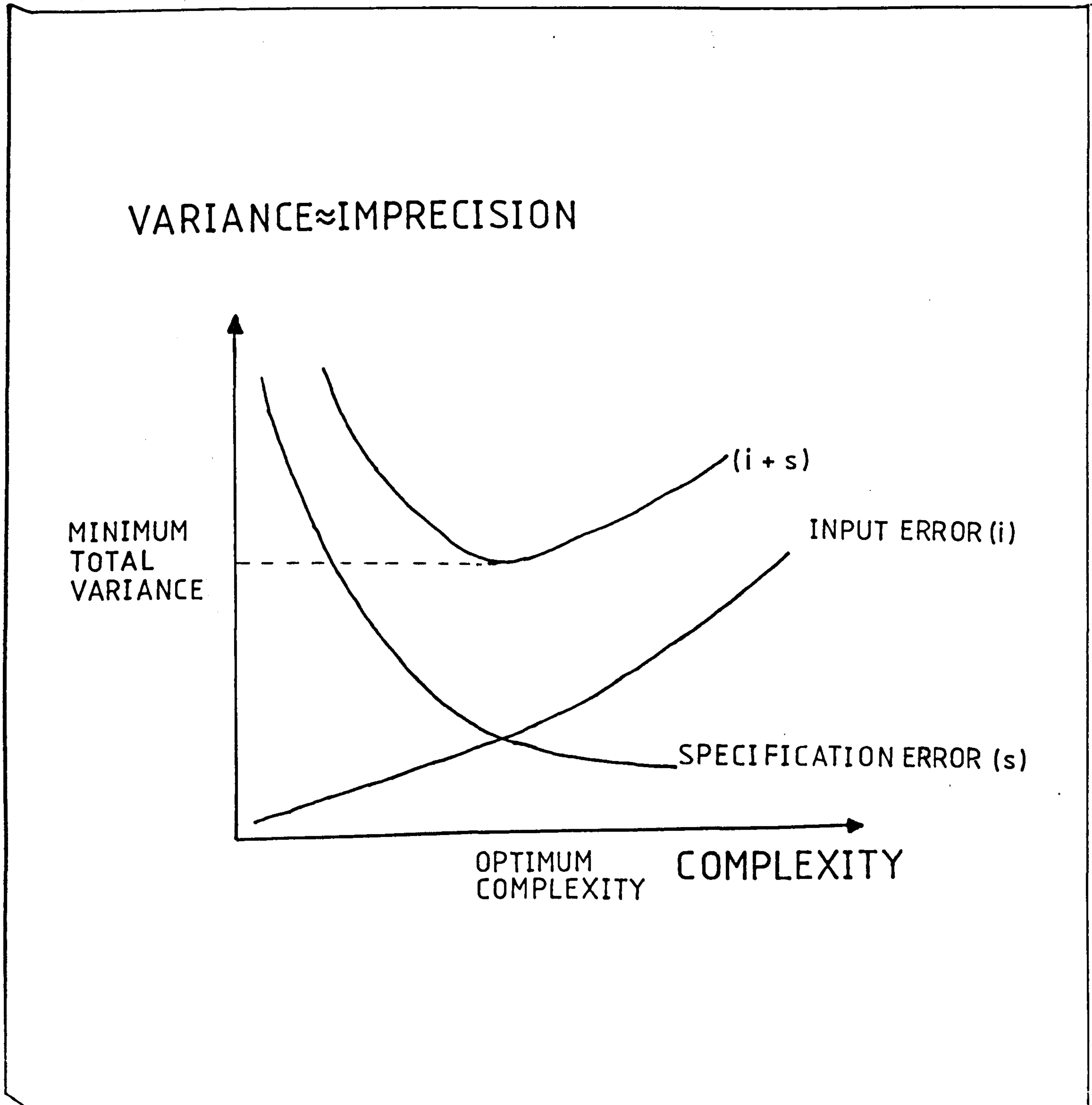


FIGURE 1.2 THE RELATIONSHIP BETWEEN ACCURACY AND MODEL COMPLEXITY
(after Alonso, 1968)

1.5 THESIS OUTLINE

Given the above, this thesis will consist of four parts. In Part One, which includes this chapter, a research framework will be developed. In Chapter Two the study's system of interest will be examined and placed in context in terms of time, place and scale. In Chapter Three the study's methodology will be examined by outlining previous work in the field and assessing the choice of modelling approaches and evaluation techniques available.

Part Two of this thesis will examine the study's data requirements. In Chapter Four the availability and suitability of existing data bases will be examined. It will be stressed how data availability has acted as a constraint on the depth of analysis, particularly with regard to modelling. In Chapter Five we will go on to detail the data base developed by the study's market research.

In Part Three the statistical modelling approaches used to forecast new station usage will be discussed. Hence in Chapter Six a series of aggregate models will be developed, whilst in Chapter Seven a set of disaggregate journey to work models will be developed. In Chapter Eight we shall go on to examine some of the issues that arise in applying the disaggregate models that we have developed.

The fourth, and final, part of this thesis will bring together the previous three parts. In Chapter Nine a number of modelling approaches will be used to forecast new station demand and their results, particularly in terms of accuracy, will be compared. A set of these forecasts will then be used as input into the evaluation framework, as detailed by Chapter Ten. Lastly, in Chapter Eleven, a series of conclusions will be drawn relating to both theoretical and empirical findings and allowing certain policy implications and recommendations for future research to be made.

CHAPTER TWO DEFINING THE SYSTEM OF INTEREST

2.1 INTRODUCTION

It should be clear from Chapter One that to a large extent our system of interest is taken as given; that is we are interested in the six new stations opened in West Yorkshire between 1982 and 1984 as well as a number (as yet unspecified) of potential new station sites. In this Chapter we shall attempt to place the West Yorkshire new station programme in context in terms of time, place and scale.

This will be done in three sections. In section 2.2 we shall examine the historical development of the West Yorkshire passenger railway network and the status of the network immediately prior to the new station programme. This will allow us to determine some of the aims of such a programme. Also in Section 2.2 we shall detail past studies of possible changes to the rail network, which will then allow us to show that a new station programme is just one of a number of options that might be adopted in order to improve the performance of the local rail network.

In Section 2.3 we shall go on to place this study in its geographical context. Firstly, we shall attempt to explain why railways have been widely adopted as solutions to urban travel problems. Secondly, recent urban railway investments in the U.K. will be discussed and following on from this it will be shown how a number of new station schemes have been developed or proposed elsewhere in the U.K. Lastly, we shall show that there have been numerous worldwide examples of recent investments in urban railways.

In Section 2.4 we shall go on to draw some conclusions from the above review. In particular it will be shown that the small scale of the West Yorkshire new station programme makes it different from many other urban railway investment programmes.

2.2 THE PAST ROLE AND CURRENT PROSPECTS OF WEST YORKSHIRE'S PASSENGER RAILWAYS

2.2.1 Historical Development

Railways have played a role in the West Yorkshire conurbation for over 150 years. The first passenger rail service in the county was from Leeds to Selby in 1834 (Joy, 1978) and this line was soon followed by many others, built by a number of different companies, so that by 1900 the County's rail network reached a peak of 700 route km and some 200 stations (see Figure 2.1). This peak system had a number of handicaps.

(i) It was multicentric being based not just on Leeds but also on Bradford, Dewsbury, Halifax, Huddersfield, Wakefield and the Five Towns area (especially Pontefract).

(ii) Due to difficult topography routes were often indirect, for example the Midland Railway route between Leeds and Bradford went via Shipley and certain stations were some distance from the communities they purported to serve (e.g. Queensbury).

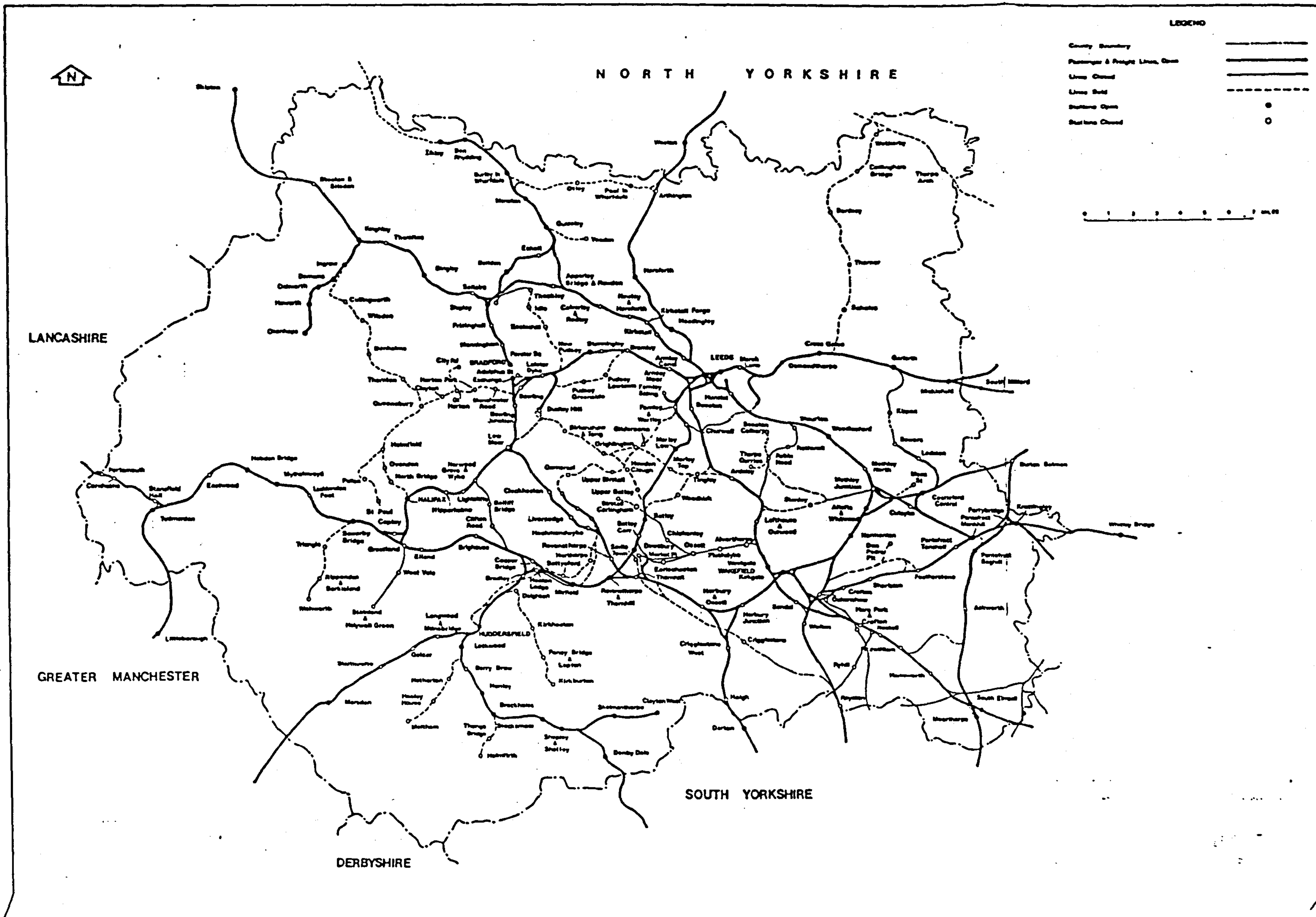
(iii) The railways were built principally for the movement of freight (especially coal) rather than passengers.

(iv) The large number of competing companies led to duplication of routes (for example the London and North Western and Lancashire and Yorkshire routes through the Spen Valley were almost identical) and separate central area termini.

As a result the railways had limited effect on the developing morphology of the conurbation, with only the Ilkley, Wetherby, Harrogate and Bradford lines producing any degree of suburbanisation (Dickinson, 1967), whilst the municipal tramways (and subsequent bus systems) and the automobile have had a much greater effect.

Because of these handicaps the passenger rail system has undergone a long period of decline. By 1945 30 stations were closed, by 1960 a further 60 stations had closed whilst the "Beeching axe" accounted for over 50 stations (of which only a half were due to line closures) (British Railways Board, 1963). However, it should be noted that the recommendations of the Beeching Report were never fully implemented; out of 18 passenger services proposed for

FIGURE 2.1 EXISTING AND CLOSED RAIL LINES IN WEST YORKSHIRE (Source, WYTCONSULT, 1977F)



closure only 12 were withdrawn with 19 stations escaping closure (Haigh, 1978). Nevertheless, the frequency of train stops in the County decreased from 4700 a day in 1960 to 2200 a day in 1970 (down 53%) with a consequent decrease in passengers from 14 m per annum to 8 m (down 43%). Similarly, WYTCONSULT (1976A) showed the number of rail passengers between 1962 and 1973 declined by 34% (a similar figure for bus was 27%), although rail receipts only declined by 19%.

2.2.2 Situation Prior to the New Station Programme

In 1981 the West Yorkshire passenger rail network consisted of 370 route km, 48 stations (4 principal intercity, 23 local staffed and 21 unstaffed) and 16 services (as shown by Figure 2.1) with the network being very similar to that of 1854. Assuming that 80% of journeys are internal to the County (WYPTE, 1979), it was estimated that in 1981 over 54,000 passengers boarded and/or alighted at West Yorkshire stations on an average weekday. The West Yorkshire Transportation Study (WYTS) rail survey (1976B) indicated that local service usage was heavily peaked, with 50-60% of demand occuring during the peak hours except for the Leeds to Sheffield and Leeds to Bradford lines, and with the journey to work accounting for about 48% of journeys. Usage was mainly centred on Leeds City (which accounts for one end of almost 50% of trips), with the Bradford/Huddersfield-Leeds-York corridors being the most heavily loaded (WYTCONSULT, 1977A).

Reliable time-series data for West Yorkshire rail patronage is difficult to obtain but Table 2.1 suggests that patronage between 1974 and 77 was declining in excess of the nationwide rate of 1.6% per annum identified by Webster (1977), possibly due to the 49% increase (20% in real terms) in rail fares in 1975. Between 1978 and 80 however there seems to have been an increase in demand, possibly in response to PTE improvements. However in 1981 and 82 there was further decline due to the effects of the recession, the increased price competitiveness of bus and the rail strikes of January/February and July 1982. This decline halted by 1983, and since then there have been large increases in demand (particularly in 1985).

Given that patronage had not increased dramatically between the mid 70s and the early 80s (and indeed has been subject to fluctuations), a number

	Financial Year Starting										
	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985
A	0.93	0.84	0.88	1.01	1.03	1.14	0.98	0.99	1.04	NA	NA
B	0.92	0.80	0.85	1.10(1)	0.76(2)	1.01	0.82(3)	0.73(3)	0.75	0.81	1.12

1974 = 1.00

TABLE 2.1 RAIL USAGE IN WEST YORKSHIRE 1974-1985

Sources:

A "Transportation Trends" 1977-80 supplemented by PTE data. Based on counts of autumn weekday usage supplied by PTS

B PTE "Annual Report and Accounts" and "Public Transport Plans" Annual Passenger - based on ticket sales data

Notes:

(1) Inclusion of South Yorkshire bound services (Leeds to Darton/South Elmsall and Huddersfield to Denby Dale

(2) Reductions in service due to fuel shortages

(3) Figures affected by industrial disputes

of problems facing West Yorkshire's passenger rail network in the early 1980s may be identified:

(i) It had low usage. Less than 1% of West Yorkshire's mechanised journeys were made by rail, accounting for only 2% of Public Transport trips and 8% of Public Transport passenger mileage (as the mean distance travelled is 17.95 km). However rail was important for journeys to work over 10 km long (10% share) and for journeys to work to Leeds in areas served by railways where it had a 6% share (WYTCONSULT, 1976A). Nonetheless Table 2.2 shows that, despite having the largest network (in terms of route km) and the third largest population, rail usage in West Yorkshire was lower than in any of the other English PTEs, with the exception of South Yorkshire.

(ii) There was poor access to the system, as only 15% of the County's population and 45% of its jobs were located within one km of a station.

(iii) The system was requiring increasing levels of subsidy. Between 1979 and 1982 the PTE's section 20 payments to BR (which cover all local services except the Pontefract Baghill-Moorthorpe and Huddersfield-Wakefield routes) increased from £5.0 to £8.9 m. All services incur a deficit, although the Leeds to Bradford, Micklefield and South Elmsall trains cover working expenses, with the deficit per passenger mile ranging from over 11.0 p for the Bradford to Keighley, Huddersfield to Wakefield and Huddersfield to Sheffield services to 1.4 p for the Leeds to Doncaster service (WYPTE, 1979). As a result train revenues only covered 39% of total costs (Sully, 1981) and it can be seen from Table 2.2 that West Yorkshire had the highest claim per passenger mile of all the six English PTEs.

(iv) As a result of the above factors there was poor utilisation of assets. Although figures are not readily available it is clear that the West Yorkshire passenger rail network had low average load factors.

(v) As we have seen, the recommendations of the Beeching Report were not fully implemented and this led to anomalies. For example on the Huddersfield line Slaithwaite and Milnsbridge stations were closed, but Marsden stayed open. Similarly, certain lines lost their intermediate stations (for example the Airedale and Leeds to Bradford lines) whilst others (for example the Harrogate and Minster lines) maintained theirs.

	1981 population	Road km	Rail route (km)	Stations	Rail (1980/1) Passenger km (millions per annum)	Rail (1980/1) Passenger journeys (millions per annum)	1974 % of Mechanised journeys	Claim payment per pass. mile (pence)
Tyneside	1.36	2375	68	45	-	16.0	6.0	5.2
Merseyside	1.50	2259	110	71		40.6	8.0	5.0
Greater Manchester	2.57	4587	197	106	314	27.5	5.0	6.4
West Yorkshire	2.00	4314	214	49	157	8.8	0.5	7.2
South Yorkshire	1.29	2947	142	25	-	-	-	5.2
West Midlands	2.62	3875	105	62	247	23.3	2.0	2.7

TABLE 2.2 RELATIVE ROLE OF RAILWAYS IN SIX ENGLISH PTES (Source: Chan, 1982)

Table 2.3 and 2.4 indicate some of the reasons for the low rail usage in West Yorkshire. Firstly, the cities and towns in the county are relatively self contained, so that 87.6% of the County's work force work in their district of residence, although there is evidence that this figure may be decreasing. Secondly, in part due to the short distances travelled, bus is a much more important Public Transport mode than rail, with rail only accounting for 4.8% of the Public Transport journey to work market. Lastly, Table 2.4 indicates that competition from private transport, due in part to the good local road network, is important. In 1981 out of seven major English city centres, Leeds was the only instance where the Public Transport share of the journey to work was less than 50%.

2.2.3 Opportunities for Change - Past Studies

Several studies designed to improve West Yorkshire's Public Transport network have been undertaken. The main emphasis has been on the role of buses in freestanding towns/cities as exemplified by the Leeds approach (Leeds City Council et al. 1969 -reappraised 1973) and the Bradford (Travers Morgan and Partners, 1974) and Huddersfield (LGORU, 1976) bus studies. However there have been a large number of studies that have looked at options for the local rail network and included studies of new stations. These will be examined in turn.

(i) The first of these was provided by the County's first Land Use and Transportation Study (LUTS), entitled the West Yorkshire Transportation Study (Traffic Research Corporation, 1969), although the area under study differed from the current county in that it included Harrogate, Saddleworth and Selby but excluded the Hemsworth area. The main emphasis of this study was on predicting inter-urban road traffic but there was a small scale rail study, involving station counts for one week in October 1966, and the calculation of passenger totals for trains crossing screen lines betwee 0700 and 0900 am (93 trains, 4742 passengers). The main problems were seen to be low usage, poor accessibility and low frequencies. A number of improvements were studied, in particular a figure of 8 rapid transit system linking Leeds with Shipley, Bradford, Garforth and Castleford with branches to Skipton, Harrogate, York, Selby and Wakefield as well as a Huddersfield to Denby Dale link. This network was then compared with a network of express buses, with the benefit-

	Living and Working in District	of which Rail	of which Bus	Living elsewhere and Working in District	of which Rail	of which Bus
Bradford	157710	850	42390	26660	600	4600
Calderdale	67970	110	16280	12040	70	2700
Kirklees	124740	550	27680	16770	240	2700
Leeds	265790	2740	74570	51850	2900	6720
Wakefield	107160	210	22290	25610	320	5850
TOTAL	723370	6350	183210	132930	4130	22570

TABLE 2.3 JOURNEY TO WORK IN WEST YORKSHIRE

Source : Census 1981

	Total Employment	Rail	(%)	Bus	(%)	Public Transport	(%)
London	1,070,170	689,330	64.4	117,990	11.0	807,320	75.4
Manchester and Salford	106,950	15,780	14.8	38,320	35.8	54,100	50.6
Liverpool	92,690	18,210	19.6	31,650	34.1	49,860	53.8
Sheffield	59,720	1,390	2.3	31,830	53.3	33,220	55.6
Newcastle	57,620	3,620	6.3	27,090	47.0	30,710	53.3
Birmingham	101,190	12,270	12.1	39,950	39.5	52,220	51.6
Leeds	59,010	4,120	7.0	24,660	41.8	28,780	48.8

TABLE 2.4 JOURNEY TO WORK TO CITY CENTRES

Source: Census, 1981

cost ratio evaluated at 1.5 and 17.0 respectively and hence the studies favoured the development of express buses. In a recommended programme of priorities it was, however, suggested that up to 1981 frequencies on the Leeds to Harrogate, Garforth and Bradford services should be improved to provide 20 minute peak headways, with the incremental benefit-cost ratio ranging from 0.66 for Horsforth services to 4.57 for Bradford services. It was also suggested that the Shipley to Bradford Forster Square line should be converted to bus- way.

(ii) Following the formation of the Metropolitan County in 1974 a second West Yorkshire Transportation Study (WYTS) was undertaken by WYTCONSULT, a joint company of Freeman Fox and Associates, Martin Voorhees and Associates and seconded County staff. Again this was a wide ranging study but did include an important rail element. The data base for the rail studies was provided by a survey on June 17th 1975 which involved the issueing of 11,274 questionnaires to departing passengers at seven main stations (Leeds City, Bradford Exchange, Bradford Forster Square, Wakefield Westgate, Wakefield Kirkgate, Huddersfield and Halifax) with a response rate of 43% (or 31% of all passengers counted) (WYTCONSULT, 1976B). It should be noted that the barrier survey missed travellers with an origin and destination outside the county, with an origin and destination outside the surveyed stations and with an origin or destination outside the County and destination or origin outside the surveyed stations. In addition a rail inventory (WYTCONSULT, 1975A) was undertaken, including station and line inventories, station accessibility studies, analysis of passenger services and loadings and matrices of station to station movements.

Important findings included the high usage of rail by people from car owning households (38% of all passengers, 73% in the morning peak), the importance of intermediate non manual, junior non manual and skilled manual users and the high deficit per passenger mile (in 1974 3.0p for rail compared to 1.0p for bus). This second WYTS also included a number of further rail studies, which are considered below.

(iii) Initial rail studies were included as part of Public Transport tests for Leeds and Bradford. For the short term (up to 1981) a rail reinstatement test was carried out for Leeds (WYTCONSULT, 1976C) which involved the development of Bradford to Garforth and Harrogate to Wakefield/Doncaster through services with a net increase of 23 in the number of stations. This network involved increased costs over 1975 services of £1.09m, whilst revenue was estimated to increase by £1.77m. However, compared to the 1981 Economic Base (which included only inter-city rail services), there was an increase in user benefits of £0.16m but an increase in deficit payments of £1.91m. The tests concluded that rail improvements were only viable for certain corridors.

(iv) As Central Area employment, car ownership and traffic congestion were forecast to grow between 1981 and 1991 the case for rail retention was seen to be stronger in the medium term. Hence the Medium Term (1991) Provisional Plan for Leeds (WYTCONSULT, 1977B) proposed the development of a South Garforth spur, a Bradford central rail link and withdrawal of services via Castleford and Normanton, with a net increase of eight stations. The schemes would involve capital costs of £1.28m (at 1975 prices) and an operating deficit compared to the Economic Base of £6.0m, but an increase in user benefits of £9.2m. However, there were large variations in the performance of lines, with the Skipton-Bradford-Leeds-South Garforth spine providing the best results and the Caldervale line the worst, whilst intercharge was seen as being crucial.

(v) In the long term Light Rapid Transit (LRT) might be considered for Leeds (WYTCONSULT, 1977C). Two LRT networks were considered. Network A (52 km, 19.5 km in street) serves the York Road, Headingley, Moortown, Belle Isle and Harehills corridors. A simpler and more segregated Network B serves only the York Road and Stanningley corridors (26 km, 1.4 in street, 37 stops). Network B, in particular, had important savings over bus but as only £25.8 m (at 1975 prices) was available for all transport capital projects in Leeds up to 1991 the infrastructure costs proved prohibitive. The safeguarding of routes, particularly at Quarry Hill flats, Colton, Hunslet and Wellington Street was, however, recommended.

(vi) For Bradford two rail reinstatement tests were carried out (WYTCONSULT, 1977D) for the medium term (1991). A subtest on the Improved Accessibility network included links to Huddersfield and Wakefield with up to nine new stations. This was disregarded as it led to a small net disbenefit for Public Transport users and increased the Public Transport deficit by £1.5m. For the Medium Term Provisional Plan for Bradford (WYTCONSULT, 1977E) the same network as that of the Medium Term Provisional Plan for Leeds was studied. Greater consideration was given to the alignment of a cross Bradford link and six associated service options. The network involved capital costs of between £3.8m and £5.6m, with an increased Public Transport deficit over the 1991 Economic Base of around £1.5m producing user benefits of only £0.6m pa, and was therefore rejected.

(vii) WYTCONSULT also analysed six countywide networks (1977F) ranging from the six station Basic network to the 109 station Maximum Accessibility network (see Figure 2.2). Analysis of the Intermediate network included a circular Leeds-Castleford-Normanton-Wakefield service, the Rationalised network included the development of through services from Todmorden to Micklefield, from Bradford to Knottingley and from Keighley to Darton whilst the Improved Accessibility network included new services such as Leeds to Todmorden via Brighouse, Bradford to Wakefield via the Spen Valley and a circular Leeds-Garforth-Castleford-Wakefield service. The Maximum Accessibility network would include a number of links to allow the development of through services from Ilkley to Scholes and from Wetherby to Horsforth. A comparison of all six networks showed that only the Basic network fails to incur a deficit, whilst the Intermediate and Rationalised networks were favoured as they generated more traffic at less cost than the existing network.

(viii) An Interim Preferred Network was also studied (WYTCONSULT, 1976A) which identified the Leeds to Skipton and South Elmsall services as having the best potential for development. The reopening of the Kippax branch and the Brighouse link was proposed, although the Spen Valley and Five Towns circular service were disregarded as they failed to cover working expenses. Overall a net gain of 15 stations was proposed.

(ix) The findings of WYTCONSULT were finally put together to form a series of recommended policies and plans (1977G). Up to 19 station openings were proposed (including the South Garforth spur) and 11 closures, due to withdrawal of passenger services between Castleford and Knottingley, Huddersfield and Denby Dale/Clayton West and the rerouteing of the York-

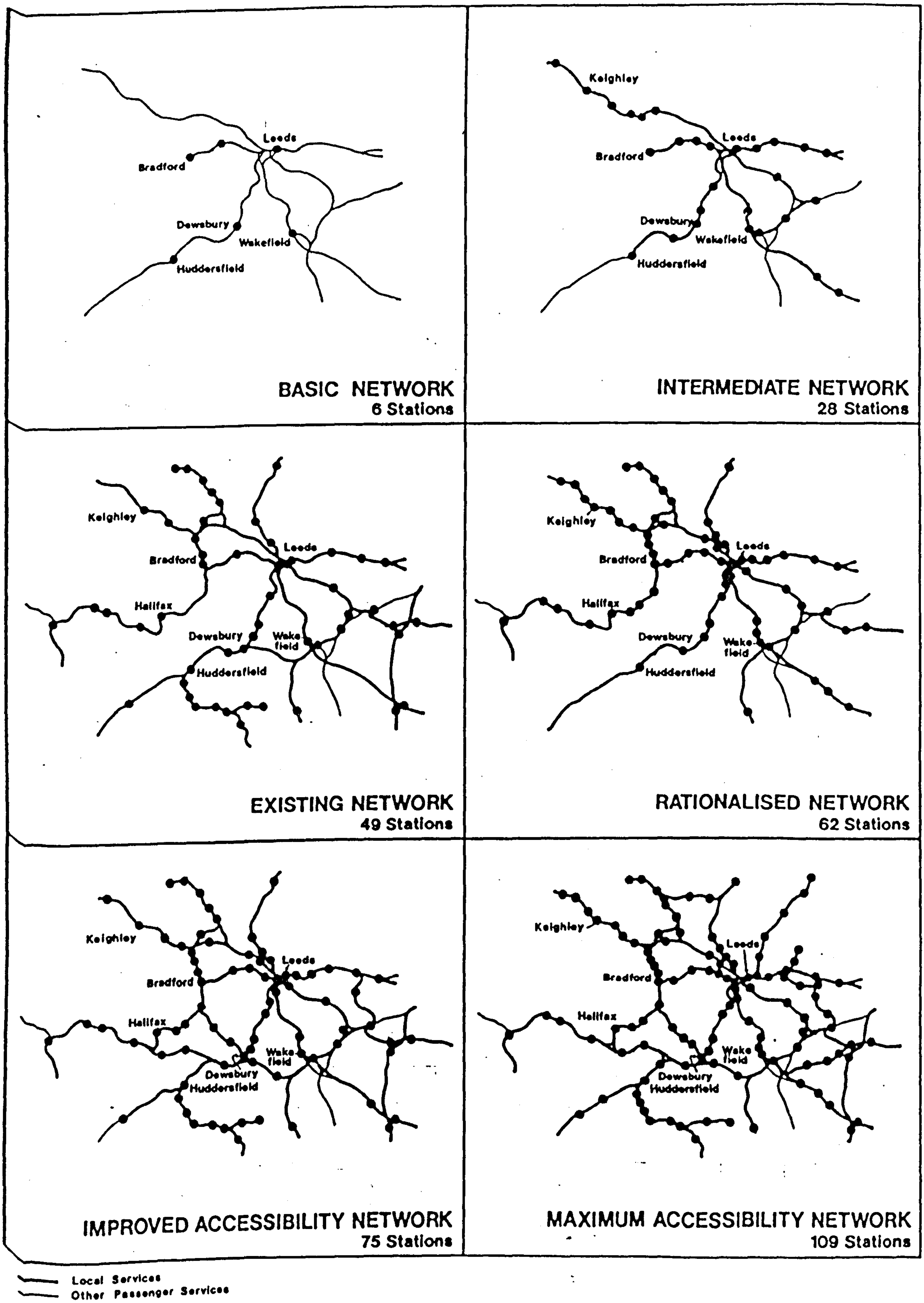


FIGURE 2.2 ALTERNATIVE RAIL NETWORKS
(Source: WYTCONSULT, 1977F)

Sheffield service via Barnsley, Wakefield and Normanton. Bus feeders were to be developed at eight stations, whilst frequencies for five services were to be improved. The total capital cost of the proposals was put at £1.3m; £0.5m for the Garforth rail spur, £0.4m for the reinstatement of a third track from Neville Hill to Cross Gates and £0.4m for the development of new stations.

(x) The findings of WYTCONSULT were used to develop an overall plan for transport in West Yorkshire. This began to be defined with the publication of the Structure Plan (WYMCC, 1978A) which helped to define the objectives of the decision-making body for Public Transport, the Passenger Transport Authority (PTA). Maximum use was to be made of existing infrastructure, whilst upgrading of Public Transport was implicit in policy statement T1.

"similar standards of accessibility to various facilities will be provided for both public and private transport journeys"

Utilisation of spare rail capacity was seen as leading to time and fuel savings and hence policy T6 stated:

"local rail passenger services will be generally maintained except where the actual cost becomes unacceptably high in relation to any benefits and numbers of passengers carried, or the economies of providing an alternative are favourable".

Provision for a new station development programme was outlined in policy T9:

"Provision will be made for additional staffed and unstaffed stations on the local rail network where operational constraints allow, sufficient demand exists and the net marginal cost would not impose a significant extra deficit on the service".

Financial support (through Section 20 of the 1968 Transport Act) was to be given to all local services except the Huddersfield to Wakefield, York to Sheffield and Huddersfield to Clayton West services, with adequate Public Transport provision being provided where BR services were withdrawn (policy T7).

A number of proposals were modified following public participation (WYMCC, 1978B). In particular the Leeds-Goole service was to be retained (rather than terminating at Castleford), the singling of the Huddersfield to

Penistone service and rerouteing via Barnsley was proposed whilst the Garforth spur was to terminate at Ninelands Lane.

It should be noted that these policies were only one strand of a range of planning proposals that included the concentration of new housing and industry at a small number of large sites, improvements to the highway network and the development of parking restraint in the central areas of Leeds and Bradford (WYMCC, 1980).

(xi) These policies were further developed in order to produce the Public Transport Plan (PTP) which included input from the administrative body and main operator of Public Transport in the County, the PTE (WYPTE, 1979). The PTP supported the reinstatement of a South Garforth spur, the opening of new stations, the use of disused railways, the identification of corridors in need of LRT (for example York Road), the relocation of existing stations, the provision of additional station car parking and the development of bus feeder services. The PTP also emphasised the need for extra track provision at the western end of Leeds City station, the modernisation of routes in certain areas, the need for standardisation of all fares, the extension of prepayment to a target 80% of fares and the improvement of information display systems at railway stations.

It is evident that all the eleven rail studies discussed above involve a new station element but there is little concensus on either the number or location of new stations that might be developed. Thus Table 2.5 shows that anything between plus 60 and minus 43 stations have been considered. Moreover, Table 2.6 shows that in comparing the sites proposed by seven studies, only two sites (Saltaire and Wrenthorpe) were always put forward.

Although we have seen that the objectives for rail transport in West Yorkshire have been stated in both the Structure Plan and the Public Transport Plan, Table 2.6 indicates that there has been no clear search procedure for the identification of new station sites. It is thus evident that a more scientific approach is required and hence in Chapter Nine a "quasi-logical" search procedure will be developed in order to identify a number of potential new station sites in addition to the six new stations opened between 1982 and 84. At this point it might be noted that non-scientific approaches have included:

		Definite openings	Possible openings	Closures	Net balance
WYTCONSULT (1976C)	- Leeds Short Term reinstatement test	-	25	2	23
WYTCONSULT (1977B)	- Leeds Medium Term provisional plan	-	15	7	8
WYTCONSULT (1977D)	- Bradford Medium Term Improved Accessibility	-	9	-	9
WYTCONSULT (1977E)	- Bradford Medium Term Provisional Plan	-	15	7	8
	- Maximum Accessibility Network	-	60	-	60
	- Improved Accessibility Network	-	26	-	26
WYTCONSULT (1977F)	- Rationalised Network	-	22	11	11
	- Intermediate Network	-	8	29	-21
	- Basic Network	-	-	43	-43
WYTCONSULT (1976A)	- Interim Preferred Network		27	12	15
WYTCONSULT (1977G)	- Recommended Network	9	11	11	9
WYMCC 1980	- Structure Plan	8	-	8	0
WYPTE 1979	- PTP	-	29	-	29
WYPTE 1984A	- Transportation Committee	9	8	-	17
Railway Development Society 1984		-	19	-	19
Transport 2000 1983		-	34	-	34

TABLE 2.5 STATION IMPLICATIONS OF SELECTED WEST YORKSHIRE RAIL STUDIES

WYTCONSULT (1976A) Interim Preferred Network	WYTCONSULT (1977G) Recommended Network	WYMCC (1980) Structure Plan	WYMCC (1982) TPP	WYPTE (1984A) Transportation Committee	WYPTE (1979) PTP	Transport 2000 (1983)
Crosshills	Parish Church	Parish Church	Crossflatts	Deighton	Ardsley	Parish Church
Steeton	Green Lane	Green Lane	Slaithwaite	Fitzwilliam	Bramley	Osmondthorpe
Utley	South Garforth	Ninelands Lane	Bramley	Crossflatts	Charlestown	Green Lane
Crossflatts	Hemsworth	Stanningley	Saltaire	Slaithwaite	Copley	Ninelands Lane
Nab Wood	Fitzwilliam	Saltaire	Lightcliffe	Bramley	Cottingley	South Garfoth
Frizinghall	Wrenthorpe	Wrenthorpe	Hawksworth	Saltaire	Crofton	Hunslet
Thackley	Stanningley	Fitzwilliam	Low Moor	Lightcliffe	Crossflatts	Methley
S. Horsforth	Crossflatts	Hemsworth	Wrenthorpe	Hawksworth	Deighton	Beeston
Kirkstall	Saltaire		Silsden	E. Garforth	Fitzwilliam	Outwood
Hawksworth			Crosshills	Gamble Hill	Frizinghall	Wrenthorpe
Marsh Lane	Steeton		Kildwick	Cottingley	Hawksworth	Fitzwilliam
Osmondthorpe	Low Moor			Frizinghall	Hemsworth	Hemsworth
South Garforth	Ninelands Lane			Low Moor	Horbury Bridge	Cottingley
Armley	Hawksworth			Methley	Kildwick	Deighton
Bramley	Deighton			Milnsbridge	Kirkstall	Milnsbridge
Laisterdyke	Milnsbridge			Stanningley	Parish Church	Slaithwaite
Hunslet	Slaithwaite			Wrenthorpe	Lightcliffe	Thornhill
Beeston	Outwood				Low Moor	Horbury
Cottingley	Osmondthorpe				Luddendenfoot	Berry Brow
Methley	Methley				Methley	Brighouse
Outwood	Bramley				Milnsbridge	Elland
Wrenthorpe					Osmondthorpe	Lightcliffe
Fitzwilliam					Outwood	Low Moor
Lightcliffe					Saltaire	Laisterdyke
Brighouse					Slaithwaite	Stanningley
Elland					Stanningley	Bramley
Stanningley					Steeton	Armley
					Thornhill	Frizinghall
					Wrenthorpe	Saltaire
						S. Horsforth
						Crossflatts
						Hawksworth
						Steeton
						Kirkstall

TABLE 2.6 NEW STATIONS PROPOSED BY SELECTED WEST YORKSHIRE RAIL STUDIES

(i) Sites that are deemed self explanatory. For example Pelaw and Kingston Park stations were opened in Tyne and Wear because of new housing development and Cathays station in Cardiff was opened to take advantage of new employment opportunities. There are few, if any, such sites in West Yorkshire.

(ii) Sites that are chosen due to political pressure. This may have played some role in the choice of Fitzwilliam and Deighton and has also meant that the PTE has attempted to seek sites that are distributed amongst the five districts in the County.

(iii) Sites that are chosen due to consumer pressure. This may have affected, to an extent, the choice of Saltaire and Slaithwaite where local community groups had campaigned for station re-openings.

2.2.4 Opportunities for Change - the Range of Options

A new station programme was just one of a number of possible improvements that could be made to the West Yorkshire rail network. The range of options also included:

(i) Improve frequency of existing services. For example in May 1979 the Leeds/Bradford to Ilkley/Keighley and the Leeds to Marsden services were upgraded. Surveys carried out by WYPTE in December 1979 indicated that patronage on the Bradford to Keighley service had doubled whilst abstraction from bus was very low (12%). We have reason to believe that this low figure may be due to factors other than the service improvements although the source data and report has not been made available to us.

(ii) Change the fare scale and structure. An early example of innovative ticketing was the Bullseye ticket, the effects of which were analysed by Emerson (1979). Table 2.7 shows that since 1981 a host of fare schemes have been introduced and this has resulted in a marked reduction in the real price of travel by Public Transport. It will be shown in later chapters that this has made time-series modelling, in particular, difficult.

(iii) Improve interchange. An example is given by Peat, Marwick, Mitchell and Co. (1975) who studied eight feeder bus routes and improved parking provision. These changes were predicted to lead to an increase of between 500 and 1000 commuters a day, with increased bus operating costs (at 1974 prices) of £86,000 pa, rail operating costs of £46,000 pa and capital costs of up to £90,000. Interestingly this report observed that:

Month/Year	Details
5/81	Trial off peak schemes for Elderley and Disabled (E & D) and children - bus only
7/81	Off-peak scheme formally introduced - bus only. Maximum fare Sundays 30p (E & D, children 15p). Mondays to Fridays 09.30-15.00 and 18.00 onwards 15p (E & D, children only)
8/81	Off-peak scheme extended to Saturdays for E & D, children - bus only
10/81	Off-peak scheme extended to Monday to Fridays 09.30-15.00 and 18.00 onwards for adults - bus only
5/82	Off-peak scheme extended to Saturdays after 18.00 for adults - bus only.
6/82	60p maximum off-peak fare introduced for rail (30p child, E & D)
4/83	Off-peak schemes introduced all day Saturdays for adults
7/83	Public Transport adult fares rounded to nearest 10 pence. Introduction of saverstrip (12 trips for the price of 10) on bus only.
1/84	Saverstrip introduced on 3 rail lines
4/84	Where normal rail fare 60 pence or less, off-peak fare reduced to 30 pence (15 pence for child, E & D). Saverstrip introduced on rest of rail network
7/84	Free travel on Sundays for E & D
1/85	Metrocard reduced in price (Weekly £5.50, Monthly £20, Annual £200) and extended to rail
4/85	Job seeker introduced. Allows ½ fare for 4 weeks at a cost of £1.50 - bus only
5/85	Introduction of Transtrip (bus only)
8/85	"Fare Deal 1" E & D free travel extended to Mon-Sat after 18.00. Day Rover prices halved (child 50 pence, Adult £1, Family £2) Saver Strip - 10% price reduction
9/85	"Fare Deal 2". Metrocard prices down - Monthly £18, Annual £180, Quarterly (£47) and Student (£40) Metrocards introduced.
9/85	"Fare Deal 3" E & D free travel extended to all off-peak periods
1/86	"Fare Deal 4". Off-peak rail fare for normal journeys costing 70p reduced to 30p. Job Seeker extended to rail. Metrocard weekly price reduced to £5. Off-peak Weekly Metrocard introduced (£2.50).

TABLE 2.7 DETAILS OF MAIN TICKETING INITIATIVES IN WEST YORKSHIRE 1981-1986

"Except where new dense housing developments have taken place adjacent to the rail route, the justification for any new (or reopened) local station is likely to depend upon effective interchange arrangements to extend the catchment area beyond the immediate walking distance of the station".

(iv) Refurbishment of stations. The main example of this was the new station at Bradford Exchange as part of the Interchange project, at a total estimated cost of £3m (1975 prices).

(v) Marketing. An early example was in June 1972 when a publicity campaign costing £30,000 was launched to promote the West Riding's rail services. It should be noted that the launching of new stations involved some marketing, including the distribution of publicity leaflets and free travel vouchers to all households within the catchment area of a new station.

(vi) Introduction of new rolling stock. Initially a report by WYPTE and BR (1982) favoured the replacement of two-thirds of the existing DMU fleet with class 141 lightweight units, although subsequently a combination of class 143 and 151 units has been favoured. The initial option was assumed to lead to a 31% reduction in the average deficit over 25 years (or around £1.7 million pa at 1981 prices) as well as allowing a smaller fleet and greater flexibility in operations.

(vii) Modified services. An example is the Huddersfield to Sheffield service which has been rerouted via Barnsley since 1983 (allowing a new station at Silkstone Common).

(viii) New services. These might involve reinstated track, such as the South Garforth spur favoured by the Structure Plan, or the use of currently freight-only lines such as the Brighouse link or the Spen Valley line. Such services would involve a number of new stations but they will not be considered in our study because their large capital requirements make them unlikely, as do the increases in fixed costs (especially if a prime user costing approach is used) and operating costs accrueing to the PTE.

2.3 CURRENT PROSPECTS OF URBAN PASSENGER RAILWAYS OUTSIDE WEST YORKSHIRE

2.3.1 The Economics of Urban Railway Investment

We shall show in the following sections that, as in West Yorkshire, there has been a great deal of interest in the rest of the U.K. and the world in investing in urban railways. In this section we shall briefly examine some of the economic reasons why there has been such interest (and in doing so we shall draw heavily from Mackie (1977)). Suppose we were to compare a rail option (for a conventional service with halts at two mile intervals, averaging 30 mph including stops) with three types of bus service averaging 30 mph (which would require express operation and possibly reserved track), 20 mph and 10 mph for a 10 mile urban corridor, operating 18 hours a day with a two hour peak in each direction on weekdays. The costs of provision of such a service might be as shown by Figure 2.3. The main differences between rail and bus costs are:

(i) Rail is capable of scale economics as it is possible to vary the mix and length of trains.

(ii) Rail costs include provision, operation and maintenance of infrastructure. Rail will thus be favoured where there is high absolute volume, a high peak ratio, low road speed and a requirement for rail infrastructure for other services (e.g. inter-city, freight). If track costs can be ignored because infrastructure of the required capacity would be provided in any case for other traffic (which is the case, to some extent, for all but the Ilkley line in West Yorkshire) then rail services may be worth providing at even relatively low corridor flows of 400,000 pa (about 1500 a day) in each direction. However if track costs are wholly specific to the service in question, then in order for a rail service to be justified vis-a-vis a 20 mph bus service flows of around 1.4 million pa (about 4700 a day) would be needed and this figure would have to increase substantially in order to match bus operations at 30 mph

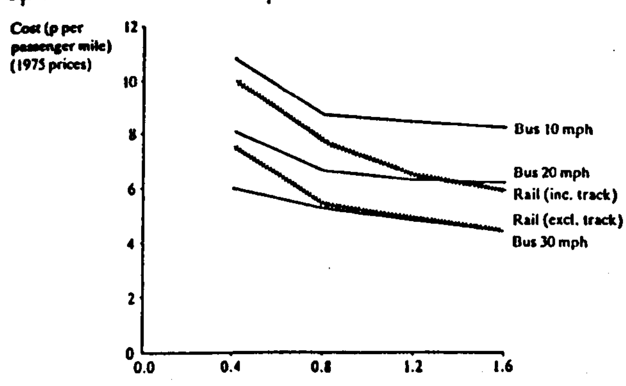


Figure 2.3 Bus and Rail Total Costs

Includes both operating and user costs. Assumes 50% of traffic in the peak

In our study it is assumed that the case for passenger rail services in West Yorkshire has been proven. We are merely examining the incremental effects of opening new stations on existing services.

2.3.2 Recent Urban Railway Investment in the U.K.

Following the Beeching report in 1963 the trend in the U.K. was towards dis-investment in urban railways, with cities such as Birmingham, Bristol and Edinburgh loosing much of their suburban network. However by the mid 1970s the emphasis had shifted from line closures and rationalisation to investment in new lines and services. These improvements were concentrated in five main centres (see Preston, 1987 for further details).

(i) In London and the South East there have been a number of major new investments of which the most important is possibly the Docklands Light Railway (DLR) with 18 new stations (Clarke and Cotton, 1983). However due to the large size of its rail network, the high proportion of underground operation, the high passenger volumes and the high amount of road congestion, for our purposes, experience on the London rail network may be considered largely irrelevant.

(ii) In Glasgow major schemes completed in 1979 were the cross-city Argylle line (which included five new stations) and the modernised underground. The effect of these schemes was studied by the Glasgow Rail Impact Study (Martin and Voorhees et al. 1982).

(iii) In Merseyside the "link and loop" schemes connecting four central area stations and allowing the reopening of the Garston branch were completed in 1978 with a total of six new stations. The impact of these schemes has been assessed by Merseyside P.T.E. and BR (1981).

(iv) In Tyneside there has been a major investment in the Tyne and Wear Metro, an LRT system that makes use of 41 km of former BR suburban railway, 4 km of new passenger line and 12.8 km of totally new line and includes 15 new stations. The system was opened in stages between 1980 and 1983 and its effects analysed by the Metro Monitoring and Development Study (TRRL et al. 1986).

(v) In Birmingham the main investment has been in the "Cross City line" which includes three new stations and vastly improved service frequencies. The effects of this scheme and other rail improvements in the West Midlands are discussed by Haywood and Blackledge (1982).

Major rail investments have been proposed for other UK cities but, as yet, have failed to be developed. These cities include Manchester (see Jones, 1984), Sheffield (Coventry, 1977) and Bristol (Cottrell, 1980), whilst a number of cities have put forward LRT proposals similar to those for Leeds (Kompfner, 1979).

Table 2.8 summarises some of the main features of the four major urban rail schemes carried out in provincial U.K. in the late 1970s (see also Walmsley, 1982). A number of points can be made:

(i) Investment in rail generally generates substantial increases in ridership, especially from optional travellers. Where city centre access was improved (Glasgow and Merseyside) patronage on affected services increased by about a quarter but where a new or vastly improved service was introduced (Tyneside and Birmingham) patronage increased four fold.

(ii) Most of the new traffic comes from bus (or other superseded Public Transport modes) although rail does attract some car users.

(iii) Interchange (especially for the Tyne and Wear Metro) is important but integration of bus and rail, even prior to the 1985 Transport Act, was difficult to achieve.

(iv) Although all four schemes make use of existing infrastructure they involve high capital costs ranging from the meso-scale (£12 million in the case of the West Midlands) to the macro-scale (£280 million for the Tyne and Wear Metro). By contrast the West Yorkshire new station programme, invisaged as involving a maximum of £1.5 million over a five year period (WYMCC, 1982), may be considered a micro-scale investment.

(v) Urban rail improvements appear to have limited effect on land use. Thus we tend to take the view of Hall and Hass-Klau (1985), rather than Knight and Trygg (1977), as the former state that:

> "the process of urban growth and decline have deep and subtle causes. Transit investment is in large measure irrelevant to these processes, though it may affect some of them at the margin."

If macro-scale investments are incapable of producing noticeable development effects then it is highly improbable that any such effects will be produced by micro-scale investments.

	Out-turn capital cost	Operating cost p.a.	New Route km	New Stations	Increased Patronage	Increased Revenue (£m) pa
Merseyside - "Link and Loop"	46.6	1.9	14.0	5	up 27%	2.6
Glasgow - "Argylle line" and	36	6.3	5.0	5	up 27%	3.5
Underground modernisation	59					
Tyne and Wear - "Metro"	280	14.9 (1983/4)	16.8	15	up 400% (P.T. up 2%)	-
West Midlands - "Cross City line"	12.3	-	-	3	up 400%	5.8

TABLE 2.8 RAIL INVESTMENT IN FOUR PROVINCIAL CONURBATIONS

Source: Walmsley, 1982

(vi) General conclusions on economic performance are difficult to draw but it does seem that most schemes fail to yield the 7% financial return normally required by central government to test public sector investments.

2.3.3 New Station Programmes in the U.K.

Table 2.9 shows some features of the main urban rail networks in the U.K. in the early 1980s. In terms of station provision it can be seen that there are wide variations. For example in terms of mean inter-station spacing the figure varies from just over 1 km in Belfast to almost 6 km in South Yorkshire (West Yorkshire 4.4 km). Similarly in Greater Glasgow there are 7.7 stations per 100,000 population compared to 1.9 in South Yorkshire (West Yorkshire 2.5), whilst there are 10.9 stations per 100 square kilometres in Merseyside compared to only 0.9 in Nottingham and Derbyshire (West Yorkshire 2.4). In terms of these three variables it can be seen that overall only the networks in Nottingham/Derbyshire, Edinburgh and South Yorkshire perform worse than the West Yorkshire network.

Partly because of these inconsistencies new stations have been opened throughout the U.K. Thus Table 2.10 shows that some 123 stations have been opened on publicly owned railways in Great Britain since 1970. Two distinct periods can be identified. Firstly, up to 1973 closures were dominant, reflecting the tail end of the rationalisation process implemented after Beeching. Secondly, since 1974 openings have outnumbered closures (excepting 1983). This reflects the importance of the 1974 Railway Act, which prevented major line closures, whilst added impetus was provided by the 1981 (Speller) Amendment to the 1962 Transport Act, which provided for experimental opening of stations. The period 1974 to 1977 may be seen as one of stability, with the first major batch of new stations being related to the four provincial investment schemes discussed in the previous section. Since then this momentum has been maintained by new station schemes, particularly in West Yorkshire, Greater Manchester and Scotland.

Table 2.11 indicates that new station programmes are likely to continue in the immediate future as up to 57 are under construction or at the planning stage with the majority being in Greater London (20 including DLR),

	1981 population (m)	Area (km²)	Stations	Network km[6]	Interstation spacing (km)	Stations/ Population (in 100,000)	Stations/ km² (in 100)
Avon	0.92	1346	25	82	3.3	2.7	1.9
South and Mid Glamorgan	0.90	1435	50	99	2.0	5.6	3.5
Cleveland/ S. Durham[1]	0.86	2346	28	80	2.9	3.3	1.2
Nottingham/ Derby[2]	1.72	4254	40	175	4.4	2.3	0.9
West Midlands	2.62	899	62	105	1.7	2.4	6.9
Merseyside	1.50	652	71	110	2.2	4.7	10.9
Greater Manchester	2.57	1287	106	197	1.9	4.1	8.2
Tyne and Wear	1.36	540	45	68	1.6	3.3	8.3
South Yorkshire	1.29	1561	25	142	5.7	1.9	1.6
Greater Glasgow[3]	2.33	7040	179	416	2.3	7.7	2.5
Belfast[4]	0.65	1318	33	36	1.1	5.1	2.5
West Yorkshire	2.00	2039	49	214	4.4	2.5	2.4
Edinburgh[5]	0.77	2306	25	110	4.4	3.2	1.1

TABLE 2.9 STATION PROVISION IN UK URBAN AREAS

Source: Baker 1984, Census 1981

1 South Durham defined as Darlington, Sedgefield, Teesdale and Wear Valley Districts

2 Excludes South Derbyshire and High Peak Districts

3 Includes all Strathclyde except Argyle and Bute District

4 Includes Carrickfergus, Castlereagh, Larne, Lisburn, Newtonabbey, North Down and Belfast districts

5 Includes Lothian region and Dunfermline and Kirkcaldy districts of Fife

6 Estimate derived using curvimeter - includes routes used by regular passenger services only.

	Open	Closed	Net balance
1970	3	75	-72
1971	5	13	- 8
1972	4	33	-29
1973	5	9	- 4
1974	1	-	+ 1
1975	5	4	+ 1
1976	9	8	+ 1
1977	1	1	0
1978	16	1	+15
1979	7	3	+ 4
1980	9	-	+ 9
1981	8	-	+ 8
1982	10	-	+10
1983	8	9	- 1
1984	9	1	+ 8
1985	23	5	+18

TABLE 2.10 NEW STATION OPENINGS IN BRITAIN 1970-85

Source: Modern Railways

	Under construction/ Planning Stage	Proposed by Local Authorities
England-Met. Countries	33	42
England-Shire Counties	4	45
Wales	14	5
Scotland	6	13
TOTAL	57	105

TABLE 2.11 PLANNED OR PROPOSED NEW STATIONS IN BRITAIN

Source: Roberts 1985A and B, 1986 A and B

Greater Manchester (3) and South Yorkshire (4) in England, South Glamorgan (6) and Mid Glamorgan (7) in Wales and Lothian (4) and Strathclyde (2) in Scotland. Moreover, an additional 100 plus new station sites are under consideration (Roberts, 1985A and B, 1986A and B) with a large proportion being in the English shire counties, such as Leicestershire and Somerset. Other proposals include those of Transport 2000 (1983) which as an alternative to the Serpell report put forward an option T which included 100 new stations to be opened by 1992 and the Railway Development Society (1984) which has put forward 410 possible new stations.

The stations opened in Britain over recent years might be placed into a number of categories. (It should be noted that these are not mutually exclusive).

(i) Stations related to a new transport system, such as the stations on the Tyne and Wear Metro or on the Dockland Light Railway.

(ii) Stations related to a new service, such as the new stations on the "Cross City line South" in the West Midlands, or Sinfin North, Sinfin Central and Peartree in Derby.

(iii) Inter-city Parkway Station such as Bristol Parkway (opened 1972), Alfreton and Mansfield Parkway (1973) and Birmingham International (1976) (see Prideaux, 1983).

(iv) Stations related to New Town development such as Basildon (1974), Newton Aycliffe (1979) and Milton Keynes Central (1982).

(v) Stations related to improved central area rail links, for example those opened due to the Argyle line and "Link and Loop" schemes.

(vi) Stations on existing services serving local transport needs. These may be:

(a) Manned, for example Watton-at-Stone opened in 1982.

(b) Unmanned, but related to a major employment centre for example IBM Halt and BSC Redcar (both opened 1978).

(c) Unmanned and related to mainly residential areas.

Of these new stations we are only really interested in category vi(c). In West Yorkshire stations of this type are distinguished by:

(i) their simple design, consisting of wooden platforms and stone shelters,

(ii) access ramps to allow disabled usage,

(iii) use of existing road bridges to enable access to both platforms,

(iv) small scale parking provision (none at Deighton and Saltaire)

(v) and consequently their low cost, with outturn costs being between £72,000 to £125,000 and much lower than previous BR estimates (which were based on more complicated designs). Thus typically a new station in West Yorkshire, at 1984 prices, costs £100,000, although this may be reduced to £60,000 where there is a single track or increased to over £150,000 if there are local access problems.

Similar stations have been developed in Greater Manchester (Modern Railways, 1985), in South Yorkshire (Silkstone Common, opened 1983) and Tyne and Wear (Dunston, opened 1984), whilst in Scotland use has been made of concrete section platforms, for example at Auchinleck and Kilmaurs.

2.3.4 Urban Railway Investment in the Rest of the World

Compared to the U.K. a number of commentators have noted that other nations, especially in continental Europe, are more willing to invest in their urban rail networks (Hellewell, 1977, European Conference of Ministers of Transport, 1980). Thus, leaving aside the definitional problem of what constitutes an urban rail system (see White, 1976, page 75 for clarification), Table 2.12 shows that in the mid 1960s there were an estimated 41 cities possessing or constructing rapid transit systems, of which only two (London and Glasgow) were in the U.K. By 1980 this worldwide figure had increased to 93 (with three U.K. cities, Birmingham, Liverpool and Newcastle upon Tyne defined as gaining rapid transit), whilst a further 24 schemes were proposed of which only one (Manchester) was in the U.K.

Thus, although the absolute role of urban rail transport has increased worldwide, we have seen that growth in the U.K. has been limited. Moreover it is clear that the cities gaining rapid transit tend to be large, with dense traffic flows, hence allowing macro-scale investments in heavy rapid transit systems (U-bahn) or high technology solutions such as the VAL system in Lilles or the UTDC system in Vancouver (see Preston, 1987 for further details). Such investments are clearly unlikely in West Yorkshire, although less capital intensive schemes such as upgrading conventional suburban railways (S-bahn) or developing LRT might be feasible in the future. Recent experience of urban railways in the rest of the world is thus largely irrelevant to this study. There are, however, some studies of micro-scale investments that are of interest; for example new station studies at Arlington Park, Illinois (Von Ehrenrook, 1975) and Silver Spring, Maryland (Winch and Smith, 1980).

	Europe (including the USSR)	(of which U.K.)	North America	Elsewhere	TOTAL
mid 60s[1]	28	(2)	9	4	41
1980[2]	54	(4)	15	24	93
proposed[2]	3	(1)	11	11	24

TABLE 2.12 CITIES HAVING OR CONSTRUCTING RAPID TRANSIT SYSTEMS

Sources: [1] "Rapid Transit Round the World". International Railway Journal. 1967.

[2] "Jane's World Railways and Rapid Transit Systems 1980-81" 22nd Edition. Editor: P. Goldsack. 1981.

2.4 CONCLUSIONS

In this Chapter our system of interest has been more precisely defined as existing and potential unmanned, low cost, new stations on existing local rail services in West Yorkshire. In placing our system of interest in context we were able to identify a number of objectives of the new station programme. Firstly, our historical analysis showed that the application of the Beeching report had been uneven in West Yorkshire leading to inconsistencies in station provision which might be remedied by a programme of reopenings. Furthermore, it was shown that West Yorkshire's local rail system suffered from a number of problems. To an extent new stations might reduce these problems by increasing accessibility to the network and thus increasing usage, improving utilisation of assets (as no extra operating resources were envisaged) and reducing the financial deficit. Thus although the Structure Plan talked of new stations provided "no significant extra deficit" was involved, in practice a positive (or at worst neutral) financial return was sought. Moreover, it is clear from the Structure and Public Transport Plans that the County Council, the PTA and the PTE were each concerned, to varying degrees, with social issues such as accessibility, safety and the environment and hence a new station programme may be seen as being consistent with a policy of maximising social benefits (or as a proxy passenger mileage) within a budget constraint.

Secondly, our geographical analysis has shown that West Yorkshire's passenger railways perform worse than those in most metropolitan counties, whilst the provision of stations is inferior to many U.K. urban areas. Again a new station programme of the type we are interested in may, in part, help to remedy this. Moreover, we have shown that the West Yorkshire new station programme should not be viewed in isolation. There are several similar schemes throughout the U.K. and they have implications for the application of this research. Also it is clear that station numbers (and location) on passenger railways in Britain are dynamic rather than static. Thus between 1975 and 1985 there was a significant (if unperceived) change in that over 100 new stations have opened on passenger railways in Britain, compared to only around 30 closures.

Thirdly, our analysis of scale has produced a number of findings. It has been shown that, as in West Yorkshire, there has been much recent investment in urban railways in both the rest of the U.K. and the rest of the world but this has tended to involve much more capital intensive schemes than the type we are interested in. Hence these projects are only of limited relevance to us, although it is interesting to note that even such macro-scale investments fail to produce significant development effects. It is, however, clear that there are a number of other changes to the local rail network that might be pursued. In particular changes in fare structure and introduction of new rolling stock are two options that have been developed concurrently with new stations, and may at times produce confounding effects.

Lastly, despite the objectives outlined above, our study of past rail studies in West Yorkshire failed to detect any agreement on the scale of the new station programme to be adopted in West Yorkshire. Hence in Chapter Nine we shall attempt to develop a search procedure that identifies potential sites whilst in Chapter Ten we shall determine which of these potential sites are consistent with the PTE's objectives.

CHAPTER 3 DEFINING THE METHODLOGY

3.1 INTRODUCTION

Having defined the aims of this research in Chapter One and our system of interest in Chapter Two, in this Chapter we shall determine the study's methodology. In the next section we shall examine previous work assessing new station location and/or usage and determine which approaches might be applicable to this study. In particular it will be shown that we need to develop statistical models that can predict patronage at potential new stations. Thus in section 3.3 we shall examine the range of modelling approaches available and make preliminary steps towards choosing the appropriate model forms, although our final decision is dependent on data availability, which will be examined in Chapter Four. The forecasts provided by the statistical models will act as the main input to the evaluation stage, the framework of which is discussed in section 3.4.

3.2 PREVIOUS WORK

Four strands of previous work have been identified as being of relevance to this study. Firstly, the provision of new stations might be thought of in terms of determining an optimal inter-station spacing pattern, an issue that has been well covered in the literature. Secondly, a number of studies have made use of simple aggregate modelling approaches to determine new station usage, including extrapolation from historical data, trip rate models, simple regression models and elasticity models. Thirdly, there have been a number of mode choice studies that have included a forecasting component. Fourthly, there have been several studies that have used market research to collect stated intentions data which may then be used to forecast new station usage. In the rest of this section the relevance of these four approaches will be examined.

3.2.1 Optimal Station Spacing

As was shown in Table 2.9 (and is further illustrated by Figure 3.1) interstation spacing varies from urban area to urban area and from rail system to rail system. This has led to interest in the problem of determining the optimal interstation spacing pattern, originating with Vuchic's (1966) thesis. Emphasis has been placed on the trade-off between in-vehicle time and access time and a number of different objectives have been studied. These include minimising travel time (Vuchic and Newell, 1968), maximising passengers (Vuchic, 1969), minimising travel cost (Hurdle and Wirasinghe, 1980) and maximising the travel time advantage of rapid transit over alternative modes (Fukayama, 1983).

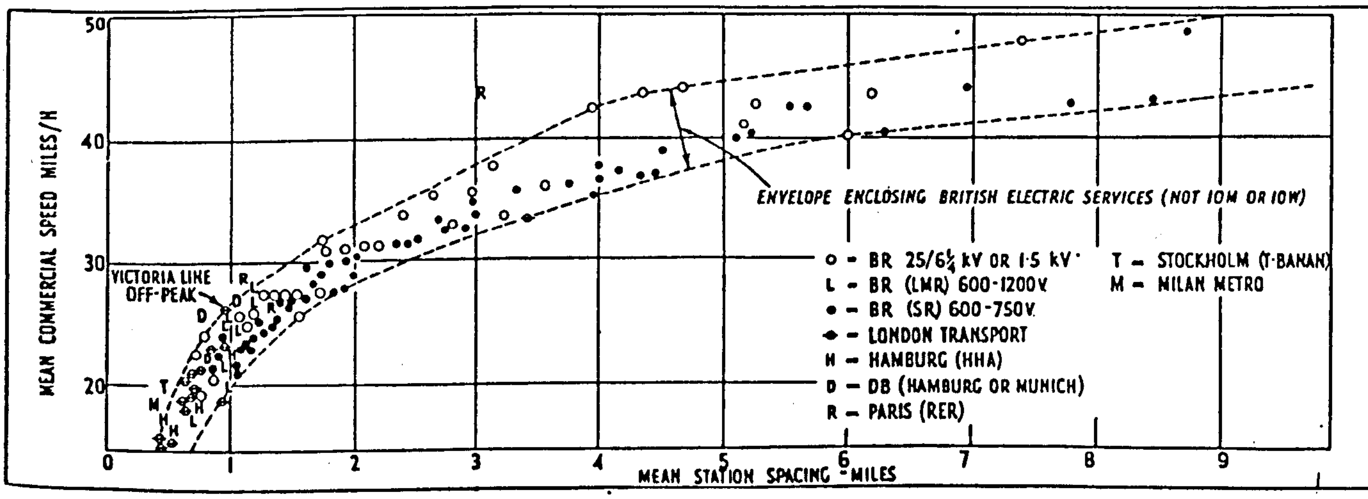


FIGURE 3.1 Mean Commercial Speeds on Typical Urban Systems in Relation to Station Spacing Source: Hughes, 1971.

However, in order for the problem to be tractable, a number of assumptions have to be made which, for our purposes, are unrealistic. These include:

(i) The case for rapid transit has already been proven and hence costs of construction and operation need not be taken into account.

(ii) There is uniform population distribution with constant trip rates throughout. In reality population density is likely to decrease with distance from the Central Business District (CBD).

(iii) There is cumulative loading towards the CBD. For many lines in West Yorkshire this is not totally the case. This assumption has been relaxed by Fukuyama (1983).

(iv) Consideration is normally only made of one access mode and only one alternative mode.

(v) Access is made by regular geometric routes, normally grid iron although Wirasinghe et al. (1977) have investigated the effect of a radial-circumferential network.

To further illustrate this approach Vuchic's (1969) graphical solution method has been applied by this study to a West Yorkshire situation, as shown by Figure 3.2. In this example we are studying the continuously built-up 10 mile Leeds-Bradford corridor (L) where train offers a cruising speed (V) of 70 mph (although in this corridor this is unlikely to be achieved), whilst the average speed of access modes (V_a) is 7 mph and for competing modes (V_c) is 40 mph. Time lost due to station stops (T_l) is estimated to be 60 seconds per stop. The interstation spacing (S**) at which average train speed is equal to V_c may then be defined as:

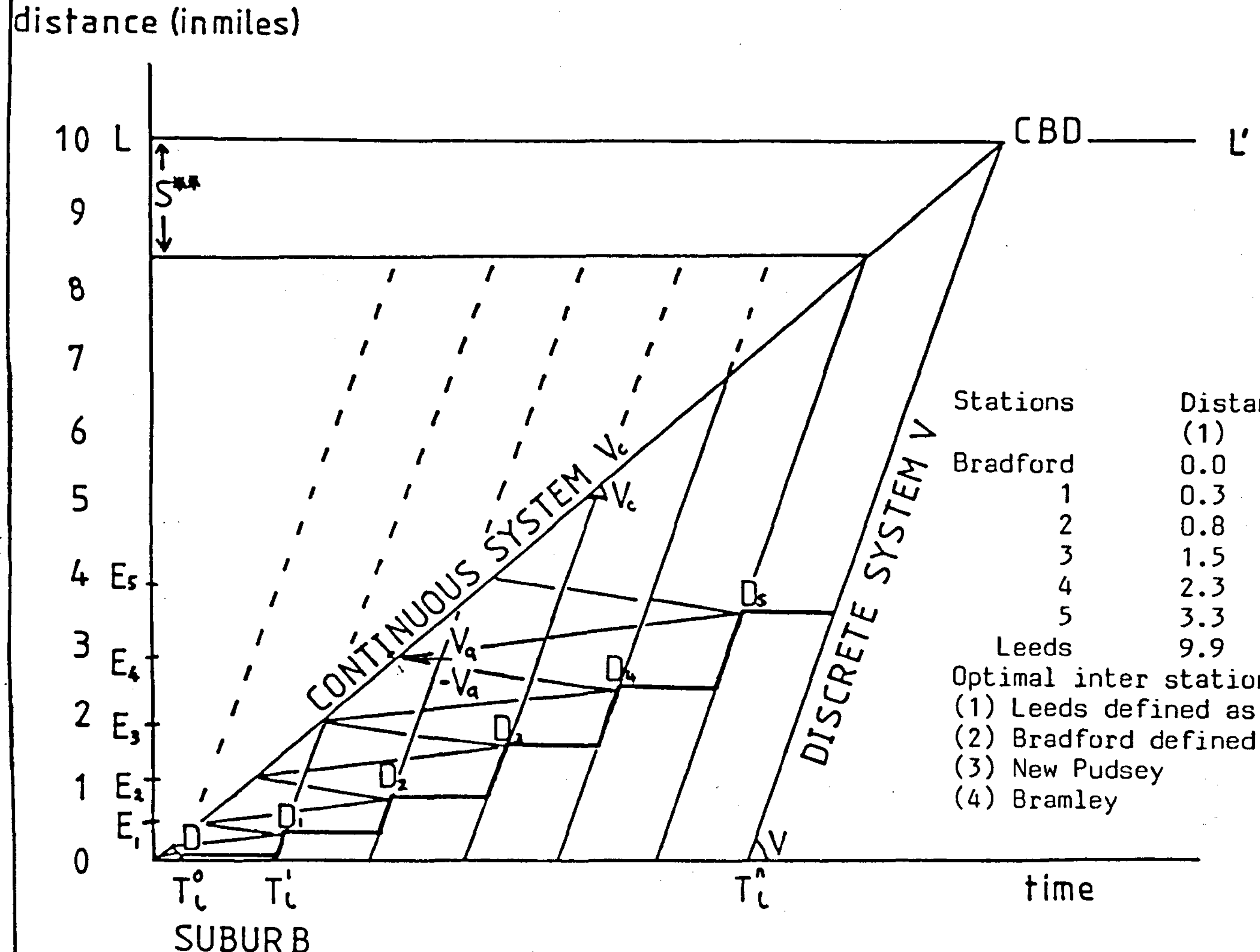
$$\frac{\nu}{B_c - B} \text{ where } \nu = \frac{T_l V_a}{2}, \; B_c = \tfrac{1}{2}(1 + \frac{V_a}{V_c}) \text{ and } B = \tfrac{1}{2}(1 + \frac{V_a}{V})$$

ν is the access distance passed during T_l and
B, B_c are the passenger shed factors for train and competing mode respectively.

A solution is determined by finding the intersection of LL' and a line of slope V_c from the origin. This intersection is designated the CBD. From the CBD a line with slope V is drawn back to the horizontal axis, giving point T_l^n. Distances representing the incremental time loss per station are marked off from T_l^n to T_l^o. Thus the number of intermediate stations (in this case five) that the rail system can have and still offer travel times to the CBD equal to or less than the competing system is determined. In order to determine the location of the stations, lines with slope V are drawn from all T_l points up to the V_c line. A line with slope V_a is drawn from the origin to line T_l^o, intersecting at station D. From D a line with slope $-V_a$ is then drawn intersecting V_c at passenger shed line E, from E a line with slope V_a is drawn intersecting T_l' at D_1. This procedure is continued, giving successive D and E points until point E_5 (in this case) is reached.

Aside from problems of accurately measuring V_a, V_c and T_l the above example illustrates a number of problems with the optimal station spacing approach:

(i) No consideration is made of historical and geographical constraints on station location.



Stations	Distance (miles) (1)	(2)	Average	Actual average
Bradford	0.0	0.0	0.0	
1	0.3	6.6	3.45	
2	0.8	7.3	4.05	4.15 (3)
3	1.5	8.4	4.95	
4	2.3	9.1	5.70	
5	3.3	9.6	6.45	6.4 (4)
Leeds	9.9	9.9	9.9	

Optimal inter station spacing on the Leeds Bradford line
(1) Leeds defined as CBD, Bradford as suburb.
(2) Bradford defined as CBD, Leeds as suburb.
(3) New Pudsey
(4) Bramley

FIGURE 3.2 OPTIMAL STATION SPACING - A GRAPHICAL SOLUTION FOR THE LEEDS-BRADFORD CORRIDOR
(after Vuchic, 1969)

(ii) If cumulative loading is assumed then interstation spacing increases as the CBD (defined initially as Leeds) is approached. In reality there are important intermediate (e.g. Bradford to New Pudsey) and reverse flows (Leeds to Bradford). As a result optimal station spacing is likely to involve a concentration of station sites in the middle stretch of the route rather than at either end.

(iii) Leeds to Bradford is the only continuously built-up rail served corridor in West Yorkshire and even in this case population densities are not uniform.

These assumptions might be relaxed by adopting complicated numerical approaches but it is clear that such an approach only has limited relevance to our study. However, the graphical solution illustrated in Figure 3.2 might have a role in identifying potential sites in continuously built-up rail corridors.

3.2.2 Simple Demand Models

A number of simple demand models have been used to forecast station usage. These include.

(i) Historical extrapolation. Demand at reopened stations is simply calculated from pre-closure usage figures, adjusted to take into account changes in catchment area population in the intervening period. This method is used by Roberts (1986C) in an attempt to estimate demand at reopened stations on the Settle-Carlisle line. This approach cannot be applied to totally new station sites (such as Crossflatts) nor where the level of rail service has changed significantly. The emphasis on constant trip rates over time means this approach has limited relevance to this study.

(ii) Trip rate models. These simply express demand at a new station as a function of population within the catchment area, based on the experience of existing stations with a similar rail service and population characteristics. An example is given by the work of Jeanes and Lesley (1984) who estimated usage at six new stations on the Wrexham-Bidston line based on the following relationship estimated for existing stations on the line:

$$N = -643 + 0.961\,H \qquad R^2 = 0.55 \qquad \text{Equation 3.1}$$

N = number of journeys per week

H = number of households within 800m of a station

Similar approaches have been used elsewhere, for example by Central Regional Council (1980) in estimating demand at Bridge of Allan and BR Southern Region in forecasting usage at Winnersh Triangle.

Although trip rate models fail to take into account a number of factors they have advantages in terms of simplicity. Hence such a model will be developed in section 5.5.1 and its accuracy assessed in section 9.3.2.

(iii) Regression Models. Such models are similar to Equation 3.1 but take into account a greater number of factors. Examples in West Yorkshire include the WYTCONSULT (1975B) model which determined rail usage as a function of the population within one km of a station, the percentage of non car owning households within one km of a station, the distance to a major employment centre, the level of service in the direction of travel and journey time. Similarly, the PTE calibrated a regression model with 1979 data for 36 small town, suburban and rural stations, an example of which took the form:

$$D = -385.2 + 0.405\,H1^{*} + 0.230\,H2 + 10.053\,S + 6.194^{*\,+}\,J \qquad R^2 = 0.47$$

Equation 3.2

D = Daily usage, H1 = Number of households with 400m of a station, H2 = Number of households within 800m of a station, S = Service frequency, J = Journey time, * = insignificant at the 95% level and $^{+}$ = wrong sign.

A number of problems with the PTE model are apparent, including the large negative intercept, the low significance of some parameter values and the wrong sign of the journey time parameter value. Nonetheless regression models clearly have a role to play and thus in section 6.2.1 the WYTCONSULT and PTE models will be rerun and possible improvements examined.

Outside West Yorkshire a number of regression approaches for forecasting rail passenger demand have been developed (see British Railways Board, 1986 for more details). Early examples include the work of Evans (1969) and Tyler and Hassard's (1973) Model for Optimising the Network of

Inter City Activities (MONICA). This latter model was however only developed for London based inter-city flows and hence only has limited relevance to this study. A similar type of model has, though, been developed by White and Williams (1976) to predict 30 flows on the Reading-Tonbridge line as follows:

$$\log T_{ij} = 4.64 + 0.59 \log P_j P_i - 0.73 \log H_j - 0.09\, CH - 1.70 \log C_{ij} + 1.59 \log S_{ij}$$

Equation 3.3

where P_iP_j = the population of towns i and j
T_{ij} = the number of rail trips between i and j
H_j = the hierarchy level of the destination
CH = dummy variable of the availability of through services
C_{ij} = the average fare level and
S_{ij} = average speed.

The models discussed so far have been based on cross sectional data but a number of models based on time-series data have also been developed. Again the main emphasis has been on inter-city flows (see for example McLeod et al., 1980, Jones and Nichols, 1981) although there have been studies of commuter flows in London and the South East (see, for example, Glaister, 1983) and in Glasgow (Stark, 1981). It will be shown in Chapter Four that opportunities for time-series modelling in West Yorkshire are restricted due to the lack of consistent data sets.

(iv) Elasticity models. An example of such a model is MOIRA (Whitehead, 1981) which is the BR model designed to supersede MONICA and can be used to predict changes in passenger flows as a result of changes to the timetable, expressed as a level of service quality index Q. This model is, however, highly aggregate in nature and can only be used to examine inter-city and long distance provincial flows and hence has little relevance to West Yorkshire.

Another example of an elasticity model is the Parkway Access Model (PAM) developed by Steer, Davies and Gleave (1984) with reference to parkway stations at Iver, Hinksey, Patcham, Tiverton and Plymouth and a new station at Springfield (Essex). Elasticities were indirectly determined from market research that examined the trade-off between road access times, rail journey time and costs and then applied to the model structure shown by Figure 3.3. The PAM may be criticised for its use of constant access elasticities and has been modified so that access time elasticity is broadly proportional to access time by adopting a generalised cost approach (Transecon International, 1986). Nonetheless even the modified PAM has limited relevance to our study as emphasis is placed on inter-city rather than local trips, mechanised access modes rather than walk and on the effect on existing demand, but in West Yorkshire existing usage of rail for local trips will be effectively zero for many potential new station zones.

It should, however, be noted that in section 9.2.3 we shall develop an incremental approach that has some similarities with the elasticity models discussed above.

3.2.3 Mode Choice Models

So far in this section we have mainly examined models that determine T_{ijk}, that is the number of trips between i and j by mode k (rail). However, in forecasting demand for rail a number of studies have merely calculated P_k, the probability of choosing rail, given that T_{ij} is known. Models of this type vary greatly in their complexity. Simple approaches include those of Kingham (1976) who used diversion curves to estimate the effect of upgrading rail services in the South West Baltimore corridor or the regression model used by Leake (1971) to determine the rail and air shares of business travel. A more complex approach is provided by the mode split models developed in the West Midlands (WMCC, 1984) which consisted of binary logit models of the choice between bus and train for non car owning households and car and train for car owning households, calibrated with data collected using choice based sampling techniques. These models were then used to evaluate the effect of the Snow Hill scheme in Birmingham, linking the Shirley, Dorridge and Stourbridge lines and involving four new stations. The most complicated approaches we have identified are studies based on families of multinomial logit models. Such studies include the Urban Travel Demand Forecasting Project (McFadden et al. 1979), which included forecasts of the impact of Bay Area Rapid Transit, and the SIGMO study which included an assessment of railway investment options in Amsterdam (Ruhl et al. 1979).

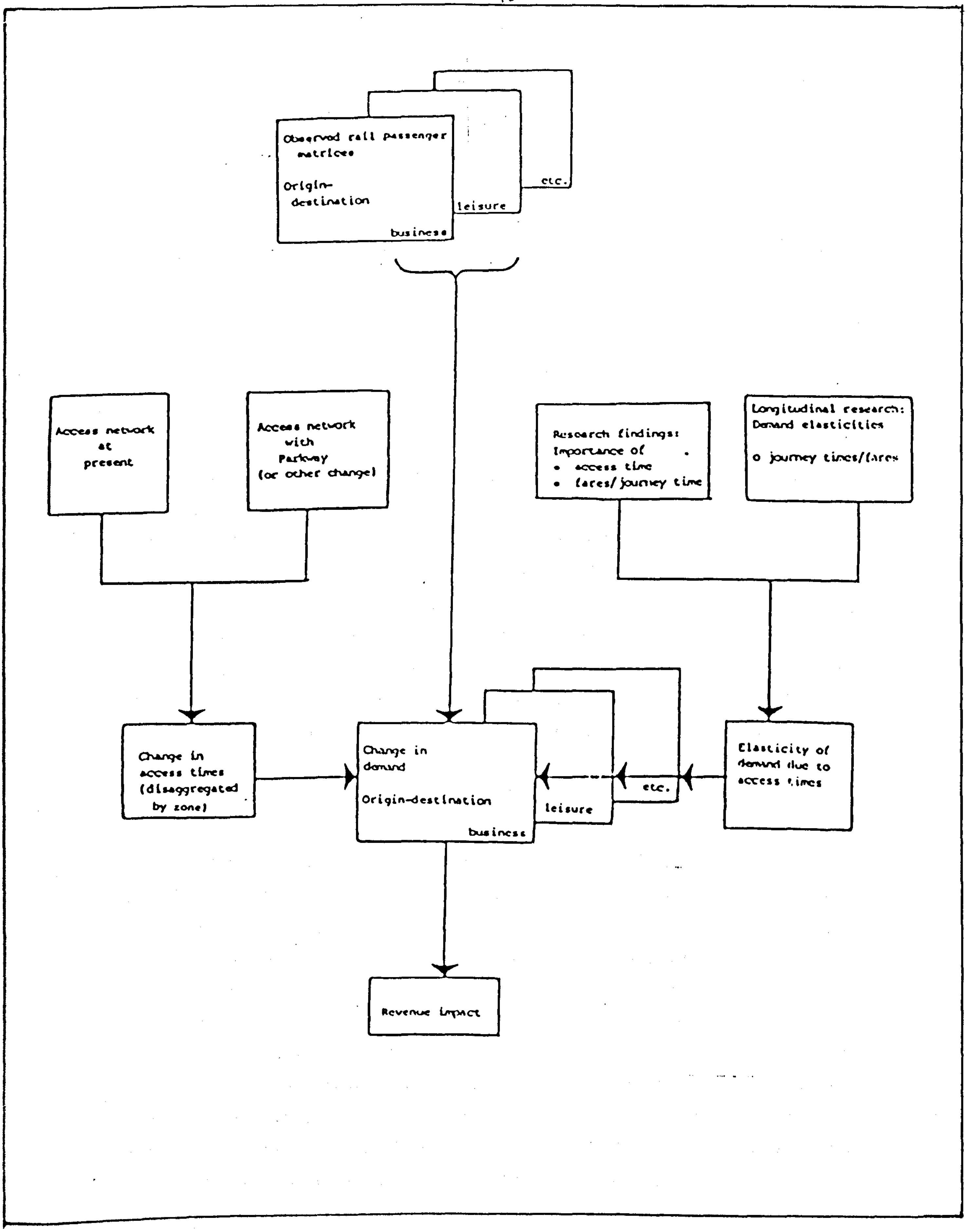


FIGURE 3.3 APPLICATION OF THE PARKWAY/ACCESS MODEL
Source: Steer, Davies and Gleave, 1984

Choice models have also been applied to the related problems of demand for feeder modes, for example Ortuzar's (1980A) study of mixed mode choice in the Garforth corridor of West Yorkshire and specific studies of park and ride (Mufti et al. 1977), kiss and ride (Demetsky and Korf, 1979) and bus feeders (Read, 1977), and station choice given that rail has been chosen (Desfor, 1975, Liou and Talvitie, 1975, Nitta and Mori, 1986, Harata and Ohta, 1986).

3.2.4 Stated Intentions Approaches

A further approach to forecasting demand at new station i is simply to use market research to ask "if a new station was opened at i, with level of service Q, how often and for what journeys would you use it?" This type of approach has been used by Hockenhull (1984) to forecast usage at South Wigston and by Farrington (1986) in a study of Newtonhill. However, work by Couture and Dooley (1981) has shown that such a simplistic approach will lead to an overestimate of usage due to non commitment bias, which in a case study of a new transit system in Danville, Illinois they showed resulted in a ratio of intended to actual users of three to one. Hockenhull's results appear to show a bias of a similar magnitude, even when assuming (after Heggie and Papoulias, 1976) that non respondents will be non users.

More sophisticated approaches have, though, been developed to examine hypothetical choice situations, which have been collectively termed stated preference (SP) techniques (see Benjamin and Sen, 1982, for a comparison of four such techniques). These techniques are in contrast to the revealed preference (RP) approaches, based on actual choices, used by the studies discussed in the previous subsection. SP approaches have been adopted in a number of practical studies both in the U.K. (Bradley et al. 1986, and the work of Steer, Davies and Gleave, 1984, already referred to) and elsewhere (Kocur et al. 1982, Bates and Roberts, 1983, Bovy and Bradley, 1985). In addition a number of studies have shown that SP models perform no worse than RP models and most often perform considerably better (Meyer et al. 1978, Louviere et al. 1981, Louviere and Hensher, 1982).

In order for an SP approach to have been adopted in this study it would have been necessary to launch a major market research initiative in order to collect a sufficient sample of individuals in West Yorkshire who would be prepared to make hypothetical trade-offs between their existing mode and rail. Such an initiative was beyond both the time and money budgets of this research project. In a subsequent study examining the demand for rail travel of a new service between Leicester and Burton on Trent, it is hoped to develop an SP approach to forecasting new station demand. The non-labour market research costs of such an approach were, however, shown to be in excess of £12,000, whilst data punching alone might involve in excess of 10 person weeks, thus confirming that resources are required that are beyond the scope of this study.

3.3 RANGE OF MODELLING APPROACHES

In Section 3.2 a number of relevant modelling approaches have been identified of which at least three, the trip rate model, the regression model and the mode choice model, were deemed worthy of further investigation in this study. This section has two main aims. Firstly, to set these existing models within the context of the range of transport demand modelling approaches available, thus allowing some of the terminology introduced in 3.2 to be more clearly defined. Secondly, by examining this range we hope to be able to suggest improvements on previous work, particularly in terms of providing a theoretical basis, hence allowing a preliminary choice of the type of models to be used, prior to discussion of data availability in Chapter Four.

3.3.1 Spectrum of Models Available

The range of modelling approaches identified by this study may best be thought of as a continuous spectrum, as Figure 3.4 shows.

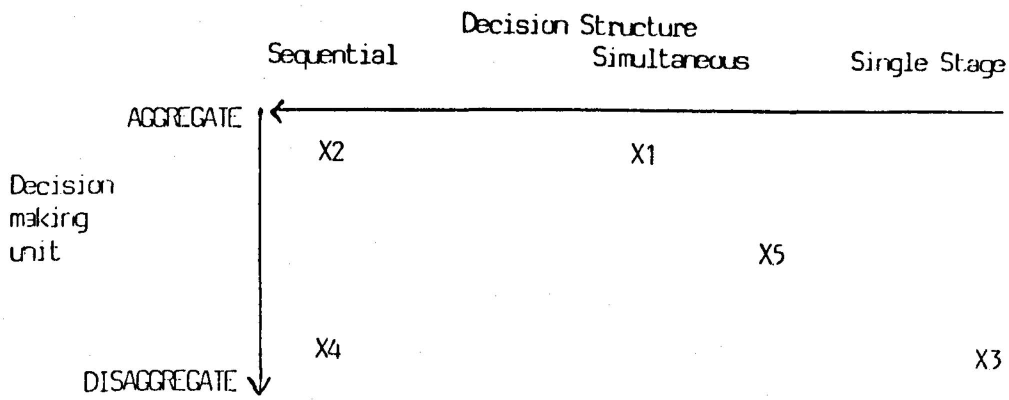


FIGURE 3.4 SPECTRUM OF MODELS AVAILABLE
X1 = Aggregate Simultaneous, X2 = Aggregate Sequential, X3 = Disaggregate Mode Split, X4 = Disaggregate Sequential, X5 = "Hybrid" approach

Two dimensions are identified as being relevant:

(i) The decision making unit. Aggregate models are based on grouped data, which in this study will normally be based on zones covering all, or part of, a station catchment area. Disaggregate models are based on information from actual decision making units such as individuals or households. However it should be noted that:

"a certain mystique has been created around differences between aggregate and disaggregate models... An aggregate model is a model estimated with a dependent variable which represents a group of observations, whilst a disaggregate model is estimated with an observation of a single occurrence. In disaggregate models some of the independent variables can be represented by aggregate data". (Richards and Ben Akiva, 1975).

(ii) The decision structure. Traditionally, as we shall see in the next section, transport modelling has been seen as consisting of four stages. The simplest approach might be to consider only one step (e.g. mode choice) in isolation. The next simplest approach might be to consider all four stages in one step, i.e. adopt a simultaneous structure. The most complicated approach might be to consider each stage in sequence, i.e. adopt a sequential structure. It is clear from Figure 3.4 that different degrees of complexity are involved with, of the models to be discussed in the next sections, the aggregate simultaneous model (X_1) being the least complex and the disaggregate sequential model (X_4) being the most complex. It is the purpose of this study, by examining the trade-off between complexity and accuracy, to determine the best point in this continuous spectrum for models that evaluate demand at new local stations.

3.3.2 Aggregate Models

In this section we shall examine the relevance of aggregate models of both sequential and simultaneous structure.

(i) Aggregate Sequential Models

The four step aggregate model might be considered the "first generation" of transport models having being developed as part of the Urban Transportation Planning Process (UTP) in North America in the early 1950s (Oi and Shuldiner, 1962) and by the early 60s was being used in Land-Use and

Transportation Studies (LUTS) in the U.K. (Lane, Powell and Prestwood-Smith, 1971). Models of this type normally consisted of the following four stages:

(a) Trip Generation. This involves the determination of the number of trips from zone i (production) and the number of trips to zone j (attraction). These trips can be disaggregated by purpose, whether home-based (residential) or non home-based, and by car availability. In West Yorkshire measures of trip attraction per unit of activity for residential, employment and educational uses have been developed by WYTCONSULT (1976D).

(b) Trip distribution. This involves the determination of a matrix of origin and destination trips, T_{ij}, given information on trip production, O_i, and trip attraction, D_j. The most common form for these models is the gravity model which takes the general form:

$$T_{ij} = A_i B_j O_i D_j f(C_{ij}) \qquad \text{Equation 3.4}$$

where A_i, B_j = balancing factors

$f(C_{ij})$ = impedance (deterrence) functions based on travel times and/or costs

Important issues include the form of the deterrence function and the nature of the balancing factors (Wilson, 1970) whilst improvements include the development of the generalised cost concept, the use of information theory (for example entropy maximisation as used by Wagon and Hawkins, 1970) and the development of links with utility theory.

(c) Mode Split models. At this stage trips T_{ij} are allocated to modes k. In fact mode split models can occur before the distribution stage, after the distribution stage or at the same time as the distribution stage. The first attempts to model modal shares were based on diversion curves of the type shown by Figure 3.5. This is based on the TRC model of Hill and Von Cube (1963) which considered measures of relative travel time, relative travel costs, relative excess travel time, income and trip purpose for five modes, producing up to 320 diversion curves. These were used to predict the modal shares of the Bloor-Danforth subway in Toronto, the Lindenwold line in Philadelphia and the Washington Metro. Building on this approach share models of the following general form have been developed (McLynn and Watkins, 1967).

$$\frac{T_{ijk}}{\sum_k T_{ijk}} = \frac{g\,(C_{ijk})}{\sum_k g\,(C_{ijk})} \qquad \text{Equation 3.5}$$

where $g\,(C_{ijk})$ = (for example) $e^{-\lambda C_{ijk}}$

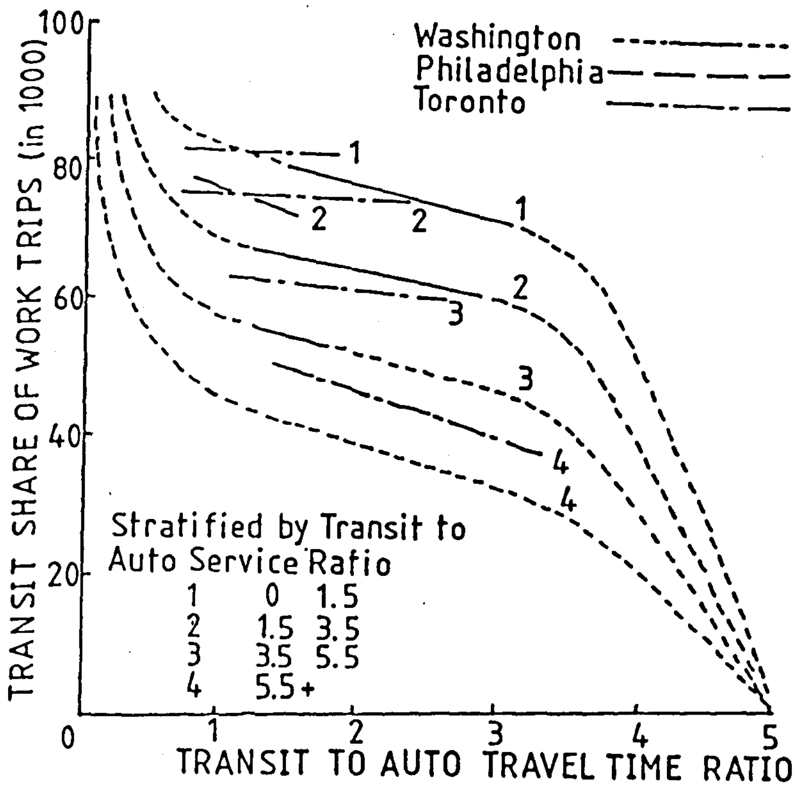


FIGURE 3.5 TRC MODEL - DIVERSION CURVES

(d) Assignment. At this stage trips T_{ijk} are allocated to routes m. Although Public Transport network assignment can be complex (see for example Lamb and Havers (1970)), in our case, where mode k = rail, very few major ij pairs have more than one route.

(ii) Aggregate Simultaneous Models

The aggregate simultaneous model considered in this section deal with the first three stages outlined above in order to determine T_{ijk}. At least two types of such models may be identified.

(a) Abstract mode models. Such an approach is based on the work of Lancaster (1966) on perfect substitutes and an example is given by Baumol and Quandt's (1968) model given by Equation 3.6.

$$T_{ijk} = a_o\, P_i^{a1}\, P_j^{a2}\, Y_i^{a3}\, Y_j^{a4}\, M_i^{a5}\, M_j^{a6}\, N_{ij}^{a7}\, f_1(H)\, f_2(C)\, f_3(D) \quad \text{Equation 3.6}$$

where

P = Population of i and j
Y = Average income of i and j
M = Measure of industrialisation of i and j
N = Number of modes serving i and j

H = Journey time
C = Journey cost
D = Service frequency

The three functions (f_1, f_2, f_3) took the form of ratios of particular mode characteristics and best mode characteristics. The model is thus not mode specific and in theory can be used to predict the performance of new modes, although in our case rail is not so much a new mode as an existing mode extended to new areas. However Crow et al. (1973) have shown that there are problems in applying models of this type to predicting usage of a new mode or drastically altered existing mode which is not "best" in any characteristic (which might be the case for rail given a coarse zoning system). In such a situation none of the other modes will be affected by the new or altered mode -a wholly unrealistic outcome.

(b) Direct demand models. These models typically are mode specific, disaggregated by purpose and tend to use generalised times and/or costs rather than separate level of service variables. A model of this form may be represented as follows (after Hensher, 1977).

$$T(ij/P_oM_o) = \alpha[s(i/P_o)\ \underline{A}\ (j/P_o)\ \underline{T}\ (iji/P_oM_o)\ \underline{C}\ (iji/P_oM_o)\ \underline{\underline{T}}\ (iji/P_\phi M_\phi)\ \underline{\underline{C}}\ (iji/P_\phi M_\phi)]$$

Equation 3.7

where

$T(ij/P_oM_o)$ = Number of round trips for mode o by purpose o,

$S(i/P_o)$ = vector of socio-economic characteristics for travellers in zone i,

$A(j/P_o)$ = vector of economic/land use characteristics in zone j,

$\underline{T}(iji/P_oM_o)$ = vector of travel time (travel costs) for round trip by mode o,
$\underline{C}(iji/P_oM_o)$

$\underline{\underline{T}}(iji/P_\phi M_\phi)$ = vector of travel times (travel costs) for each alternative mode
$\underline{\underline{C}}(iji/P_\phi M_\phi)$ (ϕ = 1,2 .... , N).

Models of this type normally take the form of multiple regression equations estimated by Ordinary Least Squares (OLS). They may be shown to have tentative links with utility theory (Beckmann and Golob, 1971). A model of this type may be affected by the identification problem of distinguishing cause and effect, also known as simultaneous equation bias. This is because when Public Transport trip rates are unusually high (low) due to an unmeasured attribute of the attractor, generator or transport system, then the system will have an unusually high (low) revenue and suppliers will respond by reducing (increasing) fares and increasing (reducing) service levels. Thus fares and service level will be correlated with the demand equation. Therefore, in addition to modelling demand, it is desirable to take into account the behaviour of suppliers (regulators and operators), using two stage least squares (TSLS) regression (Frankenna, 1978). This may be considered beyond the scope of this study, particularly as this problem will be limited at small-scale stations.

In discussing both the aggregate sequential and simultaneous models generalised cost has been mentioned and this concept requires further definition. Generalised cost is essentially a device that combines money and time costs of travel. In WYTCONSULT, 1976E, generalised cost for Public Transport was estimated as:

$$GC = IVT + \frac{a_2}{a_1}WK + \frac{a_3}{a_1}WT + \frac{a_4}{a_1}F + \frac{a_5}{a_1}X + \delta \qquad \text{Equation 3.8}$$

where

IVT	=	In Vehicle Time	F	=	Fare
WK	=	Walk Time	X	=	Transfer Time
WT	=	Wait Time	δ	=	Modal Penalty

with $a_2/a_1 = 1.7$, $a_3/a_1 = 2.3$, $a_4/a_1 = 0.515$ pence per minute (at May 1975 prices)

N.B. Generalised cost is often used as a generic term. Equation 3.8 should more accurately be referred to as generalised time.

Use will be made of the generalised cost concept in Chapter Six. Theoretically such an approach may be justified in that it brings together variables with certain common properties. These are:

(a) they are expended by travellers in definite quantities

(b) these quantities are capable of objective measurement

(c) expenditure itself does not give benefit

(d) there are possible alternative expenditures and thus, ceteris paribus, a traveller will prefer to expend a smaller quantity on a journey than a larger quantity and

(e) possibilities exist for travellers to trade-off between the various components.

A number of problems with such an approach have been identified, in particular during the debate between Grey and Goodwin (1978) on the generalised cost dilemma. For example, IVT parameters should be mode specific, in order to reflect comfort etc, waiting time should vary with mode reliability, different values should be determined for different journey purposes and different sizes of time savings, time and money components should have variable utility and values of working time should be based on observed behaviour rather than on economic theory.

Lastly in this section, it should be noted that a number of the models examined in 3.2.2 are simple examples of aggregate simultaneous models. For example the trip rate and PTE/WYTCONSULT regression models directly forecast station usage, T_{ik}, whilst the MONICA and White and Williams models, based loosely on gravity formulations, directly forecast rail flows, T_{ijk}.

3.3.3 Disaggregate Models

Disaggregate models may be described as the "second generation" of transport models and came into vogue in the early 1970s. The early application of disaggregate techniques in North America is documented by Spear (1977), whilst early case studies include those of Eindhoven (Richards and Ben Akiva, 1975) and Pittsburgh (Domencich and McFadden, 1975). Evolution of disaggregate techniques may be detected from successive conference reports (Stopher and Meyburg, 1976, Hensher and Stopher, 1979, University of Leeds, 1980, Stopher, Meyburg and Brog, 1981).

Disaggregate models may be applied to all stages of the UTP model, for example Dobson and McGarvey (1977) have studied their application to the trip generation stage, Burnett and Prestwood (1976) have analysed the trip distribution stage whilst McGillivray (1972) has examined modal split. Following our findings in 3.2.3 emphasis will be placed on mode split, especially as it will be shown in later sections that this stage has a simple choice structure compared to, say, trip distribution and as we will see in Chapter Four has the best data available.

Given the above it should be clear that disaggregate models, like aggregate models, may involve a variety of structures. These include:

(a) Independent sequential

$P(d, m : DM) = P(d : D)\,P(m : M)$ Equation 3.9

where $P(d, m : DM)$ = probability of choosing destination d and mode m out of the choice set of destination D and mode M

(b) Conditional sequential (or recursive)

$P(d, m : DM) = P(d : D)\,P(m : M_d)$ Equation 3.10

where M_d is the set of alternative modes to destination d

(c) Simultaneous

$P(d : D_m) = \text{Prob}\,[U_{d/m} \geq U_{d'/m}, \forall\, d' \in D_m]$ Equation 3.11

$P(m : M_d) = \text{Prob}\,[U_{m/d} \geq U_{m'/d}, \forall\, m' \in M_d]$ Equation 3.12

where $U_{d/m}$ = the utility from destination d given that mode m is chosen
$U_{m/d}$ = the utility from mode m given that destination d is chosen and
D_m = the set of alternative destinations by mode m

These two conditional probabilities provide insufficient information to predict the joint probability which is directly estimated as:

$P(d,m : DM) = \text{Prob}\,[U_{dm} \geq U_{d'm'}, \forall d',m' \in DM]$ Equation 3.13

It has been argued that the simultaneous structure has theoretical and computational advantages (Richards and Ben Akiva, 1975, Adler and Ben Akiva, 1976). Empirical evidence is not clear. Ben Akiva (1975) in a study of mode split in Washington D.C. favours a simultaneous structure over a recursive one, whilst Liou and Talvitie (1975) claim that conditional sequential models are superior to simultaneous structures. However, theoretically at least, such arguments may be irrelevant as Williams (1977A) has shown that within a random utility theory interpretation of transport demand models, the joint structure is simply a restricted case of the more general sequential structure and there is no reason to argue that either represent a different decision making process.

Disaggregate mode split models are models of qualitative choice where the dependent variable is discrete rather than continuous and hence linear regression does not normally apply. Early mode split models (Warner, 1962, Quarmby, 1967) were based on discriminant analysis but subsequent tests have shown that relative coefficient values are markedly different and goodness of fit is statistically inferior compared to that achieved by logit and probit models (Stopher and Lavender, 1972).

Hence models more firmly based on random utility theory have been developed. These take the general form, for binary choice, of:

$$P_{i1} = \text{Prob}\,(RU_{i1} > RU_{i2}) \qquad \text{Equation 3.14}$$

where P_{i1} = the probability that individual i will choose mode 1,
RU = Random Utility, and

$$RU_{i1} = \sum_j B_{ij1} X_{ij1} + \varepsilon_{i1} \qquad \text{Equation 3.15}$$

where X_{ij1} = the value of jth relevant attribute (explanatory variable)
B_{ij1} = parameters to be estimated
ε_{i1} = an error term which introduces a stochastic element so to take into limited account unobservable aspects and omitted factors

Different models may be developed depending on the assumed distribution of the error term. The binary logit model is based on the cumulative logistic probability function (also referred to as the reciprocal exponential or Gumbel or Weibull (extreme value, Gnedenko distributions). The binary probit model is based on the cumulative normal distribution. The Arctan probability model is based on the Cauchy distribution. These models are defined by Equations 3.16, 3.17 and 3.18 respectively.

$$P_{i1} = \frac{\exp\,(U_{i1})}{\exp(U_{i1})+\exp(U_{i2})} = \frac{1}{1+\exp(U_{i2}-U_{i1})} \qquad \text{Equation 3.16}$$

$$P_{i1} = \Phi(U_{i1} - U_{i2}) \qquad \text{Equation 3.17}$$

where Φ = the standard cumulative normal distribution

$$P_{i1} = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} (U_{i2} - U_{i1})$$ Equation 3.18

(see also Figure 3.6)

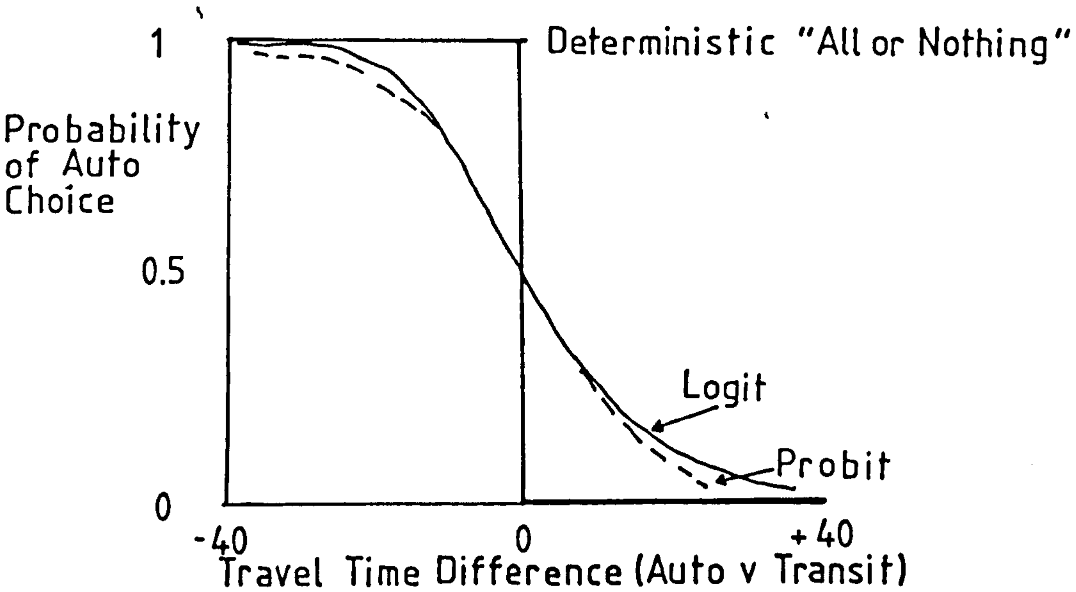


FIGURE 3.6 COMPARISON OF PROBIT, LOGIT AND DETERMINISTIC "ALL OR NOTHING" CURVES

It can be shown that the arctan model approaches the asymptotes less rapidly than the probit and logit models and is thus rarely used.

Given "strict utility" theory (Luce and Suppes, 1965) the logit model may be extended to multiple choice situations giving the Multinomial Logit Model (MNL), for all modes m as:

$$P_{i1} = \frac{\exp (U_{i1})}{\sum_{m} \exp (U_{im})}$$ Equation 3.19

MNL models are usually based on the following assumptions:

(i) The distribution of utilities for each mode follows an identical and independent distribution (IID) generated from a Weibull distribution and thus does not allow for taste variation i.e. $B_{ijm} = B_{\cdot jm}$ for all i.

(ii) The utility function is linear in the parameters with an additive disturbance (LPAD). This is known as the separability of decisions property.

(iii) The probability that a particular alternative is chosen has to be greater than zero. This is known as the positivity assumption (Ortuzar, 1980B).

Assumption (i) allows any number of alternatives to be studied (see, for example, Rassam et al. 1971) but also means that alternatives are perceived as totally distinct and independent, as the cross elasticity for demand for each mode with respect to another mode is uniform across all modes, and thus precludes the possibility of differential substitutability and complimentarity. This is known as the Independence of Irrelevant Alternatives axiom (IIA). This has important implications. Suppose that in a new station catchment area trips were split 60 : 40 between car and bus and then rail was introduced capturing 10% of the market, the IIA axiom would mean the same proportional decreases in car and bus usage resulting in a 54 : 36 ratio. However "a priori" beliefs (confirmed by the results of our market research in Chapter Five) suggest that new rail users are more likely to be abstracted from bus than car.

3.3.4 Alternatives to the Multinomial Logit Model

In order to alleviate the problems of the IIA axiom a number of alternatives to MNL models have been proposed. These include Multinomial Probit models (MNP), Extended Logit models (EL) and Hierarchical Logit (HL - also known as tree, nested or cascading logit).

(i) MNP models are derived by assuming multivariate normally distributed error (Daganzo, 1980). They have advantages in that alternatives can be dependent and taste variation is allowed. However, the Clark approximation used to estimate MNP models is inaccurate in some instances (Horowitz, 1985) and thus initially applications were limited to relatively "small" problems involving three or four options (we wish to model at least four options)

(Hausman and Wise, 1978). Hence MNP models have been rejected on computational grounds, because although there have been improvements in estimation procedures (Langdon, 1984, Dansie, 1985), there are still problems of availability of reliable software.

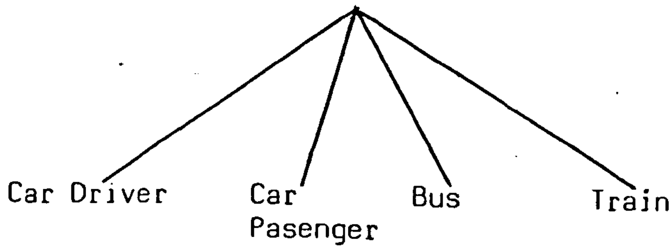
(ii) EL models may be thought of as extensions to the MNL and are based on variants of the Gumbel distribution and include the Fully Competitive model (McLynn, 1973), the Generalised Extreme Value model (McFadden, 1981), the CRA Hedonics model (Cardell and Reddy, 1977) and the DOGIT model of Gaudry and Dagenais (1979) so called because "it dodges the researcher's dilemma of choosing a priori between a format which commits to IIA restrictions and one that excludes them". Another example of an EL model is the cross correlated logit model (Williams, 1977A) which was developed directly from the HL but allowing a more comprehensive specification of correlations between alternatives. EL models have been used in only a limited number of studies and there is a practical problem in the lack of available software.

(iii) HL models. These are based on decision tree structures of the type shown by Figure 3.7. In order to keep this structure simple consideration will only be made of the mode split stage. For our purposes, the mode choice decision process might be divided into two stages: i = the choice between private and public transport and j = the choice amongst public transport alternatives. Then the perceived utility to individual k may be written as (based on work by Gunn, 1980):

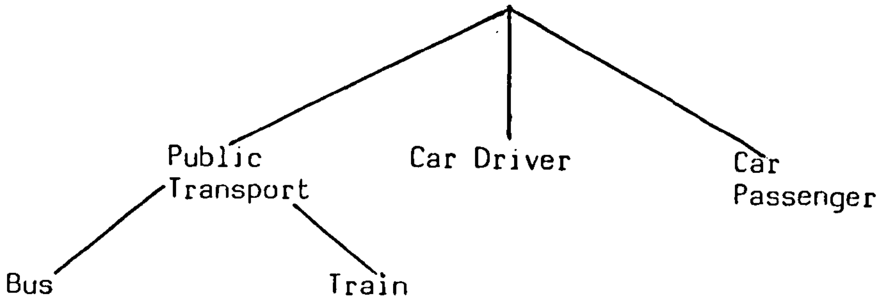
$$U^k(i,j) = U_i + U_j + U_{ij} + \varepsilon^k{}_i + \varepsilon^k{}_{ij} \qquad \text{Equation 3.20}$$

Assuming the ε terms are IID for each individual and that var (ε_i) and var (ε_j) are zero this results in the MNL. If either var (ε_i) or var (ε_j) is not zero, then there will be correlation between the U^k (i,j) terms. For example if var (ε_j) = 0 but var (ε_i) $\neq$ 0 then individuals with high values of U^k (Public Transport, Bus) will also tend to have high values of U^k (Public Transport, Train). Assuming no common measureable attributes between bus and train, an HL model may be developed with a two level choice structure as follows:

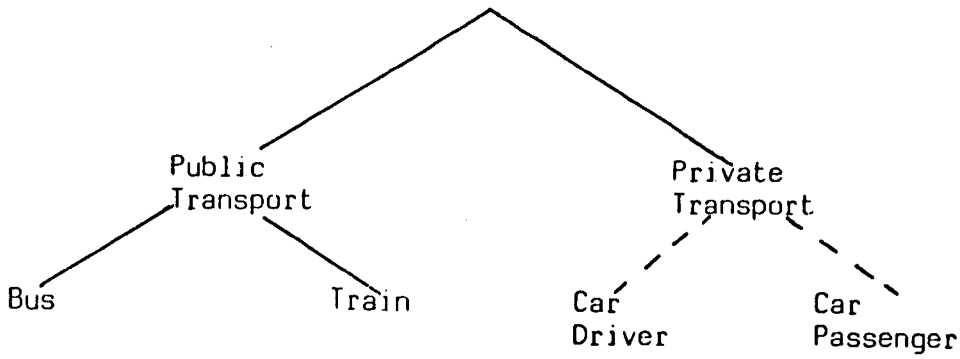
$$P_{ij} = \left[\frac{\exp(U_i + \phi U_{i*})}{\sum_i \exp(U_i + \phi U_{i*})} \right] \cdot \left[\frac{\exp(U_j)}{\sum_j \exp(U_j)} \right] \qquad \text{Equation 3.21}$$



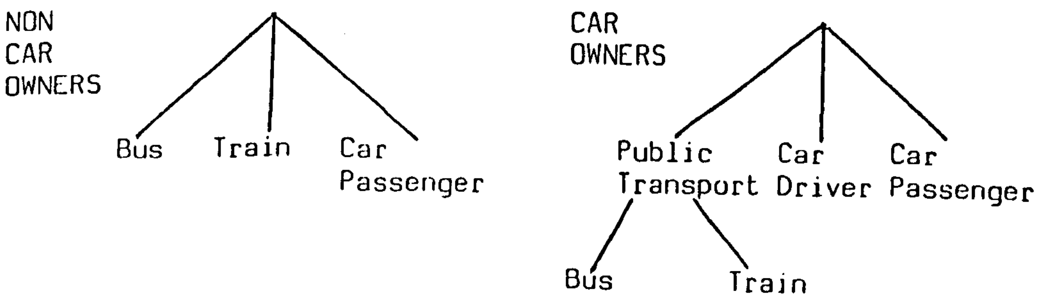
(A) MULTINOMIAL STRUCTURE



(i)



(ii)
(B) HIERARCHICAL STRUCTURES - SINGLE MARKET



(C) MIXED STRUCTURES - SEGMENTED MARKET

FIGURE 3.7 DECISION TREE STRUCTURES TO BE ANALYSED where $U_{i*} = \ln\{\sum_j \exp(U_j)\}$ (Equation 3.22) and is referred to as the expected maximum utility (EMU) or logsum and subscript i may be thought of as denoting the upper split (or nest) and subscript j the lower split (or nest).

An HL model of this form has a built-in diagnostic test in that $0 < \phi \leq 1$ (see papers by Williams and Senior, 1977A,B, 1978, Williams, 1977A,B and Williams and Ortuzar, 1979 for derivation and examples of violation). An infringement would lead to "pathological" results as:

(a) if $\phi < 0$ an improvement in the utility of one member of the nest can decrease the choice probability of that nest,

(b) if $\phi = 0$ an improvement in the utility of one or all members of the nest will not change the choice probability of the nest and

(c) if $\phi > 1$ an improvement in the utility of one option may result in the improvement of choice probabilities of other options in the nest.

HL models may be estimated using maximum likelihood (ML) in two ways:

(a) Indirect procedure (or heuristic or bottom-up calibration) which involves the sequential calibration of MNL models (through the use of composite utility functions (Equation 3.22)). Examples of this method are included in Ortuzar (1980A) and the Zuidvleugel (Van Zwam and Daly, 1981) and SIGMO (Project Bureau, 1977) studies.

(b) Direct procedure (or full information likelihood maximisation) whereby Equation 3.21 is treated as a functional expression, the parameters of which are estimated using ML. The theoretical background to such a calibration procedure is given by Coslett (1981), with practical examples including the work of Daly and Zachary (1978), Matzoros (1982) and Small and Brownstone (1982).

In comparing the two methods it can be said that the indirect procedure has advantages in terms of computational and mathematical simplicity but has a number of statistical disadvantages. Thus, although there is good consistency between the parameter values of the two methods, the standard errors with direct estimation are smaller (Matzoros, 1982). It has thus been suggested that the two methods may be used in a complimentary sense, with the fast indirect procedure producing initial estimates that may be further refined by using a direct procedure. However, due to lack of readily available software, emphasis in this study was placed on indirect estimation.

A survey of software showed that a large number of relevant packages existed, of which Basic Logit (BLOGIT) was chosen because it was well documented (Crittle and Johnson, 1980), it was readily available, having been used in the Value of Time Study (MVA et al., 1986) and has been applied widely (Hensher and Johnson, 1981).

3.3.5 Comparisons of Model Forms

In this section we shall briefly examine some of the theoretical advantages of the models discussed in the previous sections. More practical considerations, especially concerning data availability, will be examined in Chapter Four.

Aggregate models have advantages in terms of simplicity, particularly if a simultaneous rather than a sequential structure is adopted. They, however, have a number of disadvantages:

(i) The population is assumed to be located at one point in space. This can only hold if sub-groups are homogeneous.

(ii) Aggregation before model construction clouds underlying behavioural relationships. As a result aggregate models are non behavioural and their policy orientation is limited (McFadden, 1974).

(iii) An approach based on averages does not necessarily represent an individual consumer's behaviour nor the average behaviour of a group under a variety of conditions. This leads to problems of ecological fallacy (de Neufville and Stafford, 1971).

(iv) Aggregate models incur a great degree of information loss and hence data inefficiency. This is due to the fact that, using an oft-quoted example, up to 80% of variation of socio-economic variables may be intrazonal and only 20% interzonal (Fleet and Robertson, 1968).

(v) Aggregate models, particularly the traditional LUTS models, may thus be shown to provide poor forecasts (Williams and Senior, 1977A,B) and lack transferability.

Conversely disaggregate models have a number of "starting advantages" (Daly, 1982):

(i) Data, almost by definition, are based on decision making units. This gives models a firm behavioural basis and their links with utility theory can be easily demonstrated.

(ii) Disaggregate models can establish the importance of factors that exhibit greater intra- than inter-zonal variance, in particular walk and wait time which are important components of Public Transport trips (Gunn, 1982).

(iii) Disaggregate models typically adopt a probabilistic (stochastic) choice process rather than a deterministic stance.

(iv) Disaggregate models are statistically efficient, requiring as few as 100 to 200 observations for calibration (Spear, 1976) and they can also be used to model minor alternatives. For example Ortuzar (1980A) successfully modelled rail trips, kiss and ride and park and ride with 21, 21 and 10 observations choosing the respective alternatives.

(v) Other claims made on behalf of disaggregate models are that they are robust (Atherton and Ben Akiva, 1976), are less likely to be biased by correlation, are consistent with consumer surplus measures (Williams, 1977A) and can be used at a variety of planning levels (Ben Akiva, 1975).

However it should be noted that disaggregate models were greeted in the early 70s with a certain amount of euphoria and thus Richards (1980) commented "Disaggregate models have been presented if not as a Utopia certainly as a panacea". By the late 1970s disaggregate models were being treated with a greater degree of healthy cynicism (Williams and Ortuzar, 1980). For example a disaggregate mode split model based on HL may be affected by problems of data intensity (discussed in 4.3), transferability (discussed in 8.2), aggregation (discussed in 8.3) and model structure (Sobel (1980) has shown that if there are four alternatives in a choice set, the number of possible decision tree structures can be as high as 24). In addition the differences between aggregate and disaggregate models should not be overstressed. For example the aggregate mode split model, given by Equation 3.5, can be shown to be similar to the MNL model (Richards and Ben Akiva, 1975). Moreover, aggregate and disaggregate parameters may be combined as for example, with the GENMOD set of models developed for Amsterdam (LeClercq, 1980) and those developed for West Yorkshire (Hartley and Ortuzar, 1979).

3.4 EVALUATION FRAMEWORK

The results from the statistical models of the type described in 3.3 will be used to provide forecasts of new station demand and hence, given information on destination choice, ticket type and previous mode used, we will be able to determine the gross and net revenue to the PTE. This, as we will see, will be a key input to our evaluation framework. In addition we shall make use of the analysis of PTA/PTE objectives in 2.4 in order to define our evaluation measures and determine decision rules on whether a new station should be opened. In the rest of this section we shall examine existing evaluation procedures for railway investment and the ways in which Cost Benefit Analysis (CBA) may be used. We shall then itemise the costs and benefits we wish to study and discuss ways in which they may be presented. We shall then be able to determine the evaluation framework we wish to use in this study.

3.4.1 Existing Evaluation Procedures for Railway Investment

In keeping with the British Railways Board's commercial remit, financial appraisal is generally used rather than social CBA, for example in studies of resignalling, electrification and rolling stock replacement schemes (Leitch, 1977). Such appraisals are normally based on the difference between minimum renewal and various alternative options under consideration. The use of financial appraisal rather than CBA is in direct contrast to the practice on trunk road assessment in which a cost benefit analysis program, COBA, is used (Department of Transport, 1981). As a result the benefits of rail investment may have been understated vis-a-vis road (see Table 3.1).

The use of social CBA for rail investment has, however, been long established, originating with Foster and Beesley's study of the Victoria line (1963). Emphasis has been placed on the application of such techniques to the non commercial sectors of BR i.e. Provincial and, to a lesser extent, London and the South East. Initially emphasis was placed on closure (i.e. disinvestment studies) with the Department of Transport carrying out 32 studies between 1969 and 1974 (including Bradford to Ilkley and Keighley in 1971-2, with busway considered as an alternative). In addition Pryke and Dodgson (1975) were able to identify mean social benefits and specific costs

Case	Financial Return	Cost Benefit return	Rate of Cost Benefit return to financial return
Heathrow Airport	(NPV/C)	(NPV/C)	
Picadilly Extension	1.0	3.65	3.65 : 1
BR Feltham	(IRR)	(IRR)	
resignalling	13%	18%	1.39 : 1
BR London Bridge	14%	23%	1.65 : 1
resignalling			

TABLE 3.1 COMPARISON OF FINANCIAL APPRAISAL AND COST BENEFIT ANALYSIS OF RAIL INVESTMENT SCHEMES

Source: Leitch, 1977

	Social Benefits per passenger mile (p)	Specific Costs per passenger mile (p)
Central Wales[1]	1.79	3.05
Cambrian Coast[2]	2.20	-
Ashford-Hastings[3]	2.28	2.39
Sheffield-Barnsley[4]	2.12	1.25
Manchester-Buxton[5]	2.18	1.48
Manchester - Glossop/Hadfield[6] New Mills/Marple	2.37	1.75

TABLE 3.2 SOCIAL BENEFITS AND RAILWAY COSTS PER PASSENGER MILE (1971 PRICES)

Sources:

[1]Clayton and Rees (1965) [2]Department of Transport (1969)

[3]Foot and Starkie (1970) [4]Else and Howe (1966)

[5]Bristow and Rodriguez (1973) [6]Foster (1974)

per passenger mile from six studies (see Table 3.2). It has been noted that the implications of these findings have rarely been acted on (Dodgson, 1977).

Leitch (1977) notes that rail practice differs from that of roads in that:

(i) Rail cost-benefit analyses do not normally include accident savings

(ii) A different split is assumed between travellers on work and non work journeys, as typically for rail the ratio is 5:95 work: non work whilst for road it is 17:83

(iii) Small time savings are treated differently. For rail the smallest time interval recorded is 30 seconds (usually based on signal box records) whilst for COBA the interval may be as small as one second.

In addition Sugden (1972) has pointed out that:

(iv) Operator's surplus should be taken into account. It was argued by Else and Howe (1966) that this was a pecuniary spillover which should be ignored if the objective is to "achieve a better allocation of resources than would be achieved through the ordinary workings of market forces". This procedure was adopted by the Ministry of Transport in their study of the Cambrian Coast line (1969). It is argued by Sugden that, if it is assumed that prices equal marginal social cost, then Else and Howe underestimate benefits by NF (1-T), where N = Number of rail passengers before closure making a journey, F = Fare charged per journey (assuming bus and rail fares equal) and T = proportion of journeys that transfer to bus, and that a more accurate approach is provided by Clayton and Rees' (1965) study of the Central Wales line.

Point (i) to (iv) above suggest that conventional CBA may have understated the benefits of railway investment and thus might have affected policy decisions.

3.4.2 Types of Cost Benefit Analysis

CBA has been the main tool in the evaluation stage of LUTS (Neuberger, 1971) and has been applied in three ways:

(i) Fixed trip matrix assumption. The main user benefit is the net change in generalised costs:

$$UB = T_{ij} \; (C_{ij}^1 - C_{ij}^2) \qquad \text{Equation 3.23}$$

where UB = User benefit, T_{ij} = number of trips between i and j,

C_{ij}^1 = Generalised cost of travel between i and j in the do nothing situation

C_{ij}^2 = generalised cost of travel between i and j in the do something situation

This method only applies where changes have a negligible effect on the extent and pattern of trip making. In the case of new stations this is unlikely to apply except for work journeys in the short term. The extent of user benefit (or consumer surplus) by this method is given by the shaded area of Figure 3.8A.

(ii) Rule of half (Harrison, 1974). Where an investment has resulted in changes in the extent or pattern of trip making, for those individuals who have switched behaviour we have no information on their generalised cost in the before situation. Hence it is assumed that such switchers receive half the benefits of stayers because:

a) any switching behaviour represents a net benefit change,

b) for a traveller who nearly does not switch after a change the gain will be close to zero,

c) for other travellers a very small amount of change in generalised cost would induce a change in benefit and where change is large they experience virtually all of it as net benefit to them,

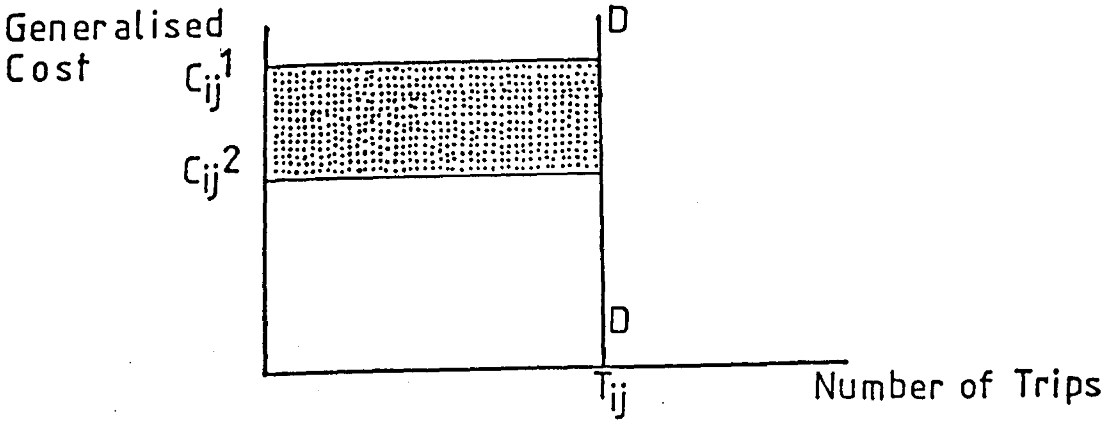
d) other travellers are spaced out along a spectrum between these two points and

e) the amount of user benefit can not exceed the amount of change. Thus

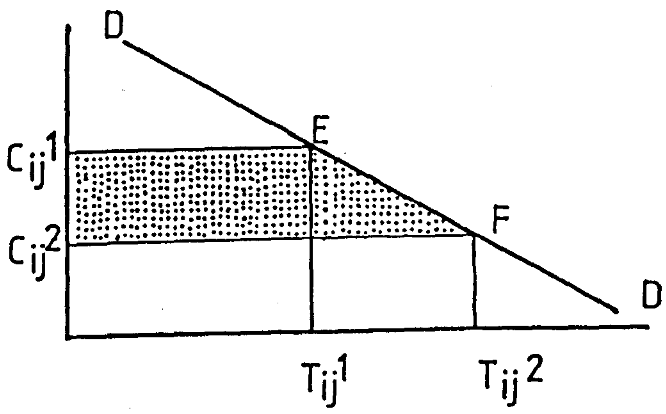
$$UB = T_{ij}^1 (C_{ij}^1 - C_{ij}^2) + \tfrac{1}{2} (T_{ij}^2 - T_{ij}^1) (C_{ij}^1 - C_{ij}^2) \qquad \text{Equation 3.24}$$

which may be shown to be

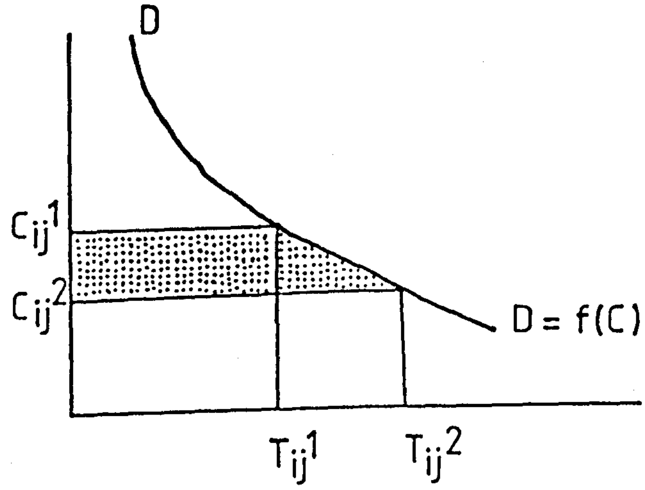
$$= \tfrac{1}{2} (T_{ij}^1 + T_{ij}^2) (C_{ij}^1 - C_{ij}^2) \qquad \text{Equation 3.25}$$



A. FIXED TRIP MATRIX ASSUMPTION



B. RULE OF HALF



C. DIRECT ESTIMATION

FIGURE 3.8 THREE MEASURES OF CONSUMER SURPLUS
(after Neuberger, 1971)

The area of benefit is shown in Figure 3.8B which also illustrates the assumption of linearity between points E and F although empirical evidence suggests this is more likely to be concave. In practice this normally does not result in more than a 10% divergence from the true benefit (Reaume, 1974). It should be noted that in this research, the demand curve DD will be that of total travel demand as in a more specific analysis of rail demand there would be difficulties in correctly locating point E. Hence by our definition stayers will be all new station users who were making the journey by other modes whilst switchers are those new station users whose journey can be considered to have been generated.

(iii) Direct estimation from the Demand Curve. It has been argued, for example by Williams (1977A), that the most appropriate evaluation measure would be to determine the shape of the demand curve and then integrate to find the area of consumer surplus (the shaded area in Figure 3.8C) i.e.

$$UB = \int_{C_{ij}^{1}}^{C_{ij}^{2}} f(c)\, dc \qquad \text{Equation 3.26}$$

For a MNL it can be shown that the path integral is:

$$UB = \frac{1}{\Delta} \ln \frac{\sum_m \exp(-\Delta C_{ij}^{2})}{\sum_m \exp(-\Delta C_{ij}^{1})} \qquad \text{Equation 3.27}$$

where Δ = dispersion parameter

$C_{ij}^{2(1)}$ = Mean generalised cost in the do something (nothing) situations.

However, in our study, use of 3.27 is limited because we do not know the mean generalised cost of all modes, m, in the before and after situations. Instead we may only consider total travel demand and thus equation 3.27 is considered equivalent to 3.23.

Our evaluation will be based on the Marshallian measure of consumer surplus, which may be defined as the difference between the price consumers actually pay for a quantity of goods and the price they would be willing to pay.

This may be justified, as given that although rail's income elasticity is positive it is likely to be less than unity whilst the average proportion of household income likely to be spent on rail fares will be low so that, using the Slutsky equation, the income effect may be shown to be negligible (Douglas, 1985, pages 280-282).

3.4.3 Costs and Benefits to be Analysed

Table 3.3 shows the cost and benefits that will be studied. The framework is designed so that both financial and social appraisals may be undertaken. Variables of which we are not certain of their relevance (denoted ? in Table 3.3) will only be briefly analysed for the six new stations that have been opened or for a sub-sample of potential stations. We shall be unable to quantify the effect of land use and environmental effects, which we suspect to be of a small scale, but a qualitative study will be undertaken. Study of distributional effects will be limited, although the effect on different operators (i.e. the PTE and BR) will be examined. In addition, following conventional procedures (see for example Gwilliam and Mackie, 1975 p.203), a non resource correction will be made to take into account changes in indirect taxation. Further details of the costs and benefits to be studied will be given in Chapter Ten.

Although our cost-benefit framework is as complete as practically possible, it should be noted that there are some omissions. For example, consideration of the effects of risk and uncertainty (see Button and Pearman, 1983, Chapter 7) will be limited to sensitivity analysis concerning the demand/revenue forecasts, project life and interest rates. Furthermore, it should be noted that choice of interest rate has been a matter for "furious" debate (Peacock, 1973), with two approaches in particular having been put forward, the social time preference discount rate and the social opportunity cost, and a certain amount of methodological controversy ensueing (Layard, 1972, pp.243-333). We shall avoid this issue and follow the conventional practice of using the rate for public investment established by the government. Lastly, we are unable to measure the option value of new stations, which is the price individuals are willing to pay for an assurance that rail travel via a new station will be available if they want it and the benefits individuals derive from the opportunity for others to use it (Pearce and Nash, 1981, pp.78-80).

	FINANCIAL		SOCIAL		
PUBLIC	Capital cost of construction	(-)			
	Increase in Rail operating costs	(-)			
TRANSPORT	Decrease in PTE Bus/Rail revenue (transfers)	(-)			
OPERATOR	Decrease in PTE/BR Rail (inc. journey time)	(-)			
	Increase in PTE/BR Rail revenue	(+)			
	Decrease in PTE Bus Operating Costs	(?)			
			RAIL	GC penalty for existing users	(-)
				GC Saving to new users	(+)
			OTHER	GC Saving to non users	(?/+)
USERS			MODES	Reduction in Accidents	(?/+)
REST OF				Fuel savings	(?/+)
ECONOMY				Land use changes	(?)
				Environmental effects	(+/-)
				Distributional effects	(+/-)

TABLE 3.3 COSTS AND BENEFITS TO BE STUDIED

+ = Expected Benefit
- = Expected Cost
GC = Generalised Cost
+/- = Uncertain of direction
? = Uncertain of relevance

3.4.4 Presentation of Costs and Benefits

At least four ways of presenting costs and benefits are commonly used (Harrison, 1974):

(i) First Year Rate of Return (FYRR). (Also known as Single Year Rate of Return). This may be defined as:

$$FYRR = (B_1 - C_1)/K \qquad \text{Equation 3.28}$$

where B_1 = First Year Benefits, C_2 = First Year Costs, K = Capital Cost of Construction

This is computationally simple and may provide a basic rule on priorities (e.g. open a new station if the FYRR is greater than the test discount rate r). Assumptions include that the rate of growth of benefits for all schemes is the same, the rate is always positive and sufficient to ensure returns at the going rate and that the capital costs of a scheme are spent quickly or have the same profile of expenditure over time. These assumptions may only partially hold for new stations.

(ii) Net Present Value (NPV) This is defined as:

$$NPV = \sum_{n=0}^{n=N} \frac{B_n - C_n}{(1+r)^n} \qquad \text{Equation 3.29}$$

where B_n = Benefits in year n, C_n = Costs in year n
N = Project life and r = discount rate

This method requires the compilation of benefits and costs for each year of preparation and of use and therefore can be computationally tedious. It is an absolute measure with schemes being acceptable if the NPV is positive.

(iii) Internal Rate of Return (IRR). This is the rate of discount which will bring benefits of a scheme to zero and may be solved from:

$$\sum_{n=0}^{n=N} \frac{B_n - C_n}{(1+i)^n} = 0 \qquad \text{Equation 3.30}$$

where i = IRR.

This measure is slightly more computationally complex and is affected by the multiple roots problem, whilst the approach may be inconsistent with the results of NPV when considering mutually exclusive projects.

(iv) Benefit-Cost Ratio (BCR). This might be defined, in similar fashion as the NPV measure, as:

$$BCR = \sum_{n=0}^{n=N} \frac{B_n/(1+r)^n}{C_n/(1+r)^n} \qquad \text{Equation 3.31}$$

It may be shown, for example by Wohl and Hendrickson (1984, page 182) that under certain conditions the above four methods, as usually defined, may give results that are neither identical nor consistent, although in practice NPV is the most commonly used measure.

3.4.5 Chosen Framework

This study's evaluation framework will thus be based on CBA, similar to that used in recent studies of U.K. rail investment as discussed in 2.3.2. The variables that will be studied will be as in Table 3.3 and they will be presented using NPV measures, although FYRR will also be used as a simple check measure. Following the conventions for railway investment, project life will be assumed to be 30 years, although there is some evidence that the life expectancy of a station based on wooden platforms may be less than this, so NPV measures over 15 years will also be analysed. Following revisions to the White Paper on Nationalised Industries (Cmnd 7131, 1978) a test discount rate of 5% (the local government rate) will be used, although to allow sensitivity analysis 7% (the current central government rate) and 10% (the previous rate) will also be used. In monetising benefits such as time savings, accident savings and fuel savings use will be made of COBA9 conventions, although disaggregate models developed in this study may provide values of time that can be utilised in the evaluation stage.

However it should be noted that CBA has been criticised on both practical (Crumlish, 1966, Self, 1972) and theoretical (Hunt and Schwartz, 1972, Peacock and Rowley, 1975) grounds. In particular there has been a tendency to use CBA as an automatic decision maker (although this is really a criticism of the users of the technique, not the technique itself), valuations are based on limited empirical evidence and may differ from those determined by the "political process", consideration of distribution is limited, project interdependence is rarely considered and intangible benefits (such as environmental effects) are not evaluated. As a result a number of alternative approaches have been developed, including multicriterion ranking (Buchanan and Partners, 1970), goal achievement matrix (Hill, 1975), planning balance sheets (Litchfield, 1971) and mathematical optimisation (Scott, 1971), but these approaches have only had limited success in overcoming the shortfallings of CBA. It may, however, be noted that:

"the case for cost-benefit analysis is strengthened, not weakened, if its limitations are openly recognised and indeed emphasised" (Prest and Turvey, 1965).

and we shall thus attempt to follow the recommendations of Leitch (1977) and use CBA to develop an evaluation framework that is comprehensive, comprehensible, rational, inexpensive and allows effective control of decisions.

3.5 CONCLUSIONS

In this Chapter the study's methodology has been more clearly defined so as to be consistent with the framework propounded in Chapter One. In terms of market research we have seen that what we might achieve, given the study's time and money budgets, is limited and this will be considered further in Chapter Five. In terms of statistical modelling, we reviewed a number of studies in 3.2 that used such models to forecast rail usage. Although work specifically referring to new local rail stations was limited it did appear that trip rate, regression and mode split models were worthy of further investigation. In 3.3 these three models were placed within the context of transport modelling in general. It was shown that a mode split model, particularly if based on disaggregate data and an hierarchical logit model, might have a number of advantages, although it seems likely that consideration of other stages (i.e. trip generation and distribution) is beyond the scope of this study. Alternatively, regression models might be developed within an aggregate simultaneous model framework. The choice of models to be used in this study clearly depends to a large extent on data availability (to be discussed in Chapter Four) and hence our final decision on model selection will be made at the end of the next chapter. The output from the chosen model(s) will then be used in the evaluation framework which will be based on conventional cost-benefit analysis, even though a number of weaknesses with such an approach have been recognised.

PART TWO: DATA BASES

CHAPTER FOUR DATA REQUIREMENTS AND AVAILABILITY

4.1 DATA REQUIREMENTS

Our system of interest and methodology, outlined in the previous two chapters, clearly require extensive data if we are to achieve the objectives of this research. Some of the main data requirements of the statistical modelling and evaluation stages of this study are shown by Table 4.1. It is evident that the data implications of the statistical modelling stage are more exacting than those of the evaluation stage. Due to resource constraints, emphasis in this study will be placed on the use of existing data sources, which may inevitably involve some compromises. This will be supplemented by small scale market research, which will be discussed in the next chapter. In particular use will be made of BR's Passenger Train Surveys, the County Council's 1984 Rail Survey, the 1981 Census and the West Yorkshire Transportation Study (WYTS) up-date, including the 1981 Corridor and Countywide surveys. These data sets will be discussed in more detail in section 4.2.

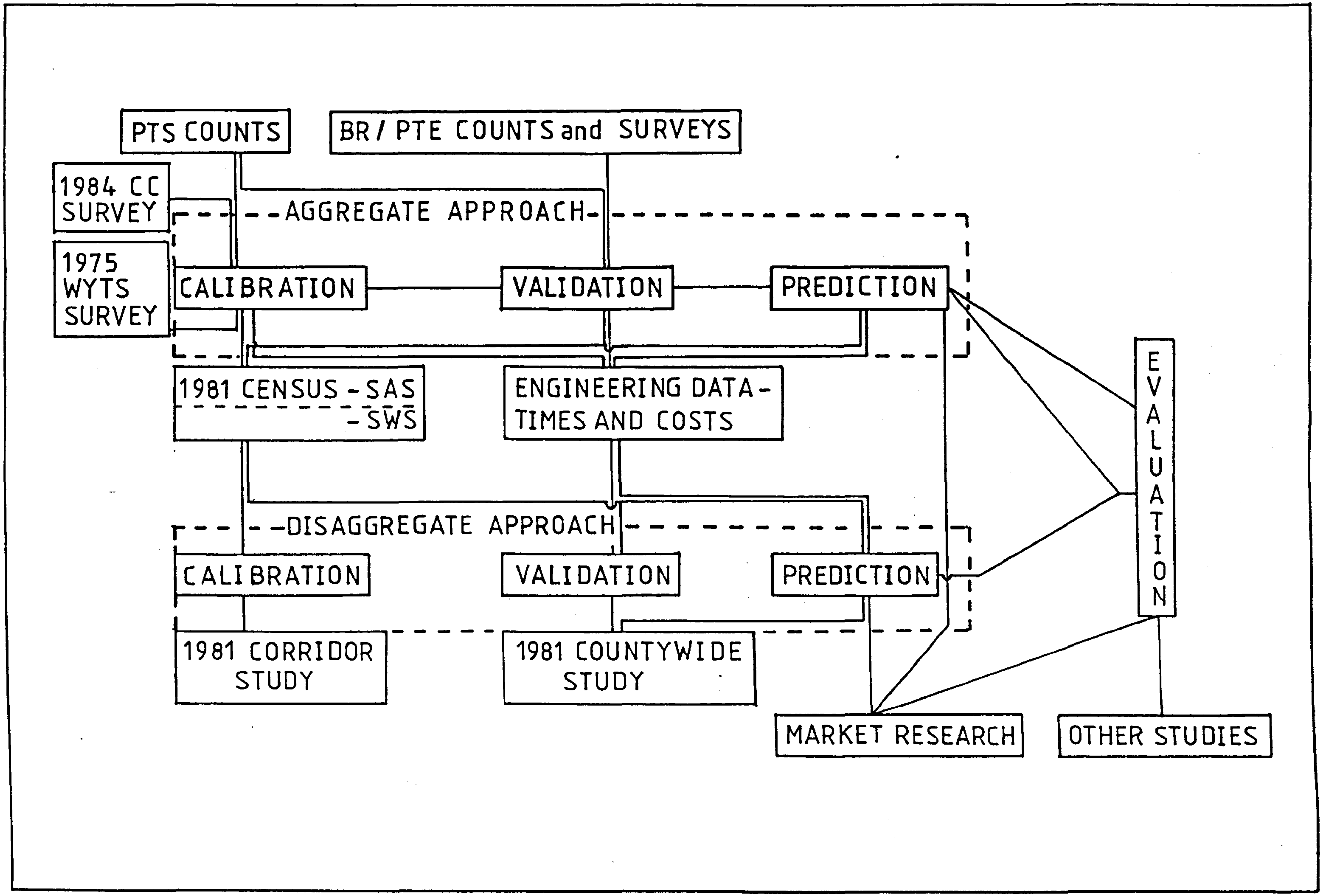
Figure 4.1 shows how these data bases and this research's methodology are inter-related. This diagram shows that a disaggregate modelling approach is dependent on individual data for calibration, validation and prediction and that this may be provided by the WYTS up-date, although, as we shall see, data for the prediction stage is limited. An aggregate modelling approach is dependent on zonal data which may mainly be provided by on-train surveys and the census. In addition it should be clear from the discussion in Chapter Three that a disaggregate approach requires more detailed data than the aggregate approach. In particular an aggregate approach can be based on zonal engineering times and costs whereas, ideally, a disaggregate approach should be based on individual's reported times and costs, although it will be shown that we will have difficulties in achieving this, particularly for the validation and prediction stages.

In the next section we shall examine the availability of the data required by the aggregate and disaggregate approaches and thus we shall, in section 4.3 be able to make a final decision on model choice.

	DATA REQUIRED	AVAILABILITY
Demand	Passenger journeys between i and j - existing stations	BR PTS 1981
	- new stations	PTE Counts/Surveys
	- potential new stations	Model forecasts
Mode Split variables	In vehicle time for rail, car driver	Aggregate model - Engineering data
	Out of vehicle time and/or passenger and	Disaggregate model - 1981 WYTS up date Corridor and County Studies
	Overall travel costs	
	Socio-economic - income	Disaggregate model - 1981 WYTS up date
	- car ownership	
	- occupation	Aggregate model - 1981 Census
Generation and Distribution variables	Trip production characteristics - population	1981 Census
	Trip attraction characteristics - employment	1971 Census of Distribution
		Local Planning Data
User characteristics	Journey purpose	Aggregate - 1984 County Council survey
		Disaggregate - 1981 WYTS up date
Financial Costs and Benefits	Capital Costs	PTE
	Operating Costs	PTE
	Gross Revenue	Demand Forecasts
	Abstraction from other modes	Market Research/Demand Forecasts
Social Costs and Benefits	Time savings/losses - new station users	Market Research (Based on COBA9 conventions)
	- existing rail users	BR PTS
	- non rail users	Urban Congestion Monitoring Project
	Accident savings	Department of Transport
	Fuel savings	Department of Transport
	Distribution effects	Market Research/Demand Forecast
	Accessibility	WYTCONSULT procedures
	Land use	Observation
	Environment	

TABLE 4.1 DATA REQUIREMENTS AND AVAILABILITY

FIGURE 4.1 RELATIONSHIP BETWEEN DATA BASES AND RESEARCH METHODOLOGY



4.2 EXISTING DATA

In terms of data availability it is generally true that rail data is good in quantity but may be poor in quality (see, for example Munby 1978A, B, Aldcott, 1981). More specifically there have been a number of "ad-hoc" surveys in West Yorkshire (as discussed in section 2.2.3) including the 1975 local Rail Survey (WYTCONSULT, 1976B), the Peat, Marwick, Mitchell and Company's (1974) study of local rail interchanges, the PTE surveys of the Airedale and Wharfedale corridors (1979), a study of Leeds to King's Cross travel (Moss and Leake, 1977), BR studies of the Leeds to Bradford (1975) and Huddersfield to Sheffield (1984) corridors and a County Council study of the Leeds to Harrogate line (1975). A number of these data sources were brought together by Hartley (1979A). These surveys, however, are only of limited use to us due to their patchy coverage and outdatedness.

Similarly use of NPAAS (National Passenger Accounting and Analysis System) ticket sales data may be limited due to a number of problems outlined by Webb (1980). In particular certain information may not be recorded, for example blank tickets and Day Rovers will not give full origin and destination (OD) information, whilst further problems exist where there is fraud, validity categories are misleading, rebooking occurs, tickets issued are not used or journeys do not correspond with the O and D shown on the ticket. In particular, the fact that around half of West Yorkshire's rail services are paytrain services and that, in 1981, only 28 out of 49 stations were manned, means that NPAAS data can not extensively be used.

4.2.1 Estimating Origins and Destinations using the Passenger Train Surveys

In order to calibrate an aggregate model information is required on the origin and destination of rail passengers at small town, suburban and rural stations for the study's base year of 1981. A possible approach might be to estimate ODs of rail travel from the information gathered by the PTS (see Preston, 1984A for further details). This is a passenger count which involves on/off counts for every train for one week in August and one week in October/November. Information is available back to, and including, 1970 for all services receiving the Public Service Obligation (PSO) grant. It has become practice that services should be studied in a three year rotation, although since 1982 WYPTE has stipulated that all local services in receipt of Section 20 payments should be surveyed annually.

Rail ODs were estimated by applying the probabilistic approach developed by Savage (1983) for bus revenue estimation, which itself was a modification of a method put forward by Davies and Gribble (1978). Use was made of the November 1981 and November 1982 PTS, using Tabulation B which gives the total on/offs for each station on an average weekday, Saturday and Sunday. A probability method of estimating the number of trips between two points was then calculated as:

$$T_{i,j} = A_{ij} \left[\frac{B_j}{\sum_{i}^{j-1} A_i - \sum_{i+1}^{j-1} B_j} \right] \qquad \text{Equation 4.1}$$

where T_{ij} = the number of rail trips between i and j,

A_{ij} = the number of people getting on at station i and still on train prior to reaching j,

B_j = the number of people getting off at station j and

$\sum_{i}^{j-i} A_i - \sum_{i+1}^{j-1} B_j$ = number of passengers on the train prior to stopping at j.

The part of the equation in square brackets is the probability of alighting (the number of successes over the total number of cases) and will be unity where a service terminates. Where a train is empty at any point and thus the divisor is zero, the probability is set to zero.

Tabulation Bs were obtained for all local rail services in West Yorkshire except for Pontefract Baghill and Moorthorpe stations. Where the service could be considered a closed system, for example Huddersfield to Wakefield, aggregate autumn weekday tabulations were used. However most services had to be disaggegated to identify:

(i) Semi fast trains, for example between Hull and Leeds with four intermediate stops.

(ii) Stations served by only certain, normally peak, services, for example Ravensthorpe.

(iii) Short haul peak workings, for example the Castleford, Garforth and Horsforth to Leeds morning services.

(iv) Special operational services, for example the early morning workings from Huddersfield to Marsden and Bradford to Hebden Bridge.

(v) Services, normally off-peak, terminating at the PTE boundary, for example the Leeds to Knottingley and Leeds to Marsden workings.

(vi) Extended services, for example some trains on the Minster/Selby lines ran to Huddersfield.

(vii) Holiday based services, for example the Leeds to Cleethorpes, Morecambe, Scarborough and Southport services. This problem is more acute on Saturdays and during the summer season.

In addition services need to be disaggregated by direction of travel. As a result 74 separate services were considered of which 20 were surveyed in 1981 and 54 in 1982. These are listed in Table 4.2.

Using a series of FORTRAN programs a matrix of flows with both origin and destination in West Yorkshire (internal flows) was developed for 47 local stations. For reasons of commercial confidentiality the full results can not be presented here. However, it can be said that 34,379 internal ons and offs were recorded, with Leeds City station accounting for 35.7% of them. In addition 9,054 journeys were recorded with either an origin or a destination outside the County (i.e. external flows), meaning that only 20.8% of passengers using local stations travelled beyond the County boundary, with Leeds City being even more dominant, accounting for 57.9% of these external flows.

There are, however, a number of problems with the probability method used:

(i) The PTS may be inaccurate due to unreliable or missing counts.

(ii) The week surveyed by the PTS may be atypical.

(iii) The method does not give information on the final destination where rail interchange occured, although this effects only 15% of local journeys (Hartley and Nash, 1980). Clearly no direct estimate of contributory revenue may be made.

(iv) Local passengers on trains stopping solely at inter-city stations were not included and hence flows to/from Leeds City, Bradford Exchange, Huddersfield, Dewsbury, Wakefield Westgate and New Pudsey are

Code	Description	Code	Description
01100	Bradford to Keighley (1)	06101	Castleford to Leeds (1)
01101	Keighley to Bradford (1)	06201	Knottingley to Leeds (1)
01200	Bradford to Skipton (1)	06300	Leeds to Goole (1)
01201	Skipton to Bradford (1)	06301	Goole to Leeds (1)
01300	Leeds to Skipton (1)	08100	Leeds to Cleethorpes (1)
01301	Skipton to Leeds (1)	08200	Leeds to Doncaster (1)
01400	Leeds to Morecambe (1)	08201	Doncaster to Leeds (1)
01401	Morecambe to Leeds (1)	09200	Leeds to Sheffield
01500	Leeds to Carlisle	09201	Sheffield to Leeds
01501	Carlisle to Leeds	10100	Huddersfield to Clayton W.
02100	Bradford to Ilkley (1)	10101	Clayton W. to Huddersfield
02101	Ilkley to Bradford (1)	10200	Huddersfield to Penistone
02200	Leeds to Ilkley (1)	10201	Penistone to Huddersfield
02201	Ilkley to Leeds (1)	10300	Huddersfield to Sheffield
03100	Leeds to Harrogate	10301	Sheffield to Huddersfield
03101	Harrogate to Leeds	11100	Huddersfield to Wakefield
03200	Leeds to Knaresborough	11101	Wakefield to Huddersfield
03201	Knaresborough to Leeds	11200	Leeds to Marsden
03300	Horsforth to Leeds	11201	Marsden to Leeds
03500	Leeds to York (A)	11300	Leeds to Huddersfield
03501	York to Leeds (A)	11301	Huddersfield to Leeds
04100	Scarborough to Bradford	11400	Leeds to Southport (E)
04200	Leeds to York (B)	11500	Stalybridge to Leeds
04201	York to Leeds (B)	11600	Leeds to Manchester (E)
04300	Scarborough to Leeds	11601	Manchester to Leeds (E)
04400	York to Huddersfield	12100	Bradford to Leeds
04500	York to Mancheter	12101	Leeds to Bradford (F)
04600	York to Bradford	12200	Leeds to Southport (F)
04601	Bradford to York	12201	Southport to Leeds (F)
05100	Leeds to Selby	12300	Hebden Bridge to Leeds
05101	Selby to Leeds	12301	Leeds to Hebden Bridge
05200	Leeds to Hull (C)	12400	Leeds to Halifax
05300	Garforth to Leeds	12401	Halifax to Leeds
05400	Garforth to Leeds	12500	Bradford to Hebden Bridge
05500	Hull to Huddersfield	12600	Leeds to Manchester (F)
06100	Leeds to Castleford	12601	Manchester to Leeds (F)

(1) Surveyed in 1981. (A) via Harrogate. (B) via Church Fenton. (C) semi fast. (D) stopping. (E) via Dewsbury. (F) via Halifax.

TABLE 4.2 O/D ESTIMATION - SERVICES ANALYSED underestimated.

(v) For certain flows, for example between Shipley and Bingley, it is difficult to detect which station is the generator and which is the attractor.

(vi) The PTS was carried out at different times, with two-thirds of services being counted in 1982 and the Clayton West branch last surveyed in August 1981. Weighting factors were developed to adjust these figures to November 1981 levels.

(vii) The mechanics of the calculation may have led to some errors due to the assumption that the probability of alighting is uniform for all passengers regardless of where they got on, rounding errors and the occurrence of unbalanced flows due to the disaggregation process used.

The OD data from the probability estimation method was compared with sample data from NPAAS and WYTCONSULT, using a χ^2 (chi-square) test. This failed to show that the PTS data was significantly different from the NPAAS and WYTCONSULT data, although this may be in part due to the lack of a truly appropriate test.

4.2.2 County Council Rail Survey 1984

As the OD information derived from the PTS seemed likely to be subject to input error, alternative data sources were examined. One such source was the OD data collected by the PTE/BR on-train survey team. However this information was not available, on a sufficient scale, during the study period. An alternative data source was provided by the WYMCC local rail survey carried out between October 8th and October 18th 1984. In total, 21,105 interview forms were distributed at all local stations on a weekday and 10,987 were distributed at two stations (Leeds City and Bradford Forster Square) on a Saturday. The response rate was around 35% on a weekday and 15% on a Saturday, with a total of 10,033 records coded.

Again due to reasons of commercial confidentiality the full results of this survey cannot be presented. However, we might expect those OD flows recorded from the 1981 PTS estimates to be greater than those of 1984 for two reasons:

(i) In 1981 the final destination for trips that involved interchange was given as the first station where a change was made. In 1984 these flows were correctly recorded to their ultimate destinations. Hence in 1984 there are a lot more OD pairs with trips recorded compared to 1981.

(ii) In 1984 only a 12 hour day was surveyed (between 0700 and 1900 hours), whilst in 1981 all trains were surveyed.

It should also be noted that the 1981 data represented an average weekday whilst the 1984 data was based on a particular weekday.

Furthermore in comparing the two data sets a number of anomalies were observed. For example between 1981 and 84 the flow between Wakefield Kirkgate and Leeds appeared to have declined by 90%. In addition the County Council believed that response rates for passengers travelling outside the County were lower than average, with a comparison with 1981 data suggesting that the 1984 survey might be underestimating external flows by around 30%. It also seems evident that short distance flows may have a lower response rate as, for example, between 1981 and 1984 flows between New Pudsey and Bradford were seen to decline by 68%. As long (external) and short distance flows seem to be underestimated it seems likely that the use of station weighting factors may have overestimated certain middle distance (internal) flows.

Despite these concerns the 1984 data set has advantages in that it allows disaggregation by journey purpose and more up to date information on ticket type to be used.

In addition in this section it should be noted that the PTE has initiated a number of "ad hoc" counts of new station usage during our survey period. Including information from the PTS this involved 31 new station counts on weekdays and 24 on Saturdays. These counts formed the basis of our analysis of demand growth over time, described in 9.4. In addition surveys were carried out on a weekday and Saturday at Fitzwilliam (1st and 5th March, 1983), Deighton (10th and 14th May, 1983) and Crossflatts (17th and 21st May, 1983). All boarding and alighting passengers were asked simple questions concerning their origin/destination, journey purpose, frequency of travel and ticket type. This data will be used to detect biases in our own market research in 5.3.2.

4.2.3 Census Data, 1981

As we have already seen, 1981 Census data can be used by both modelling approaches. For an aggregate approach the Census is particularly useful in supplying information on characteristics which may affect trip generation such as population, car ownership and occupation. These are available in the Small Area Statistics (SAS) which may be accessed by the SASPAC package via the University of Manchester Regional Computing Centre (UMRCC, 1982). New station catchment zones can be defined in terms of aggregations of whole or part Enumeration Districts (EDs). Figure 4.2 shows an example for Crossflatts. Where an ED only partly falls into a distance band the proportion of the built up area within the band is estimated using Ordnance Survey 1:10,000 or 10,560 scale maps.

Census data may also be useful in the application stage of a mode split model, calibrated with disaggregate data. In particular journey to work flows and modal shares from section C of the 1981 Special Workplace Statistics (SWS) (OPCS, 1983) may be useful. This data was supplied by WYMCC on magnetic tape and was based on a 452 zoning system classified as follows:

(a) zones 1-448 are actual zones within West Yorkshire
(b) zone 419 represents all external zones
(c) zone 450 represents no fixed workplace
(d) zone 451 represents workplace not stated
(e) zone 452 represents workplace outside U.K.

The data tape consisted of eight files:

(a) All mechanised private transport
(b) Bus
(c) Train
(d) Walk
(e) Bicycle
(f) All trips including other not stated, work at home
(g) Car driver
(h) Car pool and passenger

We will only be interested in the bus, train, car driver and car pool and passenger files. In addition we will only be intereted in rail served zones, of which 103 were identified, as shown by Table 4.3. The data was read using the TRADVV suite of matrix handling programs, in particular the M1 (to create and print a file) and the M4 (to compress a matrix) programs. This data will be used in conjunction with the incremental logit model, as described in 9.2.3.

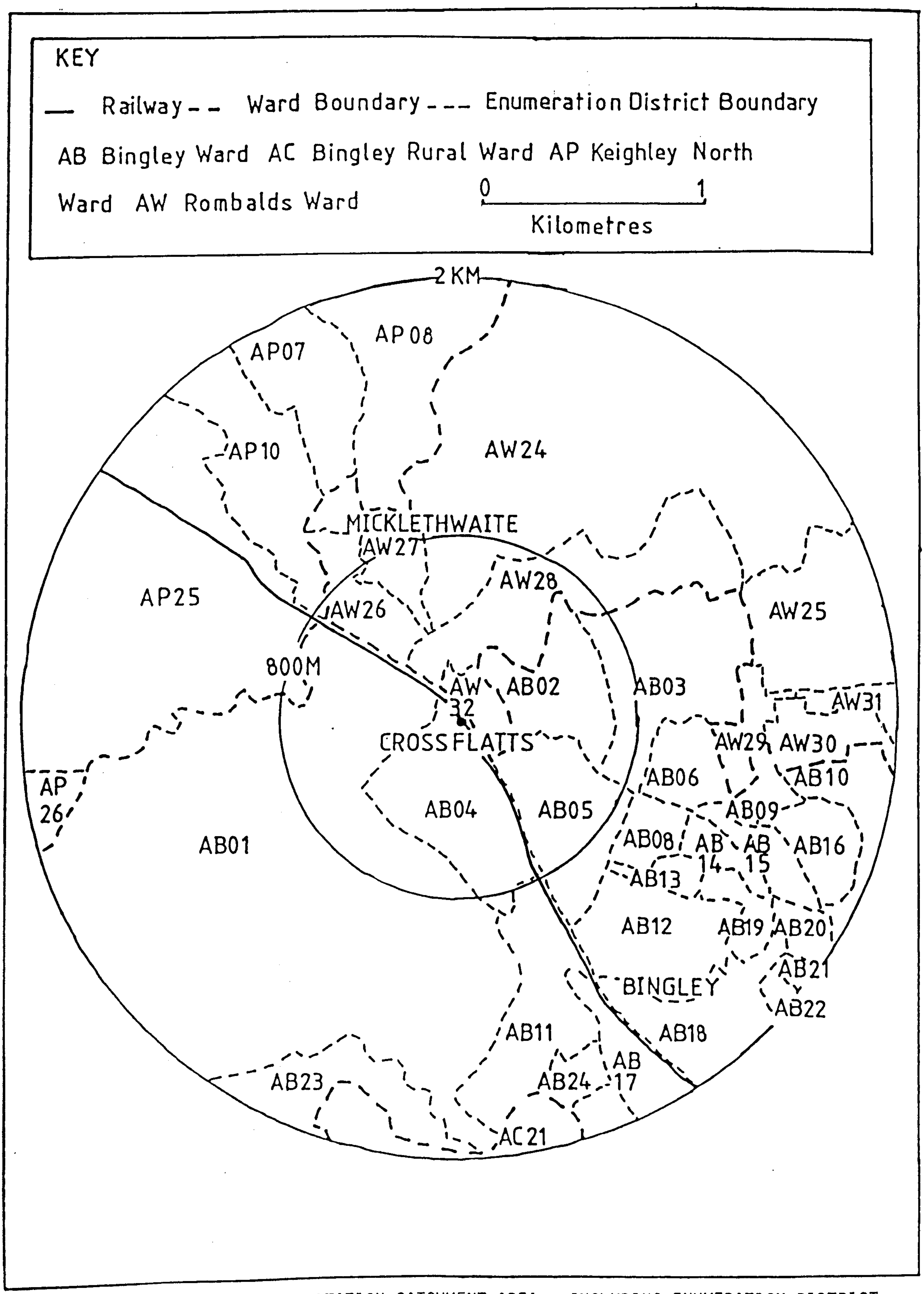


FIGURE 4.2 EXAMPLE OF A NEW STATION CATCHMENT AREA - INCLUDING ENUMERATION DISTRICT BOUNDARIES

Code	Description	Code	Description
4	Ilkley Central	5	Ben Rhydding
6	Burley in W.	7	Menston
8	Silsden	10	Steeton
11	Utley	13	Keighley N.
16	Keighley Central	17	Keighley E.
26	Crossflatts	29	Bingley
38	Shipley Glen	39	Baildon E.
40	Baildon - Charlestown	42	Saltaire
43	Shipley	51	Frizinghall
67	Bradford Central	69	Laisterdyke N.
75	Laisterdyke S.	80	Bowling
84	Low Moor E.	85	Low Moor W.
86	Low Moor N.	107	Guiseley
111	Horsforth N.	121	Horsforth S.
122	Horsforth Central	124	Hawksworth
125	Headingley W.	126	Headingley E.
146	Stanningley	147	Bramley N.
148	Bramley E.	149	Bramley S.
150	Armley	151	Burley
152	University	157	Osmondthorpe
159	Cross Gates	165	Swinnow Moor
168	Wortley W.	170	Leeds - Westgate
171	Leeds - Headrow	172	Leeds - City Square
173	Leeds - South West	174	Leeds - South East
175	Leeds - Markets	178	Halton Moor
181	Garforth W.	182	Garforth Central
183	Garforth E.	184	Micklefield
186	Wortley E.	188	Beeston
195	Morley W.	196	Cottingley
197	Morley E.	211	East Ardsley
212	Thorpe	221	Todmorden
229	Hebden Bridge	230	Mytholmroyd
246	Luddendenfoot	250	Halifax Central
251	Halifax E.	253	Hipperholme
254	Lightcliffe	256	Sowerby Bridge
258	Salter Hebble N.	259	Salter Hebble S.
260	Salter Hebble E.	296	Batley W.
297	Batley E.	301	Hanging Heaton
305	Mirfield	307	Ravensthorpe
310	Dewsbury Central	325	Deighton N.
327	Milnsbridge N.	330	Huddersfield Central
332	Deighton S.	335	Marsden
336	Slaithwaite	338	Milnsbridge S.
376	Outwood W.	377	Wrenthorpe
378	Outwood E.	387	Wakefield - Westgate
394	Wakefield - Kirkgate	395	Sandal N.
403	Sandal S.	412	Castleford
417	Normanton	442	Pontefract - M'hill
424	Knottingley W.	425	Knottingley Central
426	Knottingley E.	431	Pontefract Central
437	Fitzwilliam	444	Hemsworth
447	S. Elmsall		

TABLE 4.3 RAIL SERVED AND POTENTIALLY RAIL SERVED ZONES

4.2.4 Corridor Study, 1981

Disaggregate data was collected on mode choice for the journey to work in the Garforth, Airedale and Wharfedale corridors as part of the 1981 West Yorkshire Transport Study in conjunction with phase one of the Department of Transport's Value of Time study. This was based on household interviews with information gathered for the following variables:

(a) Journey time, cost etc. of travel to/from work by Yesterday's Mode (YM).
(b) Journey times, costs etc. of travel to/from work by up to three Alternative Modes (AMs).
(c) Ranking of modes in order of preference.
(d) Frequency of travel by YM and AMs.
(e) Transfer price increase of YM (see below).
(f) Transfer price decrease of Best Alternative (BA).
(g) The effect of different departure/arrival times on working practices.
(h) Hours worked.
(i) Income from main job.

An initial screening process excluded the following households:

(i) Where no member was employed outside the home.
(ii) Where the worker(s) had more than one place of work or frequently moved from site to site.
(iii) Where the worker(s) had to use the vehicle driven to/from work for business purposes during the day, except on an occasional basis.
(iv) Where the worker(s) travelled less than one mile or worked outside the study (defined as the three corridors plus Leeds and Bradford) (MVA, 1984).

The effect of this screening process is shown by Figure 4.3 and has serious implications for the aggregation stage of this research. The data was brought together in two tranches, with the first tranche consisting of 1,014 records and the second tranche containing 533 records, giving 1,547 responses in total. This data can be split into a number of subgroups (see 7.1) with the largest available complete RP data set, WBA9, consisting of 1,271 observations.

It should be noted that the corridor study included information on Transfer Price (TP). This is based on the hypothesis that when presented with two qualitatively different options people can not only consistently form a preference for one over the other, but can in certain circumstances, such as mode choice, also report the minimum amount of money which would just compensate them sufficiently to accept the best alternative (i.e. the less favoured option), or alternatively the maximum amount they would be prepared to pay (in addition to any other costs) to continue their first preference (i.e. yesterday's mode). The results of the TP analysis were reported by Broom et al. (1983). Unfortunately it was found that a TP approach had statistical problems. In particular it was found that an assumption of homoscedastic Normal error structure is inappropriate when used with linear utility functions, whilst it was also noted that a substantial proportion of respondents were either unable or unwilling to provide TP information (possibly reflecting non compensatory behaviour). Gunn (1984) has shown that some of the statistical problems may have stemmed from mis-specification of the TP models. In particular it was shown that analyses which use TP as a dependent variable, but restrict its sign (either by modelling the options separately or by switching the observable characteristics to reflect the difference between the chosen and rejected option) can not be made consistent with orthodox 'rational utility maximising' theory without a complex statistical approach. Such analyses result in an overstatement of the parameter usually associated with habit, whilst all other parameter values may have low significance. However it is not made clear how Gunn's correction (consistently modelling the difference between mode 1 and mode 2 rather than between rejected and chosen mode) may be applied to a multinomial choice situation, whilst problems with the high incidence of non respondents to TP questions still remains. Thus it was not possible to demonstrate TP's superiority over RP analysis and hence it was recommended not to use TP in travel contexts where RP models can be feasibly applied. Our study will heed this advice but, given our findings in 3.2.4, we note, with some regret, that the corridor data set did not include SP data rather than, or in addition to, TP information.

Apart from problems with the TP data, two further types of problem were apparent with the Corridor study:

(i) Intrinsic data problems. These include that:

(a) The study used inexperienced survey staff and this resulted in an "overall impression that the data was very variable in quality and coverage" (MVA 1984).

(b) Following on from (a) certain questions posed problems. In particular some respondents gave marginal and others fully distributed car operating costs, others had difficulty in calculating the cost of a Public Transport trip when they owned a season ticket whilst, as already hinted, a number of respondents claimed they did not have a best alternative.

(c) Few respondents regularly used alternative modes and their knowledge of alternative modes' times and costs was unlikely to be reliable.

(ii) Problems of data collection. The Corridor study data was collected, in the main, to allow a comparison between the TP and RP mode choice models and their resultant values of time. In this study we are, however, more interested in developing predictive rather than descriptive models. Hence a number of problems may be noted:

(a) The type of format of the questionnaire was constrained to be consistent with previous WYTCONSULT studies.

(b) If a person broke his/her journey to/from work then full details of the interrupted journey would be recorded, but only 'simple' home-work journeys were contained in the RP data sets.

(c) If a person did not rank the mode they used yesterday as 1 (i.e. it was not the mode they normally used) they were excluded from analysis.

(d) Wherever possible a Public Transport option was coded as the best alternative so as to avoid problems with perceived car running costs, payment to driver etc.

(e) Despite the choice of rail served corridors and the doubling of the sample rate within one kilometre of a station rail is still a minor alternative. Some implications of the sampling procedure are examined in 8.2.2.

(f) Information on modal combinations, for example park and ride, kiss and ride and feeder bus, have been subsumed into the main mode leg.

Of the above problems points (a) and (f) are relatively minor, points (b), (c) and (e) have implications for aggregation as shown by Figure 4.3 whilst some of the effects of point (d) will be examined in section 7.1.4. Thus despite these problems it will be possible to calibrate MNL and HL models of mode choice with this data set.

CORRIDOR DATA SET	No. of obs
	4598
Exclude h'holds making no trips	
Exclude non work trips	screened out
Exclude workers with more than one place of work	3360
Exclude workers who use vehicle for business purposes during the day	screened out
Exclude workers who travel less than one mile or work outside the study area	screened out
Double sampling rate within station catchment areas	1238
Exclude if bike or walk yesterday's mode	938
Exclude if income question unanswered	850

COUNTYWIDE DATA SET	No. of obs
Exclude h'holds making no trips	89507
Exclude if origin not rail served	18751
Exclude non work trips	7271
Exclude if Motor Cycle, Bike or Walk chosen mode	5468
Exclude if workers travel less than two miles	2890
Exclude if destination not rail served	877

FIGURE 4.3 SCREENING PROCEDURES USED

4.2.5 Countywide Study, 1981

Extensive surveys were carried out as part of the WYTS up-date in 58 areas throughout the county. The County Council made the household, person and trip data files available to us on tape. This was a very large data set covering around 11,000 household and almost 90,000 trips. However, only a very small proportion of these trips (0.4%) were made by rail. In order to concentrate on flows where rail was an alternative a screening process of the type shown in Figure 4.3 was enacted. Survey areas with existing stations, new stations or potential stations were identified. These are shown by Table 4.4 which shows that although, for our purposes, there is good coverage for the calibration and validation stages, with the catchment areas of 13 existing and new stations included, there is only limited data for prediction, with only two potential station catchment areas included. This suggests that for prediction alternative data sources need to be investigated.

Code	Calibration	Code	Validation	Code	Prediction
01	Ilkley	03	Crossflatts	27	Osmondthorpe
02	Keighley	06	Saltaire	49	Wrenthorpe
03	Bingley	22	Bramley		
04	Baildon	43	Slaithwaite		
06	Shipley	45	Deighton		
15	Guiseley				
31	Todmorden				
35	Halifax				
36	Batley				
40	Mirfield				
43	Marsden				
47	Central Huddersfield*				
54	Castleford				
56	Knottingley*				
58	South Elmsall				

TABLE 4.4 RELEVANT SURVEY AREAS FOR CALIBRATION, VALIDATION and PREDICTION

* later excluded as no rail trips recorded.

If only the 13 areas in the calibration data set are considered the total number of trips is reduced to 18,751 of which only 169 (0.9%) are rail trips. Of these 99 were work trips (58.6%) and hence it was only feasible to study this purpose as there were only 26 shopping trips and 28 recreational/social trips made by rail. Even so rail only accounted for 1.4% of work trips in these 13 areas. Hence the data set was further reduced (and rail's share boosted) by excluding journeys made by non-mechanised modes, journeys of less than two miles and journeys to non rail served destinations thus reducing the data set to a total of 877 observations. The implication of this screening process on the aggregation stage is again illustrated by Figure 4.3.

Although the Countywide data set was adequate in terms of the socio-economic and mode choice information it provided, there were a number of limitations. For example if car was defined as yesterday's mode no data on travel costs were provided. Also, apart from car ownership, there was no data on alternative modes and thus there was no time and cost data for alternative modes.

Given these limitations it was clear that the Corridor study data provided a better basis for calibration than the Countywide data. However it was felt that if suitable decision rules about alternative modes were formulated and if engineering times and costs were inserted, as described in 8.2.2, the Countywide data set might be used in validation and limited prediction studies. This might lead to incompatibility between the calibration and validation data sets. Possible biases due to using reported and engineering data are discussed in the next section. The validation data set consisted of 343 trips (10 by rail), whilst the prediction data set only consisted of 57 trips (none by rail). A breakdown of these observations by station catchment areas is given by Table 4.5. It should be noted that the Bramley survey area covered a large area that was not within the existing new station's catchment area but might be served by a station at Gamble Hill.

	No. of obs.		No. of obs.
Bramley (1)	21	Bramley (2)	94
Deighton	41		
Crossflatts	36	Osmondthorpe	28
Slaithwaite	70	Wrenthorpe	29
Saltaire	81		

TABLE 4.5 BREAKDOWN OF VALIDATION AND PREDICTION DATA SETS

(1) = Swinnow Road (2) = Gamble Hill

4.3 DATA LIMITATIONS AND MODEL SELECTION

In Chapter Three a number of modelling approaches were examined. In terms of dimensions identified by Reid (1979A) and Sherret (1979), such as geographic detail and scale, responsiveness to issues, comprehensiveness and resultant evaluation measures, a disaggregate approach to forecasting new station usage appears preferable. However, an aggregate approach may have some advantages in terms of comprehendability and feasibility.

In this chapter we have re-examined these modelling approaches in terms of the resources that are required and those that are available. In doing so we can examine their likely accuracy and complexity, particularly in terms of data requirements. It has already been noted that a disaggregate approach is likely to be more complex than an aggregate approach and thus ought to have higher forecasting accuracy. It may though be noted, following the framework discussed in 1.4.2 (after Alonso, 1968), that although a disaggregate approach may reduce specification error, its need for high quality data may lead to increases in input error.

Moreover in this chapter we have noted that development of a disaggregate approach is constrained by data availability. This means that we are limited to calibrating a model of mode choice for the journey to work based on RP data. In validating such a model we shall have to combine reported ($\approx$ perceived) and engineering ($\approx$ actual) data, the general implications of which have been examined by Gunn et al. (1980) and MVA et al. (1986, pages 146 to 149). In particular it appears that using reported times and costs tends to lead to an exaggeration in the difference between chosen and rejected options compared to engineering times and costs. This is likely to be due to justification bias, although it is not clear if this takes the form of making the rejected option worse or the chosen option better (or both), although the former seems the most likely. Hence it is likely that a RP model based on reported data will have greater explanatory power than one based on engineering data due to systematic misreporting and may also have significantly different coefficient values (Wardman, 1986). Moreover, for a model based on reported data for the chosen option and engineering data for the rejected option, there will be a different and even more complex pattern of biases.

The main problem with the disaggregate approach is the lack of data for the prediction stage, as the random, stratified (by area type) samples provided by the Corridor and County data sets are insufficient. Aside from using aggregate Census data, as discussed in 4.2.3, there are the following possibilities (after Lerman and Manski, 1979):

(i) Develop our own random samples based on mailback questionnaires. However, given 35 sites to study and a minimum of 30 observations of times and costs of the journey to work by mechanised modes to rail served destinations per new or potential station catchment area, this would require in excess of 1,000 usable replies based on over 12,000 initial contacts (assuming a 25% response rate and a screening-in rate of 32.5% of households (implicit from Figure 4.3)). A survey of this scale is clearly beyond the cost, time and manpower resources available to this study.

(ii) Develop choice based samples of, for example, bus, train and car users. Due to resource constraints such work by us will be limited to train (and more specifically new station) users. It has been shown by Manski and Lerman (1977) that, in order to produce consistent logit model results with a full set of such data, the maximum likelihood estimators should be amended by weighting each observation's contribution to the log likelihood by Q(i)/H(i), where Q(i) is the fraction of the decision making population selecting i and H(i) is the analogous fraction for the choice based sample. We shall briefly return to this issue in 10.3.1.

(iii) Simulation techniques might be used to produce synthetic samples of households from Census data. Examples include the SYNSAM programs used by McFadden (1979) and the work of Pownall and Wilson (1976) and Bonsall and Champernowne (1979). Such an approach would require major computational effort but will be briefly examined in 9.2.3.

From the above it seems clear that a disaggregate mode split model for the journey to work can be calibrated and validated with existing data but there may be problems in further application. As a result it is proposed to develop aggregate approaches as a "back up". Firstly, a non-work aggregate model may be calibrated, based on the 1984 County Council rail survey. Forecasts produced by this model may then be combined with disaggregate forecasts for the journey to work so as to provide a total station usage figure. This might then be thought of as being a hybrid aggregate-disaggregate approach. Secondly, we might develop a pure aggregate approach, calibrated on either PTS or the 1984 County Council data sets (or both). Such aggregate models would be simultaneous in structure and similar in form to the direct demand model of Equation 3.7. It should, though, be noted that such aggregate approaches are themselves constrained by the paucity of existing data, particularly concerning rail ODs and trip attraction characteristics.

By developing a plural approach to modelling we shall be able to examine, in a policy context, the relative advantages of disaggregate and aggregate model types, hence supplementing previous work by Watson (1975) and Liou et al. (1975). Thus we have followed the advice of Hartgen and Liou (1976) who suggested that:

"Considerable effort in the next several years should be placed on the development of a series of demonstrations designed to show that these tools (disaggregate demand models) are appropriate, and highlight the relative advantages of these tools versus conventional procedure".

Moreover, one way in which different model types may be examined is by comparing the accuracy of their demand forecasts for the six new stations that have been opened. Thus we are following the advice of Burnett and Hanson (1979) in placing emphasis on a deductive/predictive rather than explanatory/descriptive study, and the advice of Williams and Ortuzar (1980) who stated that:

"a high priority in empirical research must be the accumulation of evidence relating to response in before and after studies".

CHAPTER FIVE NEW STATION SURVEYS

5.1 INTRODUCTION

In previous chapters it has been established that our market research will be limited to surveying new station users. In particular this should enable us to achieve one of the study's main objectives by determining the proportion of users that are generated and the proportions abstracted, as well as collecting important information on trip and socio-economic characteristics of new station users.

In the rest of this section we shall examine the issues of survey design and application. In a second section we shall go on to discuss the results of the pilot survey, whilst the third and fourth sections will discuss the results of the first and second year surveys respectively. In a fifth, and final section, we shall examine two ways our market research may be applied.

5.1.1 Survey Design

As three new stations had been opened before this study began, our market research was seen as an "ex post facto" experiment relating causes to effect. A high initial priority was given to the development of market research in order to capture recall information at these three stations. Thus it was necessary to quickly determine the means of data collection and approach to respondents. A number of possible approaches were identified (after Moser and Kalton, 1971). On-train interviews were rejected as, for most stations, a large number of trains would have to be covered. Household interviews were rejected due to their high cost (see 4.3) whilst self completion questionnaires were dismissed due to problems of collection and the likely low response. It was thus decided to concentrate on brief on-platform interviews. Emphasis would be placed on surveying boarding passengers as they were more "captive" than alighting passengers. The surveys would take place on a representative weekday (normally a Tuesday or Thursday) between 07.00 and 16.00 hours and a Saturday between 08.00 and 14.00 hours during the Spring and/or Autumn. Initially the surveys were envisaged as requiring only one interviewer. This survey technique was chosen as it should result in a good response: cost ratio and would allow the collection of observational data (for example on sex and party size). Possible problems might include interviewer bias and sample bias due to the emphasis on boarding passengers and/or time of day or year chosen.

5.1.2 Questionnaire Design

Use was made of a precoded interview form similar to those used in the Glasgow Rail Impact Study (Martin Voorhees et al. 1980) and by WYTCONSULT (1976B). The form used is given in Appendix 1, whilst the nature of questions asked and their order are shown by Figure 5.1. Question 5's role as a filter is vital in determining the proportion of trips that are generated due to the new station and the proportion that may be considered as having been abstracted from other modes. Particular attention has been paid to question wording and to question order, especially with regards to general summary questions (question 8) and questions asked in conjunction (for example questions 3 through to 7 and questions 13 through 15) (Schuman and Presser, 1981). Questionnaire design was limited by the need to be brief (an interview was envisaged as taking, on average, less than two minutes), whilst it was also important that the most vital (trip) information was captured first.

5.1.3 Survey Schedule

The questionnaire was tested by surveys at Clayton West and Skelmanthorpe prior to their closure on 22nd January, 1983. The survey design was tested more rigorously by pilot surveys at Mirfield, Baildon, Marsden and South Elmsall. These stations were chosen as they had similar levels of rail service and similar population characteristics as Deighton, Crossflatts, Slaithwaite and Fitzwilliam respectively. The main surveys were carried out in two phases:

(i) An initial survey was carried out between 6 and 12 months after opening (having allowed enough time for demand to settle after an initial burst of promotional activity). The only exception was Saltaire where the survey occured just over a year after opening, as the preferred survey date (October/November 1984) clashed with the WYMCC survey and BR counts.

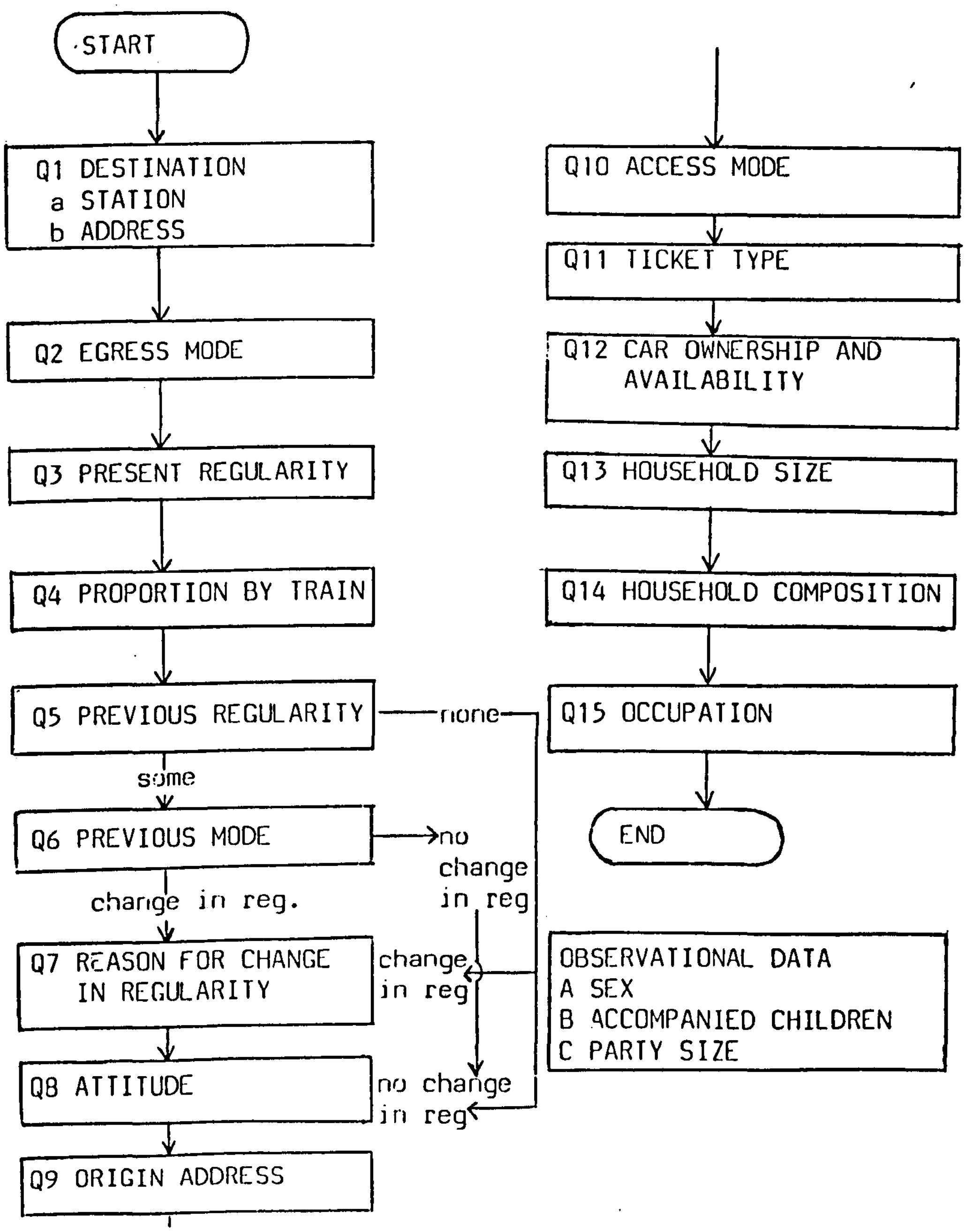


FIGURE 5.1 STRUCTURE OF NEW STATION QUESTIONNAIRE

(ii) Those stations which opened in 1982 and 83 were resurveyed a year after the initial survey.

School holidays were avoided, except for one survey at Crossflatts which aimed to highlight different travel patterns during the school holiday period. The completed survey schedule is given by Table 5.1 and consisted of eight days of pilot surveys (of which two were on Saturdays) and 21 days of main surveys (seven on Saturdays).

5.2 PILOT SURVEYS

The responses to the pilot surveys are shown by Table 5.2. The small scale survey at Clayton West and Skelmanthorpe showed that the questionnaire was well understood and interviewees were willing to respond (although as expected some respondents were unwilling to reveal their precise home address although they would give their street name and/or postcode). A 100% response rate was achieved as the low service frequency allowed interviews to be carried out on-train. The responses emphasised the local role of the Clayton West branch, with 59% of travellers proposing to use the bus after closure.

A more thorough test of the survey design was provided by the pilot surveys at Mirfield, Baildon, Marsden and South Elmsall. A number of problems were apparent:

(i) The response rate was low, with only 35% of boarding passengers being interviewed. This was partly because demand, particularly at Mirfield and South Elmsall, was greater than that anticipated at new stations and also exhibited greater peakiness than anticipated.

(ii) The method was labour intensive with only 3.7 forms collected per man hour. This was inevitable given the low service frequency at certain stations. An average of 1.7 forms were collected per train departure, with the maximum number being six.

(iii) The method gave a poor coverage of peak hour trains vis-a-vis off-peak trains. For example at Baildon during the peak hour (08.00 to 09.00 hours) 36 passengers boarded trains of which only eight were interviewed (22%), although nine out of 13 off-peak passengers were covered (69%).

PILOT SURVEYS		FIRST YEAR SURVEYS		SECOND YEAR SURVEYS	
Clayton West	Thursday 13th Jan. 1983	Fitzwilliam	Saturday 23rd April 1983	Fitzwilliam	Thursday 3rd April 1984
Skelmanthorpe	Thursday 20th Jan. 1983	Fitzwilliam	Tuesday 26th April 1983	Deighton	Tuesday 12th April 1984
		*Crossflatts	Tuesday 7th April 1983	Crossflatts	Tuesday 1st May 1984
Mirfield	Tuesday 22nd Feb. 1983	Crossflatts	Tuesday 10th May 1983	Crossflatts	Saturday 2nd June 1984
Baildon	Thursday 24th Feb. 1983	Crossflatts	Saturday 14th May 1983	Slaithwaite	Tuesday 15th May 1984
Marsden	Saturday 26th Feb. 1983	Deighton	Saturday 21st May 1983	Slaithwaite	Saturday 12th June 1984
Marsden	Tuesday 1st March 1983	Deighton	Tuesday 24th May 1983	Bramley	Thursday 25th April 1985
South Elmsall	Saturday 12th March 1983	Slaithwaite	Saturday 4th June 1983	Bramley	Saturday 27th April 1985
South Elmsall	Tuesday 15th March 1983	Slaithwaite	Tuesday 7th June 1983		
		Bramley	Saturday 19th May 1984		
		Bramley	Tuesday 22nd May 1984		
		Saltaire	Thursday 18th April 1985		
		Saltaire	Saturday 20th April 1985		

TABLE 5.1 SURVEY SCHEDULE

* School holiday period

	No. of forms collected	No. of people surveyed	No. of people boarding	Response rate (%)
Clayton West/Skelmanthorpe	23	25	25	100
Baildon	16	17	50	34
Marsden	35	40	110	36
S. Elmsall	59	66	183	36
Mirfield	39	58	175	33
TOTAL	172	206	543	38
(excl. Clayton West/Skelmanthorpe	149	181	518	35)

TABLE 5.2 RESPONSE RATE TO PILOT SURVEYS

(iv) The method may have introduced a bias in that respondents were interviewed in order of arrival, but early arrivals may have different characteristics from late arrivals. This may be overcome through a systematic survey of late arrivals through the use of on-train interviews. Evidence of sampling bias was suggested by the fact that the female:male ratio of respondents was 62:38.

(v) There was a problem with partially completed interviews, which formed 9% of responses. This might be limited by ensuring that the most essential questions were answered before the interview was terminated.

(vi) It was evident that certain information was not being collected that should be. In particular it was shown to be necessary to insert additional questions concerning the return mode and alternative modes.

It was thus clear that the method of interviewing boarding passengers had been severely tested by the pilot surveys. However, less than 3% of those contacted had declined to be interviewed and the survey had generally been well received. It was felt that response rates could be dramatically improved by the use of a second interviewer during peak periods, who might also interview late arrivals on the train. Given these minor amendments it was proposed to persevere with interviewing boarding passengers.

The pilot surveys at Mirfield, Marsden, Baildon and South Elmsall also allowed us to test a number of aspects of survey analysis.

(i) Coding. The coding of most variables proved straightforward except for:

(a) Origin and Destination addresses. It was decided to make use of the Post Code system that consisted of up to seven alpha-numeric characters (GPO, 1982). The main advantages of using this method are its accessibility (four directories cover most of West Yorkshire), its thoroughness and the high standard of mapping available.

(b) Following PTE conventions, revenue was calculated by using information on ticket type in conjunction with the Local Fares Directory (BR Divisional Passenger Manager, 1982) and Selective Prices Manual. Revenue from season tickets (Multiriders) was calculated as nine-tenths the normal fare, where Metropasses or Railcards were used it was calculated as half the normal fare and if Privilege tickets were used it was calculated as a quarter the normal fare. Revenue from Day Rover was estimated at one-third the ticket cost.

(c) Detailed questions were asked concerning occupation, mainly as a proxy for social class and hence income distribution. Social class was determined by identifying the chief economic supporter of a household using criteria established by the Census (OPCS, 1979) and using the 17 socio-economic groups and the 39 socio-economic classes, as defined by the 1968 Standard Industrial Classification. This gave a simple five-fold division of social class similar to that of the Census (Table 5.3).

	Census		Pilot Surveys
I	Professional	I	Managerial
II	Intermediate	II	Professional
III(N)	Skilled Non Manual	IIIN	Skilled Non Manual
III(M)	Skilled Manual	IIIM	Skilled Manual
IV	Partly Skilled	IV	Un/Semi-Skilled
V	Unskilled		

TABLE 5.3 CLASSIFICATION OF OCCUPATION

(ii) Processing. Having coded the data it was punched into the Amdahl computer and run in conjunction with SPSS (Statistical Package for the Social Sciences) (Nie et al. 1975). In later surveys the SAS (Statistical Analysis System) package (Helwig, 1983) was used, thus allowing comparisons between the two packages. To counter the low response rate grossing up factors for the peak and off-peak were developed as follows:

$$G = [z - \frac{z}{s}(d + \frac{d}{n} - n)]/(s - d - n) \qquad \text{Equation 5.1}$$

where z = the total number of passengers boarding trains, s = the total number of passengers interviewed, d = the number of selected passengers that became ineligible (interviews incomplete) and n = the number of non responses.

If the peak is broadly defined as being between 07.00 and 09.00 hours on weekdays and the off-peak as all other time periods respective grossing up factors of 4.35 and 2.14 were calculated.

(iii) Implications for Sample Size. The pilot survey may also be used to determine the sample size for the main survey. This might be based on a back of envelope calculation:

$$N = \left[\frac{zs}{d}\right]^2 \qquad \text{Equation 5.2}$$

where s = the standard deviation of the pilot survey variable, d = the tolerable margin of error and z = the z table value corresponding to the level of confidence.

Using Equation 5.2 it can be shown that, at a 95% level, to determine household size within a margin of error of ± 0.5 requires a sample size of around 29, but if the margin of error is to be reduced to ± 0.1 then the sample size needs to be increased to around 724. For more continuous variables, for example straight line distance from origin address to new station, the tolerable margin of error will be wider. Thus for origin distance to be within ± 200 m would require a sample size of 194, whilst if the margin of error was reduced to ± 100 m the sample size would have to be increased to around 774.

These results suggest that we should be aiming for a sample size exceeding 700 observations. This might only be achieved by re-surveying new stations.

(iv) Data analysis. Frequencies and cross-tabulations of all the key variables were produced. The results are presented elsewhere (Preston, 1983) but some of the key features included:

(a) 29% of travellers had changed the regularity with which they made their stated journey over the last year, just over one-third being due to employment changes, with personal reasons and changes in residence also being important.

(b) 12% of travellers had changed mode over the past year, two-thirds switching from bus and one third from car.

(c) Points (a) and (b) show that the rail market is very dynamic. Over 40% of travellers did not use rail with the same regularity as one year ago.

(d) A number of long distance rail trips were recorded. As a result the mean fare was calculated as £1.39, although the median fare of £0.70 is more representative.

(e) Differences between stations were evident. An example is given by origin distance, defined as the straight line distance between the origin address and the station, as shown by Table 5.4.

	0-300m	301-600m	601-800m	800m-1km	1.1km-2.0km	3.1km+
Mirfield	0	5.1	20.5	43.6	23.1	7.7
Baildon	37.5	37.5	6.3	6.3	12.5	-
Marsden	39.4	22.9	17.1	14.9	-	5.7
South Elmsall	5.1	5.1	20.3	18.6	10.2	40.7

TABLE 5.4 ORIGIN DISTANCE (%) - PILOT SURVEY

There is a clear distinction between Mirfield and South Elmsall, with less than one-third of travellers originating within 800m, compared with Baildon and Marsden where over three-quarters of travellers originate within 800m (at Clayton West and Skelmanthorpe 82% of travellers originated within 800m). This distinction is partly due to local geography. The town of Mirfield is well to the north of the station, with the station also having subsidiary catchment areas in Lepton and Upper Hopton. Similarly South Elmsall station also serves the neighbouring villages of South Kirkby, Moorthorpe, Upton and Ackworth. This distinction is also reflected by access mode. At Baildon and Marsden on-foot access accounts for 100% and 83% of passengers respectively. At Mirfield walking only accounts for 61% of passengers whilst car accounts for 28%. At South Elmsall walking only accounts for 41% of passengers, whilst bus accounts for 43%. In terms of origin distance and access modes we would expect new stations to be more like Baildon/Marsden than Mirfield/South Elmsall.

(vi) A number of differences were apparent between weekday and Saturday travel. On a Saturday travel was concentrated on the major conurbation centres, was mainly for shopping (58%) and seemed to be optional in nature. There was some evidence that usage by retired persons and people from non-car owning households was greater on a Saturday, whilst the number of accompanied children and party size was greater than on a normal weekday.

5.3 FIRST YEAR SURVEYS

5.3.1 Response rate

The response to the first year surveys is shown by Table 5.5. In all 588 interview forms were collected, involving 675 passengers and representing a

		No. of forms collected	No. of people surveyed	No. of people boarding	Response rate (%)
Crossflatts	H	31	38	61	62
Crossflatts	H	33	41	56	73
Crossflatts	S	35	44	45	98
Deighton	W	48	53	59	90
Deighton	S	23	24	25	96
Fitzwilliam	W	32	37	41	90
Fitzwilliam	S	30	38	39	97
Slaithwaite	W	34	37	38	97
Slaithwaite	S	44	48	66	73
Bramley	W	55	60	84	71
Bramley	S	70	90	100	90
Saltaire	W	96	100	141	70
Saltaire	S	57	65	92	71
TOTAL (A)		588	675	847	79.7

TABLE 5.5 FIRST YEAR SURVEYS - RESPONSE RATE

H = Holiday W = Weekday S = Saturday

Fitzwilliam	W	35	36	39	92
Deighton	W	43	48	62	77
Crossflatts	W	53	55	89	62
Crossflatts	S	25	28	31	90
Slaithwaite	W	38	44	58	76
Slaithwaite	S	44	53	61	87
Bramley	W	41	46	69	67
Bramley	S	54	72	102	71
TOTAL (B)		333	382	511	74.7
OVERALL TOTAL (A+B)		921	1057	1358	77.8

TABLE 5.6 SECOND YEAR SURVEYS - RESPONSE RATE

response rate of 80%. When combined with the second year interviews (Table 5.6) a total of 921 interviews were completed representing 1057 passengers and giving a response rate of 78%. Thus the sample sizes projected in 5.2 were achieved whilst the response rate with respect to the pilot surveys was more than doubled, due to the use of a second interviewer during peak periods and the more amenable demand profile. The response rate did fall below 75% at Crossflatts on weekdays, Slaithwaite on Saturdays, Bramley on weekdays and Saltaire on weekdays and Saturdays, as these were the stations where the demand was heaviest and/or most peaked.

The data collected was coded using the procedures established in the pilot survey. The coding inventory is given in Appendix 2. The data was analysed using SPSS and SAS programs. Weighting factors were developed so as to take into account the differential coverage of peak and off-peak passengers (for our purposes the peak was defined as being between 07.00 and 09.00 hours on a weekday and the off-peak was defined as all other time periods). Rather than use Equation 5.1 weighting factors were simply calculated as:

(S/N) where S = number of people gêtting on trains and N = number of people surveyed — Equation 5.3

The weighting factors used for the first year surveys are given by Table 5.7 and for the second year surveys by Table 5.8. In total it can be shown that the weighting factor for peak trips was 1.54, whilst for off-peak trips it was only 1.18.

5.3.2 Testing for Bias

It should be evident from previous sections that the survey method chosen might lead to a number of biases.

(i) Those passengers interviewed may represent a different population to those passengers boarding during the survey times who were not interviewed, particularly as earliest arrivals were more likely to be sampled. The strength of this bias was tested by comparing the results of our survey with data collected by BR for Fitzwilliam, Deighton and Crossflatts in May 1983 (where a 100% sample was achieved). This was done by the use of a X^2 (chi-square)

		PEAK			OFF PEAK		
		Covered	Total ons	Weighting factor	Covered	Total ons	Weighting factor
Crossflatts	H	24	41	1.71	14	20	1.43
Crossflatts	W	29	44	1.52	12	12	1.00
Crossflatts	S				44	45	1.02
Deighton	W	18	20	1.11	35	39	1.11
Deighton	S				24	25	1.04
Fitzwilliam	W	15	18	1.20	22	23	1.05
Fitzwilliam	S				38	39	1.03
Slaithwaite	W	21	21	1.00	16	17	1.06
Slaithwaite	S				48	66	1.38
Bramley	W	35	58	1.67	25	26	1.04
Bramley	S				90	100	1.11
Saltaire	W	52	90	1.73	48	51	1.06
Saltaire	S				65	92	1.42
TOTAL (A)		194	292	1.51	481	555	1.15

TABLE 5.7 FIRST YEAR SURVEYS - WEIGHTING FACTORS

Fitzwilliam	W	15	16	1.07	21	23	1.10
Deighton	W	21	27	1.29	27	35	1.30
Crossflatts	W	35	62	1.77	20	27	1.35
Crossflatts	S				28	31	1.11
Slaithwaite	W	15	27	1.80	29	31	1.07
Slaithwaite	S				53	61	1.15
Bramley	W	26	48	1.85	20	21	1.05
Bramley	S				72	102	1.42
TOTAL (B)		112	180	1.61	270	331	1.23
TOTAL (A+B)		306	472	1.54	751	886	1.18

TABLE 5.8 SECOND YEAR SURVEYS - WEIGHTING FACTORS

test based on the development of contingency tables (see Table 5.9).. Altogether 11 OD pairs could be compared where the observed (O) and expected (E) values exceeded five. A null hypothesis (H_0) was established as being that the usage figures are independent of the data set chosen (i.e. the two figures are broadly similar), whilst the alternative hypothesis (H_A) was that the usage figures are associated with the data set chosen. The χ^2 statistic was calculated as, letting the results of our own survey being O and the results of the BR survey of boarding passengers being E, as

$$\chi^2_{(calc)} = \sum_i \left[\frac{(O_i - E_i)^2}{E_i}\right] = 9.46 \qquad \text{Equation 5.4}$$

However with 10 degrees of freedom and the 5% decision criterion $\chi_{0.05}$ = 18.307 and as this is greater than $\chi^2_{(calc)}$ we can not reject the null hypothesis. It may be concluded that, at least for this variable, the emphasis on interviewing earliest arrivals has not led to a noticeable systematic bias.

Station From	To	Own survey (O)	BR survey (E) Ons	BR survey (E) Offs	Ons $\frac{(O-E)^2}{E}$	Offs $\frac{(O-E)^2}{E}$
Fitzwilliam	Wakefield (W)	8	8	6	0.00	0.67
Fitzwilliam	Leeds (W)	13	10	21	0.90	3.05
Crossflatts	Leeds (W)	22	29	28	1.69	1.29
Crossflatts	Bradford (W)	17	27	27	3.70	3.70
Crossflatts	Leeds (S)	17	17	15	0.00	0.27
Crossflatts	Bradford (S)	13	8	7	3.13	5.14
Deighton	Huddersfield (W)	26	28	19	0.14	2.58
Deighton	Leeds (W)	12	10	14	0.40	0.29
Deighton	Wakefield (W)	14	14	12	0.0	0.33
Deighton	Huddersfield (S)	14	23	40	2.13	16.90
Deighton	Leeds (S)	6	8	23	0.50	12.56
TOTAL					9.46	46.78

TABLE 5.9 CONTINGENCY TABLE OWN SURVEY COMPARED TO BR SURVEY

(ii) Boarding passengers may represent a different population from alighting passengers. The asymmetry of demand at new stations is reflected by the on:off counts collected by the PTS. From Table 5.10 it can be seen that at all stations boarding passengers exceed alighting passengers. This may be, in part, due to the fact that in West Yorkshire the morning peak is better served by rail than the evening peak.

Bramley	0.934	Fitzwilliam	0.631
Crossflatts	0.922	Saltaire	0.844
Deighton	0.782	Slaithwaite	0.906

TABLE 5.10 WEEKDAY OFF : ON RATIO

Source: PTS 1983, 1984

Whether boarding passengers represent a different population from alighting passengers can again be tested by making comparisons with the BR survey, only this time letting alighting passengers be E. This gave an χ^2(calc) of 46.78 which is greater than $\chi^2_{0.05}$ at 10 degrees of freedom, hence suggesting that the null hypothesis can be rejected and that there is thus a possibility of bias. However, the service at one site (Deighton) on the Saturday surveyed by BR was disrupted by engineering work. If these two observations are excluded the χ^2(calc) reduces dramatically to 17.32 which however is still greater than $\chi^2_{0.05}$ at 8 degrees of freedom which is 15.507. Hence there is still a suggestion that assuming that boarding and alighting passengers are the same population may lead to some bias. In order to evaluate how severe this bias is a question concerning return mode has been included.

(iii) Interviews were carried out between 07.00 and 16.00 hours on a weekday and 08.00 and 14.00 hours on a Saturday, but the characteristics of boarding passengers outside these time periods might be different from those that we interviewed. Unfortunately there is no complete data set available that will allow comparisons. However, from the PTS, we can estimate that at new stations on a weekday 80% of joining pasengers board trains during our interview time period, with this figure decreasing to 71% on a Saturday. The main categories of travellers we have missed are passengers working in the new station catchment area (this is a feature at Saltaire, where 45 offs were recorded between 07.00 and 09.30 and, to a lesser extent, Deighton and

Bramley) on a weekday and evening leisure/social trips on a weekday and a Saturday.

(iv) The weekdays chosen for study may be unrepresentative. However Tuesdays and Thursdays were chosen as being representative, with, for example, the PTS showing that 19.8% of weekday usage at new stations occured on a Tuesday. It should be noted that Sundays were not surveyed as the new stations had a very limited service.

(v) Finally there may be a bias due to seasonal factors. Our surveys took place, outside school holidays, during the spring when demand was felt to be most representative. There is a possibility that at other times of the year, particularly during the summer and public holidays, demand will be dramatically different from that surveyed, but again data do not exist to allow comparisons.

5.3.3 First Year Results (All results are weighted by the grossing up factors presented in Table 5.7).

The results of our first year surveys will now be presented by examining the variables listed in Appendix 2 in turn. Emphasis will be placed on drawing distinctions between weekday and Saturday usage and between the different new stations surveyed.

From Table 5.11 it can be seen that weekday usage is highly peaked, with 41% of passengers boarding trains during a single hour (0730 to 0830) and 64% boarding trains during a two hour period (0700 to 0900). It is noticeable that demand falls off sharply after 10.00. By contrast usage on a Saturday is more constant, with a slight indication of a peak around lunch time.

In terms of destination, Leeds dominates weekday usage accounting for 42.5% of trips, although Bradford and Huddersfield are other important destinations accounting for around 17% and 11% of usage respectively. On a Saturday these three centres are even more dominant, accounting for almost 77% of usage. On a weekday only 10% of trips cross the West Yorkshire border, rising to 11% on a Saturday. A significant proportion of these trips were to Doncaster (from Fitzwilliam) although some long distance inter-city trips, to destinations as diverse as King's Cross, Watford, Glasgow and

	7.00-7.30	7.31-8.00	8.01-8.30	8.31-9.00	9.01-10.00	10.01-11.00	11.01-12.00	12.01-13.00	13.01-14.00	14.01-15.00	15.01-16.00	(No. of valid cases)
Weekday	10.6	19.1	21.9	12.7	13.0	6.6	4.2	4.0	3.5	2.8	1.4	(424)
Saturday			8.0	5.0	15.5	13.6	16.4	18.9	22.6			(323)

TABLE 5.11 TIME OF TRAVEL (%)

	On Foot	Bus	Car	Taxi	Other	(No. of valid cases)
Weekday	88.9	8.3	1.7	0.7	0.5	(424)
Saturday	91.4	5.0	1.9	0.3	1.5	(324)

TABLE 5.12 EGRESS MODE (%)

	Work	Education	Shopping	Leisure	Personal Business	Firms' Business	Social	Other	(No. of valid cases)
Weekday	55.3	9.1	18.8	9.1	3.0	1.2	3.5	-	(430)
Saturday	6.5	0.9	63.2	18.0	3.7	0.9	6.5	0.3	(323)

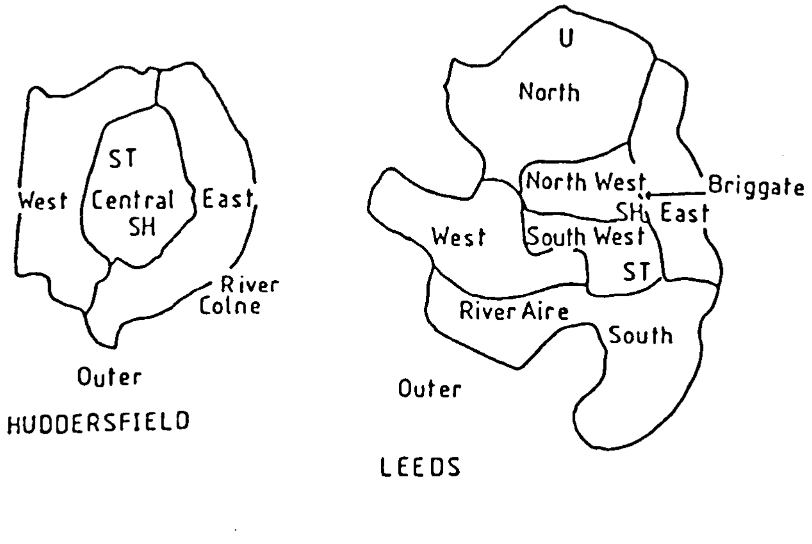
TABLE 5.13 JOURNEY PURPOSE (%)

	None	Once	2 - 4 times	5 - 10 times	11 - 15 times	More than 15 times	(No. of valid cases)
Weekday	14.4	5.7	13.0	7.3	2.6	57.1	(424)
Saturday	26.8	9.3	29.4	13.4	1.3	19.8	(313)

TABLE 5.14 FREQUENCY OF TRAVEL IN LAST MONTH (%)

	New Station	Employment	Residence	Personal	Destination change	(No. of valid cases)
Weekday	19.4	50.7	20.8	6.0	3.0	(134)
Saturday	65.5	8.8	15.9	8.8	0.9	(113)

TABLE 5.15 REASON FOR CHANGE IN FREQUENCY



KEY: SH=Centre of Main Shopping Area ST= Station U= University

HUDDERSFIELD

Central	31.6
East	23.7
West	26.3
Outer	18.4
No. of valid cases	48

LEEDS

Central	- North West	26.1
	- South West	27.0
	- East	12.6
North		15.3
West		3.6
South		5.4
Outer		9.9
No. of valid cases		139

BRADFORD

Central	32.6
North	9.6
South West	11.5
South East	23.1
Outer	23.1
No. of valid cases	65

FIGURE 5.2 CENTRAL AREA DESTINATIONS OF NEW STATION USERS
(Unless stated otherwise figures refer to percentages)

Edinburgh, were recorded. There are obviously differences in destination choice for individual stations, but Leeds is the main destination except from Slaithwaite and Deighton (where it is Huddersfield) and Crossflatts during school holidays (where it is Bradford). At Fitzwilliam and Saltaire, both on a Saturday, Doncaster and Bradford respectively rival Leeds as the main destination.

In terms of destination address Figure 5.2 shows how, for Leeds, Bradford and Huddersfield, final destinations are concentrated within the Central Business District. Due, in the main, to the fact that most shoppers were unsure of their precise destination only 252 destination addresses in the three major centres could be located, representing 37% of the sample. In Leeds over 53% of destinations were concentrated in the area bounded by the River Aire, Briggate, the Headrow and Westgate, with similar patterns of concentration in Bradford and Huddersfield. Destinations further afield may be due to major trip attractors such as the Universities and Colleges in Leeds (concentrated in the North zone of that city and attracting 15% of trips) and Bradford (concentrated in the South East zone of that city and attracting 11% of trips). The relatively high proportion of trips beyond the central area in Bradford may reflect the better interchange with bus compared to Huddersfield or Leeds. However it should be noted that overall around 90% of passengers walk from the destination station to their final address (Table 5.12).

Table 5.13 shows that on a weekday work is the main journey purpose accounts for 55% of trips, with shopping the next most important purpose (19% of trips), followed by education and leisure (9% each). By contrast on a Saturday work accounts for less than 7% of trips, whilst shopping accounts for 63% and leisure accounts for a further 18%. There are some differences between stations. For example, due to local amenities, Saltaire has been successful in attracting leisure trips accounting for 16% of trips on a weekday and 33% of trips on a Saturday. Similarly at Crossflatts during the school holidays no education trips were recorded, whilst leisure trips accounted for 20% of passengers.

On a weekday 17% of those interviewed will definitely not return by rail, increasing to 19% on a Saturday. The two stations where this phenomenon was most common were Deighton and Slaithwaite. Of those not returning by rail around two-thirds would use bus.

Table 5.14 shows that on a weekday around 57% of passengers make journeys to their stated destination four times a week or more (i.e. they are regular travellers), with the other main categories being those who travel approximately weekly (13%) and infrequently (14%). By comparison on a Saturday only 20% of passengers regularly travel to their stated destination, compared to 29% who are weekly and 27% who are infrequent travellers. Of those passengers travelling at least once a month to their specified destination, only 81% stated they made all such journeys by rail on a weekday, with this figure decreasing to 65% on a Saturday. The use of alternative modes was most noticeable at Crossflatts and Slaithwaite. Of those passengers not making all their stated journeys by rail on a weekday around 48% use car and 39% use bus, whilst on a Saturday the percentages change to 36% and 55% respectively. These results illustrate that rail demand at new stations is asymmetric and exhibits important interfaces with other modes, particularly bus.

One important result of our market research was that we were able to determine the proportions of travellers abstracted from other modes and the proportions generated. This was based on an analysis of previous mode used (obtained from question 7 of the interview form shown in Appendix 1), modified to take into account situations where a change in regularity (detected by comparing responses to questions 5a) and 6) meant that only a proportion of trips was abstracted. Thus:

$$\text{Generated trips} = \sum^{n} 1 - \frac{\text{Previous Regularity}}{\text{Present Regularity}} \qquad \text{Equation 5.5}$$

and, mode by mode,

$$\text{Abstracted trips} = \sum^{n} \frac{\text{Previous Regularity}}{\text{Present Regularity}} \qquad \text{Equation 5.6}$$

where n = number of observations. The ratio of previous regularity to present regularity was set to a maximum of 1 so as to accommodate any observations where previous regularity exceeded present regularity.

Excluding those who made their stated journey less than once a month, it was shown that on a weekday 37% of trip makers have changed the frequency that they made their trip since the new station opened, increasing to 49% on a Saturday. However, equations 5.5 and 5.6 need to be adjusted because, as Table 5.15 shows, on a weekday only 19% of changes in trip frequency were stated as being due to new stations, although this increased to over 65% on Saturdays, reflecting the effect new stations have in generating optional travel. On a weekday the main exogenous cause for changes in frequency of the stated journey was changes in employment, whilst on a Saturday the main cause was changes in residence. We were unable to exactly measure the importance of changes in destination choice, but they appear to be relatively minor. Thus equations 5.5 and 5.6 were adjusted by assuming that trips generated due to reasons other than the new station would have been abstracted from existing modes in the same proportions observed for all non generated trips. To a limited extent this adjustment might cause generated trips to be underestimated because changes in residence and/or employment may themselves be related to new stations. We were unable to investigate this complex causal change in detail although comments from interviewees suggested that this kind of change was not occuring, at least in the short term.

Given the above adjustment Table 5.16 shows that, although there is considerable variation between stations, overall on a weekday only around 9% of trips might be considered to be generated, compared to over 56% abstracted from bus, 15% abstracted from other rail stations and 17% abstracted from car. On a Saturday the percentage of generated trips increases to over 30% mainly at the expense of rail (5%), car (13%) and to a lesser extent bus (52%).

These results indicate that abstraction from bus is much greater than the 12% anticipated by the PTE following their Airedale and Wharfedale surveys or indeed the 31% used by Cottham (1985). It appears that our estimates of generated travel are lower than those of the PTE (1984B) because the PTE have made use of our unadjusted results, which did not take into account exogeneous factors, and suggested that overall 42% of Monday to Saturday new station usage was generated, 37% was abstracted from bus, 8% abstracted from train, 11% was abstracted from car and 2% was abstracted from other modes. The corresponding adjusted figures are 13%, 56%, 13%, 16% and 2%.

		Bus	Train	Car	Other	Generated
Fitzwilliam	(W)	57.6	-	13.7	7.6	21.1
	(S)	46.7	-	16.4	-	36.9
Deighton	(W)	52.2	16.2	15.9	3.2	12.6
	(S)	64.7	9.6	6.3	5.8	13.5
Crossflatts	(W)	46.1	23.7	23.0	4.4	2.9
	(S)	26.4	13.2	23.1	-	37.2
Slaithwaite	(W)	39.8	7.0	32.2	-	21.0
	(S)	41.2	2.3	14.6	2.3	38.7
Bramley	(W)	78.5	4.6	12.6	-	4.2
	(S)	63.5	1.8	4.6	0.5	29.6
Saltaire	(W)	58.5	22.3	10.9	-	8.3
	(S)	56.4	4.9	14.4	-	24.3
Mean	(W)	56.6	15.2	16.7	2.6	8.9
	(S)	51.5	4.6	12.5	0.9	30.5
	(M-S)	55.8	13.4	16.0	2.3	12.5

TABLE 5.16 PREVIOUS MODE USED (%) (ADJUSTED)

Analysis of previous studies, for example of the Cross City line in the West Midlands and Watton-at-Stone in Hertfordshire, as well as our own initial results (see Preston, 1984B), suggested that as a rule of thumb around half of new station users might be expected to come from Public Transport, with the remainder being split almost equally between generated trips and those abstracted from private transport. However our final adjusted results suggest that if applied for a Monday to Saturday in West Yorkshire such a rule would underestimate the percentage of abstractions from Public Transport by almost 20% and roughly double the effect of generated trips.

Table 5.17 shows that most travellers say they have chosen to travel by rail from a new station due to the speed of the journey, but it is also apparent that more qualitative factors, in particular comfort and convenience, play an important role.

Table 5.18 shows that new stations tend to have highly localised catchment areas, for on a weekday 63% of travellers originate from within 800m of new stations, with over 86% originating within 2 km. The corresponding figures for a Saturday are almost 67% and over 83% respectively. The precise origin locations of new station users are shown by Figures 5.3 to 5.8. At all six sites there are distinctive clusters around the station but subsidiary catchment areas are imporant for example: Hemsworth and Ackworth for Fitzwilliam; East Morton, Riddlesden and Micklethwaite for Crossflatts; Linthwaite for Slaithwaite; Bradley for Deighton and Lucy Hall for Saltaire. At certain sites it is noticeable that the catchment area extends back furthest in the opposite direction to the main travel destinations, resulting in catchment areas that are elliptical rather than circular. This phenomenon is most marked for Crossflatts, Fitzwilliam and Saltaire. It is also noticeable that for the two sites, Crossflatts and Saltaire, which are close to existing stations (Bingley and Shipley respectively) the catchment area boundary appears to be clearly defined in the area of potential overlap and does not extend beyond the half way mark.

Table 5.19 shows how origin distance is related to access mode. Although walk is everywhere dominant two sub-groups may be identified:

i) Urban sites (Bramley, Deighton and Saltaire) where the mean access distance is between 400 and 750 m and walk accounts for between 80 to 100% of travellers' access modes.

	Speed	Cost	Reliability	Nearness to home	Comfort & convenience	Other	(No. of valid cases)
Weekday	59.5	9.4	3.5	4.3	17.0	6.3	395
Saturday	58.6	7.9	0.8	6.7	18.4	7.5	239

TABLE 5.17 ATTITUDE TO RAIL (%)

		0 to 300m	301 to 600 m	601 to 800 m	801m to 1 km	1,001 km to 2 km	2,000 km to 3 km	Over 3 km	(No. of valid cases)
Slaithwaite	W	8.8	43.7			27.0	3.0	17.3	
	S	34.1	9.1	6.8	11.4	20.5	4.5	13.6	
Crossflatts	H	38.9	18.9	6.3		17.1	2.9	16.0	
	W	22.1	22.0	5.8	3.8	26.0	2.9	17.4	
	S	43.6	10.3		5.1	17.9		23.1	
Deighton	S	56.5	17.4	4.3		17.4		4.3	
	W	29.2	33.3	4.2	6.3	18.8	2.1	6.3	
Fitzwilliam	W	34.7	42.3	3.3			19.7		
	S	23.3	16.7		3.3	16.7		40.0	
Bramley	W	10.6	36.2	19.1	10.0	15.0		9.1	
	S	12.0	40.1	13.2	9.6	2.4	4.7	17.9	
Saltaire	W	20.7	17.3	25.1	14.5	13.2	2.0	7.1	
	S	43.9	12.3	35.1	1.8	7.0	-	-	
Weekday		22.6	27.3	13.1	8.2	16.3	3.5	9.1	(429)
Saturday		32.4	20.2	14.1	5.8	11.0	2.1	14.4	(327)

For those stations re-surveyed in the second year the figures are:

Weekday	20.2	34.6	8.3	5.1	17.5	4.3	10.0	(278)
Saturday	25.7	24.0	8.4	9.2	11.4	3.6	17.7	(211)

TABLE 5.18 ORIGIN DISTANCE (%)

		Access mode				Mean Access	Mean Egress
		Walk	Bus	Car	Other	distance (m)	distance (m)
Slaithwaite	W	76.3	-	23.7	-	903	903
	S	84.2	5.3	10.5	-	782	782
Crossflatts	H	82.8	10.6	6.5	-	1002	540
	W	68.3	3.2	28.5	-	1009	525
	S	84.6	7.7	7.7	-	982	558
Deighton	S	86.4	13.6	-	-	437	516
	W	91.5	4.3	4.3	-	619	769
Fitzwilliam	W	80.3	6.7	13.0	-	786	1184
	S	56.7	20.0	23.3	-	2392	1687
Bramley	W	79.9	1.6	18.5	-	743	757
	S	87.6	2.5	3.7	6.2	715	548
Saltaire	W	89.2	0.8	9.2	0.8	738	734
	S	100.0	-	-	-	431	781
Mean	W	82.8	3.3	13.7	0.2	807	745
	S	86.4	4.8	7.1	1.7	848	768

TABLE 5.19 ACCESS MODE (%), DISTANCE AND EGRESS DISTANCE

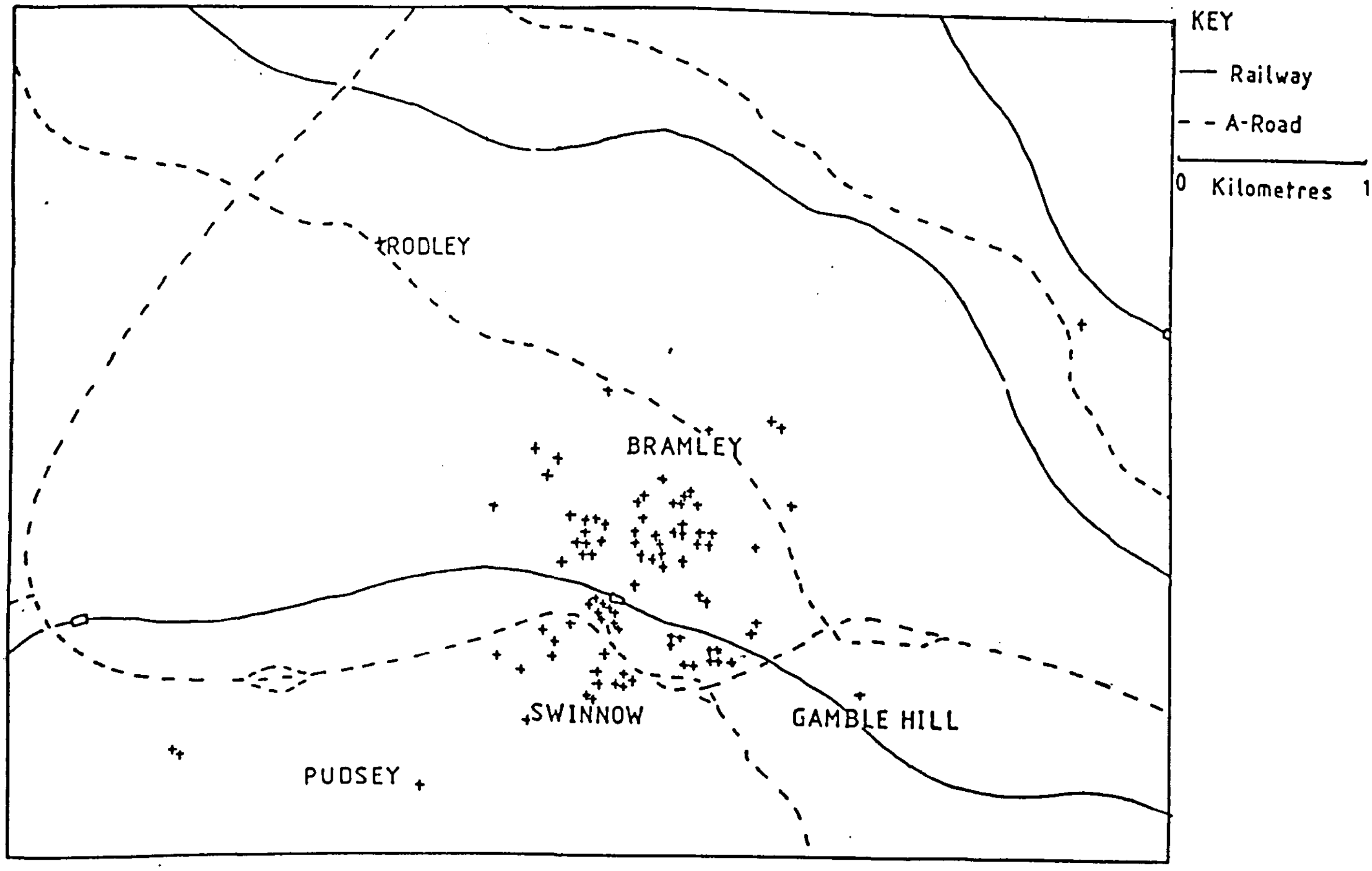
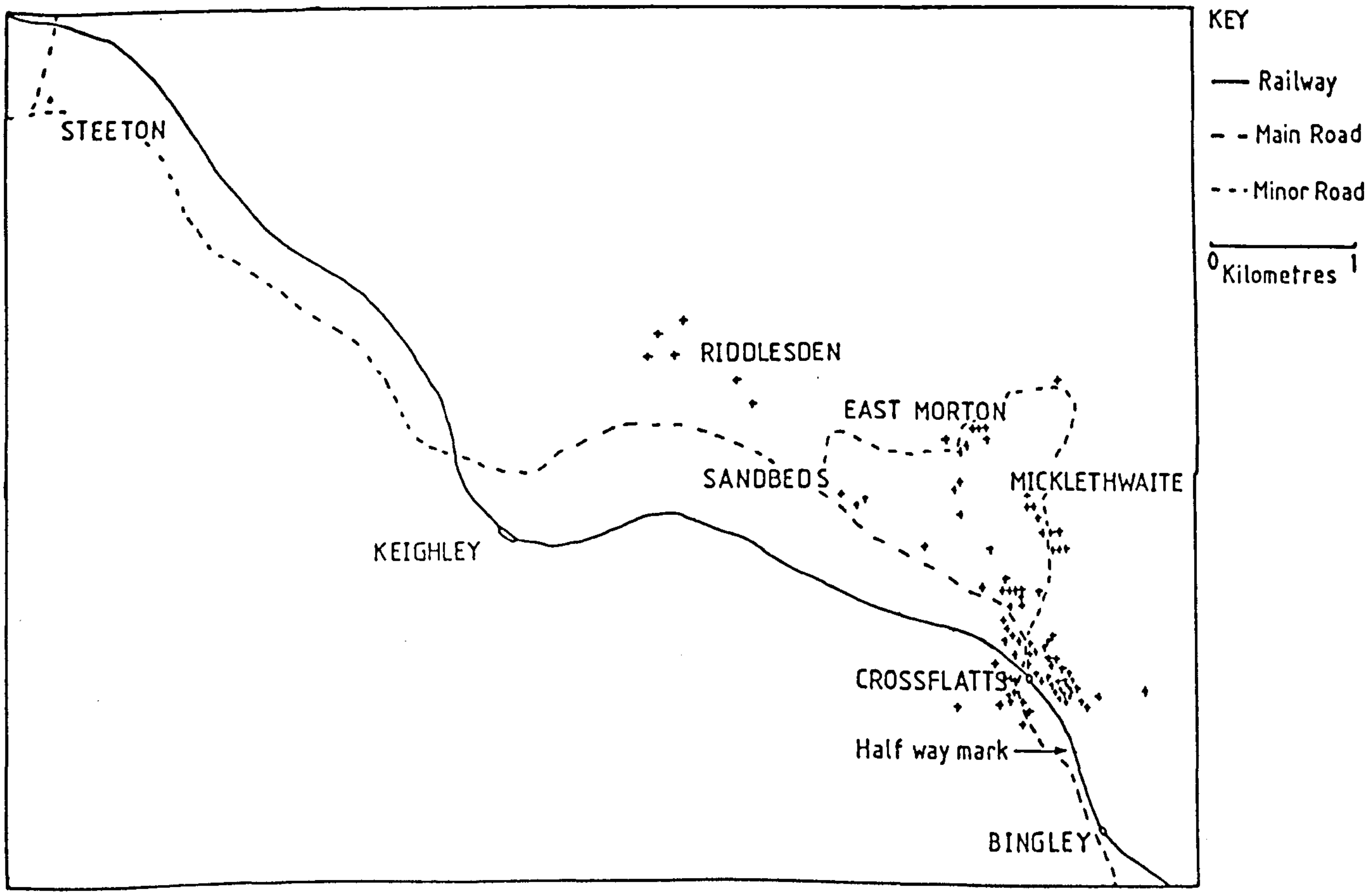


FIGURE 5.3 ORIGIN OF NEW STATION USERS - BRAMLEY (above) + = origin of respondents

FIGURE 5.4 ORIGIN OF NEW STATION USERS - CROSSFLATTS (below)



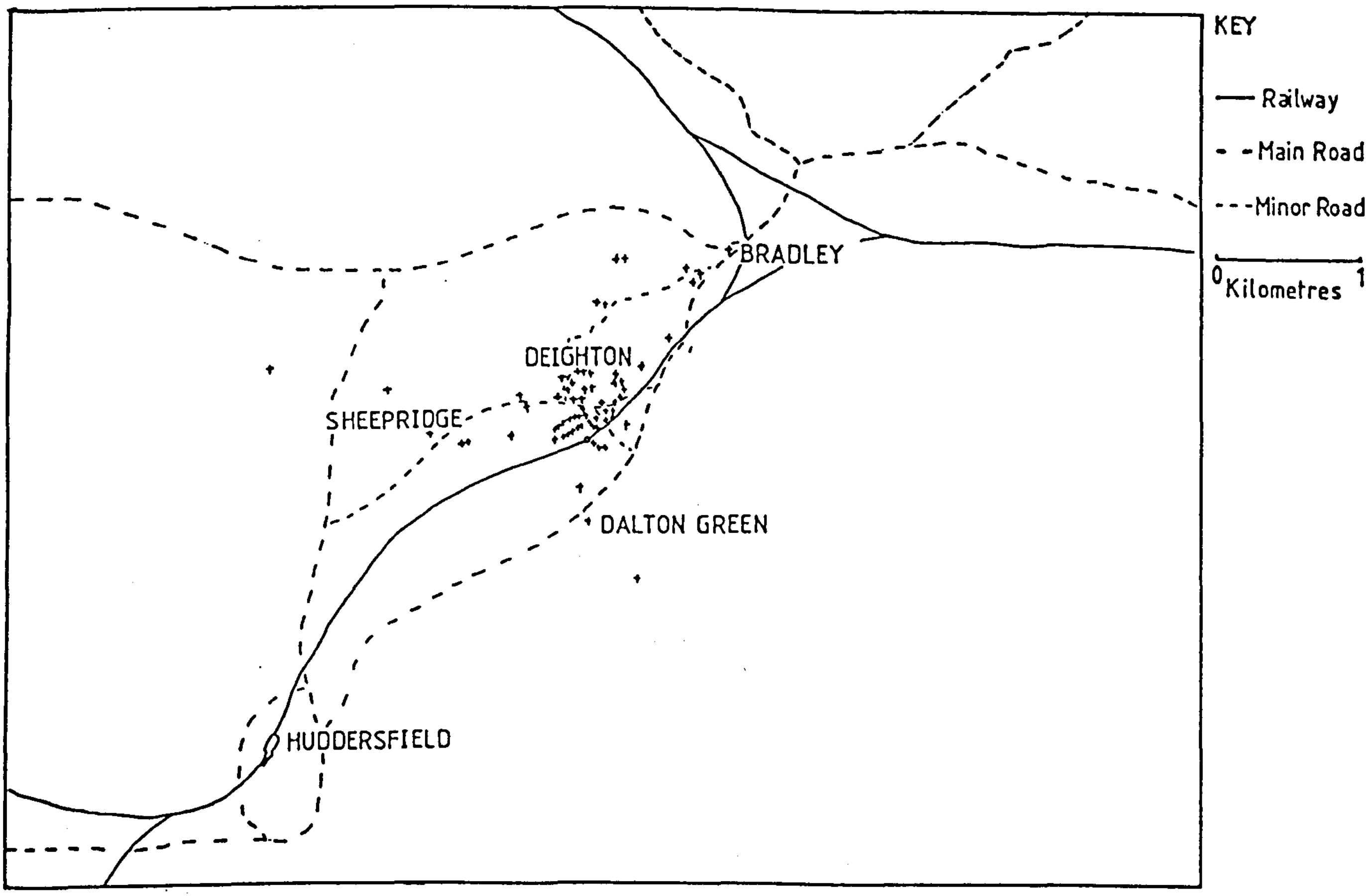
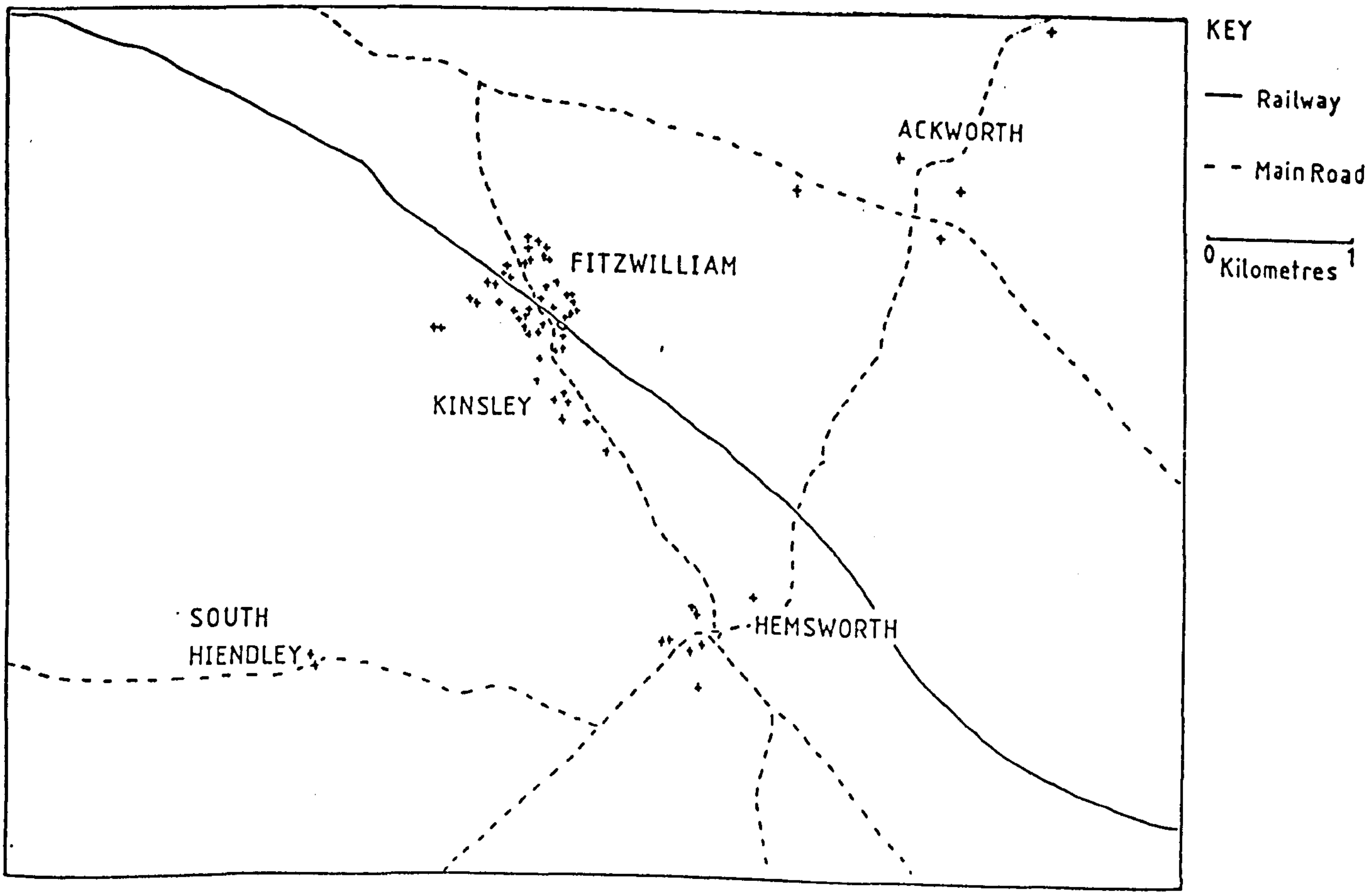


FIGURE 5.5 ORIGIN OF NEW STATION USERS - DEIGHTON (above)

FIGURE 5.6 ORIGIN OF NEW STATION USERS - FITZWILLIAM (below)



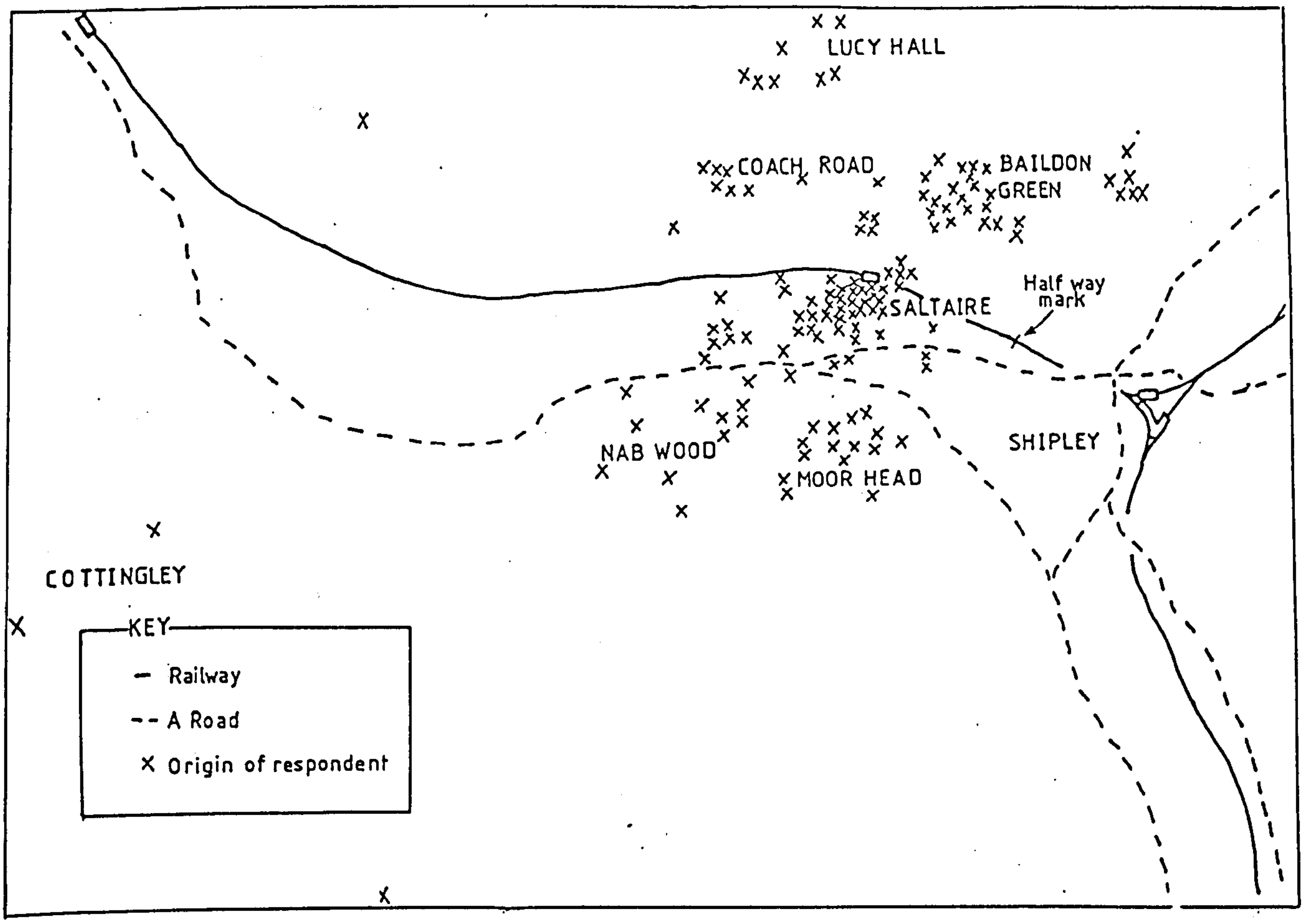


FIGURE 5.7 ORIGIN OF NEW STATION USERS - SALTAIRE (above)

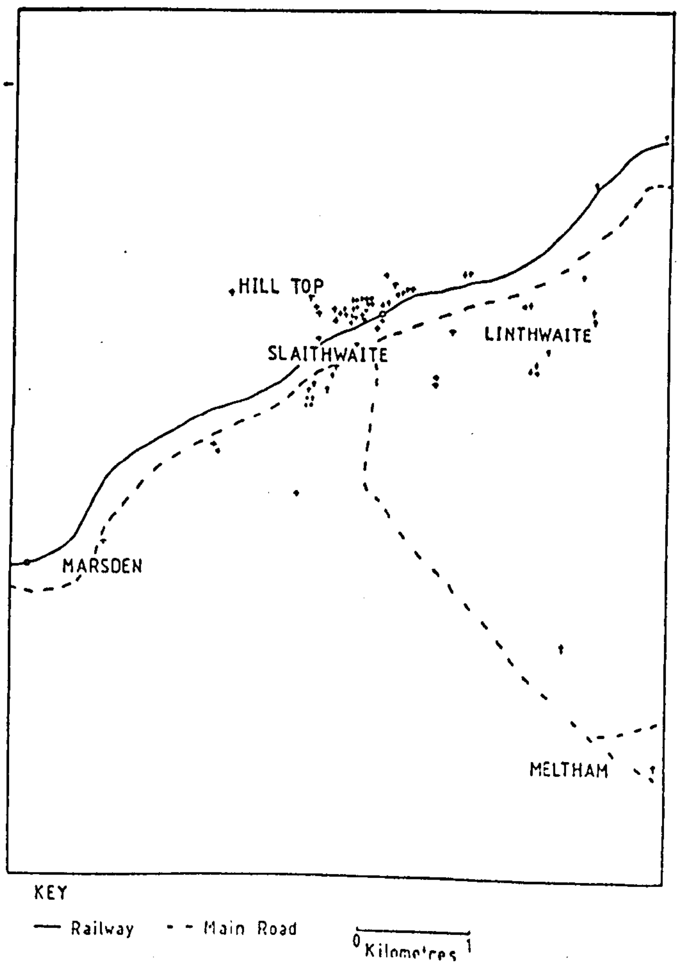


FIGURE 5.8 ORIGIN OF NEW STATION USERS - SLAITHWAITE (left)

ii) Semi urban or rural sites (Crossflatts, Fitzwilliam and Slaithwaite) where the mean access distance exceeds 750 metres, with between 15 and 45% of travellers making use of mechanised access modes, the majority of whom use car, especially as a passenger, as only 28 instances of car parking facilities being used were recorded.

It is noticeable that car is twice as important as an access mode on a weekday as on a Saturday, with walk being even more important on a Saturday than on a weekday. Perhaps surprisingly mean access (and egress distance) is greater on a Saturday than on a weekday, but this is partly due to the outlier effect of Fitzwilliam. On average it can be seen that access distances exceed egress distances with mean values being between 745 m and 848 m.

A multiplicity of ticket types were available for journeys from new stations, with the availability and range of these ticket categories changing during the survey period (see Table 2.7). On a weekday 44% of new station users made use of standard adult tickets, with only 9% making use of child's fare and the County Council concessionary fare for OAPs and the disabled (although this figure doubles if accompanied children are taken into account), whilst a further 4% made use of BR Rail Cards. However, 11% made use of off-peak bargain fares, with a further 20% making use of discounted fares of one type or another. The type and volume of discounted fares depends on the year in which the new stations were surveyed. For those stations surveyed in 1983 (Crossflatts, Deighton, Fitzwilliam and Slaithwaite) the only discounted fare available was through Multi-rider which accounted for 10% of weekday trips. At Bramley in 1984 the Saverstrip facility became available, accounting for 30% of weekday trips, with Multi-rider only accounting for 3%. At Saltaire in 1985 the Metrocard facility was available, accounting for 22.6% of weekday trips, with Saverstrip accounting for only 7.3% and Multi-rider only 1.5%.

By contrast, on a Saturday only 24% of users made use of standard adult fares, with 17% making use of child's and County Council concessionary fares (30% if accompanied children are taken into account). The main ticket type category on a Saturday was the off-peak fare which accounted for 38% of trips (even though Fitzwilliam was surveyed before the off-peak scheme was extended to Saturdays). The three discounted ticket types mentioned above accounted for only 9% of trips, although Day Rover accounted for an additional 4%.

The effect of destination station and ticket type on fare paid is shown by Table 5.20. On a weekday 67% of travellers pay 60 pence or less with only 17% paying a £1 or more, with the corresponding figures for a Saturday being 84% and 10%. However, as Table 5.21 shows, there are also considerable differences between stations in terms of the mean fare paid, with the lowest mean fare being at Slaithwaite (where most of the journeys are to Huddersfield) and the highest mean fares being at Deighton and Fitzwilliam. It should be noted that these figures are susceptible to outliers, for example the mean fare at Deighton on a weekday was boosted by a trip to London, whilst the mean fare at Bramley on a Saturday was boosted by several long distance trips. It can be seen that although, in aggregate, our results are similar to the PTE's estimated mean fare of 60 pence, there are important differences between stations. In addition it should be noted that the mean fare on a Saturday is higher than that on a weekday due to the greater propensity to make very long distance trips.

The first year surveys also provided some information on the socio-economic characteristics of new station users. Table 5.22 shows the occupation of the head of household of travellers. A problem occured in that a number of households could not be categorised, due to the head of household being unemployed, in full time education or retired. These formed the "other" category which, for example, at Deighton on a weekday accounted for 38% of travellers (the Deighton Estate is known to be a local unemployment "blackspot"). Nonetheless it is evident that on a weekday the main group of users are clerical workers, although skilled and unskilled manual workers are also important. Saturday usage, by contrast, appears to be dominated by members of manual worker's households (70% of travellers). It is clear, by comparing our results with those from WYTCONSULT (1976A), that the socio-economic profile of new station users is similar to that of rail users in West Yorkshire as a whole and is different from U.K. rail commuters as a whole, and especially from London and the South East, where occupation groups 1/11 and 111N are dominant (Johnson and Nash, 1981). These results would have important implications if we were interested in equity issues.

In terms of car ownership it can be shown that on a weekday 35.5% of travellers come from non-car owning households, compared to 41.5% on a Saturday. Of those coming from car owning households, on a weekday 48.3%

	0 to 30p	31 to 60p	61 to 99p	£1+	(No of valid cases)
Weekday	18.2	49.1	15.7	17.1	(428)
Saturday	42.5	41.3	5.8	10.4	(327)

TABLE 5.20 FARE PAID (%)

	Weekday	Saturday	Monday to Saturday
* Slaithwaite	52.5	51.4	52.3
* Crossflatts	57.1	78.1	60.6
* Deighton	68.3	38.4	63.3
* Fitzwilliam	58.6	79.5	62.1
Bramley	54.5	98.2	61.8
Saltaire	61.4	58.9	61.0
Mean	59.3	69.7	61.0

TABLE 5.21 MEAN FARE PAID (PENCE) (1984 PRICES)

* Factored up by 6% to take into account fare increase of July 1983.

Social Class		Own Market Research Weekday	Saturday	Total weekly use	Other Studies (to nearest %) WYTCONSULT 1975	Southern Region 1978	Rail Commuters National Travel Survey 1975/6
I/II	Professional/	9.0	4.5	8.3	12		
	Managerial			*10.1		31	26
IIIN	Clerical	33.6	12.1	30.0	22		
(skilled non manual)				*36.4		60	49
IIIM	Skilled	27.8	53.0	32.0	34		
	Manual			*38.8		6	15
IV/V	Un/Semi-	11.2	16.7	12.2	14		
	Skilled			*14.8		3	11
Other		18.4	13.6	17.6			
(Number of valid cases)		(277)	(198)				

TABLE 5.22 OCCUPATION (%)

* Excludes "other" category.

claimed to have a car available (i.e. 31.2% overall), whilst on a Saturday only 40.4% made that claim (i.e. 23.6% overall).

Table 5.23 shows that there are considerable socio-economic differences between users at each of the new stations. Crossflatts users have the highest car ownership rates, high economic activity rates and are mainly in "white collar" occupations, and as such are most similar to the traditional view of rail commuters. At the other extreme, Deighton has low car ownership, low economic activity and a high proportion (of those in work) in "blue-collar" occupations. The rest of the new stations are somewhere in between these two extremes; Saltaire has some characteristics in common with Crossflatts, Fitzwilliam is similar to Deighton due to the dominance of "blue-collar" occupations (Fitzwilliam is a mining village), whilst Bramley and Slaithwaite might be considered to be somewhere in the middle of the socio-economic scale.

Lastly from observational data we were able to note that on a weekday the male/female split was 0.559:0.441, changing to 0.438:0.562 on a Saturday, so that overall 50.7% of those sampled were male. In addition we were also able to observe that on a weekday the mean party size was 1.35, which increased during the school holidays to 1.67 and on a Saturday was 1.96.

5.4 RESULTS OF SECOND YEAR SURVEYS (All results are weighted by the grossing up factors presented in Table 5.8)

As has already been outlined, five out of the six new stations opened in West Yorkshire were resurveyed in order to determine any changes in the pattern of usage. By and large many of the results re-iterated those found in the previous section and hence will not be discussed in detail. However some interesting differences were detected.

(i) We were able to compare the counts of boarding passengers in the second year surveys with the first year surveys. Although comparisons of just two days, at slightly different times of year, are liable to a great deal of variability, such "snap-shots" may tell us something about patronage growth. (More detailed time-series information is presented in Section 9.4). It appeared that weekday patronage at both Crossflatts and Slaithwaite had

		Household car ownership	Household size	No. in full time employment	Activity rate (%)	Manual workers (%)
Slaithwaite	W	0.77	2.92	1.08	36.9	43.5
	S	0.68	3.24	1.16	35.8	69.4
Crossflatts	H	1.23	3.06	1.67	54.6	23.9
	W	1.22	3.29	1.67	50.8	21.1
	S	1.08	3.27	1.81	55.4	33.3
Deighton	W	0.57	3.16	1.05	33.2	38.3
	S	0.45	2.77	1.04	37.5	77.3
Fitzwilliam	W	0.62	3.22	1.26	39.1	48.5
	S	0.78	3.36	1.75	52.1	92.9
Bramley	W	0.74	2.69	1.75	65.1	56.3
	S	0.72	3.57	1.66	46.5	82.8
Saltaire	W	0.88	3.10	1.83	59.0	42.4
	S	1.12	3.85	1.42	36.9	40.0

TABLE 5.23 MEAN SOCIO ECONOMIC CHARACTERISTICS OF NEW STATION USERS

grown substantially, especially during the peak. By contrast weekday patronage at Bramley seemed to have declined substantially. This was mainly due to the decline in peak journeys to Leeds, which may be associated with problems of obtaining a seat on already overcrowded trains. The main change in Saturday usage was at Crossflatts where there was, in relative terms, a substantial decline, with a slight decline also occuring at Slaithwaite. Overall our results suggests that weekday usage has increased by 14% but Saturday usage has decreased by 8%, so that overall Monday to Saturday usage has increased by 10%. These results may be taken to suggest that the learning process is gradual with it taking time for the advantages, or disadvantages, of a new station to become apparent to all potential users.

(ii) It is clear that the results concerning previous mode used will be different in the second year of usage. It should be noted that question 7 of Appendix 1 was revised so as to refer to "one year ago" rather than "before this station was opened" (questions 5 and 8a were similarly revised). We were thus able to develop Table 5.24 along the same lines as Table 5.16. This shows that on a weekday over three-quarters of users would have used the new station one year ago, with over one-eighth being abstracted from bus and less than one-tenth being generated. On a Saturday the situation changes in that less than two-thirds of users would have used the new station one year ago, with 29% being abstracted from bus (although we suspect that some people in this category were reacting to off-peak fare changes) and less than 5% of trips being generated. These results suggest that, in aggregate, weekday usage is beginning to exhibit a stable pattern but Saturday usage (as one might expect) is more volatile.

(iii) It is, though, clear that there have been some important changes in overall usage, and this is reflected by the origin distance travelled by those surveyed in the second year. In particular, it is noticeable from Table 5.26 that on a weekday people are travelling longer distances to get to the station as 39.5% come from beyond one km, compared to 31.8% for the same stations in the first year surveys. By contrast, on a Saturday people seem to be travelling shorter distances to get to the station, as only 19.3% travel more than one km compared to 32.7% for the same stations in the first year surveys. As weekday usage seems to be increasing and Saturday usage decreasing, this suggests that those people living on the margins of a new station catchment

	Generated	As before	Bus	Car driver	Car passenger	(No. of valid cases).
Mean weekday	8.9	76.8	13.0	1.3		(317)
Mean Saturday	4.8	63.5	29.1	-	2.6	(194)
Mean Monday to Saturday	8.2	74.6	15.7	1.5		

TABLE 5.24 PREVIOUS MODE (%) - SECOND YEAR SURVEYS

		0 - 300 m	301-600m	601-800m	801m-1km	1,001km-2km	2.001km-3km	Over 3km
Fitzwilliam	W	28.9	14.1	2.9	-	-	25.9	28.1
Deighton	W	23.9	40.5	7.1	2.4	14.3	-	11.9
Crossflatts	W	39.8	5.3	5.3	-	13.4	17.5	18.5
	S	40.0	16.0	-	12.0	20.0	-	12.0
Slaithwaite	W	13.5	17.0	12.0	5.7	25.5	4.2	22.0
	S	47.7	15.9	15.9	-	4.5	2.3	13.6
Bramley	W	12.0	32.0	16.0	20.0	16.0	-	4.0
	S	29.6	25.9	33.3	3.7	7.4	-	-
Mean weekday		24.5	21.2	8.9	5.9	14.7	8.9	15.9
Mean Saturday		37.0	21.2	22.5	3.9	8.5	0.7	6.2

TABLE 5.25 ORIGIN DISTANCE (%) - SECOND YEAR SURVEYS

	Walk	Bus	Car Driver	Car Passenger	Other	(No. of valid cases)
0 to 800m	96.7	-	0.6	2.5	0.2	(487)
801m to 2km	73.9	8.3	7.6	9.6	0.6	(157)
Over 2km	33.3	19.0	16.7	27.4	3.6	(84)

TABLE 5.26 ORIGIN DISTANCE BY ACCESS MODE (%) (First year surveys - Weekdays and Saturdays)

area are the same people for whom the economic benefits of a new station are most marginal.

5.5 APPLICATIONS OF MARKET RESEARCH

In this section our findings from market research will be used to, firstly, develop a simple trip rate model, and, secondly, to determine the extent of time savings accrueing to new station users. These findings will be discussed in turn:

5.5.1 A Trip Rate Model

From our analysis of origin distance travelled to a new station (Tables 5.18, 5.25) we have determined two important thresholds as 800 m and 2 km. As Table 5.26 shows the three resultant distance bands are related to access mode. Passengers coming from within 800 m of a new station dominantly walk (97%), whilst although walk is still the dominant access mode for passengers originating within 800 m to 2 km of a station, mechanised modes are used by around 26%. For passengers coming from beyond 2 km, although a significant proportion still walk, mechanised access modes are dominant, with car accounting for over 44% of passengers.

Given information on new station usage (from the P.T.S.), on origin distance (from Table 5.18) and on population (from the 1981 Census via SASPAC) it was possible to develop trip rates for the 0 to 800 m and 801 m to 2 km distance bands. These results are given by Table 5.27, which maintains the distinction between weekdays and Saturdays. It can be seen that the trip rates on Saturdays tend to be higher than those on a weekday, although they also exhibit greater variability. Overall, for the 0 to 800 m distance band, we estimate the mean trip rate as being 126.12 weekly trips (i.e. ons and offs) per thousand population, with a standard deviation of 37.62. For the 801 m to 2 km band the mean weekly trip rate declines to 26.40 trips per thousand population, with a standard deviation of 14.10. In addition it should be remembered that we estimate that around 13.2% of weekly usage originates from beyond 2 km.

Clearly use of such trip rates fails to take into account a number of

	Population		Trip Rate	
	0 - 800m	801m - 2km	0 - 800m	801m - 2km
Fitzwilliam	3594	2914	16.75	0.0
Deighton	4017	9990	14.45	2.19
Crossflatts	2989	5764	20.03	6.20
Slaithwaite	2817	5450	24.23	6.40
Bramley	10072	16554	14.79	3.41
Saltaire	5206	9134	30.78	7.70
		Mean	20.17	4.32
		Standard Deviation	5.81	2.69

(A) WEEKDAY

Fitzwilliam	As		11.46	7.07
Deighton	above		25.70	2.30
Crossflatts			20.80	4.63
Slaithwaite			33.55	10.48
Bramley			18.67	2.09
Saltaire			41.39	2.25
		Mean	25.26	4.80
		Standard Deviation	9.86	3.99

(B) SATURDAY

TABLE 5.27 TRIP RATE MODEL
See text for explanation.

factors such as the socio-economic characteristics of the catchment area population, the attractiveness of destinations, the level of rail service and competition from other modes. An attempt was made to take into account the effect of bus competition by excluding the population within the catchment area of bus stops. Using trade-offs identified by Whiteing (1977) areas within 150 m of a bus stop were excluded, except where the service exceeded 200 buses in both directions per day in which case areas within 250 m were excluded. However this failed to significantly reduce the variability of the trip rates (as measured by the coefficient of variation), particularly within the 0 to 800 m band, although the trip rates more than doubled on average.

Lastly it should be noted that overall new station patronage has increased since the initial usage figures given in Table 5.27. These results would imply that the mean weekly trip rate from the 0 to 800 m band would increase to 156.28 per thousand population (up 24%) with a standard deviation of 36.29, whilst the weekly trip rate from the 801 m to 2 km band would increase to 31.16 per thousand population (up 18%) with a standard deviation of 14.11. It is interesting to note that these modifications have reduced the variability of our results.

5.5.2 Measuring Benefits to New Station Users

Another application of our market research would be to determine the level of time and cost savings accrueing to new station users. This was done for all respondents to the first year surveys by adding a number of additional variables to the data set. These were:

(i) Rail access time. This was estimated by considering access distance and access mode. Access distance (defined as origin distance multiplied by 1.2 for bendiness) was converted into time by utilising the mean speeds of access modes used in the WYTCONSULT studies (Lupton, 1976) i.e. walk 5 km/hr, bus 20 km/hr and car 28 km/hr.

(ii) Rail wait time. This was calculated from the difference between the time of interview and the time of departure of the train caught by the interviewee. There are at least two possible sources of error:

a) although interviews were carried out in order of arrival of respondents, not all interviewees were interviewed immediately on arrival at the station.

b) the survey procedure means that there was likely to be a slight bias towards those passengers that arrived earliest at the station, and thus had a high wait time. This was reduced by some on-train surveys of passengers who had arrived within one minute of departure time. To a certain extent these two errors should cancel each other out.

Where a change of trains took place an estimate of the interchange wait time was derived from the British Rail timetable. Where a transfer to bus took place wait time was increased by half the service interval of the feeder bus (i.e. it was assumed there was no co-ordination).

(iii) Rail in-vehicle time. This was estimated from the British Rail timetable.

(iv) Rail cost. This was already included in the data set as fare paid, which was calculated from the Local Fares Directory and Selective Prices Manuals, given information on ticket type.

(v) Rail egress time. Egress distance was calculated as the straight line distance from the destination station to the destination address multiplied by 1.2 for bendiness. This was converted into minutes using the same procedure as for rail access time.

(vi) Best Alternative. In order to determine the mode that would have been used prior to the new station opening the following procedure was used:

a) For abstracted journeys this was simply defined as the previous mode used.

b) For generated journeys, if an alternative mode was used for some journeys to the stated destination then this mode was assumed to be the mode that would have been used. Otherwise, if a different return mode was used, this was assumed to be the mode that would have been used. Otherwise, if a car was stated to be available, then car driver was assumed to be the mode that would have been used. Otherwise, the minimum cost Public Transport mode, which in most cases was bus, would be used. This procedure was adopted as it made maximum use of existing information.

The breakdown of the choice of best alternatives is given by Table 5.28. As might be expected the majority of trip makers (almost two-thirds) would have used bus as their best alternative, with 19% using train and only 14% using private transport.

(vii) Best Alternative access time. Access distance was measured from OS maps or Geographia Town Plans as the straight line distance between the origin address and the nearest relevant bus stop or BR station, again multiplied by 1.2 to take into account bendiness. This was converted to minutes by using the same procedure as for rail access time. For private transport this variable was set to zero.

(viii) Best Alternative wait time. This was calculated as a function of headway (this procedure was also used in Chapters 6.2.1 and 8.2.2) which for rail was:

Wait time = 3.0 + 0.185 Service Interval (from WYTCONSULT, 1976A).
and for bus was
Wait time = 1.46 + 0.26 Service Interval (from Bradford Bus Study (Travers, Morgan and Partners, 1974)).
An upper limit (excluding interchange) was set as 15 minutes. This was particularly important for bus travellers from Fitzwilliam to Leeds. Interchange was dealt with as for rail wait time. For private transport this variable was again set to 0.

(ix) Best Alternative in-vehicle time. For rail and bus this was derived from the British Rail and West Yorkshire P.T.E. timetables respectively. For private transport this was calculated as a function of road distance, based on the link flow speeds determined by WYTCONSULT (1975C) and given in Table 5.29.

(x) Best Alternative cost. For rail and bus this was taken from the BR local fares directory and the PTE regional fare scale respectively, with consideration being made of the availability of concessionary and bargain fares (based on ticket type purchased when using a new station). For private transport, operating costs were based on fuel costs only, being determined by the average price of a gallon of petrol and the fuel consumption of a typical family car (44 km per gallon in urban conditions).

		Bus	Car Driver	Car Passenger	Train
Fitzwiliam	(W)	83.7	3.4	12.8	-
	(S)	63.8	12.9	12.9	10.4
Deighton	(W)	82.7	-	4.4	12.8
	(S)	64.3	-	10.4	25.2
Crossflatts	(W)	62.5	2.9	2.9	31.8
	(S)	62.3	7.4	7.0	23.3
Slaithwaite	(W)	56.8	8.3	14.3	20.6
	(S)	56.1	6.9	2.3	34.6
Bramley	(W)	85.2	-	6.9	7.9
	(S)	77.2	2.4	3.5	16.9
Saltaire	(W)	56.1	5.3	21.4	17.3
	(S)	61.5	6.8	7.0	24.7
Mean (Mon-Sat)		66.4	4.6	9.6	19.3

TABLE 5.28 BEST ALTERNATIVE (%) ('Other' modes excluded)

Rural congested	77
Rural uncongested	86
Urban congested	30
Urban uncongested	63

TABLE 5.29 LINK FLOW SPEEDS (in km/hr)

	Central areas	Non central areas
Long stay	96	-
Short stay	37	20

TABLE 5.30 AVERAGE PARKING CHARGES (Pence, April 1984 prices)

(xi) Best Alternative parking charges. These were determined from information collected by WYMCC in Leeds (reported by Heydecker (1985)) and are given by Table 5.30.

(xii) Best Alternative egress time. This was calculated as the distance from the destination station, nearest relevant bus stop or off-street public car park to the destination address, again adjusted to take into account bendiness. This was again converted to minutes using the same procedure as for rail access time.

In order to determine time savings variables (i) to (v) above will be brought together to form rail generalised cost and variables (vii) to (xii) will be brought together to form the best alternative generalised cost. In order to do this we will need to make use of values of time that might be derived from the Department of Transport or obtained from our own disaggregate model. As a result we will not estimate time savings until section 10.3.1 where up to five different benefit measures will be tested.

5.6 CONCLUSIONS

Market research has been based on interviewing boarding passengers on a representative weekday and Saturday at six new stations in the year after opening and where possible in the following year also. This method proved to be cost effective with 921 responses, although there were some implications in terms of bias.

The main finding of the pilot survey was that rail travel in West Yorkshire is very dynamic, with only 60% of travellers making their journey by rail with the same regularity as one year ago.

The main findings of the first year surveys were that:

(i) New stations have well defined and highly localised catchment areas and provide mainly for local travel needs.

(ii) Most new station users were abstracted from existing Public Transport, although the proportion abstracted from private transport was significant on a weekday, as was the proportion generated on a Saturday.

(iii) There are, however, significant differences between new station users. For example at Crossflatts users are predominantly in "white collar" occupations, whilst at the other extreme, Deighton users are mainly in "blue collar" occupations (if employed).

(iv) There are notable differences between weekday and Saturday usage in terms of journey purpose, propensity to make long distance trips, party size, the number of accompanied children and the occupation of users.

The second year surveys indicated that there may have been some growth in overall usage. Furthermore, there are indications that the people for whom the benefits of travelling via a new station are most marginal are those people who live on the edge of new station catchment areas.

Lastly, we have been able to make use of our market research in order to establish a simple trip rate model (which will be used further in section 9.3.2), whilst we have extended the data collected so as to be able to establish the benefits to new station users (which will be done in section 10.3.1).

PART THREE STATISTICAL MODELLING

CHAPTER SIX AGGREGATE APPROACH

6.1 INTRODUCTION

In 4.3 it was shown that, having examined the range of modelling options available and their implications in terms of resources, one possible way of producing forecasts of new station usage was through the development of an aggregate simultaneous model. This would be based on zonal data, based on the definitions of new station catchment areas established in Chapter Five. In the next section we shall begin to develop an aggregate simultaneous model, firstly by rerunning the PTE and WYTCONSULT regression models and secondly by developing our own approach based on OD data from the 1981 PTS (as described in section 4.2.1). In a subsequent section we shall attempt to update our approach by using data collected by the County Council in 1984. We will then go on to validate the models we have produced by comparing their forecasts with actual usage at new stations.

6.2 DEVELOPMENT OF AN AGGREGATE SIMULTANEOUS MODEL

6.2.1. Re-run of Previous Models

We have seen in section 3.2.2 that both the PTE and WYTCONSULT had developed simple regression models that might be used to predict new station usage. We were unable to obtain the original data sets used to develop these models, thus it was necessary to reconstruct the data from documentary sources before the two models could be rerun.

(i) PTE 1979 regression model. This involved calibrating a model for 36 rural, small town and suburban West Yorkshire stations using average autumn weekday usage figures for 1979, based on the PTS, as the dependent variable and the number of households within 800 m (estimated crudely by the PTE from O S maps) and the number of weekday trains (from the BR timetable) as the independent variables. A simple model of this form was as follows:

Usage = -258.5* + 0.237 Households* + 9.359 Service $R^2 = 0.496$

* = not significant at the 95% level $\bar{R}^2 = 0.465$

Equation 6.1

Interestingly if four outliers were dropped (Cross Gates, Garforth, Ilkley and Keighley) the R^2 increased to 0.761, although still only the service parameter value was significant at the 95% level.

(ii) WYTCONSULT 1975 regression model. This model was based on the following independent variables; households within one km of a station (listed in WYTCONSULT, 1976A), the level of service in the direction of travel, the distance to a major employment centre and journey time (WYTCONSULT, 1975B). However, it is not clear how these last three variables were specified. We have therefore had to make our own definition of the nearest major employment centre (either Leeds, Bradford or Huddersfield in most cases) and then measure the relevant journey time (from the BR timetable) and distance. The dependent variable was the five day one-direction usage figures derived from the 1975 WYTCONSULT rail survey, which we have already seen (in section 2.2.3) omits certain types of travellers. This model was calibrated for 40 stations (the same 36 as the PTE model plus Clayton West, Skelmanthorpe, Pontefract Baghill and Moorthorpe). Our re-run took the form:

Usage = 460.52* + 32.48 Service + 0.08 Population - 16.10 Proportion of No

Car Owners* - 14.59 Time* + 36.26 Distance*+ $R^2 = 0.450$

$R^2 = 0.369$

Equation 6.2

* = insignificant at the 95% level + = wrong sign

From equation 6.2 it can be seen that four out of six parameter values are insignificant at the 95% level, whilst the distance parameter value is of the wrong sign, possibly due to collinearity with the journey time variable.

Equations 6.1 and 6.2 suggest that simple regression models have a number of practical disadvantages. In particular, with the exception of frequency of rail service and population, most of the variables are insignificant and both models explain less than half of the variation in station usage. Moreover, it is difficult to interpret certain parameter values, especially the intercept (Jeanes and Lesley (1984) in applying a variation of the PTE model obtained negative usage figures). Hence a number of improvements to these models might be suggested:

(i) Develop 1981 (rather than 1975 or 1979) as the base year for calibration, due to the availability of PTS and Census information, and as the most suitable year prior to the new station programme (1982 was affected by industrial disruption).

(ii) Rather than consider station usage as a whole as the dependent variable, break up usage into its constituent flows. This will have a secondary benefit of increasing the number of observations in the calibration data set.

(iii) Develop a gravity type formulation by including variables that reflect generation characteristics (such as origin population), attraction characteristics (such as destination employment), level of service and distance. Thus we shall be developing an aggregate simultaneous model of the form shown by equation 6.3:

$$T_{ij} = f(O_i, D_j, SE_i, GC_r, GC_c) \qquad \text{Equation 6.3}$$

where T_{ij} = Number of trips between stations i and j on an average weekday

O_i = Population at origin zone i,

D_j = Measure of attraction of destination zone j,

SE_i = Socio-economic characteristics of population at i,

GC_r = Generalised cost (or level of service) of rail between i and j.

and GC_c = Generalised cost (or level of service) of competing modes between i and j (car and/or bus).

(iv) Consider other functional forms in addition to the linear additive model. These improvements will be implemented in the next sections.

6.2.2 Data for the Aggregate Simultaneous Model

The dependent variable was defined as the number of trips from i to j or j to i per average autumn weekday (FLOW). This variable was derived from the probability estimation method described in section 4.2.1, using the November 1981 and 1982 PTS. Table 6.1 lists the independent variables that were considered. In addition, we shall find the generalised cost (or more accurately generalised time) approach (discussed in section 3.3.2) useful. For our purposes this will be defined as:

Variable name/type	Definition	
DEPENDENT VARIABLE		
FLOW	Number of trips by rail from i to j and j to i (i.e. ons and offs). From 1981/2 PTS	
INDEPENDENT VARIABLES		
OPOP	Usually resident population within a straight line distance of 800 m of the origin station, derived from the 1981 Census via SASPAC.	
OPOP3	Usually resident population within 800 m and 2 km of the origin. This too was derived from the 1981 Census. Where catchments overlapped population was reallocated to the nearest station.	
DPOP	Number of workplaces within a straight line distance of 800 m from the destination station. This was derived from information given by WYTCONSULT (1976A) and updated with information from Local Authority Planning Departments.	
RJT (BJT, CJT)	Rail (Bus, Car) journey time (in minutes)	
RS	Rail frequency (Number of weekday trains, both directions)	
BS	Competing bus frequency (Number of weekday buses, both directions)	
OBS (DBS)	Number of bus stops within 800 m of the origin (destination) station	
SP	Average rail speed between origin and destination stations (in km per minute)	
PA	Parking cost dummy variable (1 = if destination Leeds/Bradford, 0 = else)	
BINT	Bus interchange dummy variable (1 = interchange required, 0 = else)	
RF(BF)	Rail (Bus) Fare (in pence)	
GCRA	Generalised Cost of Rail	See
GCBU	Generalised Cost of Bus	text for
GCCA	Generalised Cost of Car	definitions
GCOTH	Composite measure of inter modal competition	
DRX	Number of excess workplaces (workplaces minus economically active population) within 800 m of destination station	
ORX	Number of excess workers (economically active population minus workplaces) within 800 m of origin station	
OEA	Ratio of the number of workers to total population within 800 m of origin station	
NCO	Proportion of non car owning households within 800 m of origin station	
RSOC	Number of residents in social classes 1 and 11 (Managerial and Professional) within 800 m of origin station (from 1981 Census, 10%)	
RSOC2	Number of residents in social classes 1, 11 and 111 (Managerial, Professional, Clerical) within 800 m of origin station (from 1981 Census, 10%) divided by total population	
In addition the following variables were used in calculating independent variables.		
OW(DW)	Median access (egress) distance in metres	
OWA (DWA)	Proportion walking to/from origin (destination) station	
OB (DB)	Proportion using bus to/from origin (destination) station	
OC (DC)	Proportion using car to/from origin (destination) station	
UDIST	Urban road distance between origin and destination stations (in km)	
RDIST	Rural road distance between origin and destination stations (in km)	
RAD	Rail distance between origin and destination station (in km).	

The prefix L denotes a logarithm has been taken.

TABLE 6.1 VARIABLES USED IN THE DEVELOPMENT OF THE AGGREGATE MODEL

Generalised Cost Rail (GCRA) = RJT + WK + WT + RF/VOT Equation 6.4

Generalised Cost Bus (GCBU) = BJT + BWK + BWT + BF/VOT Equation 6.5

Generalised Cost Car (GCCA) = CJT + Operating Cost/VOT + Parking Cost/VOT Equation 6.6

VOT denotes value of time which, from Highway Economic Note 2 (Department of Transport, 1980), for behavioural non-working in-vehicle time was given as 65.3 pence per hour (at average 1979 prices). In accordance with Department of Transport guidelines this was factored up to November 1981 prices, using the Retail Price Index, to give 74.4 pence per hour, with walk and wait time valued double this figure. It should be noted that there is only limited empirical evidence for assuming that walk and wait times have the same parameter values. For example disaggregate studies by Bradley et al. (1986) and Daly and Zachary (1977) have suggested that wait time may be valued more highly than walk time.

The data requirements of the three generalised cost measures will now be discussed in turn:

(i) <u>Rail generalised cost</u>

a) In-vehicle time. This was derived from the 1981-2 BR Passenger Timetable. It was not possible to make detailed consideration of delays even though it was known that around 9% of local trains were over five minutes late (West Yorkshire P.T.E., 1983)

b) Walk time. This was calculated using the median area radius for the origin (OW) and destination (DW) stations derived from WYTCONSULT (1976A). Consideration of access/egress mode was taken into account and distance was converted into time using the same mean speeds as in section 5.5.2.

c) Wait time was based on the same headway function as in 5.5.2. The headway interval was defined as the number of weekday trains divided by the service period (normally 17 hours). This approach fails to take into account the problems of unreliability and of scheduled delay.

d) Fare was derived from the Selective Prices Manual and the Local Fares Directory for 1981. All internal flows were assumed to make use of the Multi-rider facility (0.9 times the ordinary single fare) whilst all external flows were assumed to make use of day returns. These simplifications were relaxed in later analysis.

(ii) Bus generalised cost

a) In-vehicle time. This was derived from the PTE District timetables (September 1980 series) with again no consideration of delays.

b) Walk time. No data source for this variable existed. This was estimated by dividing the walk time of rail by the number of bus stop pairs on competing bus services that pass through the station catchment area.

c) Wait time. This was based on the same headway function as in 5.5.2 with the normal service period again defined as 17 hours.

d) Fares were based on the PTE regional fare structure (i.e. they were mileage related). An upper limit of 55 pence was used, based on the cost of a monthly Metrocard (£22). The assumption of regular usage is consistent with the assumption made for rail.

e) Interchange. An interchange penalty of four minutes (over and above any additional waiting time), where applicable, was assumed, in line with the findings of LGORU (1973).

(iii) Car Generalised Cost

a) In-vehicle time was calculated from road distance using the link flow speeds given in Table 5.29.

b) Operating Costs. In accordance with Department of Transport procedures only fuel costs were taken into account. A simplified approach was based on the fuel consumption for an average family saloon car (Ford Escort 1300L), which was 44 km per gallon for urban conditions and 62 km per gallon for rural conditions. Petrol was priced at £1.58 per gallon. (These figures were derived from the November issue of Drive magazine (Automobile Association, 1981)).

c) Parking costs. WYTCONSULT studies suggested that parking charges were only significant in Leeds and Bradford (1977A), and that up to half of parking costs may be paid for by the employer so that mean 1981 out of pocket charges may be as low as 30 pence.

The generalised cost approach adopted has a number of weaknesses:

(i) It uses a non local value of time. Local valuations might be provided by WYTCONSULT (30.9 pence per hour at 1975 wage levels) or from the RP models developed with the West Yorkshire corridor data set and reported by MVA et al. (1986), which gave values of time for Public Transport, Car, Walk and Wait of 144, 222, 330 and 186 pence per hour respectively (at mid 1985 prices), although these results had wide confidence intervals. In addition it will be shown in Chapter Seven that this study will be able to develop its own values of time.

(ii) The weightings for out of vehicle time and interchange are based on inadequate evidence.

(iii) Certain modal attributes such as flexibility, reliability and comfort have not been taken into account.

(iv) No consideration was made of the value of working time although typically up to 6% of rail travel and 15% of car travel may be so classified (Leitch, 1977).

(v) It fails to take into account differences between journey purpose and between times of the day.

(vi) Use is made of estimated times and costs rather than actual times and costs.

These simplifications can only be justified on the basis of expediency.

Initially our data set consisted of 120 flows of over 25 per weekday, but data on the attraction of destinations, egress mode, egress distance and road speeds did not readily exist for locations outside West Yorkshire. As a result only 99 flows from 40 stations could be studied. It should be noted that the data set included a number of different types of station, including possible traffic attractors (such as Keighley, Dewsbury, Halifax and Wakefield Kirkgate), large scale traffic generators (for example commuter stations such as Cross Gates, Garforth, Ilkley and Horsforth) and lightly used stations (such as Altofts, Ravensthorpe and Stocksmoor). Excluded from analysis were the main central area rail termini (Leeds City, Bradford Forster Square and Interchange, Huddersfield and Wakefield Westgate) where inter-city flows are likely to be important.

6.2.3 Specification of the Aggregate Simultaneous Model

This section involves examining the data in order to determine what form the aggregate simultaneous model should take. It should be noted that our work is subject to pretest bias (Wallace, 1977) due to the use of the same data set to calibrate and specify a model. Due to limited data this problem was unavoidable.

Models were developed by adopting a classic multiple regression approach, whereby the Gauss-Markov theorem is assumed to apply, so that the ordinary least squares (OLS) estimator of each coefficient is the best linear unbiased estimate (BLUE). All calculations were performed using the SAS computer package, with use being made of the REG procedure unless specified otherwise.

In the rest of this section we shall briefly consider firstly the functional form of the model; secondly, we shall attempt to estimate time and money parameters for rail; thirdly, we shall consider which socio-economic variables are most relevant; fourthly, we shall determine how to best include competition from other modes and lastly we shall look at the possibility of using proxy variables and trip rates.

(i) Functional Form. Our initial starting point was the gravity-type model formulation given by equation 6.3. Leaving aside the exact specification of the attributes for a moment, we may hypothesise that a log linear (double log or multiplicative) or semi log model will be superior to a linear (additive) model, as previous analysis (for example in 6.2.1) suggests that interaction between independent variables may be important. Table 6.2 shows the goodness of fit for the three functional forms applied to an initial model run.

	R^2	Adjusted R^2 ($\bar{R}^2$)
Log linear	0.4890	0.4673
Semi log	0.4447	0.4211
Linear	0.3990	0.3734

TABLE 6.2 COMPARISON OF GOODNESS OF FIT OF FUNCTIONAL FORMS

To compare the double log and semi log forms it is sufficient to compare the R^2 (Granger and Newbold, 1976), and it can be seen that the double log form is superior. A direct comparison between the double log and linear form is not as valid (Rao and Miller, 1971). A statistical test may be provided by the log likelihood ratio: (as used, for example, by Mills, 1978).

$$LLR = \frac{-N}{2}(M.S.E.) + (\lambda - 1) \Sigma \log Y \qquad \text{Equation 6.7}$$

where M.S.E. = mean square error, $\Sigma \log Y = (\log Y).N$, N = Number of observations and $\lambda = 0$ if log and = 1 if linear. Thus for the linear model:

$$LLR = \frac{-99}{2} \log 18027 = -485.08$$

and for the double log model:

$$LLR = \frac{-99}{2} \log 0.441 + \log 4.454\ (99) = -107.36.$$

This reinforces the impression that a double log form is superior to the linear form.

(ii) Specification of Rail parameters

An attempt was made to estimate the parameter values of rail fare, in-vehicle time, wait time and walk time separately. In order to avoid collinearity between the fare and in-vehicle time, in-vehicle time was divided by distance (RAD) in order to produce a speed variable the negative coefficient of which could be considered a time parameter (this approach was adopted by Tyler and Hassard, 1973). Our attempts were unsuccessful in that only the wait time parameter value was significant at the 5% level. Because of this we decided to use a generalised cost type measure.

(iii) Relevant socio-economic variables

As Table 6.1 indicates a number of socio-economic variables were tested of which the most important were; OPOP (population within 800 m of the origin station), OPOP3 (population within 801 m and 2 km of the origin station), RSOC (the proportion of OPOP in social classes 1 and 11 (Managerial and Professional)) and DRX (the number of excess workplaces within 800 m of the destination station).

(iv) Competition from other modes

As with rail an attempt was made to estimate parameter values for the separate attributes of bus and rail but the only coefficients that were significant at the 5% level were the bus interchange dummy variable (BINT), the car parking charge dummy variable (PA) and the measure of the number of bus stops (OBS), and even these were not always of the right sign. In addition attempts were made to specify bus and car generalised cost with respect to rail in order to determine whether a difference or a ratio formula should be used (this issue was examined by Quarmby, 1967). As the difference formulation led to a bus cross elasticity of the wrong sign, a ratio formulation was preferred.

(v) Use of proxy variables and a trip rate approach

Dummy variables were used to model destination choice by introducing variables for Leeds, Bradford and Huddersfield (with other destinations acting as an excluded variable) which took a value of 1 if that particular destination is being modelled. In addition, dummy variables were introduced to take into account distance (taking a value of 1 if the distance travelled is greater than 6 miles, 0 otherwise) (after Hartley, 1979B) and differences between bus and rail fare, frequency and journey time. However, the introduction of dummy variables failed to lead to significant improvements in goodness of fit.

In addition the dependent variable was redefined as LRATE = LOG (FLOW/OPOP) as such a trip rate approach should reduce collinearity problems and gain one degree of freedom. This, however, failed to produce major increases in the significance of the remaining parameter values and leads to a deterioration in goodness of fit, as measured by $\bar{R}^2$.

6.2.4 Calibration of the Aggregate Simultaneous Model

As a result of the work outlined in section 6.2.3 our initial preferred Aggregate Simultaneous Model (ASM) took the form (t-statistics in brackets)

$$\begin{aligned} \text{LFLOW} = \; & \underset{(3.025)}{5.496} + \underset{(2.617)}{0.380}\ \text{LOPOP} + \underset{(1.733)}{0.164}\ \text{LOPOP3} + \underset{(2.034)}{0.246}\ \text{LRSOC} + \\ & \underset{(6.678)}{0.269}\ \text{LDRX} - \underset{(-2.269)}{1.341}\ \text{LGCOTH} - \underset{(-4.307)}{1.239}\ \text{LGCRA} \quad R^2 = 0.539 \quad \bar{R}^2 = 0.509 \end{aligned}$$

Equation 6.8

where LGCOTH = LOG (GCRA/(GCRA + GCBU + GCCA))

A key issue here is the specification of the LGCOTH variable with at least four specifications possible:

LGCOTH = LOG(GCRA/(GCRA + GCBU + GCCA)) Equation 6.9

LGCOTH = LOG(GCRA/(GCBU + GCCA)) Equation 6.10

LGCOTH = LOG(GCRA + GCBU + GCCA) Equation 6.11

LGCOTH = LOG(GCBU + GCCA) Equation 6.12

These four specifications are evaluated by Table 6.3.

Equation	LGCOTH Parameter value (t-stat)	LGCRA Parameter value (t-stat)	Model R^2	LGCOTH Condition Index	LGCRA Condition Index
6.9	-1.341 (-2.269)	-1.239 (-4.307)	0.539	57.74	90.22
6.10	0.541 (1.382)	-1.625 (-3.478)	0.523	82.37	121.47
6.11	1.341 (2.269)	-2.580 (-3.652)	0.539	83.88	205.56
6.12	0.765 (2.288)	-2.002 (-4.163)	0.539	75.93	125.64

TABLE 6.3 COMPARISON OF LGCOTH SPECIFICATIONS

(Condition index = square root of the ratio of the largest eigenvalue to each individual eigenvalue)

Table 6.3 suggests that specification 6.10 is dropped due to its lower R^2 and 6.11 is dropped due to the greater amount of collinearity, as measured by the condition index. There is, however, little to choose between specifications 6.9 and 6.12. Specification 6.12 has advantages in terms of the simplicity of the derivation of elasticities and the improved significance of the LGCOTH term. This is counter balanced by the greater significance of the LGCRA parameter in 6.9 and the lower degree of collinearity between LGCRA and LGCOTH. As a result we may initially accept equation 6.8 (and LGCOTH specification 6.9) as all variables are significant at the 10% level and of the right sign.

The use of the generalised cost formulations in equation 6.8 make the derivation of elasticities difficult. As the mean of the GCRA may be split proportionately as fare 0.451, in vehicle time 0.226 and out of vehicle time 0.323 direct elasticities can be derived from the LGCRA parameter value as:

Fare elasticity	-0.55
Journey time elasticity	-0.28
Walk/Wait elasticity	-0.40

However the variable LGCOTH includes rail generalised cost, as well as that of bus and car. From the parameter value of LGCOTH it may be estimated that the elasticity of rail generalised cost is -0.610, whilst the cross elasticity with respect to car generalised cost is 0.101 and to bus generalised cost is 0.483. Given the breakdown of GCCA and GCBU this implies price cross elasticities with respect to car and bus of 0.078 and 0.200 respectively, which are significantly lower than the 0.183 and 0.696 respectively estimated by Nash and Whiteing (1984), although this may partly reflect the problems we have had in modelling competition from other modes.

If the two components of rail generalised cost are assumed additive then direct rail elasticities become:

Generalised Cost Elasticity	-1.849
Fare elasticity	-0.825
Journey time elasticity	-0.418
Walk/Wait time elasticity	-0.600

The fare elasticity is on the high side and suggests that local rail travel is exhibiting elasticities associated with long distance leisure travel (range -0.7 to -1.4) rather than short distance commuter travel (range -0.1 to -0.3) (Nash, 1982). However, our findings are consistent with previous work carried out in West Yorkshire by Hartley (1979) with a different data set and methodology and suggests that the -0.37 rail price elasticity used in the METS model of West Yorkshire (Department of Transport, 1984) is an underestimate. The journey time elasticity is consistent with the range of -0.3 to -0.6 identified by Wabe (1969) and Hepburn (1977) in separate studies of commuting in London. The out of vehicle time elasticity is similar to that of -0.77 identified by Daly and Zachary (1977), although in our study the composition is walk -0.402 and wait -0.198 (due to the fact that in our data set walk time has been estimated as being, on average, much longer than wait time), whilst in their study it was walk -0.14 and wait -0.63.

Thus it can be concluded that, at least in terms of direct elasticities, the parameter values of equation 6.8 give plausible results. In addition it is interesting to note that the parameter value of OPOP is 2.32 times greater than that of OPOP3, which is lower than our findings for the trip rate model of 5.5.1

(where we found rates in the 0-800 m zone to be 4.78 times greater than in the 801 m to 2 km zone).

6.2.5 Some problems with the Aggregate Simultaneous Model

In this section we shall examine a number of statistical problems with an ASM of the form of equation 6.8.

(i) Explanatory power. The explanatory power of such a model is limited as the R^2 is only 0.54. Analysis of residuals shows that only 41 (out of 99) actual flows lie within the 95% confidence limits of the model's predictions. There are six flows with residuals of over 200 (Cross Gates, Garforth, Ilkley, Horsforth, Keighley to Leeds and Keighley to Bradford) perhaps indicating rail heading. The exclusion of these outliers led to an increase in the R^2 to 0.66.

(ii) Multicollinearity. It was suspected that the model may be affected by collinearity (correlation between the independent variables), particularly between the origin population variables and between the generalised cost variables. These suspicions were confirmed in part by Tables 6.4 and 6.5.

It can be seen that the use of the composite variable GCOTH in Table 6.5 has decreased the correlation between the generalised cost variables, as shown in Table 6.4. The absolute values of the coefficients in Table 6.5 are not a cause for concern, but if we apply the rule of thumb that multicollinearity is likely to be a problem if the correlation between two variables is greater than the correlation of either or both variables with the dependent variable (Kennedy, 1979) a large number of variables may be said to be collinear, with RSOC, DRX, GCRA and GCOTH particularly affected.

(iii) Heteroscedasticity. Visual inspection of scatterplots suggested that heteroscedasticity (non constant variance of the error term) is not a great problem, having been reduced by the double log form. This was tested for by using a Park-Glesjer test whereby the absolute value of the residual was regressed against each independent variable. This indicated that only LGCRA may be affected, with a parameter value of -0.237 and t-statistic of -1.793 (and even this is not significant at the 95% level). A possible correction might be provided by adopting a form of weighted least squares (Stewart and Wallis, 1981, pages 254 to 256).

	GCRA	GCCA	GCBU	GCOTH
GCRA		0.804 (0.0001)	0.832 (0.0001)	-0.451 (0.001)
GCCA			0.749 (0.0001)	-0.780 (0.0001)
GCBU				-0.758 (0.0001)

TABLE 6.4 CORRELATION MATRIX - GENERALISED COST VARIABLES

	LFLOW	OPOP	OPOP3	RSOC	DRX	GCRA	GCOTH
LFLOW		0.336 (0.0006)	0.309 (0.0019)	0.035 (0.729)	0.535 (0.0001)	0.038 (0.707)	-0.343 (0.005)
OPOP			0.573 (0.0001)	0.097 (0.340)	-0.012 (0.904)	-0.028 (0.782)	0.118 (0.256)
OPOP3				-0.330 (0.0009)	0.099 (0.327)	-0.006 (0.951)	0.065 (0.524)
RSOC					-0.194 (0.054)	-0.129 (0.202)	0.111 (0.274)
DRX						0.544 (0.0001)	-0.575 (0.0001)
GCRA							-0.454 (0.0001)

TABLE 6.5 CORRELATION MATRIX - INDEPENDENT VARIABLES

Both Tables give the Pearson product moment correlation coefficient and the significance probability of the correlation.

An alternative correction may be to specify a different functional form for the generalised cost of rail variable. For example using a semi log form (GCRA) and specifying LGCOTH as LOG(GCBU + GCCA) removes heteroscedasticity, as measured by a Park-Glesjer test. Such a correction gives a model of the form: (t-statistics in brackets)

$$\underset{(-0.790)}{LFLOW = -1.468} + \underset{(2.866)}{0.423\ LOPOP} + \underset{(1.542)}{0.147\ LOPOP3} + \underset{(2.047)}{0.248\ LRSOC} +$$

$$\underset{(6.507)}{0.291\ LDRX} + \underset{(1.721)}{0.507\ LGCOTH} - \underset{(-3.971)}{0.007\ GCRA} \quad R^2 = 0.532,\ \bar{R}^2 = 0.502 \qquad \text{Equation 6.13}$$

Equation 6.13 gives an intercept with an antilog of 0.23, but it is not significant. Given the mean value of GCRA it can be shown that the direct generalised cost elasticity of rail is around -1.45 (i.e. 28% lower than the constant elasticity from equation 6.8). A model of this form has some advantages over a constant elasticity formulation, but it still fails to improve on the goodness of fit.

6.3 UPDATING THE AGGREGATE SIMULTANEOUS MODEL

6.3.1 Re-running the Aggregate Simultaneous Model

It was apparent that the OD data estimated from the PTS through a probability estimation method involved some input error. A more reliable data set might be provided by the 1984 County Council survey, discussed in section 4.2.2, even though this too was known to contain a number of anomalies. In addition it should be noted that, all other things being equal, internal flows were lower in 1984 than in 1981. Thus despite lowering the cut-off point to 20 only 80 flows could be analysed (compared to 99 previously).

A second criticism of our work in developing an ASM was that fares had been crudely represented, whilst use of 1981 data had meant that off-peak and Saverstrip fares were not available.

The 1984 Survey provided information on ticket type used. The various ticket types could be collapsed into four main groups: ordinary, Saverstrip/Multirider, staff and Day Rover for two weekday periods: peak (up to 0930 and 1500 to 1800 hours on weekdays) and off-peak (all other time periods).

Thus for each OD pair a mean rail fare could then be calculated as: (in pence).

$$\alpha_t \, (\phi \, F_t + S \; 0.83 \, F_t + RD + P \; 0.25 \, F_t) \qquad \text{Equation 6.14}$$

where	t	=	time of day (1 = peak, 2 = off-peak)
	α_t	=	proportion of rail travellers in time period t
	F_t	=	Ordinary fare in time period t (F_2 = 30 if journey less than or equal to 8 miles, = 60 if greater than 8 miles)
	ϕ	=	proportion of travellers using ordinary tickets
	S	=	proportion of travellers using Saverstrip or Multirider
	D	=	proportion of travellers using Day Rover
	P	=	proportion of travellers using Staff Passes
	R	=	F_t if F_t less than 66, = 66 if F_t greater than or equal to 66.

Similarly a mean fare for bus could be calculated as:

$$\beta_t \, (\phi \, F_t + S \; 0.83 \, F_t + \lambda \, M \; 55) \qquad \text{Equation 6.15}$$

where	β_t	=	proportion of bus travellers in time period t
	F_t	=	standard fare in time period t (F_2 = 30)
	S	=	proportion of travellers using Saverstrip
	ϕ	=	proportion of travellers paying the standard fare
	M	=	proportion of travellers using a Metrocard
	λ	=	0 if journey less than 6 miles, = 1 else.

N.B. No information was available on the uptake of Day Rover tickets.

There was lack of data to determine α and β but this was eventually estimated from rather dated information collected for the Bradford Bus study (Travers, Morgan and Partners, 1974) where $\alpha_1 = 0.61$, $\alpha_2 = 0.39$, $\beta_1 = 0.54$ and $\beta_2 = 0.46$. This was consistent with subsequent figures released by the P.T.E. for 1984-85 which suggested that 52% of weekday travel by bus was during peak periods compared to over 60% for rail. In practice, though, it it likely that the values for α and βs will vary greatly between OD pairs. For bus S was set at 0.16 and M at 0.43, as Pickup (1984) indicates that 15% of bus users make use of Metrocard, whilst information from WYTCONSULT (1977A) suggests that 34.75% of bus journeys are greater than six miles. As a result where $\lambda = 0$, ϕ = 0.84, otherwise $\phi = 0.41$. It should be noted that we have not taken into account the effect of concessionary fares, which means we have assumed that at individual station catchment areas the uptake of such fares is uniform across both modes.

Given this new data for 80 flows, with updated fare information we were able to recalibrate the double log (equation 6.8) and semi log (equation 6.13) models, as shown by Table 6.6. From this Table the following observations can be made.

(i) There are significant changes in parameter values as measured by a z-test. For the double log model there are significant changes at the 95% level for all parameters, except LRSOC, with the change in LGCOTH being particularly severe. For the semi log model all parameter values showed significant changes, except for the intercept, LRSOC, LDRX and LGCOTH. This might suggest that this model formulation possesses greater stability than the double log model.

(ii) There are reductions in the significance of most parameter estimates. For both recalibrated models the intercept, LRSOC and LGCOTH are insignificant at the 95% level.

(iii) There has been a marked decline in overall goodness of fit. Thus the R^2 for the double log model has decreased from 0.54 to 0.39 and for the semi log model from 0.53 to 0.41. This leads on from point (ii) above. It should be noted that only the basic gravity model variables i.e. the attraction, generation and distance measures (LDRX, LOPOP and LGCRA), have retained their significance.

The changes noted above may be attributed to two main causes:

(a) The effect of using new rail and bus fare variables and using 1984 prices. The effect of such changes were tested by calibrating a model on 1984 OD data and 1981 prices, as shown by Table 6.7. Comparison of this table with column 1 of Table 6.6A shows that there is little change, although the R^2 decreases slightly from 0.388 to 0.380, and there are, at the 95% level, significant changes in the LGCRA, LDRX and intercept values.

b) The effect of using OD data where trips that involved interchange were not broken down into their constituent parts (as they had been with the initial data set). It is possible to identify a number of trips in the 1984 data set where interchange must have occurred, along with the likely interchange points, as follows:

A) DOUBLE LOG MODEL

	Recalibrated model	Original model	z-ratio
Intercept	4.669 (1.702)	5.496 (3.025)	2.319
LOPOP	0.549 (2.754)	0.380 (2.617)	-6.316
LOPOP3	0.098 (0.736)	0.164 (1.733)	3.734
LRSOC	0.232 (1.426)	0.246 (2.034)	0.641
LDRX	0.292 (4.885)	0.269 (6.678)	-2.870
LGCOTH	-0.380 (-0.457)	-1.341 (-2.269)	-8.707
LGCRA	-1.095 (-2.303)	-1.239 (-4.307)	-2.381
R^2	0.388	0.539	
$\bar{R}^2$	0.338	0.509	
No of obs.	80	99	

B) SEMI LOG MODEL

Intercept	-1.386 (-0.590)	-1.468 (-0.790)	-0.255
LOPOP	0.574 (2.934)	0.423 (2.866)	-5.701
LOPOP3	0.070 (0.533)	0.147 (1.542)	5.527
LRSOC	0.219 (1.370)	0.248 (2.047)	1.381
LDRX	0.284 (4.758)	0.291 (6.507)	0.875
LGCOTH2	0.413 (0.892)	0.507 (1.721)	1.714
GCRA	-0.008 (-2.236)	-0.007 (-3.791)	3.326
R^2	0.406	0.532	
$\bar{R}^2$	0.357	0.502	

TABLE 6.6 COMPARISON OF RECALIBRATED AND ORIGINAL ASMs

	Recalibrated model	z-ratio (Compared to column 1, Table 6.6A)
Intercept	3.760 (1.424)	2.131
LOPOP	0.533 (2.670)	0.484
LOPOP3	0.089 (0.665)	0.428
LRSOC	0.221 (1.346)	0.423
LDRX	0.265 (4.225)	2.700
LGCOTH	-0.561 (-0.605)	1.300
LGCRA	-0.902 (-2.303)	-2.797
R^2	0.380	

TABLE 6.7 EFFECT OF USING 1981 FARE MEASURE

A) DOUBLE LOG MODEL		B) SEMI LOG MODEL	
Intercept	2.808 (1.055)	Intercept	-2.470 (-1.019)
LOPOP	0.639 (3.362)	LOPOP	0.665 (3.524)
LOPOP3	0.090 (0.688)	LOPOP3	0.066 (0.506)
LRSOC	0.164 (1.026)	LRSOC	0.152 (0.957)
LDRX	0.305 (5.159)	LDRX	0.300 (5.091)
LGCOTH	-0.542 (-0.616)	LGCOTH 2	0.398 (0.837)
LGCRA	-0.935 (-2.008)	GCRA	-0.007 (-1.906)
R^2	0.448		0.458
$\bar{R}^2$	0.403		0.414

TABLE 6.8 EFFECT OF INTRODUCING AMENDMENTS FOR INTERCHANGE TRIPS

	Number of trips
Leeds City	1578
Bradford Forster Sq./Interchange	180
Wakefield Westgate	154
Keighley/Shipley	184
Huddersfield	34
TOTAL	2130

Thus 2,130 trips (or 9.2% of all trips) were detected as involving interchange. This is less than the 15% detected by Hartley and Nash (1980), but this may be due to the fact that certain OD pairs are served by a mixture of direct and indirect services (e.g. Bradford to London) but which in our case we have to consider as a direct service throughout.

Trips that involved interchange were therefore reallocated to the relevant station pairs. Thus the double and semi log models were rerun with this new OD data to give the results given by Table 6.8. By comparing this Table with Table 6.6 it can be seen that the R^2 increases slightly from 0.388 to 0.448 for the double log model and from 0.406 to 0.458 for the semi log model. Once again all parameter values are insignificant at the 95% level except the three basic gravity model parameters (LOPOP, LDRX and LGCRA). Visual inspection confirms that the majority of parameter values are significantly different from those estimated in our original models (equtions 6.8 and 6.13). Indeed the difference between the original parameter estimates and those recalibrated in Table 6.8 is greater than the difference between the original and recalibrated values of Table 6.6 for 11 out of the 14 parameters. Table 6.8 again seems to suggest that the semi log specification is superior to that of the double log.

In concluding this sub-section the following points, might be made.

(i) Re-estimation of equations 6.8 and 6.13 with the 1984 data set leads to a significant change in most parameter values and a reduction in the overall goodness of fit. This suggest that our models lack temporal transferability. There is some evidence that the semi log model may be more stable than the double log model.

(ii) The more sophisticated approach to rail and bus fares, given by equations 6.14 and 6.15, produced only a slight improvement in goodness of fit.

(iii) Amendment of the 1984 data set to take into account the effect of trips where interchange occurred (and thus be more like the 1981 data set) leads to significant improvements in the overall goodness of fit. This may be because most changes of train occur at Leeds which, as measured by DRX, is also the most attractive destination and thus the reallocation of interchange trips and the resultant increase in the number of trips modelled causes increases in the size and significance of the LOPOP and LDRX parameter values.

(iv) Use of ODs from both the 1981 PTS estimated data set and the 1984 County Council data involved input error. "A priori" one would expect reported ODs to be more reliable than estimated ODs but contrary to expectations the PTS estimated data produced the best fitting model.

6.3.2 Developing a Non-Work Aggregate Model

One advantage of the County Council's 1984 data is that it allows us to disaggregate by journey purpose. For example an aggregate non-work model might be developed, which could then be applied in conjunction with the disaggregate model of work journeys (developed in Chapter Seven). In this section we shall attempt to develop such a model.

In order to do this the following changes to our modelling procedure were introduced:

(i) The dependent variable (FLOW) becomes the number of non-work trips. This still includes a number of journey purposes, as shown by Table 6.9.

(ii) The variable of attraction characteristics of the destination (DRX) needs to be replaced. As from Table 6.9 it is evident that shopping is the most important non-work journey purpose, it was felt that a measure of retail attraction should be used. This might be based on retail employment (REMP) (Martin and Dalvi, 1976), floorspace (FSPACE) (Huff, 1963) or sales (SALES) (Black, 1966). However the most recent data source, at a usably disaggregate level, was provided by the 1971 Census of Distribution (Department of Industry, 1975). More recent (1984) information on retail floorspace is available, but only at a district level (Department of Environment, 1985).

(iii) Bus and rail frequency were amended so as to include only services in the off-peak (defined as 0900-1500 hours and 1800 hours and beyond). This also affects wait time calculations.

(iv) Bus and rail fares were limited to their off-peak values as of October

1984 (i.e. bus = 30p maximum, rail = 60p maximum (30p if distance travelled less than or equal to eight miles)).

(v) Our minimum flow size was reduced to 10, but despite this our calibration data set only consisted of 64 observations.

	Numbers	%
Shopping	1833	40.8
Social/Recreation	1421	31.5
Employer's Business	825	18.4
Personal Business	401	8.9
Other	12	0.3
TOTAL	4492	

TABLE 6.9 NON-WORK JOURNEYS BY RAIL ON A WEEKDAY IN WEST YORKSHIRE

Source: West Yorkshire County Council 1984 Rail Survey.

Our initial results, summarised by Table 6.10, involved rerunning the double and semi log models of Table 6.6 with the dependent variable becoming the number of non-work trips and the independent variable DRX replaced by measures of retail attraction. It can be seen that the t-statistics for the three retail attraction variables are broadly similar, but our preference would be to use REMP as it has the highest t-statistic. There is little difference between the goodness of fit of the double and semi log models, with both having low R^2s. A number of model runs showed that retail attraction was the only significant parameter whilst the generalised cost parameter values had particularly low significance and, in one case, was of the wrong sign. As a result a number of modifications were introduced:

(i) The only modal attributes that were included were rail and bus service frequency. This is due to the fact that service frequency is the only measurable modal attribute that varies significantly during off-peak periods, as for journeys of up to eight miles bus and rail fares are the same whilst journey time differentials are reduced during the off-peak.

(iii) Inspection of residuals showed that the initial model underestimated flows from medium sized town stations to inter city stations. These flows were thought to be atypical in that they included trips feeding onto the inter city

	Double log model t-statistic	Double log model R^2	Semi log model t-statistic	Semi log model R^2
LSALES	2.138	0.163	2.136	0.163
LFSPACE	2.141	0.163	2.139	0.164
LREMP	2.156	0.164	2.156	0.165

TABLE 6.10 EFFECT OF RETAIL ATTRACTION VARIABLES

	1	2	3	4	5
INTERCEPT	-0.985 (-0.421)	-1.298 (-0.557)	-3.580 (-2.321)	-2.680 (-1.239)	-3.893 (-2.474)
LOPOP	0.462 (2.618)	0.506 (2.775)	0.562 (3.230)	0.467 (2.625)	0.575 (3.924)
LREMP	0.162 (1.973)	0.210 (2.843)	0.252 (4.051)	0.171 (2.043)	0.218 (3.059)
LGCOTH	0.864* (1.230)				
LGCOTH2				0.263 (0.643)	
LGCRA	-1.072 (-1.224)	-0.564 (-1.030)			
GCRA				-0.001 (-0.504)	
LRS			0.574 (3.315)		0.614 (3.455)
LBS			-0.250 (-2.048)		-0.224 (-1.796)
LRJT					0.129* (1.002)
LGCBU		0.417 (1.130)			
IC	1.077 (4.596)	1.061 (4.542)	0.996 (4.534)	1.046 (4.434)	0.990 (4.607)
INTOPP	-1.215 (-7.134)	-1.217 (-7.133)	-1.247 (-8.077)	-1.217 (-7.017)	-1.283 (-8.091)
R^2	0.662	0.660	0.709	0.654	0.714
$\bar{R}^2$	0.626	0.624	0.678	0.618	0.678

TABLE 6.11 NON WORK MODELS

* = wrong sign

network, as witnessed by the longer than average access distances (WYTCONSULT, 1976A estimated that 59% of passengers at these stations come from beyond one km) and a higher than average proportion of travellers on employer's business. To represent this a dummy variable IC was set to 1 for these flows.

(iii) Flows from local stations to destinations other than the most attractive (as measured by REMP) destinations were overestimated. This might be due to the intervening opportunities provided by more attractive destinations (Schneider, 1959). Flows might be overpredicted if:

(a) They involve passing through an attractive destination. For example Crossgates to Bradford is overpredicted due to the intervening opportunity at Leeds.

(b) There is a more attractive destination in one direction than another. For example flows from Guiseley and New Pudsey to Bradford are overestimated due to the competing attraction of Leeds. Where either of these phenomena occured a proxy variable, INTOPP, was set equal to the number of competing/intervening opportunities.

Clearly the definition of this variable is somewhat arbitrary. We attempted to model the effect of destination choice explicitly by including distance and/or attraction variables to represent competing opportunities, but these results were highly insignificant.

As can be seen from Table 6.11 the IC and INTOPP variables were highly significant, of the right sign and increased dramatically the goodness of fit. Moreover, as can be seen from our statistically preferred model given by column 3, the inclusion of the two proxy variables ensures that the five other parameter values are significant at the 95% level, of the right sign and of plausible magnitude. For example the rail service elasticity of 0.574 is considerably higher than the 0.26 used for all trips in the METS model (Department of Transport, 1984). This is as expected as demand for off-peak or non-work travel should be more elastic than demand for peak or work travel (Kemp, 1973).

6.4 VALIDATION

6.4.1 Validation of the all trips Aggregate Simultaneous Model

An initial attempt was made to validate the ASM given by equation 6.8 with reference to the first four new stations that were opened (Fitzwilliam, Deighton, Crossflatts and Slaithwaite) using total usage figures from the November 1983 PTS, combined with information on destination choice from our own market research. Results are shown by Table 6.12.

From	To	Predicted flow divided by actual flow
Fitzwilliam	Leeds	1.54
Deighton	Leeds	3.60
Deighton	Huddersfield	1.08
Crossflatts	Leeds	2.44
Crossflatts	Bradford	2.25
Slaithwaite	Leeds	2.52
Slaithwaite	Huddersfield	0.79
Root Mean Square Error (RMSE)[1] measure		48.6
Absolute Deviation (AD)[2] measure		0.896

TABLE 6.12 PASSENGER FLOWS PREDICTED BY THE ASM

1 Defined as $\sqrt{\dfrac{\dfrac{\Sigma (F-A)^2}{n}}{n}}$ 2 Defined as $\dfrac{\dfrac{\Sigma |F-A|}{n}}{\dfrac{\Sigma A}{n}}$

where F = Forecast usage, A = Actual usage, n = number of observations

It may be shown that the ASM overpredicts flows to Leeds and Bradford by a factor of 2.4, whilst overall flows to Huddersfield are only 0.93 their "true" value, which suggests that the attractiveness of Leeds and Bradford is overestimated compared to Huddersfield. Overall it can be seen that our estimates are only within ± 49 of values, with an AD of approximately ± 90%.

A problem arises however in producing an aggregate station usage figure as not all flows can be modelled because of the lack of information on flows to destinations outside the County and the inability to model "actual" flows of values of less than 25 per day. Aggregation may, however, be possible using information on destination choice at adjacent existing stations to determine a weighting factor. This still leaves a problem in defining the number of flows that should be modelled. This is illustrated by Table 6.13.

	Predicted usage divided by actual usage (1983)		
	A 7 flows	B 11 flows	C PTE model 1
Fitzwilliam	1.31	1.67	2.25
Deighton	2.72	3.07	4.60
Crossflatts	2.32	2.51	3.33
Slaithwaite	1.22	1.22	1.54
RMSE measure	110.4	130.8	218.0
AD measure	0.871	1.066	1.837

TABLE 6.13 TOTAL STATION USAGE PREDICTIONS BY THE ASM METHOD

(see text for definitions)

1 WYPTE, 1984B.

The forecasts in column A are based on the seven flows listed in Table 6.12, whilst in Column B four additional flows are included (Fitzwilliam to Wakefield Westgate, Deighton to Dewsbury and Crossflatts to Keighley and Shipley). However, the inclusion of these four flows increases the AD from around 87% to 107%. The four flows that were included were, in reality, below the 25 per day threshold used in the calibration stage and thus should not be modelled. Thus for potential sites only flows that are expected to exceed 25 (i.e. to first or second order destinations) should be modelled.

From Table 6.13 it can be seen that the ASM model, based on seven flows, consistently overpredicts initial new station usage, with a RMSE of 110 trips and an AD of 87% but this represents a marked improvement on the forecasts used by the PTE, which also overpredicted usage, with a RMSE of 218 trips and an AD of 184%. We might expect that these models overpredict initial (first year) new station usage, as they were calibrated for stations that have been open for many years. The issue of patronage growth over time at new stations is examined in Chapter Nine, with Table 9.6 including "actual" usage figures for up to three different years for all six new stations and comparing them with the ASM forecasts.

Aside from the issues of patronage growth, the ASM fails to make more accurate predictions due to:

(i) The theoretical problems of using multiple regression models with cross sectional data (some of which were discussed in 6.2.5) and problems related to the generalised cost framework used (as discussed in 6.2.2).

(ii) The problems of measurement error and, in particular input error due to the limited quality of the data sets used.

(iii) The complex nature of interaction between explanatory variables. This can be seen by studies such as those of Houston (Black and Black, 1982) and London (AMV, 1972) that have identified large numbers of variables (combined through cluster or factor analysis) that may be used to determine Public Transport catchment areas.

6.4.2 Validation of the Non-Work Model

The non-work model (NWM) given by column 3 of Table 6.11 may also be validated in a similar manner, given usage data from the PTS and information on the proportion of non-work journeys from our own market research. The results are given by Table 6.14.

	Forecast number of non work trips divided by actual number
Bramley	2.03
Crossflatts	2.43
Deighton	1.42
Fitzwilliam	1.68
Saltaire	1.54
Slaithwaite	1.19
RMSE measure	55.9
AD measure	0.674

TABLE 6.14 NUMBERS OF NON-WORK TRIPS PREDICTED BY THE NWM

As with the all trips ASM, it can be seen that our forecasts exceed our estimates of actual non-work trips at all stations. Overall the RMSE measure is around 56 trips, with the AD being around 67%. Again one might expect patronage growth over time that would reduce the degree to which the NWM overpredicts.

6.5 CONCLUSIONS

In this chapter we have been able to make improvements on the type of regression models used by WYTCONSULT and the PTE. Firstly, an all trips ASM was developed with OD data derived from 1981/2 PTS data. A number of specifications were tested of which the double log model, given by Equation 6.8, was preferred, although some use has also been made of the semi log model, given by Equation 6.13. These models were subsequently updated using data from the 1984 WYMCC rail survey. As there were significant changes in parameter values this work suggests that the ASM can not be considered transferable over time. Contrary to expectations the 1981 model gave a better fit than that for 1984, with there being a suggestion that the best way to model trips involving interchange is that they should be broken up into their constituent legs. Secondly, we were able to develop an aggregate non-work model (NWM) the final form of which is given by column 3 of Table 6.11. Although both the all trips ASM and the NWM are affected by measurement and specification error and hence exhibit only moderate goodness of fit statistics, it was shown that the accuracy of their forecasts exceeded that achieved by the PTE.

The all trips ASM and the NWM (in conjunction with a disaggregate mode split model for work trips, developed in the next chapter) will both be used in Chapter Nine in order to produce forecasts of usage at existing and potential new stations, thus allowing comparisons of the forecasting accuracy of the different methods.

CHAPTER SEVEN DISAGGREGATE MODELS (1) : CALIBRATION

7.1 INTRODUCTION

In Section 3.2.3 it was shown that one possible approach to forecasting new station demand may be based on developing a mode split model and in subsequent sections it was shown that such an approach might be based on disaggregate (i.e. individual) data and that a disaggregate approach has a number of practical and theoretical advantages over an aggregate approach, thus leading to potentially greater forecasting accuracy. In particular, it was felt that a disaggregate approach would be able to establish the importance of factors that exhibit greater intra-zonal than inter-zonal variation, such as walk and wait time for Public Transport modes. However, it was also shown in Chapters Three and Four that a disaggregate approach leads to increases in complexity, particularly with regards to data requirements. Nonetheless, in section 4.2.4 it was shown that a usable data set for calibrating a disaggregate mode split model for the journey to work was provided by the 1981 West Yorkshire Corridor study. It should, though, again be stressed that this data set was primarily collected in order to establish values of time rather than to develop predictive models of mode choice.

In this chapter a number of multinomial (MNL) and hierarchical (HL) logit models, as discussed in section 3.3, will be developed. In the rest of this section previous work carried out by the value of time study (MVA et al. 1986) will be re-analysed and an initial MNL model developed. A number of improvements to this model will be suggested and hence in a second section a single market HL model will be developed, calibrated indirectly by using the BLOGIT package. This model, however, only includes level of service (LOS) attributes and hence in a third section a number of ways of incorporating socio-economic (SE) attributes will be examined. In particular a market segmented approach will be adopted based on developing HL models for members from car owning households and MNL models for members from non car owning households. Finally, having developed a number of model types, two will be chosen for further analysis.

7.1.1 MNL Model based on Tranche 1 Data

The corridor data set came in two tranches and initial analysis by MVA et al. (1984) was based on Tranche 1 data, consisting of 482 observations. The initial preferred model is shown by Table 7.1 which, although having a multinomial formulation, was calibrated by examining the binary choice between yesterday's mode (YM) and the best alternative (BA). The choice of this model was based on informal measures of goodness of fit (which are examined in more detail in 8.1.1) and interpretative reasonableness. This model resulted in the predicted success table (again see section 8.1.1 for more details) shown by Table 7.2. The values of time inferred by the Tranche 1 data initial preferred model are shown by Table 7.3. Apart from the broad confidence intervals it is noticeable that the model's estimates of a value of IVT are in excess of those of the Department of Transport (around 87 pence per hour at the time of the surveys) and are almost double the values used in the aggregate model of Chapter Six. It can also be seen that wait time is valued at 1.3 times Public Transport IVT, whilst walk time is valued at 2.3 times Public Transport IVT. This suggests that walk time has a greater disutility than wait time, which is contrary to the evidence of some studies, such as those of Ortuzar (1980B) and Daly and Zachary (1978). In terms of the size and significance of the coefficients estimated it appears that a revealed preference (RP) model of the type shown by Table 7.1 attaches greater importance to time rather than cost (the model has a lower cost coefficient than the models developed by Ortuzar and Daly and Zachary). This is probably because the RP data set is dominated by traders who are time savers. It is interesting to note that the transfer price model, based on data where traders are cost savers gives a cost coefficient more in line with previous studies (Broom et al. 1983).*

7.1.2 Further Analysis (Tranche 1 and 2 Data)

Work with the complete Corridor data set was carried out by Fowkes (1983) who developed 11 models: WBM4, WBM5, WBM9, WHM9, WRM9, WHT9, WRT9, WBT9, WBA4, WBA5 and WBA9, based on the following notation:

W - West Yorkshire data

H - Home work only

*It should though be noted that comparisons of coefficients should be treated with caution (see MVA, 1984, sections 8.3 and 8.4).

	Parameter value	t-statistic
ASC Train	-0.433	-1.24
ASC Car Driver	2.355	3.61
ASC Car Passenger	0.203	0.35
ASC Bike	1.889	1.49
Wait time	-0.045	-2.63
Walk time	-0.080	-4.98
IVT - Train and Bus	-0.034	-2.67
IVT - Car Driver and Passenger	-0.055	-3.55
IVT - Bike	-0.128	-3.65
Total cost	-0.018	-3.89
Log likelihood	-107.74	
Rho Squared (adjusted)	0.30	

TABLE 7.1 TRANCHE 1 - INITIAL PREFERRED MODEL (-WBM4)

ASC = Alternative Specific Constant, IVT = In Vehicle time

Source: Table 6.3, Column 5. MVA et al. 1983

	Predicted choice	Reported choice	% correct
Car Driver	236	240	97
Car Passenger	60	68	88
Bus	83	110	75
Train	24	41	58
Bike	12	15	80
Total	419	482	87

TABLE 7.2 TRANCHE 1 - INITIAL PREFERRED MODEL - PREDICTED SUCCESS TABLE

	Value of time	Confidence Intervals
Public Transport IVT	116.6	11.4 - 221.8
Private Transport IVT	180.4	46.1 - 314.6
Wait time	153.3	15.7 - 290.2
Walk time	269.5	107.8 - 431.2

TABLE 7.3 TRANCHE 1 - INITIAL PREFERRED MODEL - VALUES OF TIME (pence/hour)

R - Round trip
M - Monetary TP given
B - Both M and R
T - Any TP given
A - Biggest RP data set available
4 - Tranche 1 data
5 - Tranche 2 data
9 - Both tranches

Using this notation the initial model shown by Table 7.1 is referred to as WBM4.

We wish to use as our starting point the largest available RP data set, WBA9, which, for practical purposes, has 1238 observations, if walk mode is included. Each observation consisted of eight records with variables as listed in Table 7.4. The first model (WBA9A) was, with observations where walk was the main mode excluded, developed with 1028 observations. This is 14 more than the model developed by Fowkes who excluded, for reasons of evaluating values of time, cases where the cost variable for both alternatives was given as 0. The results of this model are given by Table 7.5. The parameter values do not seem to be significantly different from Fowkes' WBA9 model.

It can be seen that the addition of Tranche 2 data has led to an increase in the value of time estimates. For example Table 7.5 implies a Public Transport in-vehicle time value of 253 pence per hour, with the corresponding figure for Private Transport being 261 pence per hour. More plausible (i.e. lower) values were obtained if only round trips were analysed (WRM9), as 462 observations gave a Public Transport in vehicle time value of 110 pence per hour (127 pence per hour for Private Transport). Moreover, the adjusted rho-bar squared measure was 0.36 (compared to 0.24 for model WBA9A).

7.1.3 Some Suggested Improvements

A number of amendments might be made to the WBA9A model given by Table 7.5. We have already seen above that use might be made of alternative data sets, such as WRM9, but the use of such sub-samples has important implications when we come to aggregate forecasts. In addition we

Details of Data used:
WBA9A DAT
Each observation includes 8 records as follows

RECORD 1 FLAGS AND SOCIAL ECONOMIC DATA

			No. of columns
Flag 1			6
Flag 2			1
Flag 3	HH Income Given	Y = 1, N = 0	1
Flag 4	Age Given	Y = 1, N = 0	1
Flag 5	Personal Income Given	Y = 1, N = 0	1
Flag 6	TP (Money) Increase Given	Y = 1, N = 0	1
Flag 7	TP (Money) Decrease Given	Y = 1, N = 0	1
Flag 8			1
Flag 9			1
Flag 10			1
No of Cars in Household (including company cars)			1
No of Company Cars in household			1
No of Motor-cycles and Pedal Cycles in Household			1
No of Car Licences in Household			1
Household Income (Mid-point of band) if given			3
Sex of Traveller M = 0, F = 1			1
Age (Mid-point of band) if given			2
Whether Travel Pass owned - Y = 1, N = 0			1
Whether Car Licence owned - Y = 1, N = 0			1
Whether Priority Vehicle User - Y = 1, N = 0			1
Personal Income (Mid-point of band) if given			3
Mode Switches very inconvenient Y = 1, N = 0			1
Mode Switch very convenient Y = 1, N = 0			1

RECORDS 2-7 MODE DATA

RECORD 2 FOR TRAIN
RECORD 3 FOR BUS (SCHEDULED AND OTHER)
RECORD 4 FOR CAR DRIVER, GV DRIVER, TAXI
RECORD 5 FOR PASSENGER IN HH AND OTHER VEHICLES
RECORD 6 FOR MOTOR CYCLE AND PEDAL CYCLE
RECORD 7 FOR WALK

On each of these records data are as follows:

Rank of this Mode for Traveller (1-4, 0 = not given)	3
If mode was YM = 1, BA = -1, Neither = 0	3
Wait time (mins)	3
Walk time (mins)	3
In Vehicle time (mins)	3
Total Time = Walk + Wait + IVT (mins)	4
No of Public Transport vehicles	2
Ticket Type (Misspecified as always 1)	2
Fare (pence)	4
Private Vehicle Occupancy (Min = 1)	2
Whether Private Parking Space Y = 1, N = 0	2
Parking Charge (pence)	4
Other Driving Costs (Pence)	4
Passenger Payment (pence) if cost for passenger, not contribution to driver costs)	4
Total Cost = Last 3 items + Fare (pence)	5

RECORD 8 TP VALUES AND DEPENDENT VARIABLES

TP increase		4
TP decrease		4
Train Chosen	Y = 1, N = 0	1
Bus Chosen	Y = 1, N = 0	1
Car Driver Chosen	Y = 1, N = 0	1
Car Pass Chosen	Y = 1, N = 0	1
Bike Chosen	Y = 1, N = 0	1
Walk Chosen	Y = 1, N = 0	1

TABLE 7.4 DETAILS OF CORRIDOR DATA SET

	Parameter value	t-statistic
ASC - Train	-0.415	-1.768
ASC - Car Driver	1.496	4.049
ASC - Car Passenger	-0.510	-1.481
ASC - Bike	2.605	3.223
Wait time	-0.094	-4.885
Walk time	-0.092	-6.154
IVT - Train and Bus	-0.059	-4.551
IVT - Car Driver & Passenger	-0.061	-3.941
IVT - Bike	-0.186	-4.551
Total Cost	-0.014	-6.512
Log likelihood	-200.278	
Rho Squared (adjusted)	0.25	

TABLE 7.5 RESULTS OF MODEL WBA9A (TRANCHE 1 and 2 DATA)

might wish to:

(i) Let train replace bus as the base alternative, thus giving the ASCs a more relevant interpretation.

(ii) Exclude bike from the choice set, as it is unlikely to be a relevant alternative to rail.

(iii) Include SE variables, such as car ownership and income, in addition to LOS variables. It should be noted that previous work by MVA (1984, p38) on stratification by sex, income and age failed to produce significant results. Nonetheless we shall attempt to develop models including SE variables in section 7.3.

(iv) Experiment with choice set specification.

(v) Test for violation of the IIA axiom.

These last two points will be examined in more detail in the following sub-sections.

7.1.4 Choice Set Specification

Hensher (1979) has noted that:

".... Choice set determination ... is the most difficult of all the issues to resolve. It reflects .. the dilemma which a modeller has to tackle in arriving at a suitable trade off between modelling relevance and modelling complexity. Usually, however, data availability acts as a yardstick".

It is normal practice to assume that everybody has all alternatives available (see, for example, Ortuzar 1980A). However, with the data set we are working with this is not the case, as each individual gave up to four alternatives, but with the majority, as Table 7.6 shows, giving only two alternatives.

	Rank	0	1	2	3	4	8
Mode	Train	549	79	188	70	8	1
	Bus	103	151	493	131	17	
	Car Driver	339	478	55	19	3	1
	Car Passenger	577	143	142	21	11	1
	TOTAL		851	878	241	39	3

0 = Not Available

TABLE 7.6 RANKING OF ALTERNATIVES - 1981 CORRIDOR DATA SET

Table 7.6 highlights two problems with the rankings (note observations where walk or bike were ranked as alternatives have been excluded). Firstly, there are three alternatives which are erroneously ranked 8. Secondly, there are more observations ranked 2 than are ranked 1, indicating that a number of observations did not include a mode ranked 1.

At least three choice set specifications might be used (Table 7.7):

(i) As we have already seen previous work concentrated on modelling the binary trade-off between YM and BA. This specification was used partly so as to overcome problems with the IIA axiom (MVA, 1984, p.13). Column 1 of Table 7.7 shows model WBA9B (i.e. incorporating amendments (i) and (ii) above) with such a choice set specification. Compared to model WBA9A there have been some changes in the ASC parameter values (although two out of three are insignificant at the 95% level) and a slight deterioration in goodness of fit (as measured by the adjusted rho-bar squared).

(ii) As an alternative to the YM/BA specification use might be made of an ordered logit model based on the sequential expansion of the data set (Bates and Roberts, 1983). If n alternatives are ranked 1, 2,...,n, then it is possible to view this as (n-1) choice decisions the first being to choose the preferred alternative from the n available, the second being to choose the preferred alternative from the remaining n-1 alternatives etc. Under the assumptions of MNL these choices are regarded as independent and are treated as separate observations in a logit estimation (851 in Table 7.7 Column 2 refers to the number of individuals considered). This procedure is offered by a modified version of the BLOGIT package.

A model of this type is shown by column 2 of Table 7.7. Compared with column 1 it can be seen that most parameter values have undergone changes. Moreover, although the ASC value for bus is now highly significant at the 95% level, that of car passenger is still insignificant as has become the value for walk time, whilst the IVT parameter for car driver and car passenger is of the wrong sign. In addition, judging by the adjusted rho-squared and percent right

	YM/BA	Choice Set Ranked	All choices
ASC - Bus	0.449* (1.870)	1.358 (8.038)	-0.088* (-0.393)
- Car Driver	1.824 (4.334)	1.107 (8.244)	1.691 (4.380)
- Car Passenger	-0.092* (-0.232)	-0.146 (-0.927)	-0.503* (-1.358)
Wait time	-0.100 (-5.033)	-0.048 (-4.765)	-0.097 (-5.229)
Walk time	-0.089 (-5.960)	-0.007* (-1.197)	-0.092 (-6.642)
IVT - Train and bus	-0.059 (-4.399)	-0.007 (-2.028)	-0.039 (-3.409)
IVT - Car Driver and Passenger	-0.059 (-3.737)	0.028+ (8.890)	-0.072 (-5.498)
Total Cost	-0.013 (-6.355)	-0.001* (-1.744)	-0.012 (-5.760)
Log Likelihood	-275.915	-882.324	-349.687
Adjusted Rho Squared	0.242	0.099	0.207
% right	83.0	44.0	78.7
Number of observations	952	851	978

TABLE 7.7 RESULTS OF MODELS WBA9B

* = insignificant at the 95% level

+ = wrong sign.

measures, there seems to have been a marked deterioration in goodness of fit. This may be due to data unreliability (related to misperceptions of times and costs) increasing with depth of ranking.

(iii) The number of choices may vary across individuals (i.e. from 1 to 4). This is done by making use of the NAA parameter in BLOGIT. A model of this form is shown by column 3 of Table 7.7. Compared to column 1 it may be seen that there are marked changes in the parameter values, particularly the ASCs (with two out of three still insignificant at the 95% level) and of the IVT parameters (with Private Transport's value becoming almost double that of Public Transport). In addition there are suggestions that there has been a slight decline in goodness of fit.

Clearly all three choice set specifications have disadvantages and, as we shall see in sections 8.2 and 8.3, are not always consistent with the data sets used for validation and prediction. Although specification (iii) involves a poorer fit, it is in principle the best formulation. However it should be noted that, as we are dealing with mode split, the number of alternatives is small (four), hence limiting the scale of choice set specification problems. Such problems would increase if, for example, we were to study destination choice.

7.1.5 Violation of the IIA Axiom

In Section 3.3.3 we suggested that MNL models of the type shown in Table 7.7 might violate the IIA axiom. This may be tested through BLOGIT by using a generalised likelihood ratio test (McFadden et al. 1976). If we have three alternatives (bus, car driver and passenger) to which we add a new mode (train) with information X_{1T} and X_{2T} (total time and cost) we can estimate two models. Firstly $\tilde{M}$ where:

$$
\begin{aligned}
\tilde{V}_T &= \tilde{B}_T + \tilde{B}_1 X_{1T} + \tilde{B}_2 X_{2T} && \text{(Train)} \\
\tilde{V}_B &= \tilde{B}_B + \tilde{B}_1 X_{1B} + \tilde{B}_2 X_{2B} && \text{(Bus)} \\
\tilde{V}_C &= \tilde{B}_C + \tilde{B}_1 X_{1C} + \tilde{B}_2 X_{2C} && \text{(Car Driver)} \\
V_P &= \tilde{B}_1 X_{1P} + \tilde{B}_2 X_{2P} && \text{(Car Passenger)} \quad \text{Equation 7.1}
\end{aligned}
$$

with log likelihood = $\tilde{L}$

Secondly we restrict the above coefficients, except B_T, to their first estimated values. We then estimate the model M* where:

$$\begin{aligned} V^*_T &= B^*_T + B_1 X_{1T} + B_2 X_{2T} \\ V^*_B &= B_B + B_1 X_{1B} + B_2 X_{2B} \\ V^*_C &= B_C + B_1 X_{1C} + B_2 X_{2C} \\ V^*_P &= B_1 X_{1P} + B_2 X_{2P} \end{aligned} \qquad \text{Equation 7.2}$$

with log likelihood = L*

If L* is not significantly different from $\tilde{L}$ then the IIA condition holds. This can be easily tested as the log likelihood statistic

$$-2 \{ L^*(\underline{B}) - \tilde{L} (\tilde{\underline{B}}) \} \qquad \text{Equation 7.3}$$

is asymptotically distributed χ^2 with M degrees of freedom, where M is the number of elements of the parameter space that have been restricted (Wilks, 1962). A test of this sort gave a χ^2 of restriction of 85.860 whilst $\chi^2_{0.05}$ at 4 degrees of freedom = 9.488. This suggests that the null hypothesis can be rejected and thus the probability of choosing rail is not independent of the parameter values that have been restricted.

To avoid violation of the IIA axiom it is proposed to develop HL models as described in sections 7.2 and 7.3.

7.1.6 Conclusions

In this section we have shown that previous work in developing models with the West Yorkshire 1981 Corridor data set was limited to logit models of the binary choice between YM and BA. A number of amendments to this work can, for our purposes, be considered. In particular it has already been shown that amendments to choice set definitions may be considered. In the next section we shall develop HL models so as to overcome some of the problems due to violation of the IIA axiom, whilst in section 7.3 we shall go on to consider how SE variables may best be considered.

7.2 SINGLE MARKET HIERARCHICAL LOGIT MODEL

7.2.1 Overview

It was shown in Chapters Three and Four that we could calibrate an HL mode split model, similar to that given by equations 3.20 to 3.22, for the journey to work by using the Corridor data set. Initially this model will be based on a decision tree structure of the type shown by Figure 3.7 (B) (i). This model will be calibrated indirectly in two stages. Firstly, a lower hierarchical split (or Public Transport sub-nest) will be estimated, examining the choice between bus and train. Secondly, an upper hierarchical split will be estimated, examining the choice between car driver, car passenger and Public Transport.

7.2.2 Lower Hierarchical Split

Initially a Public Transport sub-nest was developed by analysing a maximum of 179 cases from the WBA9 data where the choice of YM/BA was between bus and rail. It should be noted that in the data set if a Public Transport mode was given as an alternative it was always coded as the BA and hence, for this sub-split, choice set specifications (i) and (iii) in 7.1.4 are the same.

Following the work in the previous section models were developed with LOS attributes only. The results are shown by Table 7.8. Models 1 to 3 include an ASC for bus which, however, is always insignificant at the 95% level. Model 4 incorporates alternative specific IVT coefficients which imply an in-vehicle time value for train at 109 pence/hour and for bus 125 pence/hour (and thus is broadly in line with the results of WBM4 given by Table 7.1). By contrast model 5 has a generic IVT coefficient which implies a Public Transport in-vehicle time value of only 95 pence/hour (and thus more in line with the Department of Transport value). In models 3 and 6 we exclude the total cost coefficient from the lower split as we make a similar assumption to that of Ortuzar (1980A) in that cost for bus and rail may be considered common attributes.

In terms of goodness of fit, as measured by the adjusted rho-bar squared, there appears to be little difference between models 1 and 2 and between 4 and 5, although models 3 and 6 do seem to be inferior. Our initial preference is to make further use of model 4 (WBA9C-4) as all the parameter values are significant (including the alternative specific IVT parameter values) and consistent with previous models. However it should be noted that a log likelihood ratio test (based on equation 7.3) would favour the use of a generic

	1	2	3	4	5	6	7
ASC - Bus	1.313*	0.716*	0.277*				
	(1.903)	(1.699)	(0.675)				
Wait time	-0.142	-0.140	-0.146	-0.132	-0.130	-0.140	-0.101
	(-3.430)	(-3.413)	(-3.476)	(-3.025)	(-3.247)	(-3.430)	(-2.463)
Walk time	-0.166	-0.173	-0.182	-0.184	-0.185	-0.189	-0.152
	(-4.640)	(-4.868)	(-4.684)	(-5.221)	(-5.269)	(-4.964	(-3.864)
IVT - Bus	-0.104*			-0.092			
	(-1.732)			(-3.024)			
IVT - Train	-0.072			-0.080			
	(-3.433)			(-2.295)			
IVT - Generic		-0.105	-0.103		-0.070	-0.089	-0.044
		(-3.465)	(-3.361)		(-3.375)	(-3.976)	(-2.513)
Total Cost	-0.054	-0.050		-0.044	-0.044		-0.052
	(-2.810)	(-2.660)		(-2.490)	(-2.508)		(-2.479)
Log likelihood	-61.386	-62.025	-60.212	-63.322	-63.538	-60.442	-23.333
Adjusted Rho-Squared	0.483	0.477	0.447	0.466	0.465	0.445	0.595
Adjusted Rho-bar Squared	0.465	0.463	0.434	0.451	0.452	0.435	0.576
% Right	78.4	77.9	76.3	77.5	77.5	76.3	84.0
Number of observations	179	179	164	179	179	164	87

TABLE 7.8 RESULTS OF MODELS WBA9C - LOWER SPLIT MODEL

IVT parameter (model 5) as 0.432 is less than $\chi^2_{0.05}$ with one degree of freedom (= 3.841).

Lastly in model 7 of Table 7.8 we have developed a lower split model with just 87 observations from the round trip data set, WRM9. This resulted in an apparent increase in the goodness of fit, but a disturbing feature is the instability of parameter values, for example the IVT-generic value in model 7 (-0.044) is much lower than in previous models. This might suggest that travellers giving round trip information are more likely to be cost savers than the total sample population. Moreover it should be noted that the values of IVT generic parameters in models 2 and 5 are not within the range implied by the alternative specific values in 1 and 4. This instability may be related to correlation between IVT and cost.

An attempt was also made to develop a private transport sub split (and hence based on a decision tree structure of the type shown by Figure 3.7(B)(ii)). However, for a model developed with 584 observations, all the parameter values were insignificant at the 95% level, except for total cost, and the parameter for car driver's in-vehicle time had the wrong sign. This confirmed our a priori preference for a decision tree structure like Figure 3.7(B)(i).

7.2.3 Upper Hierarchical Split

Having calibrated lower split models we go on to estimate the upper level of the hierarchy, initially using the coefficients from WBA9C-4 (Table 7.8). This involves computing a new variable, the Expected Maximum Utility (EMU) of the composite Public Transport choice, calculated as:

$$EMU = \ln \sum_j \exp (U_j) \qquad \text{Equation 7.4 (see also 3.22)}$$

where U_j = utility measure of Public Transport mode j.

An important issue is what should be done with observations where there is no Public Transport mode in the choice set. This issue might be seen as part of the "trader/non trader" problem (Daly, 1978). The deletion of observations will reduce the accuracy of estimates of attribute-values and bias the estimates by specifically deleting individuals for whom unmeasured attributes are important. Daly thus argued that there is no justification for the exclusion of non traders from the parameter estimation process. As we have no information for non traders, the best that we could do was to set the EMU value for such observations to an arbitarily large negative number. However, not surprisingly, the resultant model failed to converge after the maximum number of iterations. As there were only 31 of these non traders it was decided to delete these observations and consequently we developed our initial preferred model, as shown by Table 7.9, column 1 (WBA9D-1).

It can be seen that for this model the EMU parameter is between 0 and 1 and significantly different from zero and one and hence passes the in-built diagnostic test. In addition all the parameters are significant at the 95% level except that of IVT, which implies an implausibly low value of in-vehicle time of 47 pence/hour. Again this might reflect correlation between IVT and Cost. Our goodness of fit statistics imply a reasonable degree of explanation.

Subsequent modifications included the development of a total time coefficient in place of the OVT and IVT parameters (model 2, Table 7.9). A log-likelihood ratio test gives a χ^2 test statistic of 2.434 which is less than $\chi^2_{0.05}$ at 1 degree of freedom (= 3.841) suggesting that the gain of one degree of freedom is worth the deterioration in log-likelihood.

In model 3, Table 7.9 we assume that costs for bus and rail are common attributes (as in models 3 and 6 of Table 7.8) and thus Public Transport costs are included in the Total Cost variable in the upper split. A problem here is which value to use (out of several possible such as the minimum, the average or the preferred alternative) for the variable Total Cost for the Public Transport alternative. This was calculated as follows:

(i) For the chosen option if this was known (i.e. in the case of members of the lower split).

(ii) For the option in the choice set, if only one Public Transport alternative was given

(iii) For the first preference probability (FPP) option, as given by the appropriate lower split model, for the limited number of cases where both Public Transport modes are given as alternatives.

	1	2	3	4	5	6
ASC - Car Driver	2.742 (5.860)	2.362 (5.030)	2.230 (4.997)	1.630 (3.701)	1.616 (3.692)	2.253 (2.552)
ASC - Car Passenger	0.804 (1.962)	0.380 (0.921)*	0.579 (1.384)*	-0.012 (-0.055)*	-0.075 (-0.182)*	-0.227 (-0.256)*
EMU - Public Transport	0.205 (2.762)	0.276 (4.258)	0.151 (2.309)	0.415 (4.379)	0.436 (4.864)	0.287 (3.469)
OVT - Car Driver and Passenger	-0.067 (-2.698)		-0.064 (-2.597)	-0.058 (-2.296)		-0.067 (-2.857)
IVT - Car Driver and Passenger	-0.011* (-0.743)		-0.022* (-1.492)	-0.038 (-2.595)		-0.020 (-1.372)*
Total Time - Car Driver and Passenger		-0.031 (-2.397)			-0.043 (-3.429)	
Total Cost - Car Driver and Passenger	-0.014 (-6.252)	-0.014 (-6.255)				-0.013 (-4.144)
Total Cost - All modes			-0.011 (-5.490)	-0.012 (-5.666)	-0.012 (-5.682)	
Log likelihood	-199.77	-200.294	-207.068	-198.426	-198.685	-62.418
Adjusted Rho Squared	0.779	0.777	0.770	0.779	0.779	0.843
% right	87.8	87.7	87.7	87.7	87.7	92.2
Number of observations	907	907	907	907	907	429

TABLE 7.9 RESULTS OF MODELS WBA9D - UPPER SPLIT MODEL

OVT = Walk and Wait time

However, comparison of model 3's log-likelihood with that of model 1 indicates that there has not been an improvement in goodness of fit. A more consistent approach is to calculate the EMU based on model 6 of Table 7.8 which excludes cost from the Public Transport sub-split (even though we have shown in previous models at this level that the cost coefficient is significantly different from zero and hence bus and rail costs should not be considered common attributes). This has been done for model 4 of Table 7.9. Comparison of this model's log-likelihood with that of model 1 indicates that there is a slight increase in goodness of fit. In model 5 a further amendment is made by creating a total time variable for car driver and passenger, with a log-likelihood ratio test showing that the gain of one degree of freedom is again worth the deterioration in goodness of fit.

Lastly, by developing an EMU based on model 7 of Table 7.8, we were able to develop an upper split model based on WRM9 data, shown by model 6 of Table 7.9. This model, as far as we can tell from our goodness of fit statistics, leads to an improvement in the degree of explanation. However, apart from the cost coefficients, the instability of the coefficient values for models 1 to 6 is a cause for concern, with the constant values and the EMU parameter being particularly unstable (although to some extent changes in the EMU parameter value are explained by the different Public Transport sub-nest specifications used).

It should also be noted that a feature of models 2 to 6 is the insignificance (at the 95% level) of the ASC for car passenger. One possible solution might be to subsume passenger into the car driver main mode but this still resulted in an insignificant IVT parameter value, although the implied value of in-vehicle time increases to 116 pence per hour. However the log likelihood of such a model (still based on 907 observations) is -213.770 which is greater than the log-likelihood of models 1 to 5 in Table 7.9, suggesting a deterioration in goodness of fit.

7.2.4 Conclusions

In this section we have been able to develop a number of single market HL models. Our initial preferred model is based on a lower split model given by WBA9C-4 (Table 7.8) and an upper split model given by WBA9D-1 (Table

7.9) which results in all parameter values significant at the 95% level (except for IVT - Car Driver and Passenger), of the right sign and an apparently high degree of goodness of fit, as measured by both the adjusted rho squared and percent right measures. Subsequently, we have been able to suggest some minor amendments to the initial preferred model, although these are unlikely to significantly change the model's predictive abilities (for example Talvitie and Kirshner (1978) showed that there is a tendency for different formulations of the same model type to give similar forecasts). The models developed in this section have, however, been based solely on LOS attributes. A major improvement would be to incorporate SE variables and this will be done in the next section.

7.3 INCLUSION OF SOCIO-ECONOMIC ATTRIBUTES

7.3.1 Overview

The models developed so far have been based primarily on generic LOS variables. Desirably, individual utility should depend only on generic variables (GVs) because:

"Individual utility depends on the constellation of physical experiences associated with an alternative, and can not depend on labels .. attached to alternatives by the planner". (McFadden, 1976).

However, as we have seen, average utility may depend on Alternative Specific Variables (ASVs) which have an identifiable correspondence between choice alternatives and represent the influence of unobserved GVs. We thus make the distinction between attributes that are inherent in the alternatives themselves and attributes that define the unit environment. Thus in discrete choice modelling the inclusion of SE variables may be justified by neoclassical choice theory as representing utility function variables that act as constraints on utility maximisation (Truong and Hensher, 1980, Hensher and Johnson, 1981 pages 121 to 128).

Socio-economic and environmental attributes may be entered into a model specification in three ways:

(i) Pure ASVs. These are the result of an interaction between alternative specific dummy variables (ASDV = 1 for particular alternative, = 0 otherwise) and an attribute contained in the X and/or S vector set for an individual. They may be interpretated as exerting a pure shift effect on choice. If there are J alternatives in the choice set, ASVs may be specified to a maximum of J-1 alternatives. Demetsky and Korf (1979) have shown that care must be taken in choosing the alternative to which a variable is specified, as a strong descriptor for one mode may appear superfluous when applied to a different mode.

(ii) Interactive (or indirect) ASVs, which influence choice via the interactive influence on another variable. For example we can develop a cost divided by income variable, with the behavioural interpretation being that individuals with different income circumstances value cost savings or outlays differently.

(iii) Market segmentation whereby the population of interest is divided into groups on the basis of SE characteristics (Dobson, 1979). This involves estimating separate models for each segment. An alternative is to use dummy variables to develop segment-specific coefficients (see MVA et al. 1986, Appendix 4).

7.3.2 Effect of Pure ASVs

It was possible to test the effect of two groups of SE variables, namely car availability/ownership and income. These variables must be defined with respect to the dependent variable and thus have some testable causal relationship as Table 7.10 attempts to show. In addition it should be remembered that SE variables may be entered in either a discrete or continuous form.

Dependent variable	Independent variable: Car ownership/ availability	Income
Bus	-	-
Train	?	+
Public Transport	-	?
Car Driver	+	+
Car Passenger	-	?

TABLE 7.10 EXPECTED EFFECT OF ASVs

+ = positive, - = negative, ? = uncertain of effect

Initial inclusion of pure ASVs in the upper split model only (as given by WBA9D-1) produced significant coefficients of the right sign but led to the EMU parameter becoming insignificant and in some cases outside the prescribed range. This suggests that an HL structure might be inappropriate unless the utility function of the lower nest was amended to include SE variables and the EMU recalculated. Although difficult to generalise, this work also suggested that carefully chosen discrete SE variables may be preferable to continuous variables.

An example of an HL model incorporating SE variables is given by Table 7.11. Model 1 of Table 7.11 gives the lower split model extended to incorporate pure income and car ownership ASVs, whilst model 2 gives the comparable model for the upper split. Although the EMU value is significant and between 0 and 1, it can be seen that in-vehicle time for car driver and passenger is of the wrong sign whilst in model 1 the car ownership parameter value is insignificant at the 95% level. In terms of the adjusted rho-squared and percent right measures it appears that for the upper split inclusion of income and car ownership leads to a deterioration in goodness of fit compared to WBA9D-1, although there appears to be a slight improvement for the lower split, compared to WBA9C-4.

7.3.3 Inclusion of Indirect ASVs

Table 7.11 also shows the effects of including indirect ASVs as models 3 and 4 include a variable Total Cost divided by Income (Cost/Inc). However, the parameter value for such a variable is insignificant at the 95% level for the lower split model (although significant at the 90% level), whilst dividing by income does not seem to have increased the significance of the cost parameter value in the upper split. In terms of goodness of fit it appears that inclusion of a cost divided by income variable appears to lead to a slight improvement for the upper split and a slight deterioration for the lower split.

7.3.4 Development of a Market Segmented Approach

Figure 7.1 shows the different markets for bus and rail between car owning and non car owning households for which full SE information was available. This shows that the propensity to use rail for the journey to work is

	1	2	3	4
ASC - Car Passenger		-2.037 (-6.641)		0.604* (1.342)
ASC - Car Driver		-2.187 (-8.123)		2.297 (4.798)
EMU - Public Transport		0.242 (6.116)		0.272 (3.665)
IVT - Train	-0.156 (-2.764)		-0.109 (-2.427)	
IVT - Bus	-0.107 (-2.899)		-0.100 (-3.276)	
IVT - Car Driver and Car Passenger		0.167+ (11.809)		-0.028 (-1.818)
Walk time	-0.236 (-4.747)		-0.210 (-4.394)	
Wait time	-0.152 (-3.162)		-0.139 (-3.134)	
Total Cost (2)	-0.075 (-3.275)	-0.008 (-4.564)		
Cost/Income (2)			-0.908* (-1.688)	-0.492 (-6.142)
OVT - Car Driver and Car Passenger		-0.077 (-6.754)		-0.063 (-2.467)
Income (2)	0.019 (2.122)	0.008 (2.105)		
Car ownership (3)	0.858* (1.680)	-0.843 (-2.109)		
Number of observations	164	850	164	827
Log likelihood	-49.44	-334.66	-56.87	-177.68
Adjusted Rho-Squared	0.557	0.602	0.490	0.785
Adjusted Rho-bar Squared	0.537	0.600	0.474	0.478
% right	80.9	79.2	77.8	87.9

TABLE 7.11 RESULTS OF MODELS WBA9E - SE VARIABLES INCLUDED

(1) Lower split - Train, Bus; Upper split - Car Driver, Passenger

(2) Lower split - Train; Upper split - Car Driver

(3) Lower Split - No Cars - Bus; Upper split - Two or more cars - Passenger

two times and the propensity to use bus is seven times as great for trip makers from non car owning households compared to those from car owning households. It should be noted that only 15% of trips were by individuals from non car owning households, suggesting that such individuals may have been under represented in the sample.

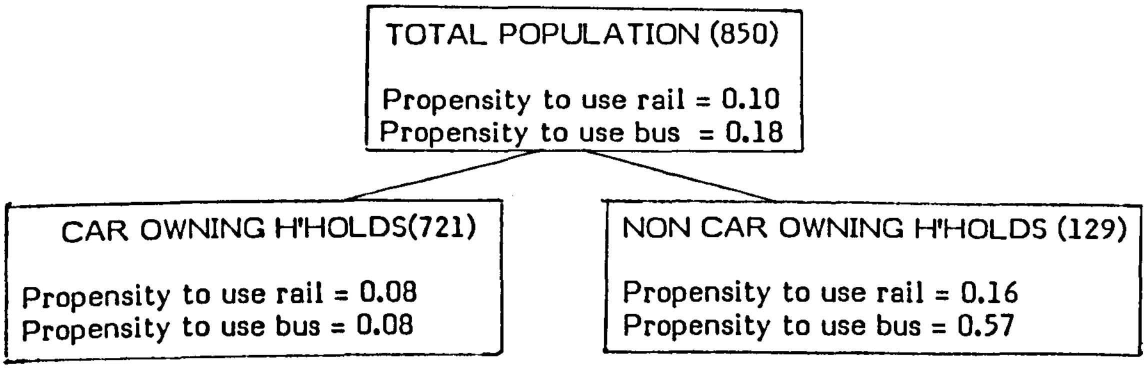


FIGURE 7.1 MARKET SEGMENTS IN THE CORRIDOR DATA SET

As Table 7.12 shows we initially developed HL models for car owning (CO) and non car owning (NCO) households, with models 1 and 2 being the lower split results and models 3 and 4 the upper split results for CO and NCO household members respectively. It can be seen that despite the small sample sizes all the parameter values in the lower split models are significant at the 90% level (although two parameter values are insignificant at the 95% level). For CO households the upper split model (model 3) produces a significant EMU parameter within the relevant range, although the ASC parameter value for car passenger and the OVT parameter value are insignificant at the 95% level. However, it is evident that for members from NCO households the HL structure does not apply as model 4 indicates that the EMU parameter value is insignificant and "pathological" (i.e. outside the prescribed range) whilst all other coefficients are insignificant. This is likely to be because the choice car driver is not available to NCO households.

A number of alternatives might be persued in developing a mode choice model for NCO households.

(i) Ignore the role of car passenger and base models on the binary bus/train split given by model 2 of Table 7.12. However, almost 27%

	1	2	3	4	5	6
ASC - Car Passenger			-0.339* (-0.596)	1.381* (1.729)	-0.844* (-1.305)	
ASC - Car (2) Driver			1.597 (2.786)	-0.053* (-0.029)	0.427* (1.004)	
IVT - Train	-0.111* (-1.785)	-0.189 (-2.154)				
IVT - Bus	-0.118 (-2.605)	-0.116 (-2.093)				
IVT - Car Driver (3) and Passenger			-0.064 (-3.178)	-0.002* (-0.062)	-0.029* (-1.339)	-0.007* (-0.041)
Walk time	-0.191 (-2.565)	-0.137 (-2.906)			-0.090 (-2.630)	-0.0514 (-2.538)
Wait time	-0.276 (-3.997)	-0.190 (-2.906)			-0.071 (-2.335)	-0.039* (-1.553)
OVT - Car Driver and Passenger			-0.059* (-1.481)	-0.040* (-1.083)		
Total Cost (1)	-0.067 (-2.196)	-0.064* (-1.822)	-0.013 (-4.176)	-0.011* (-1.075)		
EMP - Public Transport			0.377 (4.966)	-0.042+* (-1.024)		
Availability - Passenger					-3.012 (-4.643)	-3.517 (5.761)
Number of observations	97	67	721	129	173	173
Log likelihood	-28.62	-21.53	-126.188	-34.03	-72.73	-77.23
Adjusted Rho Squared	0.574	0.483	0.803	0.571	0.500	0.469
Adjusted Rho-bar Squared	0.551	0.441	0.803	0.561	0.491	0.463
% right	81.9	79.3	90.3	81.9	72.4	70.8

TABLE 7.12 RESULT OF MODELS WBA9F - MARKET SEGMENTED APPROACH

(1) Lower split - Bus and Train: Upper split - Car Driver and Passenger

(2) In Model 5 this is ASC-Bus

(3) In Model 5 this is in-vehicle time - Car Passenger, Bus and Train

of trips were made by car passenger (i.e. greater than rail) and therefore its exclusion is difficult to justify.

(ii) Develop a three-way (car passenger, bus and train) MNL model for NCO households (there was only one case where a person from a NCO household was a car driver). However, initial attempts to develop a MNL model produced insignificant parameter value for the ASCs, IVT and total cost, with the latter value being of the wrong sign. In part this may reflect the difficulty in handling car passenger as

"Even within the context of the journey to work, the mode "car passenger" is never satisfactorily handled" (MVA et al. 1986, p54).

In order to get round these problems the following amendments were made:

(i) The data set was expanded to include observations with only partial SE data and to include observations where no alternative mode was given. This expanded the data set to 173 observations.

(ii) Car passenger was specified as an alternative for all 173 observations;

(iii) Where car passenger had thus been introduced to the choice set a variable Availability-Passenger was set to 1. This may be interpretated as an indicator of those households who do not have a contact with CO households (and thus we expect this coefficient to be negative).

(iv) Where car passenger had been introduced to the choice set engineering times (based on link flow speeds) and costs (fuel costs and parking charges divided by occupancy) were introduced.

The resultant MNL model is shown by column 5 of Table 7.12 (WBA9F-5). From this model it can be seen that the only parameters significant at the 95% level are the out of vehicle times and the Availability-Passenger dummy variable, and in particular it should be noted that when cost was introduced its parameter value was highly insignificant. Model 6 re-estimates model 5 without the ASCs but a log-likelihood ratio test gives a χ^2 test statistic of 9.0 which is greater than $\chi^2_{0.05}$ at 2 degrees of freedom (= 5.991) suggesting the gain in 2 degrees of freedom does not compensate for the deterioration in log-likelihood.

7.3.5 Conclusions

In this section we have shown that introducing SE attributes into an HL formulation is not straightforward. Inclusion of car ownership and/or income as pure or indirect ASVs can result in significant parameter values (at least at the 90% level) but does not appear to lead to large increases in goodness of fit. As a result we have developed a segmented market model which consists of an HL model for members from CO owning households, based on model WBA9F-1 (Table 7.12) for the lower split and WBA9F-3 for the upper split, and a MNL model for members from NCO households, based on model WBA9F-5. It is evident that such a segmented approach may have advantages over the single market model (based on WBA9C-4 and WBA9D-1) in terms of the percentage of Public Transport trips correctly predicted. This may be calculated as being 61% for the single market model, whilst for the segmented approach it is 62% for members from CO households and 77% for members from NCO households (giving 65% overall).

7.4 SUMMARY

Using the work of MVA et al. (1986) on the 1981 West Yorkshire Corridor data set as a starting point it has been possible to calibrate a series of disaggregate mode split models. The log-likelihood ratio test carried out in 7.1.5 indicates that the IIA axiom may be violated and hence an HL structure is preferable to a MNL structure. This was confirmed by the fact that when HL models were developed the EMU parameter, ϕ, was significant and within the prescribed range.

Initially a single market HL model was developed, based on a lower split formulation given by model WBA9C-4 and an upper split formulation given by model WBA9D-1. Such a model did not, however, include SE variables and it was found that the most appropriate way of including such variables was by segmenting the market in terms of car ownership. Thus a market segmented approach was adopted, consisting of an HL model for members from CO households, based on models WBA9F-1 (for the lower split) and WBA9F-3 (for the upper split), and an MNL model for members from NCO households based on model WBA9F-5. It is this market segmented model which will, in the main, be used for further analysis in the next chapter as, although it has a number of problems, in particular the low significance of certain parameter values, it also has a number of advantages. In particular the parameter values, at least for CO households, seem plausible as the implied in-vehicle value of time for private transport users is £2.95 per hour, although for Public Transport users it is between £0.99 and £1.06 per hour which seems to be consistent with the recommended values of MVA et al. (1986, p169). In addition, as we have seen, such a model structure has advantages in terms of the number of Public Transport choices it correctly replicates, whilst we shall see in 8.3 that a segmented structure has some advantages in terms of aggregation.

In this chapter we have encountered a number of theoretical problems (a point which we shall return to in 11.2). In particular we have had problems in devising appropriate goodness of fit statistics, particularly when comparing models based on different structures and numbers of observations. This point will be examined further in 8.1.1. It is also difficult to interpret parameter values. Hence in section 8.1.2 we shall attempt to show how elasticity measures may be derived. Lastly, as was noted in 4.2.4, the data set used was of mixed quality and in particular there were problems, for our purposes, due to the choice set specification used, incomplete attribute sets and suspected correlation between in-vehicle time and cost variables. Thus the amount of research effort that could be spent on the calibration of disaggregate models was limited by data constraints, as well as the need to produce forecasts of potential new station usage and a subsequently detailed evaluation within a three year time period.

CHAPTER EIGHT DISAGGREGATE MODELS (2) : FURTHER ISSUES

8.1 INTRODUCTION

In this Chapter we examine some of the issues raised by the work in the previous chapter. In the rest of this section we shall describe in detail the goodness of fit measures that have been used and highlight some problems. We shall also examine the problem of obtaining elasticity measures from our disaggregate models. We shall then go on to examine problems in applying our preferred models. In a second section we shall examine the issue of transferability and will test the segmented market models' spatial transferability by making use of information on trip making behaviour for five of the six new station areas, provided by the Countywide data set described in section 4.2.5. In a third section we shall go on to examine the issue of aggregation and will assess the accuracy of the market segmented models' predictions of the number of work trips by rail.

8.1.1 Goodness of Fit Measures

Problems with determining suitable goodness of fit measures for disaggregate choice models estimated using Maximum Likelihood have been well documented (see, for example, Stopher, 1975 or Horowitz, 1985B) and stem from the fact that a goodness of fit measure comparable to R^2 in OLS does not exist. However a "pseudo R^2" goodness of fit measure, rho-squared, has been defined as:

$$\rho^2 = 1 - \frac{1^*(\hat{\theta})}{1^*(o)} \qquad \text{Equation 8.1}$$

where $1^*(\hat{\theta})$ is the maximised value of the model log likelihood and $1^*(o)$ is the value of the null log likelihood, typically evaluated such that the probability of choosing the i^{th} alternative is exactly equal to the observed aggregate share in the sample of the ith alternative (achieved by the constant only model with i-1 ASCs specified).

It has, however, been noted that although ρ^2 behaves correctly at the limits (i.e. 0 and 1) it does not have an intuitive interpretation between the limits (Hauser, 1978). Moreover:

"Those unfamiliar with the ρ^2 should be forewarned that its values tend to be considerably lower than those of an R^2 index (of regression analysis) and should not be judged by the standards for "good fit" in ordinary regression analysis. For example values of 0.2 to 0.4 for ρ^2 represent an excellent fit". (McFadden, 1976).

Because a ρ^2-like index can, in principle, be computed relative to any null hypothesis it is important to choose an appropriate one. We have followed the procedure of MVA et al. and calculated L* (C) as the maximised value of the log likelihood of the base model which takes into account both market shares and knowledge of the restricted choice set. This is done by specifying a total time variable and then excluding choices where total time = 0. We thus developed our adjusted rho squared measure as:

$$\rho^2 = 1 - \frac{l^*(\hat{\theta})}{l^*(c)} \qquad \text{Equation 8.2}$$

Equation 8.2 is more powerful than 8.1 as it is comparable between models with different data samples (and hence different market shares). Equation 8.2 may be further adjusted to take into account degrees of freedom.

$$\bar{\rho}^2 = \frac{1 - L^*(\hat{\theta}) \mid \sum_{q=1}^{Q} (J_q - 1) - K}{L^*(c) \mid \sum_{q=1}^{Q} (J_q - 1)} \qquad \text{Equation 8.3}$$

where J_q = the number of alternatives faced by individual q and
K = the total number of individuals in the model.

In addition a predicted success measure, such as the First Preference Recovery (FPR) or percentage correctly predicted or percent right, may be developed (Gunn and Bates, 1980). This is defined as:

$$FPR = \sum_q P_q{}^m \qquad \text{Equation 8.4}$$

where $P_q{}^m$ denotes the maximum probability of choosing any 1 of the options in the case of individuals q.

This might then be compared to a number of other measures such as the Chance Recovery (CR), predicted using the equally likely model or the market share recovery rate (MSR) from the null model. However these measures should be used with caution as:

(i) Under certain conditions, these measures can lead to the selection of an incorrectly specified model when compared to a correctly specified model, even when the effects of random sampling errors are negligible (Horowitz, 1985B).

(ii) A high value of FPR may lead to model rejection. A model can only be accepted as reasonable and informative if the FPR and the expected rate ($ER = \sum_q P_q$ where P_q is the calculated probability associated with the best option for individual q) are similar and are both larger than CR (Ortuzar, 1980B, pp36-38).

(iii) A test which weights each correct prediction equally will not be suitable for circumstances where some options (for example rail) are more important than others.

(iv) It is possible that a model might be good in prediction with respect to the estimation sample, while not necessarily predicting well the outcome of any policy change, defined in terms of movement in one or more of the individual variables. Hensher and Bullock (1979) prescribe a before and after assessment procedure as the best test of predictive strength. We shall carry out a form of such a test, with reference to the policy of new station openings, in the next chapter.

In BLOGIT predicted success tables are based on the summation of individuals' modal choice probabilities rather than allocating a probability of 0 or 1, depending on whether the modelled choice probability is less than a CR or MSR type measure or not. Thus each entry (N_{ij}) in the central matrix of the table is the probability of individual q selecting alternatives j summed over all individuals who actually select i i.e.

$$N_{ij} = \sum_{q \varepsilon Q_i} P_q(j|A_q) \qquad \text{Equation 8.5}$$

where A_q is the set of alternatives out of which individual q chooses and Q_i is the set of individuals in the sample who actually choose alternative i.

Column sums (predicted counts) are thus calculated as:

$$N_{\cdot j} = \sum_{i \varepsilon A_q} [\sum_{q \varepsilon Q_i} P_q(j|A_q)] \qquad \text{Equation 8.6}$$

Row sums (observed counts) are calculated as:

$$N_{i\cdot} = \sum_{q \varepsilon Q_i} [\sum_{j \varepsilon A_q} P_q(j|A_q)] \qquad \text{Equation 8.7}$$

$N_{ii}/N_{\cdot i}$ indicates the proportion of the predicted counts which actually choose that alternative. This might then be compared with the proportion predicted by CR or MSR measures to form a predicted success index, as described by Hensher and Johnson (1981, pp 52 to 56).

From the above discussion it should be clear that the use of the two most commonly used goodness of fit measures, rho-squared and predicted success tables, should be treated with care. Moreover the above discussion has been confined to MNL models, the use of HL models clearly compounds the problems. Thus it is apparent that:

"There are no hard and fast rules governing the choice of "best" model and there is a general problem in defining the trade-offs between goodness of fit and simplicity". (Gunn and Bates, 1980).

In this study, as we are not interested in the mechanics of goodness of fit statistics per se, we have attempted to make the best possible use of the above conventional measures, as well as likelihood ratios, t-statistics and informal tests of parameter value plausibility.

8.1.2 Elasticity Measures

Elasticity measures are of great importance in providing planning information. In this sub section we shall derive direct point elasticities for the single market HL model (WBA9C-4 and WBA9D-1). Dunne (1984) discusses how such an analysis might be extended to producing arc elasticities.

The derivation of elasticities for an MNL model is relatively straightforward (and in doing so our work will follow that of Lowe, 1981). Let the probability that individual n chooses mode i be $Pr_n(i)$ and let the value of the k^{th} variable for individual n be x_{ink}. Then a direct point elasticity of $Pr_n(i)$ with respect to x_{ink} is

$$E_{x_{ink}}^{Pr_n(i)} = \frac{\partial\, Pr_n(i)}{\partial\, x_{ink}} \cdot \frac{x_{ink}}{Pr_n} \qquad \text{Equation 8.9}$$

We know from definition that:

$$Pr_n(i) = \frac{\exp(B_k\, x_{ink} + \sum_{j \neq k} B_j\, x_{inj})}{\exp(B_k\, x_{ink} + \sum_{j \neq k} B_j\, x_{inj}) + \sum_{m \neq i} \exp(\sum_j B_j x_{mnj})} \qquad \text{Equation 8.10}$$

and hence:

$$\frac{\partial Pr_n(i)}{\partial\, x_{ink}} = \frac{B_k \exp(\sum_j B_j x_{inj}) \sum_m \exp(\sum_j B_j x_{mnj}) - B_k (\exp \sum_j B_j x_{inj})^2}{(\sum_m \exp(\sum_j B_j x_{mnj}))^2} \qquad \text{Equation 8.11}$$

$$= \frac{B_k \exp(\sum_j B_j\, x_{inj})}{\sum_m \exp(\sum_j B_j x_{mnj})} \cdot \left[1 - \frac{\exp(\sum_j B_j x_{inj})}{\sum_m \exp(\sum_j B_j x_{mnj})} \right] \qquad \text{Equation 8.12}$$

$$= B_k \, Pr_n(i) \; (1 - Pr_n(i)) \qquad \text{Equation 8.13}$$

Hence:

$$E_{x_{ink}}^{Pr_n(i)} = B_k \, x_{ink} (1 - Pr_n(i)) \qquad \text{Equation 8.14}$$

In addition it may be shown that the elasticity of $Pr_n(i)$ with respect to an attribute of mode l (i.e. cross elasticity) is

$$E_{x_{ink}}^{Pr_n(i)} = - B_k \; x_{lnk} \, Pr_n(l) \qquad \text{Equation 8.15}$$

In order to determine a direct elasticity average probability and variables values could be inserted into the micro elasticity value given by equation 8.14 (Richards and Ben Akiva, 1975) but such an approach will lead to biases (Dunne, 1984) due to aggregation problems (see section 8.3 for further details). Alternatively an aggregate elasticity can be formed as weighted averages of the individual elasticities given by 8.14, with the weight being the probability that an individual would choose the mode in question (Domencich and McFadden, 1975). Hence:

$$E_{x_{ik}}^{\bar{P}r(i)} = \sum_n Pr_n(i) \; E_{x_{ink}}^{Pr_n(i)} \Big/ \sum_n Pr_n(i) \qquad \text{Equation 8.16}$$

However for HL models calculating an elasticity in a subsequent equation for an attribute that appears in a previous equation, we need to include an allowance for the previous choice outcomes. A number of sources (Hensher and Johnson, 1981, p.93, McFadden, 1979, pp313-316, Williams and Senior, 1977C, p.39) give formulae for a variety of different sequential structures. It may be shown that if a simple nested structure, as provided by the single market HL model is adopted, then the elasticity of train in the lower split may be simply calculated as:

$$E_{x_{ik}}^{Pr_t} = \frac{\sum_n Pr_n(t) \, B_k x_{ink} \, (1 - Pr_n(t))}{\sum_n Pr_n(t)} \qquad \text{Equation 8.17}$$

where $Pr_n(t)$ = Probability that individual n chooses train in the lower split.

The elasticity of train in the upper split is then calculated as:

$$E_{x_{ik}}^{Pr_r} = \frac{\sum_n Pr_n(r)\ \phi\ B_k\ x_{ink}\, Pr_n(t)\ (1-Pr_n(r))}{\sum_n Pr_n(r)} \qquad \text{Equation 8.18}$$

where ϕ = EMU parameter and

$Pr_n(r)$ = Probability that individual n chooses train in the upper split.

These elasticities were computed through the use of FORTRAN programs and gave the results shown by Table 8.1.

	Upper split	lower split	Total
Walk	-0.101	-0.182	-0.283
Wait	-0.058	-0.092	-0.150
IVT	-0.087	-0.140	-0.227
Fare	-0.130	-0.213	-0.343

TABLE 8.1 RAIL'S DIRECT AGGREGATE WEIGHTED ELASTICITIES

It should be remembered that these elasticities were calculated for the journey to work only and allowed only for change of mode and hence we would expect them to be lower than the elasticities for the ASM for all trips. Comparison with 6.2.4 shows this to be the case.

8.2 TRANSFERABILITY OF DISAGGREGATE MODELS

By transferability we mean the usefulness of a model developed to describe behaviour in one context to describe behaviour in another context. This might involve a spatial, temporal or cultural dimension. We have seen (in section 3.3.4) that it has been claimed that disaggregate models have advantages over conventional models due to their greater powers of transferability. In this section this claim will be examined firstly by reviewing previous empirical studies and secondly by using the market segmented models (WBA9F-1, -3 and -5), calibrated for the Corridor data set, to explain mode choice for a sub-sample of the Countywide data set.

8.2.1 Previous Studies of Transferability

Empirical evidence on the transferability powers of disaggregate models is mixed. Although, due to lack of data, we are unable to explicitly study the issue of transferability over time, it may be noted that several temporal studies (Hensher and Johnson, 1977, Silman, 1979, Talvitie et al, 1980) indicate that disaggregate models offer strong transferability over short planning periods. By contrast studies by Train (1978) and Ortuzar et al. (1986) note a lack of model stability over time.

Evidence on transferability over space is similarly inconclusive. Atherton and Ben Akiva (1976) found similarity of parameters for models developed in Washington D.C., Los Angeles and New Bedford (Mass.) but Talvitie and Kirshner (1978) found statistically significant differences in studies of Washington D.C., Minneapolis/St. Pauls and San Francisco, as did Koppelman and Wilmot (1982) in a study of three sectors in Washington D.C. Ou and Yu (1983) review a number of such studies and come to the unstartling conclusion that:

"with proper selection of urban areas the transferability of demand models is possible under similar empirical and policy conditions".

They also show that the coefficients of SE variables are less transferable than those for LOS variables. Similarly Harvey (1985) notes that in general, even for work mode choice, model transferability is difficult to demonstrate, but certain variables do consistently appear significant in mode choice estimations (although not necessarily with the same or reasonably similar coefficients).

Aside from the "do-nothing" case there are at least four ways of accommodating transferability problems (Atherton and Ben Akiva, 1976):

(i) Adjustment of the ASCs, on the grounds that there is no theoretical basis for transferring terms which account for all the dimensions not explicitly explained by the model. Koppelman et al. (1985) show that such an approach can account for almost one half of the difference between full transfer and local estimation.

(ii) Enrichment by re-estimation of the coefficients with a small disaggregate sample (Coslett, 1981). Koppelman et al. (1985) suggests that the amount of data needed for updating is less than one-fifth of that needed for model development.

(iii) Re-estimation of the constant terms and of a scalar to weight all other coefficients so that the ratios between them are unchanged (i.e. the marginal rate of substitution between attributes is unchanged). Various ways of doing this are investigated by Gunn et al. (1985) who recommend scale parameters for functional sub-groups of variables.

(iv) Bayesian updating using original coefficients and the coefficients from a small disaggregate sample. Apart from the parameter estimate this requires knowledge of the related variance-covariance statistics.

In terms of simplicity we favour the use of method (i). An interesting pragmatic approach to overcoming the problems of transferability was developed by the SIGMO study of travel demand in the Amsterdam conurbation (Ruhl et al. 1979, Holder and Wood, 1980) which made use of factors that balance observed and predicted trips in terms of trip lengths and area to area movements.

8.2.2 Testing the Transferability of the Market Segmented Model

As described in section 4.2.5 in order to validate the market segmented model use was made of 343 observations obtained from the Countywide West Yorkshire Transportation Study update. These observations were work trips between new station catchment areas (prior to new stations being opened) and rail served destinations. It was shown in 4.2.5 that this data set was deficient in terms of car costs, the availability of alternative modes and the times and costs of alternative modes. The modal split for this data and the range of alternatives is shown by Table 8.2.

It can be seen that in the validation data set bus accounts for almost 44% of work trips, compared to just 18% in the calibration data set. This is partly due to the reduced car ownership in the validation data set, with 33% of trip makers coming from NCO households compared to just 15% in the calibration data set. It may also reflect the greater emphasis on trips to the CBDs of Leeds, Bradford and Huddersfield which accounted for 68% of trips.

	Chosen	%	Available but not chosen
Train	10	(3.4)	0 in before situation, 333 in
Bus	150	(43.7)	193 after situation
Car Driver	125	(36.4)	15
Car Passenger	58	(16.9)	285

TABLE 8.2 MODAL CHOICE AND AVAILABILITY - VALIDATION DATA SET

In terms of alternatives bus and car passenger were assumed to be available to all trip makers. In the before situation, due to problems dealing with mechanised access modes, rail was assumed to be unavailable but in the after situation (i.e. new station opened) this mode too will be assumed to be available to all trip makers. In the case of car driver it was assumed to be only available to driving licence holding members of CO-households, with 89% of those with a car available making use of it in the journey to work.

Car costs and times were calculated as in section 6.2.2 as was in-vehicle time, wait time and fare for rail and bus, although walk time was calculated from the distance from the fine zone centroid to the rail station (or the nearest relevant bus stop), assuming a mean walk speed of 5 km/hour (Lupton, 1976).

Given the limited number of observations where rail was the chosen mode or an alternative in the before situation and the limited number of overall observations we were unable to recalibrate the NCO MNL model (111 observations) or the CO lower split model (75 observations) but we were able to recalibrate the CO upper split model as shown by Table 8.3, model 1 (TRAN-1).

Comparison of TRAN-1 with WBA9F-3 (Table 7.12) shows that there have been significant changes in most parameter values. Moreover, ASC-Car Driver, IVT and OVT coefficients have changed signs, whilst the ASC-Car Driver, OVT and EMU parameter values are insignificant at the 95% level (although the last two values are just significant at the 90% level). Given the differences between the calibration and validation data sets it is not surprising that the market segment model is found to lack transferability. In particular in the validation data set the number of CO household members who have car passengers available as an alternative has been overstated and the engineering times and costs of car passenger possibly understated compared to the calibration data set. As a result the ASC for car passenger is highly negative, reflecting the difficulty of even members from CO households to arrange lifts. From the above it should be evident that, using the statistical tests prescribed by Koppelman and Wilmot (1982), the hypothesis that model WBA9F-3 is transferable to the Countywide data set can be rejected. For example the chi-squared test of restriction is 241.33 compared to $\chi^2_{0.05}$ of 12.592 with 6 degrees of freedom, as the log-likelihood of the fully restricted model was -177.068.

	TRAN-1		TRAN-2	
	Parameter value	(t-stat)	Enriched ASC	(t-stat)
ASC - Passenger	-4.076	(-3.426)	-2.476	(-11.263)
ASC - Car Driver	-0.638+	(-0.533)*	1.141	3.706)
IVT - Car and Passenger	0.263+	(4.051)	-0.064	-
OVT - Car and Passenger	0.944+	(1.663)*	-0.059	-
Cost - Car and Passenger	-0.103	(-6.105)	-0.013	-
EMU	0.237	(1.745)*	0.377	-
Number of observations	232		232	
Log likelihood	-56.403		-121.835	
Adjusted Rho Squared	0.681		0.312	
% right	91.9		71.3	

TABLE 8.3 RESULTS OF TRANSFERABILITY OF MODEL WBA9F-3

+ = wrong sign * = insignificant at the 95% level.

In model 2 of Table 8.3 (TRAN-2) we restrict the time and cost parameters to be as in WBA9F-3 and thus enrich the ASCs. As expected the ASC for car passenger is still highly negative whilst that for car driver is positive with both parameters significant at the 95% level. However, it is clear that compared to TRAN-1 there has been a marked deterioration in log likelihood, whilst the number of choices correctly replicated decreases from 92 to 71%. However by comparing log likelihoods it is clear that TRAN-2 is an improvement on both the fully restricted and base (see 8.1.1) models and hence will be used in later analysis.

8.3 AGGREGATION OF DISAGGREGATE FORECASTS

Aggregation is the process by which individual choice estimates are expanded over the population of interest in order to obtain a reliable, unbiased forecast of group behaviour (Westin, 1974). A problem arises because, in many cases, the function of averages of variables is not the same as the average of functions (see, for example, Fowkes and Wardman, 1985). Mathematically we can write:

$$E\,(f(x)) \neq f(E(x)) \qquad \text{Equation 8.19}$$

when f is a non linear function (as is the case for the logit model).

Aggregation thus becomes a problem in situations where data for a complete set of individuals are not available and use has to be made of zonal averages. In this section we shall first assess the extent of the aggregation problem in previous studies and secondly examine how this problem affects the prediction of usage at five new stations.

8.3.1 Previous work on the Aggregation Problem

The extent of errors arising from the aggregation problem has been well researched. Westin (1974) has shown that elasticities based on a mean individual will overestimate the true aggregate elasticity by the ratio E(p) (1-E(p)) to E[p(1-p)] (where p = probability of choosing mode m). For data collected in the Glasgow-Edinburgh corridor aggregate elasticities were shown to be overestimated by 28.5% in absolute terms, (Watson, 1975). Similarly McFadden and Reid (1975) using San Francisco data showed that aggregate elasticities were overestimated by 22% and this could be shown to be equivalent to $\sqrt{\sigma^2_{ij}+1}$ where $\sigma_{ij}^2 = \beta' A\beta$ and β' = individual parameter, β = aggregate parameter and A_{ij} = intrazonal covariance. Further studies by Koppelman (1976A), Talvitie et al. (1980), Landau (1978) and Reid (1979B) have found similar, or in some cases higher, errors, and have attempted to compare them with other error types, in particular related to transferability and model misspecification.

Koppelman (1976B) has suggested a number of alternatives to the "naive" application of disaggregate parameters with aggregate data (see Figure 8.1).

(i) The naive aggregate method may be modified to take into account choice set availability. For example in the market segment model NCO households do not have car driver available as a choice.

(ii) Enumeration procedures, representing explicit theoretical relations between aggregate and disaggregate demand, may be developed. These include:

(a) Complete enumeration which averages the choice probabilities for all individuals in the prediction group, thus avoiding the aggregation problem. The practicality of this approach is limited by its data intensity.

(b) Sample (or partial) enumeration averages the choice probability for a sample or subset of individuals in the prediction group. This approach may be based on at least three types of sampling strategy (Lerman and Manski, 1979): random, exogeneous or choice based (as described in section 4.3). Provided sampling error is minimised these approaches should act as a good approximation to complete enumeration whilst only requiring a fraction of the data.

(c) Pseudo-sample enumeration techniques have been developed that rely on synthetic household samples constructed by randomly sampling from the postulated distribution of LOS and SE data. This approach was also discussed in 4.3.

(iii) Classification procedures, as proposed by Koppelman and Ben Akiva (1977), assign members of aggregate groups to identifiable classes and use the average variable value to predict aggregate choice shares for each class. The segmented market model might be seen as a simple variant of this type of approach.

(iv) Summation or integration procedures may be used with either a known, estimated or assumed distribution (Kanafani, 1983), whereby conditional disaggregate choice probability estimates the probability density function for the independent variables.

(v) Statistical differential procedures explain aggregate shares as a function of the moments of the distribution. An aggregate function is obtained by linearising the disaggregate choice function by use of a Taylor series expansion and then taking expectations across aggregate prediction groups (Talvitie, 1973).

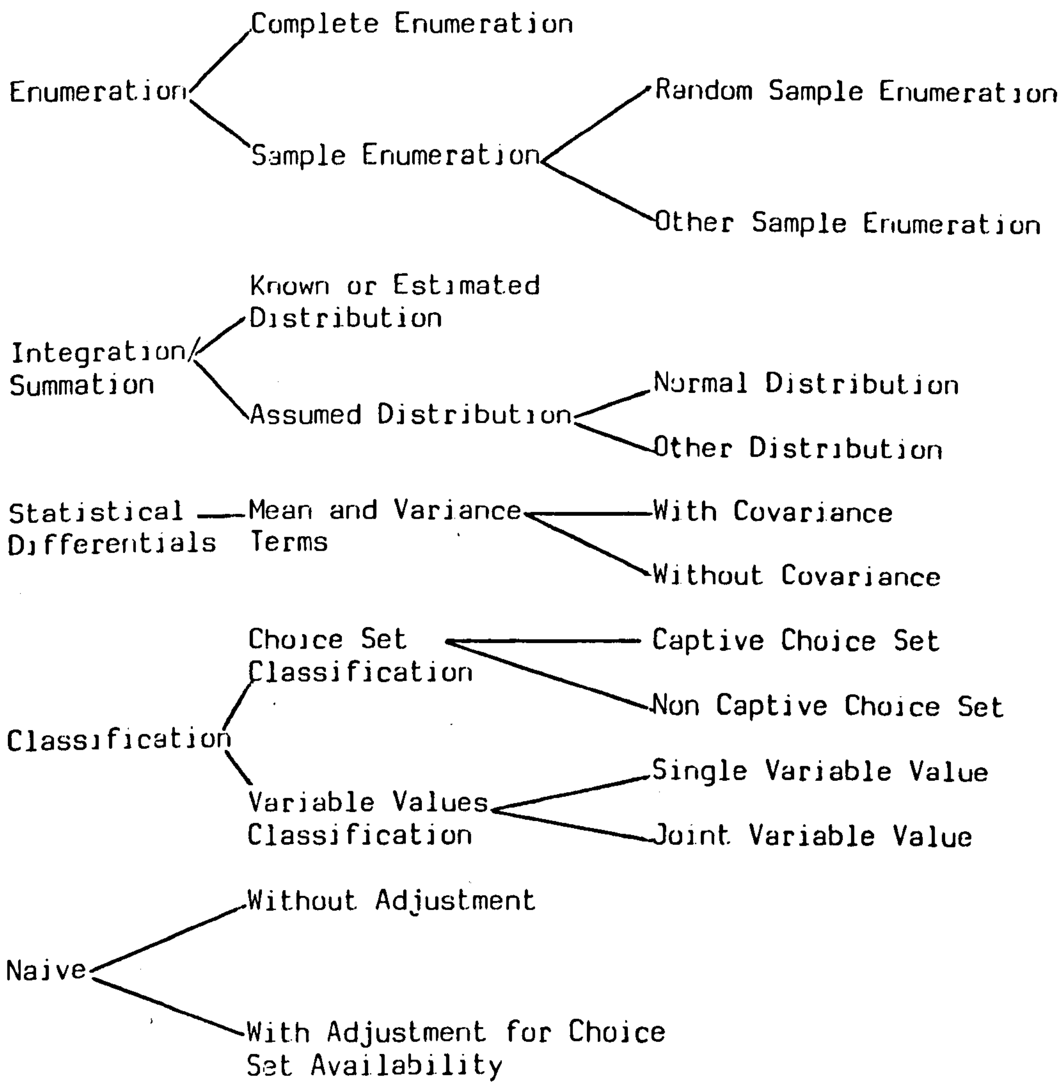


FIGURE 8.1 TAXONOMY OF AGGREGATION PROCEDURES
(After Koppelman, 1976B)

It should be noted that the above categories are not exclusive. The naive procedure may be seen as a special case of summation or integration procedures (the distribution is assumed to be concentrated at a point), statistical differential procedures (the series is truncated after the first term) or classification procedures (using one class only).

	Naive	Taylor series	Total Share Prediction by City	by Auto ownership	by Utility scales
Per cent RMS Error	40.0	121.0	17.9	21.7	3.1
No. of cells	-	-	17	4	4

TABLE 8.4 AGGREGATION ERROR BY FIVE METHODS

Empirical work by Reid (1979B), shown by Table 8.4, indicates that classification techniques reduce aggregation error and that under certain conditions the statistical differentials procedure can be expected to increase aggregation error. It was also shown that error grows substantially with aggregate size and that classification based on the differences of utility scales of major choice alternatives gives at least a five times lower error to class cell-count ratio than other classification criteria. Dunne (1985) has shown that statistical differential and integration (or density function) approaches performed well for predictions from a binary logit model but, in a multiple choice framework, these approaches are likely to be undependable and intractable respectively. These findings confirm the earlier work of Koppelman (1976A,B) who also found that aggregation error by the incremental prediction procedure is substantially lower than that by the direct procedure for all aggregation methods and for all policy changes. As a result for situations where enumeration procedures are not possible, due to lack of data, use will be made of an incremental approach. This will be described in 9.2.3.

Lastly, it should be noted that disaggregate parameters might be used in developing aggregate mode split models as occured with the GENMOD set of models in Amsterdam (Le Clercq, 1980) or the work of Hartley and Ortuzar (1979) in West Yorkshire. Such an approach again illustrates how differences between disaggregate and aggregate approaches must not be overstressed (Wootton, 1980).

8.3.2 Disaggregate Predictions of the Number of Work Trips by Rail

In order to estimate demand at five new stations use was made of the updated market segmented approach based on TRAN-2 (Table 8.3), WBA9F-1 and WBA9F-5 (Table 7.12). This approach was used in conjunction with the validation data set created from the Countywide data, as described in 4.2.5, which consisted of 249 observations (excluding Bramley-Gamble Hill). Rail was assumed to be available to all individuals. We were thus able to adopt a sample enumeration approach to aggregation. Predictions were produced by applying the HL model for CO households and the MNL model for NCO households by analysing the predicted success tables produced by BLOGIT runs, with the parameters restricted to the relevant values. The results are shown by Table 8.5. For the 10 cases where rail was chosen in the before situation a sub routine was developed to calculate the probability of using a new station (P_{NS}) as:

$$P_{NS} = 1/(1 + \exp(U_{OS} - U_{NS})) \qquad \text{Equation 8.20}$$

where U_{OS} = utility of travelling via existing station (calculated by using WBA9F-1 for co households, WBA9F-5 for NCO households)

U_{NS} = utility of travelling via new station.

Table 8.5 shows that the predicted rail share of the journey to work market to rail served destinations by mechanised modes varied from 14% at Crossflatts to 22% at Bramley. Using information from the 10% Journey To Work Tables from the 1981 Census it was possible to calculate the number of work journeys to all destinations originating within 800 m of a new station (excluding motor bike, pedal cycle, walk and other means). In order to estimate the number of rail work trips originating within 800 m a global weighting factor of 0.226 was applied which takes into account

(a) those work trips to a destination not served by rail

(b) non home based worktrips

(c) return trips

This factor was derived directly from the sampling procedure used.

Station		Mode Split Before	Mode Split After	Rail's Modal share	No. of mechanised work trips within 800 m (to work only)	Total predicted number of rail work trips	Actual number of rail work trips
Deighton	CP	4	4.4	0.1439	1090	43	39
	CD	3	3.0				
	B	34	27.7				
	T	0	5.9				
Crossflatts	CP	6	3.6	0.1338	1070	76	78
	CD	20	19.4				
	B	2	2.0				
	T	8	11.0*				
Slaithwaite	CP	18	11.7	0.1528	780	52	50
	CD	31	27.2				
	B	21	20.4				
	T	0	10.7				
Bramley	CP	1	1.4	0.2190	2350	145	156
	CD	12	11.7				
	B	8	3.3				
	T	0	4.6				
Saltaire	CP	11	8.8	0.1987	1170	84	139
	CD	36	29.7				
	B	36	29.7				
	B	32	26.4				
	T	2	16.1				
						RMSE	25.2
						AD(%)	16.0

TABLE 8.5 DISAGGREGATE PREDICTIONS OF THE JOURNEY TO WORK FOR FIVE NEW STATIONS

* of these 6 continued to be made via Bingley

CP = Car Passenger, CD = Car Driver, B = Bus, T = Train

In order to predict the total number of work trips use was made of the information on the proportion of trips originating within 800 m collected by market research. Clearly in most instances such information will not be available and the weighting factors adopted may introduce a greater degree of error.

Despite the small sample size and the caveat(s) given above it can be seen that for four out of five stations the method works very well, with overall predictions being around ± 16% (with a RMSE of 25). However, much of the overall error is caused by Saltaire where predicted work trip demand is only 60% of actual demand. This is, in part, likely to be due to the model's inability to include work trips attracted to Saltaire from other destinations. Overall, it can thus be seen that there is a tendency to underestimate the number of trips (in contrast to the aggregate approach) which may also, especially in the long run, be related to the inability to take into account generated trips. This suggests that our disaggregate approach, based only on a mode split model, may only be appropriate for short term applications. If we wished to analyse behaviour in the longer term our disaggregate approach would need to include trip generation and distribution stages.

A disaggregate approach gives additional information on:

(i) Previous mode. From the results given by Table 8.5 it was possible to derive information on the previous mode used, as shown by Table 8.6. This shows that although former bus users are the main group our disaggregate method overstates abstraction from private transport compared to our market research.

Number of new stations users	Rail	Bus	Car Driver	Car Passenger
41.3 of which	4.0(9.7%)	17.2(41.6%)	11.0(26.6%)	9.1(22.0%)

TABLE 8.6 PREVIOUS MODE USED - FIVE NEW STATIONS

(ii) Time savings. Individual switchers were initially identified by allocating users to rail if the choice probability of rail was greater than the choice probability of any other mode. This is equivalent to the Chance Recovery rate (CR). As Table 8.7 shows this only allowed us to identify a small number of switchers. As an alternative switchers were identified if the choice probability of rail was greater than the market share (0.224 for NCO households, 0.140 for CO households). This is equivalent to the Market Share Recovery rate (MSR). As Table 8.7 shows this came within one of the absolute number of switchers, as given by Table 8.5.

	No of. Observations	Rail's Mean Utility	Previous Mode Mean Utility	Time Savings (minutes)
CR				
NCO households	9	-1.574	-2.458	30.48
CO households	7	-2.888	-3.870	15.34
MSR				
NCO households	22	-1.705	-1.974	9.27
CO households	18	-3.033	-3.666	9.89

TABLE 8.7 TIME SAVINGS IMPLIED BY THE MARKET SEGMENTED MODEL - FIVE NEW STATIONS

Examination of the changes in utility shows that, applying the relevant IVT parameters (-0.029 for NCO and -0.064 for CO households) and noting that the NCO model does not include a cost parameter, the CR method indicates a mean generalised time saving for all users of 23.86 minutes. By contrast the MSR method indicates a mean generalised time saving of only 9.56 minutes. This is because the MSR method includes observations which have lower probabilities of choosing rail (compared to the CR method) and hence the differences in utility between rail and the previous mode are less marked.

8.4 CONCLUSIONS

In this chapter we have examined some of the problems of applying disaggregate HL and MNL models. It has been shown that conventional goodness of fit statistics must be interpreted with care and that precise comparisons of models with different numbers of observations, different structures and different choice sets are difficult. It has also been shown that direct point elasticities can be derived from an HL formulation and some results have been presented. We have also shown how our disaggregate mode split models may be affected by problems of transferability and aggregation. The possible magnitude of such problems and the large number of potential corrections have been noted from the literature. In applying the market segmented model to the 1981 Countywide data set we found that our model lacked spatial transferability and was thus updated by enriching the ASCs of the upper split model for CO households (model TRAN-2, Table 8.3). The updated market segmented model was then used to forecast the number of work trips to and from a new station, with aggregation being based on a sample enumeration approach, although in the next chapter we shall see that such an approach can not always be used. Our validation results show that the updated market segmented model is accurate in what it sets out to do (predict work trips by rail within a given area) but this is still a long way from providing a forecast of total station usage. In the next chapter we shall attempt to redress this fact. It should be noted that our results suggest that the updated market segmented model, calibrated on 1981 data, has some powers of temporal transferability as it makes reasonably accurate predictions of the number of work trips made by rail in 1983 (1984 for Saltaire).

PART FOUR : EVALUATION

CHAPTER NINE DEMAND FORECASTS

9.1 INTRODUCTION

This chapter will attempt to make use of the models developed in previous chapters in order to predict usage at potential station sites in West Yorkshire, as well as the six new stations already opened. These results will then act as input to our final analytical chapter in which new stations are evaluated in financial and social terms.

In the rest of this section we shall determine how potential station sites might be identified. In the next section we shall produce forecasts of usage at potential sites by using three methods; an aggregate approach and two disaggregate approaches (in fact a hybrid of aggregate and disaggregate models) based on sample enumeration and naive aggregation techniques respectively. In the following section the accuracy of these techniques, along with the simple trip rate model presented in 5.5.1 and the PTE regression model, will be assessed and the relative merits of each approach analysed. We shall then go on to look at the important issue of changes in demand at new stations over time.

It again should be stressed that this research attempts to make the most efficient use of existing data sources, even though these data sets may themselves be biased. Moreover, the emphasis of this work is on the practical rather than the theoretical. Our aim in this chapter is to develop relatively cheap and simple methods of accurately forecasting usage at new stations, which are, after all, relatively cheap and simple investments.

9.1.1 Identification of Potential Sites

In Section 2.2.3 we noted that a number of previous studies had put forward a wide range of potential new station sites in West Yorkshire but we were unable to identify a common rationale for site selection except for cases where the sites were self explanatory or were chosen due to political and/or consumer pressure. In this sub-section we shall attempt to develop a more scientific approach based on a "quasi-logical" search procedure designed to identify potential new station sites, in addition to the six stations opened between 1982 and 1984. This procedure consisted of the following stages:

(i) Exclude track not financially supported by the PTE under section 20 of the 1968 Transport Act or where further support seemed unlikely (for example Springwood Junction to Denby Dale, Altofts Junction to Leeds North Junction). After discussions with North Yorkshire County Council and WYPTE this was relaxed for the sections beyond Keighley and Knottingley stations.

(ii) Exclude track within 1 mile ($\approx$1.6 km) of an existing station.

(iii) Exclude track where engineering constraints would make the cost of a station prohibitive (for example the Bramhope, Morley and Wyke tunnels).

(iv) Exclude track passing through non built up areas. This was initially defined as less than 1000 population within 800 m, although in 11.4 this definition will be updated.

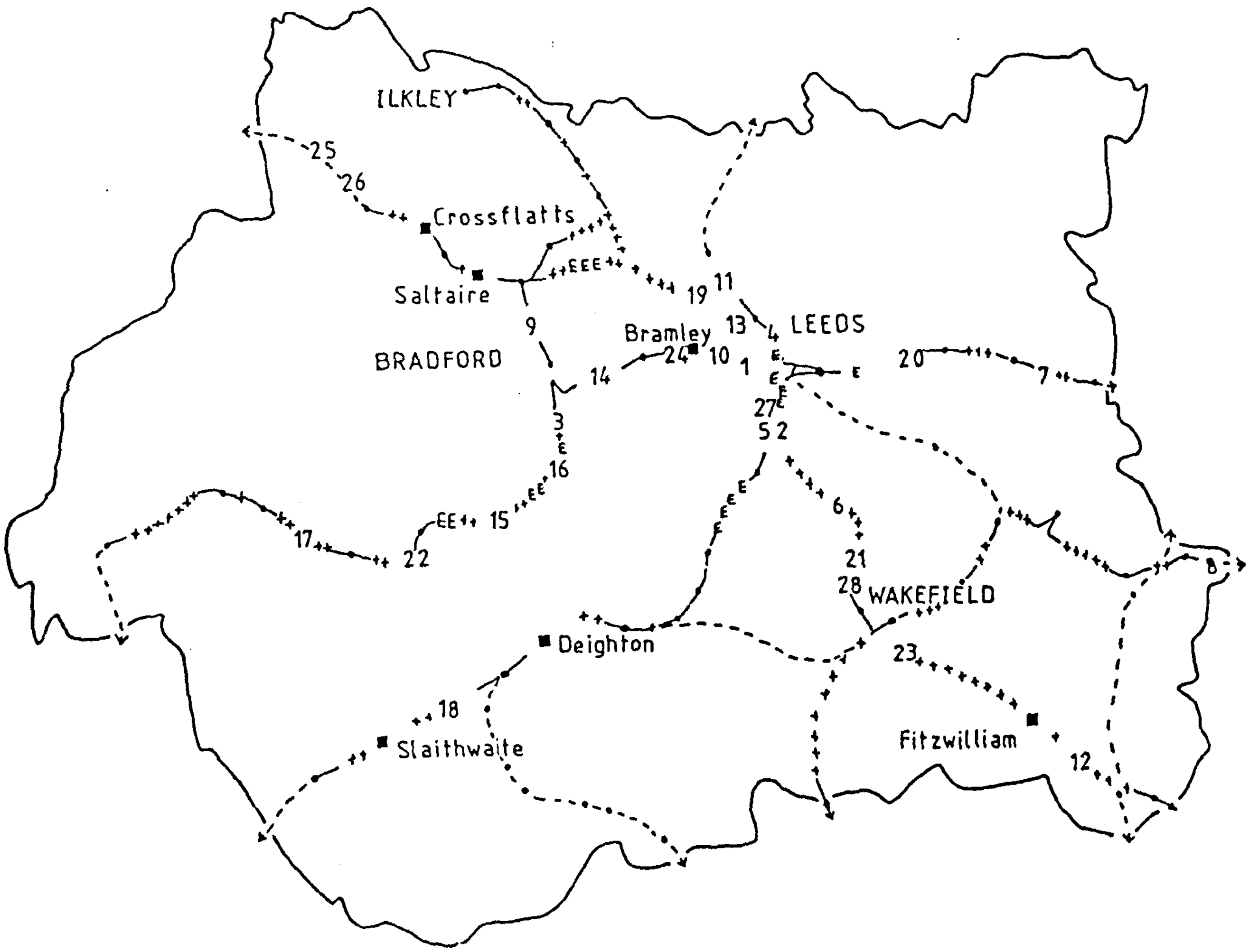
As Figure 9.1 shows this resulted in the identification of 29 sites of which one (Leeds Marsh Lane/Parish Church) may be excluded because its central area location would mean it was a trip attractor rather than a generator, whilst it would also be of a different design and cost compared to other new stations. Thus 28 potential sites were identified for further analysis.

9.2 PREDICTING DEMAND AT POTENTIAL STATIONS

9.2.1 Aggregate Approach

These forecasts were based on the ASM all-trips model given by Equation 6.8. Altogether for the 28 potential stations 41 flows were examined. The approach used was similar to the validation procedure of section 6.4.1. In particular it should be noted that:

(i) In order to produce total weekday usage results, modelled O/D flows were weighted by factors that were derived from information on destination choice at nearby existing stations with a similar level of service.



KEY

- • existing station
- ▪ new station opened 1982-84
- — section within 1 mile of existing station
- - - section not financially supported by the PTE, or support likely to be withdrawn
- +++ section alongside non built up area
- EEE section affected by engineering constraints

Potential stations:
1. Armley 2. Beeston 3. Bowling 4. Burley 5. Cottingley 6. East Ardsley 7. East Garforth 8. East Knottingley 9. Frizinghall 10. Gamble Hill 11. Hawksworth 12. Hemsworth 13. Kirkstall 14. Laisterdyke 15. Lightcliffe 16. Low Moor 17. Luddendenfoot 18. Milnsbridge 19. Newlay 20. Osmondthorpe 21. Outwood 22. Salter Hebble 23. Sandal 24. Stanningley 25. Steeton 26. Utley 27. Wortley 28. Wrenthorpe

FIGURE 9.1 IDENTIFICATION OF POTENTIAL NEW STATION SITES IN WEST YORKSHIRE

(ii) Some "a priori" information on characteristics of new station usage is required, particularly with regard to access distance and access modes, in order to determine access walk time. Excepting special cases (for example Hemsworth and Steeton where mechanised access modes are more likely to be used), we have had to make use of mean data from the new stations that have already been opened.

The results of this forecasting approach are given in column 1 of Table 9.1. In addition this approach has been used, at the request of BR Eastern region, to model potential usage at Crosshills and Cononley, in North Yorkshire, whilst at the request of WYPTE it was used to examine the effect of the rerouteing of the Leeds-Goole service. Further afield Hockenhull (1984) has used the aggregate model to assess patronage at five sites around Leicester (South Wigston, Blaby, East Goscote, Syston and Sileby), whilst Alderson (1984) applied the model to Dunston (Tyne and Wear).

In a similar manner it is possible to produce forecasts of the number of non-work trips on a weekday by applying the aggregate NWM (non-work model) given by column 3 of Table 6.11. These results are listed in Table 9.1, column 2. It should be noted that because the all-trips model was calibrated with 1981 data and the NWM was calibrated with 1984 data the number of work trips can not be determined by subtracting column 2 from column 1 of Table 9.1. However it can be seen that for all but two sites the number of all-trips (1981) exceeds the number of non-work trips (1984).

9.2.2 Disaggregate Approach - Sample Enumeration Method

We have seen in section 8.3.2 how the West Yorkshire Transportation Study update (Countywide Survey) was used to produce predictions of the number of work trips at five out of the six new stations that have already been opened. However only three out of the 58 survey areas covered coincided with potential station catchment areas; these were Bramley - Gamble Hill, Osmondthorpe and Wrenthorpe and involved 151 work trips to/from rail served destinations. It should be noted that at Wrenthorpe we have strong reason to believe the sample is biased, as no members from NCO households are included.

	Aggregate approach 1 All trips	(Rank)	2 Non work trips	"Disaggregate" Approach 3 Sample enumeration	4 Naive method work trips	5 all trips	(Rank)
Armley	371	1	134		264	398	3
Garble Hill	370	2	172	419	260	432	2
Newlay	358	3	106		60	166	14=
Burley	305	4	99		170	269	8
Osmondthorpe	293	5	142	363	350	492	1
Hawksworth	288	6	126		110	236	10
Kirkstall	236	7	116		180	296	6
Wortley	235	8	52		80	132	18
Stanningley	222	9	129		152	281	7
Salter Hebble	193	10=	98		68	166	14=
Wrenthorpe	193	10=	58	133	48	106	20=
Frizinghall	192	12	89		158	288	11
Sandal	191	13	45		52	130	19
Beeston	190	14	66		160	226	12
E. Garforth	188	15	116		266	382	4
Outwood	186	16	42		34	76	27
Cottingley	185	17	61		176	237	9
Bowling	182	18	75		256	331	5
Lightcliffe	161	19	72		34	106	20=
Utley	134	20	39		42	81	25=
Laisterdyke	132	21	96		82	178	13
Low Moor	125	22	38		126	164	16
Milnsbridge	119	23	70		74	154	17
Hemsworth	115	24	50		36	86	23
Steeton	112	25	55		36	89	22
E. Ardsley	103	26	47		34	81	25=
E. Knottingley	66	27	66		18	84	24
Luddendenfoot	47	28	53		8	61	28

TABLE 9.1 FORECASTS OF AVERAGE WEEKDAY USAGE

As in section 8.3.2 the updated market segmented model was used to determine the number of work trips from within 800 m, as shown by Table 9.2 (which is comparable to Table 8.5). The total number of work trips was derived by applying the same weighting factor (based on the proportion of work trips originating within 800 m) as used for Bramley. This figure may then be added to the total number of non-work trips predicted by the aggregate model in order to produce a total usage figure, which is given by column 3 of Table 9.1.

It may be seen from Table 9.3 (which is comparable to Table 8.6) that the vast majority (94%) of people using these three potential sites are predicted to be abstracted from bus. If the results of Table 8.6 and 9.3 are put together it can be seen that 67.7% of new station users are predicted to be abstracted from bus, 4.9% from rail, 14.1% from Car Driver and 13.3% from Car Passenger. This is broadly in line with the findings of our market research and those of other studies.

9.2.3 Disaggregate Approach - Naive Method

For 25 out of the 28 potential stations a source of disaggregate data did not exist. However, we have seen in section 4.2.3 that the 1981 Census Special Workplace Statistics - Section C provides aggregate information on journey to work flows.

It was proposed to use this data in conjunction with the incremental logit (IL) model, as used by Kumar (1980) in a study of the effects of an extension of the Shaker-Green line in Cleveland (Ohio). This might be expressed as

$$P'_T = \frac{P_T \exp(S'_T - S_T)}{\sum_m P_M \cdot \exp(S'_M - S_M)} \qquad \text{Equation 9.1}$$

where P'_T = proportion choosing train in after situation
P_T = proportion choosing train in before situation
S'_T = utility of train in after situation
S_T = utility of train in before situation

		Mode split		Rail's modal share	No. of mechanised work trips within 800 m	Total predicted number of rail trips
		Before	After			
Gamble Hill						
	CP	19	16.8	0.302	3260	277
	CD	22	22.6			
	B	53	26.2			
	T	0	28.4			
Osmondthorpe						
	CP	2	1.8	0.250	2400	191
	CD	7	6.5			
	B	19	12.7			
	T	0	7.0			
Wrenthorpe						
	CP	5	3.3	0.173	1350	75
	CD	16	17.7			
	B	8	3.0			
	T	0	5.0			

TABLE 9.2 DISAGGREGATE PREDICTIONS OF THE JOURNEY TO WORK FOR THREE POTENTIAL SITES

	Previous mode			
Number of New Station Users	Rail	Bus	Car Driver	Car Passenger
40.4 of which	0	38.1 (94.3%)	0.5 (1.2%)	1.8 (4.5%)

TABLE 9.3 PREVIOUS MODE USED - THREE POTENTIAL SITES

M = any mode in choice set (Train, Bus, Car Driver, Car Passenger).

However, as far as our work is concerned, there are at least two problems:

(i) Equation 9.1 is suitable for a MNL model but not an HL model.

(ii) In the case of new stations P_T is likely to equal 0.

In order to get round these problems we have adapted the work of Koppelman (1983) who developed the Extended Incremental Logit (EIL) model. For simple HL models of the type developed in Chapter Seven this would take the form:

$$P'_{PT} = \frac{P_{PT}\ \{\exp(S'_{NT}-S_{XT})+\exp(S'_{XT}-S_{XT})\}^{\phi}}{P_{PT}\ \{\exp(S'_{NT}-S_{XT})+\exp(S'_{XT}-S_{XT})\}^{\phi}+\{1-P_T\}} \qquad \text{Equation 9.2}$$

where P'_{PT} (P_{PT}) = Proportion choosing Public Transport in the after (before) situation

S'(S) = utility measure in the after (before) situation

XT = old Public Transport mode (bus), NT = new Public Transport mode (rail).

ϕ = EMU parameter

The lower split shares would then be

$$P'_{NT} = \frac{\exp(S'_{NT}-S_{XT})}{\exp(S'_{NT}-S_{XT})+\exp(S'_{XT}-S_{XT})} \cdot P'_{PT} \qquad \text{Equation 9.3}$$

and

$$P'_{XT} = \frac{\exp(S'_{XT}-S_{XT})}{\exp(S'_{NT}-S_{XT})+\exp(S'_{XT}-S_{XT})} \cdot P'_{PT} \qquad \text{Equation 9.4}$$

As, in most cases, we would assume no change in the utility of the existing Public Transport mode, $\exp(S'_{XT}-S_{XT})$ simplifies to 1.

For completeness private transport's share in the after situation may be defined as:

$$P'_M = \frac{P_M}{P_{PT}\{\exp(S'_{NT}-S_{XT}) + \exp(S'_{XT}-S_{XT})\}^{\phi}+\{1-P_{PT}\}} \qquad \text{Equation 9.5}$$

which is equivalent to

$$P_M \frac{1-P'_{PT}}{1-P_{PT}} = 1 - P'_{PT} \qquad \text{Equation 9.6}$$

The derivation of a model of this form is given by Appendix 3. The approach given by equations 9.2 to 9.6 needs to be modified in cases where rail has a significant market share in the before situation ($P_{NT} > 0$), as Appendix 3 shows.

The EIL approach has advantages in that it reduces the data requirements of the disaggregate approach as we only need (in the case of $P_{NT} = 0$):

(i) Information on the modal shares P_M and P_{PT} (= P_{XT}). This is provided by the Census SWS.

(ii) The utilities S_{XT} and S'_{NT}. Use was made of engineering times and costs, consistent with the conventions used in 5.5.2 (where we estimated times and costs of the previous mode used by new station users), 6.2.2 (where we estimated times and costs as part of our aggregate approach) and 8.2.2 (where we added times and costs to the 1981 Countywide Survey in order to form a validation data set). Details of the resultant data set are given by Table 9.4.

A further problem was encountered in that although car ownership information was available for each zone in the SWS data set, it was not available for individual flows. A way round this might be to use the microsimulation techniques discussed in section 4.3. This would involve allocating the probability of each particular trip being made by a CO or NCO

		Column No
Line 1		
(i)	Origin code (ORIC)	1 - 3
(ii)	Destination code (DEST)	5 - 7
(iii)	Proportion of non car owners choosing bus (NCOBUS)	9 - 12
(iv)	Bus access time, calculated as rail access time divided by number of competing bus stop pairs (BACC)	14 - 17
(v)	Bus egress time, calculated as rail egress time divided by number of competing bus stop pairs (BEGG)	19 - 22
(vi)	Bus wait time, calculated as a function of service frequency, as in section 2.1. Includes interchange calculated as 1/2 service interval, where relevant (BWT)	24 - 27
(vii)	Bus journey time, from published timetables (BIVT)	29 - 31
(viii)	Bus fare, derived from regional fares scale (BCOST)	33 - 35
Line 2		
(i)	Proportion of non car owners choosing rail via an existing station (NCOTR)	9 - 12
(ii)	Train access time, calculated as distance from new station zone centroid to existing station by proportion using each access mode by mean modal speeds (TRACC)	14 - 17
(iii)	Train egress time, calculated as distance to destination zone centroid from destination station by proportion using each egress mode by mean modal speeds (TREGG)	19 - 22
(iv)	Train wait time, calculated as a function of service frequency as in section 5.5.2. Maximum value set to 15 minutes (excluding interchange) (TRWT)	24 - 27
(v)	Train journey time, from published timetables (TRIVT)	29 - 31
(vi)	Train fare, from regional fares scale (TRCOST)	33 - 35
Line 3		
(i)	New station access time, calculated as distance from new station zone centroid to new station by proportion estimated to use each access mode by mean modal speeds. Assume 90% of passengers walk, 10% use mechanised access modes (NSACC)	14 - 17
(ii)	New station egress time, calculated as distance to destination zone centroid (NSEGG)	19 - 22
(iii)	New station wait time, calculated as a function of service frequency (NSWT)	24 - 27
(iv)	New station journey time, estimated from published timetables (NSIVT)	29 - 31
(v)	New station fare, estimated from regional fares scale (NSCOST)	33 - 35
(vi)	Number of journeys to work by members of non car owning households within catchment zone. SWS zonal flows (obtained via SASPAC) (NCOJTW)	37 - 39

Lines 1, 2 and 3 are then repeated for car owning households as follows:

Line 4		
(i)	COBUS Proportion of car owners choosing bus	9 - 12
(ii)	BACC2	14 - 17
(iii)	NEGG2	19 - 22
(iv)	BWT2	24 - 27
(v)	BIVT2	29 - 31
(vi)	BCOST2	33 - 35
Line 5		
(i)	COTR Proportion of car owners choosing train via an existing station	9 - 12
(ii)	TRACC2 N.B. Access modal split adjusted so as to reflect higher uptake of Private Transport	14 - 17
(iii)	TREGG2	19 - 22
(iv)	TRWT2	24 - 27
(v)	TRIVT2	29 - 31
(vi)	TRCOST 2	33 - 35
Line 6		
(i)	NSACC2 Again access time modal split adjusted so as to give better uptake of mechanised modes (30%)	14 - 17
(ii)	NSEGG2	19 - 22
(iii)	NSWT2	24 - 27
(iv)	NSIVT2	29 - 31
(v)	NSCOST2	33 - 35
(vi)	COJTW Number of journeys to work by members living within new station catchment zone. Again adjusted so as to be compatible with Census JTW flows.	37 - 39

For single market approach data set is as lines 1, 2 and 3 except:

(i) NCOBUS becomes the proportion of all trip makers using bus
(ii) NCOTR becomes the proportion of all trip makers using rail via an existing station
(iii) NSACC assume same access modal split as for aggregate model
(iv) NCOJTW becomes number of journeys to work within new station catchment zone

TABLE 9.4 NAIVE METHOD DATA SET

household through the use of a random number generator. If this random number was above (or below) a predetermined level this trip would be specified to a CO household. However, in order to develop such an approach information is required on the propensity of members from CO households to use different modes (which may be determined from the Corridor data set) and the propensity to make work trips of different lengths (which is not readily available). A crude attempt was made to synthesise CO and NCO trips but this led to an unacceptable increase in measurement error, mainly due to the fact that it seemed likely that we were overestimating the propensity of members from NCO households to make work trips by mechanised modes (as this propensity was assumed to be the same as for CO households). This would lead to an upward bias in our forecasts. As a result the market segmented model could not be used, and use was made of the single market model (based on models WBA9C-4 and WBA9D-1).

Using the EIL, P'_{NT} was calculated for 94 flows and then the total number of journeys to work from each new station i was calculated as:

$$T_i = \sum_j P'_{NT} \cdot T_{ij} \qquad \text{Equation 9.7}$$

The production of a total station usage figure requires knowledge of:

(i) The ratio of journeys from work by rail to journeys to work. We have assumed this equals 1 (i.e. demand is symmetrical).

(ii) The extent of new station catchment areas. This has normally been defined as being 800 m in radius and where possible coincides with SWS zonal boundaries. In order to produce a total station usage figure use has been made of the proportion of travellers originating within 800 m of a station, collected by our market research. Similarly, market research was used to define destination station catchment areas.

(iii) The number of non-work trips at each station. This was provided by the aggregate non-work model.

In column 4 of Table 9.1 the numbers of work trips estimated by this naive approach are given. By and large these figures are less than the total number of trips predicted by the aggregate model, although there are four exceptions.

Columns 2 and 4 of Table 9.1 have been added so as to produce a "disaggregate" estimate of the total number of trips, shown by column 5 of Table 9.1. It should be noted that using this approach fails to explicitly model education trips. Our market research shows that such trips are likely to be less than 9% of weekly usage for most flows, although at certain sites, for example Frizinghall, education trips may be expected to be more important. The most reliable estimates of travel by schoolchildren might be supplied by Local Education Authorities.

Despite this caveat, the forecasts in column 5 of Table 9.1 can still be compared with the aggregate forecasts in column 1. It can be shown that the mean value of the "disaggregate" forecasts is 202.4 trips, which is slightly higher than that of the aggregate forecasts (196.1 trips). However this masks a considerable degree of variation as a RMSE measure of the difference between the two forecasts gives a figure of 85.5 trips. It can be shown that 58.6% of this deviation was due to the "disaggregate" forecast being greater than the comparable aggregate forecast. In terms of ranking there are only two sites which are given the same rank by both methods, whilst there are some major differences, with the largest difference being of 13 places. Overall the mean absolute difference in rankings is 4.92.

9.3 COMPARISON OF THE FORECASTING APPROACHES USED

9.3.1 Comparison of the Disaggregate Forecasts

For up to eight sites at least three disaggregate approaches may be compared, as shown by Table 9.5.

i) Use of the disaggregate data provided by the Countywide survey in order to develop a sample enumeration approach (Column A).

ii) The use of the EIL to develop a naive approach based on zonal engineering times and costs, as discussed in the previous section (Column B).

iii) The use of the EIL to develop a naive approach based on zonal times and costs derived from the household data provided by the Countywide survey, which is a combination of engineering and reported data (column C).

The model used in these comparisons was the segmented market

	A	B	C	D
	Full choice set Sample enumeration	Incremental	Logit (IL)	"Actual"
		Naive- Engineering zonal data	Naive- Reported/Engineered zonal data	
Deighton	0.144	0.343	0.185	0.131
Crossflatts	0.139	0.297	0.156	0.137
Slaithwaite	0.153	0.121	0.142	0.147
Saltaire	0.199	0.212	0.201	0.329
Bramley	0.219	0.383	0.209	0.236
RMSE	0.059	0.148	0.064	
Absolute deviation	0.171	0.686	0.239	
Gamble Hill	0.302	0.282	0.314	
Wrenthorpe	0.173	0.126	0.193	
Osmondthorpe	0.250	0.521	0.225	
Weighted MEan	0.207	0.273	0.215	

TABLE 9.5 MARKET SHARE OF MECHANISED WORK JOURNEYS TO/FROM 800 m OF NEW STATIONS

HL/MNL model, for columns A and C and the single market HL model for column B. The "actual" market share figure was derived from our market research estimates of the number of rail work trips originating within 800 m of a new station divided by the number of total work trips originating within 800 m (from the Census) weighted by 0.226 (see section 8.3.2), and is given by column D.

Comparisons of columns A and C illustrate the effect of using aggregate data with a disaggregate model. As expected (see section 8.3.1) this results in a tendency to overpredict new station usage, although the relative increase in market share is only 4%. However compared to "actual" shares the AD measure has increased by 7 percentage points. This is in line with the findings of Talvitie et al. (1980) but less than that in many other studies (Westin, 1974, McFadden and Reid, 1975, Reid, 1979).

The market shares in model B involve a number of changes in particular with regards to the data set used, the model formulation used and the aggregation procedure adopted (in particular the zonal system used is slightly different). Compared to column A this results in a relative increase of overall market share of 31.9% (a 27.0% increase compared to column C). The increase in error with respect to column C, as measured by the AD measure, is equivalent to 44.7%, which in the main may be attributed to measurement error. This suggests that for column B around 25% of error might be specification error, 65% might be measurement error and 10% might be aggregation error.

9.3.2 Comparison of Aggregate and Disaggregate Forecasts

The accuracy of up to five forecasting methods may be tested for the six new stations already opened.

(i) The trip rate model outlined by Table 5.27. In applying this model it was assumed that data would be available for the five stations other than the one being evaluated. In reality, this is not the case, except for the last station built (Saltaire). Moreover, this approach is based on information collected for existing new stations (and there is evidence that at least initially, ceteris paribus, the level of demand at new stations is different from that at old stations).

ii) The aggregate all trips model given by Equation 6.8. This model has already been tested in Table 6.13.

(iii) The market segmented HL/MNL model of work trips, aggregated using the sample enumeration approach, combined with the aggregate model of non work trips. Data is, however, only available for five out of the six new stations.

(iv) The single market HL model of work trips, aggregated using a naive approach based on the IL, again combined with the aggregate model of non-work trips.

(v) PTE forecasts which were based on a simple regression model that included the number of households within 800 m and the number of weekday trains.

These five approaches are compared by Table 9.6. It can be seen that the forecasts of the trip rate model are within ± 42% of initial usage with a RMSE of around 71 trips. The caveats noted above should, however, be taken into account. In addition, it should be noted that this method fails to replicate the correct ranking of new stations and in this respect performs worse than the other approaches (see Table 9.7).

Of the four remaining approaches the most accurate, at least initially, is the market segmented HL/MNL approach. This method may be shown to give predictions on average around 34% above the initial usage with an RMSE of around 79 trips. It should be noted, though, that the work model alone was within ± 16% of initial work trips, whilst the non work model gave predictions, on average, 67% greater than initial non work usage. Given the weekday purpose breakdown in Table 5.13 this suggests the non-work model contributes 77% of the forecasting error of this approach.

By contrast the single market HL model gives predictions some 54% above initial usage, with a RMSE of around 97 trips. Most of this deterioration of accuracy might be attributed to the zonal data used. It can thus be seen that this approach is only slightly more accurate than the aggregate all-trips model which gave predictions some 63% above initial usage, with a RMSE of around 108 trips.

With the exception of the trip rate model (for three sites) and the

	B	C	1	2	3	4	5
					Disaggregate or hybrid approach		
	Second year usage	Nov. 1985 usage	Trip rate model	Aggregate model	Market segmented model - sample enumeration	Single market model - naive	PIE forecast
Fitzwilliam	1.61	1.95	1.39	1.31	-	1.84	2.25
Deighton	2.03	2.11	1.76	2.72	1.28	1.98	4.60
Crossflatts	1.57	1.83	0.80	2.32	1.48	2.03	3.33
Slaithwaite	0.78	1.31	0.67	1.22	1.05	0.92	1.54
Bramley	1.04	1.07	1.47	1.35	1.71	1.75	3.54
Saltaire	-	1.34	0.58	1.49	0.99	1.11	2.56
RMSE - Initial usage			71.3	108.3	78.6	96.9	335.8
- Second year usage				(63.7)	(84.7)	(77.1)	(291.8)
- Nov. 1985 usage				(48.8)	(86.3)	(71.7)	(275.9)
AD - Initial usage			0.422	0.630	0.343	0.540	1.936
- Second year usage				(0.363)	(0.373)	(0.312)	(1.392)
- Nov. 1985 usage				(0.211)	(0.331)	(0.238)	(1.015)

TABLE 9.6 FORECASTED WEEKDAY USAGE AT NEW STATIONS
(All figures (except RMSEs and ADs) divided by initial usage)

	Initial usage	Trip rate model	Aggregate model	Market segmented	Single market	PIE forecast
Fitzwilliam	6	4	6		5	6
Deighton	5	2	4	5	4	3=
Crossflatts	4	5	3	3	3	3=
Slaithwaite	3	6	5	4	6	5
Bramley	2	1	2	1	1	1
Saltaire	1	3	1	2	2	2
Nos correct		0	3	1	0	1
$\Sigma|F-A|$		12	4	4	8	7

TABLE 9.7 ACTUAL AND FORECASTED RANKING OF SIX NEW STATIONS

market segmented model (for one site), the forecasting methods discussed above all overpredict initial usage. They may, however, be seen as an improvement to the simple PTE forecasts which overpredicted usage by around 193%, with a RMSE of 336 trips. In the light of this experience the PTE have scaled down their forecasts. If the appropriate factor is used from Table 9.6 (= 0.341) it can be shown that such an approach still results in an AD of around 25%, with a RMSE of 41 trips, with initial usage being seriously underestimated at some sites. Such an approach clearly lacks a theoretical basis and should only be used when no other forecasting approach is available.

The fact that our models tend to overpredict initial usage may be due to the fact that new station demand builds up over time, as will be shown in section 9.4. In Table 9.6 later usage figures, derived from the November 1984 and 1985 PTS are presented. Compared to the November 1984 usage figures the accuracy of the ASM all trips model and the two hybrid approaches is broadly comparable, although compared to the November 1985 figures it is apparent that the ASM all trips model has some advantages. It is noticeable that the unadjusted PTE model still overpredicts usage by over 100%. It should be noted that our forecasts have not been adjusted to take into account reductions in real fares over the period 1983-85 which, with the introduction of Metrocard and the various off-peak schemes, are non trivial. The ASM all-trips forecasts will be most sensitive to such price reductions, whilst it should be noted that neither the non-work model used in both "disaggregate" approaches nor the MNL model for NCO households used in the market segmented approach include a cost parameter. Clearly a trip rate model can not take account of such changes, unless it was recalibrated. The effect of Metrocard on predicted patronage will be examined further in 10.3.4. Despite this it is interesting to note that by November 1985 the "disaggregate" approaches underpredict usage at four sites. This may, in part, reflect the inability of our disaggregate mode split modes to take into account the effect of generated work trips.

9.3.3 Comparison of Aggregate and Disaggregate Models

We are now in a position to make some comparisons between the aggregate and disaggregate methods that we have used. However it should be noted that:

(i) We are comparing different model formulations. The approaches used are based on different dependent and independent variables, different internal specifications and different overall structure. In particular the disaggregate approach is based on logit models of mode choice which, as a closed function, can only consider abstracted trips. We hypothesise that, at least in the short run, there are unlikely to be any generated work trips (this is confirmed by our market research). The aggregate model might be seen as dealing with trip generation, distribution and modal split simultaneously and is, to an extent, unbounded. The forecasts of such a model might be seen as being more indicative of a longer term equilibrium.

(ii) The various methods used were based on different data sets;

a) The sample enumeration method was based on the West Yorkshire Transportation Study update which involved household interviews. The times and costs of the preferred mode were reported, but those of the alternatives were based on engineering data.

b) The naive method is based on zonal engineering data concerning Public Transport times and costs and Census data regarding existing journey to work patterns.

c) The aggregate approach was based on Census data regarding origin and destination zonal characteristics and on zonal engineering data concerning Private and Public Transport times and costs.

Hence we are comparing specific models rather than methods and are unable to generalise to the same extent as previous studies (Watson, 1975, Hartgen et al. 1975). Nonetheless, the following comments might be made:

(i) A purely disaggregate approach is highly accurate in what it does. For example Table 8.5 shows that this method's forecasts of work trips were on average within 16% of the actual figure (and if Saltaire, which may be considered a special case, is excluded this figure falls to less than 6%). This is similar to the level of accuracy achieved by Harata and Ohta (1986) in a study of the application of disaggregate nested logit models to railway station and access mode choice in Greater Tokyo.

(ii) However the disaggregate models we have developed are limited in their extent. They can only predict the number of work trips by rail originating within a pre-defined area (which may or may not coincide with the station catchment area). If time had permitted, and the data had been readily available, the disaggregate approach might have included a model of non-work trips as well as generation and distribution stages.

(iii) By contrast the aggregate approach appears to be inaccurate, giving predictions 54% above initial usage. If however patronage growth over time occurs, and in the next section we have evidence that over five years new station patronage may grow by as much as 75%, then under this assumption the predictions of the aggregate approach would be within $\pm$ 26% of the actual usage (although the RMSE will have only fallen to 71 trips).

(iv) However the aggregate models we have developed are more comprehensive than our disaggregate models. In order for the "disaggregate" approach to produce a total station usage figure we require information on the areal extent of catchment areas for work travellers and an estimate of the number of non-work trips. Such aggregation procedures have led to an increase in forecasting error, as measured by the AD, of between 214 and 338%. This first figure refers to the results of the sample enumeration technique which is limited to instances where household interview data is available, but, as such data is limited, the more inaccurate naive method (given by the second figure) will normally have to be used, unless fresh data can be collected. As a result the "disaggregate" approach may only be 14% less inaccurate in predicting first year station usage than the aggregate approach, whilst if real patronage growth exists then the aggregate approach may even be more accurate.

(v) It is often claimed that a disaggregate approach is less data intensive than an aggregate approach (Spear, 1976, Hartgen et al. 1975). However, from our experience, in order to calibrate and validate a disaggregate model some 1294 observations were required, whilst the aggregate model required only 110 cases, although each case contains at least 25 observations. Moreover, the aggregate data is generally easier to collect and in this case already existed. Thus, aside from the need for greater expertise, the "disaggregate" approach involved at least treble the research effort compared to the aggregate approach.

(vi) A disaggregate approach makes available a potentially wider range of behavioural parameters, and in particular is able to assess in more detail the effect of out of vehicle time, as can be seen from Table 9.8.

	Aggregate - all trips	Disaggregate - single market work trips only
Price	-0.825	-0.343
In vehicle time	-0.418	-0.227
Wait	-0.600 (Wait and Walk combined)	-0.150
Walk		-0.283

TABLE 9.8 COMPARISON OF ELASTICITIES

As we have already noted, the elasticities implied by the disaggregate model are lower than those implied by the aggregate model as they refer to work trips only, which previous research (Kemp, 1973) and "a priori" reasoning have shown to be relatively inelastic. In addition a disaggregate approach allows us to calculate a value of time. However, it is interesting that the implied behavioural non-work (i.e. commuting and leisure trips) values of in-vehicle time are in most cases significantly higher than those being used in current practice (MVA et al, 1986 pp174-5).

(vii) Lastly, we have seen that a disaggregate approach provides important information on abstraction from previous modes and, if a sample enumeration approach is used, on the extent of time savings to new station users (see section 8.3.2). However, information on destination choice is limited (and may be better reproduced by the aggregate approach) whilst, as we have already noted, our disaggregate approach is unable to take into account the effect of generated trips.

9.4 DEMAND GROWTH OVER TIME

To an extent our analysis so far has been static in that changes in patronage over time at new stations has not been considered in detail. Past evidence for this in West Yorkshire is not conclusive. For example between 1974 and 1979 patronage at New Pudsey grew by 228% but patronage at Baildon fell by 23%. Time-series information on patronage growth at recent new or resited stations is limited, but Table 9.9 shows information collected for eight sites in Merseyside, South Yorkshire and the West Midlands. This Table shows that only one site (Chapeltown) exhibits continual absolute growth. However, we need to take into account exogeneous factors and this has been done crudely by finding the difference of the percentage change in station usage and the percentage change in system usage. This shows that five

Year	1		2		3		4		5		6	
	A	B	A	B	A	B	A	B	A	B	A	B
Garston[1]	2089	0	2766	14	3059	35	3565	62	3855	71	3601	49
Cressington[1]	867	0	1030	1	1070	12	1280	30	1155	19	2295	142
Aigburth[1]	1160	0	1499	11	1852	51	2665	112	1993	58	3104	145
St. Michaels[1]	1417	0	1448	-16	1373	-12	1620	-4	1556	-4	1919	12
Five Ways[2]	7401	0	6997	6	8957	35	7356	20				
University[2]	15201	0	12087	-8	12714	-2	10384	-11				
Longbridge[2]	13893	0	10144	-15	11086	-6	9210	-13				
Chapeltown[3]	2028	0	3109	24	3305	15	3453	22				

Notes

[1] reopened 1978. Figures refer to total joining/alighting per average weekday. 1983 figure (year 6) affected by extension of service to Hunts Cross. Source: Merseyside PTE

[2] opened 1979. Figures refer to total joining per week. Source: West Midlands PTE. Market Research Unit. "Annual Statistics Report. 1982-3. Rail Supplement". April 1984

[3] resited 1981. Figures refer to total joining/alighting per week. From 1983 onwards passengers using rerouted Huddersfield-Sheffield trains excluded.

TABLE 9.9 CHANGES IN DEMAND AT EIGHT STATIONS

A = Absolute usage figure.

B = Percentage growth of station patronage minus percentage growth of system patronage.

of the eight sites have exhibited positive relative growth over the survey period. The exceptions are provided by St. Michaels, University and Longbridge all of which may have been affected by decreased employment opportunities within their catchment areas. It is interesting to note that over five years the stations on the Garston branch exhibited real patronage growth of around 47%.

There is thus some evidence for the PTE's hypothesis that new station demand will grow by around 10% pa for up to five years after opening, but such evidence is not overwhelming. It was thus decided to investigate the effect of patronage growth at the six new stations already opened in West Yorkshire, as BR has carried out up to 31 counts covering, in the case of Fitzwilliam, a maximum period of 33 months. An overall pattern of growth is indicated by a simple regression using all the count data available (Table 9.10A). This suggests that patronage growth is around 2% per month on a weekday and 1.6% on a Saturday. The goodness of fit of such a regression is, however, low.

In order to improve the goodness of fit of such a regression the counts were combined into indices for three monthly periods. This greatly reduces the number of observations and emphasises the role of those stations that have been opened the longest. In addition a dummy variable (PROMO) was introduced for the first quarter in order to demonstrate the effect of promotional activity. Although only significant, at the 95% level, on a Saturday, the value of this dummy suggested that promotional activity boosted patronage by around 20%, as shown by Table 9.10B. This Table also indicates that weekday patronage was growing by around 2.6% per month (1.8% on Saturdays).

These results will, however, be affected by exogeneous factors, in particular the numerous fare changes during the period 1982-4. In order to take into account the effect of external factors indices were adjusted so as to take into account the system wide trend (based on total number of passengers on local services) and a measure of seasonal variation (based on Bellamy, 1978, and information provided by WYPTE). The results of such an adjustment are given by Table 9.10C. The adjustments suggest that, compared to 9.10B, the weekday patronage growth has declined to 2.1% per month (1.3% on Saturdays). This implies additive patronage growth of around 24% per annum.

	Intercept	Time	Promo	DW		R^2
Weekday	89.036	2.174[1]		1.648		0.333
	(10.080)	(3.672)	No. of observations		=	28
Saturdays	87.019	1.606		1.212		0.219
	(8.335)	(2.484)	No. of observations		=	23
(A)						
			Promo			
Weekday	73.270	7.693[2]	19.696	2.003		0.842
	(7.287)	(4.506)	(1.373)	No. of obs.	=	6
Saturday	80.603	5.490	19.055	2.709		0.891
	(14.159)	(5.680)	(2.347)	No. of obs.	=	6
(B)						
Weekday	-17.328	6.293[2]	11.855	2.101		0.810
	(-1.835)	(3.924)	(0.880)			
Saturday	-11.284	4.094	10.021	2.273		0.877
	(-2.426)	(5.182)	(1.510)			
(C)						

<u>TABLE 9.10</u> <u>PATRONAGE CHANGE AT SIX NEW STATIONS</u>

1 Monthly period

2 Three monthly period

DW = Durbin Watson statistic.

9.5 CONCLUSIONS

In this Chapter the forecasting accuracy of up to five modelling approaches has been assessed for the six new stations that have already been opened in West Yorkshire. Based on the absolute deviation from initial usage these approaches may be ranked in order of decreasing accuracy as:

1. Market segmented disaggregate model with aggregate non-work model (NWM).
2. Trip rate model
3. Single market disaggregate model with aggregate NWM.
4. All-trips ASM
5. PTE Regression Model.

Thus the market segmented approach was applied to five out of the six existing new stations, but due to lack of data could only be applied to three of the potential sites. Therefore for one of the existing new stations (Fitzwilliam) and 25 of the potential sites use was made of the single market approach based on zonal engineering times and costs and a naive aggregation procedure based on the use of EIL models (which was shown to greatly increase inaccuracy). It will be the demand forecasts from these models that will be used as the main input to the evaluation framework outlined in the next Chapter.

Our results show a trip rate model can provide reasonably accurate forecasts of initial usage at minimal research cost. However, such an approach does not provide information on destination choice or previous mode used which is needed to determine gross and net revenue. In addition, as we have already shown, such a model is not responsive to policy changes.

It was noted that the advantages of the "disaggregate" (or more strictly "hybrid") approaches over the aggregate approach further diminish if there is patronage growth over time, due to the former's limited capacity to incorporate generated trips. Evidence of real demand growth over time at new stations is not conclusive but it was shown that it may be reasonably expected to witness growth over the first five years of between 0 and 75% overall.

CHAPTER TEN EVALUATION RESULTS

10.1 INTRODUCTION

Having developed a full range of forecasts for existing and potential new stations in West Yorkshire, we are now in a position to put into practice the evaluation framework outlined in 3.4. This chapter will consist of three main sections. In the next section we will examine the main financial costs and benefits that affect new stations, whilst in the following section we will identify the main social costs and benefits, including re-analysis of the market research data presented in 5.5.2 in order to assess the extent of benefits to new station users. In the last main section we will consider a number of costs and benefits that have proved difficult to measure, although these will be evaluated for selected sites only.

The next two sections will include at least three sub-sections:

(i) "Actual" evaluation of the six new stations already opened, based on the usage data collected by BR and from our own market research.

(ii) "Modelled" evaluation of the six new stations, based on the demand forecast information provided by the updated market segmented approach, except for Fitzwilliam where the single market approach had to be used (see Table 9.6).

(iii) "Modelled" evaluation of the 28 potential station sites, based on the single market approach's demand forecasts (see Table 9.1), except for Gamble Hill, Osmondthorpe and Wrenthorpe, where the market segmented approach could be used (Table 9.2).

As described in 3.4.5, in presenting costs and benefits emphasis will be placed on using NPV measures, based on discount rates of 5, 7 and 10%, project life of 15 and 30 years and high and low growth rate assumptions. In addition, in order to make simple comparisons, use will be made of FYRR measures (see 3.4.4 for definitions).

10.2 FINANCIAL COST-BENEFIT ANALYSIS

10.2.1 Actual Financial Costs and Benefits at Six New Stations

For the six new stations that have already been opened in West Yorkshire, the following financial costs and benefits (to the PTE and BR taken together) may be taken into account:

(i) Capital cost of construction. Outturn costs were provided by WYPTE (1984B) and are shown in column 1 of Table 10.1.

(ii) Recurrent costs of maintenance and administration. Based on experience at existing local stations this was calculated by the PTE, at 1984 claim prices, to be, on average, £1700 per station per annum, as shown by Column 2 of Table 10.1. As the new stations selected were those which could be accommodated within existing timetables there was no identifiable effect on train service expenses. However the recurrent cost figure did include a notional amount to cover increased fuel and braking costs due to the additional stop.

(iii) Gross revenue. This was calculated by multiplying mean fare paid (from our market research) by patronage (from the PTS). No account was taken of Sunday traffic and in order to take into account holidays only 50 weeks (= 300 days) were considered. This is shown in column 3 of Table 10.1.

(iv) Net revenue. This was calculated as gross revenue times the sum of the proportion that was generated and was abstracted from private transport (from our market research).

(v) Existing revenue lost. The addition of a new station will lead to a time penalty estimated by the PTE as two minutes (Cottham, 1985), although examination of BR timetables before and after new stations are opened suggests that on average the penalty is only one minute. The PTE argue that passengers are unlikely to react to such a small change (or are unlikely to perceive the change) and hence existing revenue lost is zero. Thus following the approach adopted by the PTE in columns 5 and 6 of Table 10.1, we have assumed existing revenue lost is zero.

However, ignoring the impact of time penalties is inconsistent when, as we shall see in section 10.3, time penalties are evaluated as a social cost. In order to determine existing revenue lost the number of passengers on trains

	1 Capital Cost	2 Recurrent costs	3 Gross revenue	4 Abstracted revenue	5 Balance (3-2-4)	6 FYRR	7 FYRR-PTE	8 Existing revenue lost	9 Balance 5-8	10 FYRR
Bramley	124500	1700	37988	29176	7112	0.057	0.302	3279	3833	0.031
Crossflatts	80000	2200	41199	25670	13329	0.167	0.273	2275	11054	0.138
Deighton	77000	1500	33286	21952	9834	0.128	0.214	2153	7681	0.100
Fitzwilliam	86000	1500	22946	11877	9569	0.111	0.122	809	8760	0.102
Saltaire	118000	1700	40731	27790	11241	0.095	NA	2826	8415	0.071
Slaithwaite	106000	1700	19634	8196	9938	0.094	0.222	480	9458	0.089
TOTAL	591500	10300	195784	124661	61023	0.103	0.232	11882	49201	0.083

TABLE 10.1 "ACTUAL" FINANCIAL FYRR FOR SIX NEW STATIONS (1984 PRICES)

	3	4	5	6	9	10
Bramley	30616	22130	6786	0.055	3507	0.027
Crossflatts	29362	11978	15184	0.190	12909	0.161
Deighton	13926	8132	4294	0.056	2141	0.028
Fitzwilliam	20521	13372	5649	0.066	4840	0.056
Saltaire	42377	18685	21992	0.186	19166	0.162
Slaithwaite	22234	5774	14760	0.139	14280	0.135
TOTAL	159036	80071	68665	0.116	56843	0.096

TABLE 10.2 "MODELLED" FINANCIAL FYRR FOR SIX NEW STATIONS (1984 PRICES)

prior to stopping at a new station was estimated from the OD data derived by the probability method (see section 4.2.1) updated and converted into passenger miles using data from "Transportation Trends" (WYMCC, 1985). Given the average speed of a service (from the BR Timetable) this could be converted into a mean in-vehicle time per passenger figure. The effect of an increase of one minute in journey time was estimated by applying the journey time elasticity of the aggregate all-trips model (-0.418). The number of passenger miles lost can then be converted into a total revenue lost figure by applying the mileage related farescale. The results of taking into account existing revenue lost are shown by columns 8 and 9 of Table 10.1. It should be noted that the effect of this revenue loss to the PTE is reduced by the fact that, assuming switchers are the same as the total train using population, 60% of travellers are likely to switch to bus (from the 1984 County Council survey).

From Table 10.1 it can be seen that our estimates of net financial benefits as measured by a FYRR figure are less than half those of the PTE. The main reason for this is that the PTE assessed generated traffic as all new trips whilst in our study a stricter definition of generated trips as all new trips that could be attributed to the new station was used, thus excluding trips generated by exogeneous factors such as changes in employment or residence. Further differences are due to the fact that the PTE used a system wide mean fare estimates. Comparison of columns 5 and 6 with 8 and 9 shows that the addition of six new stops decreased existing revenue by £11,822 pa causing a decrease in net benefits of 19% and an absolute decrease in the overall FYRR of 2%.

In developing NPV figures for the six new stations two assumptions concerning demand growth over time were tested.

(i) No growth. This assumes that patronage remains stable at the initial usage figure.

(ii) High growth. Following the findings of 9.4 this assumes that initial usage increases additively by 24% per annum up to the fifth year and then reaches stability.

Findings from other areas (such as the Garston branch in Merseyside) suggests that in reality demand growth will be somewhere between these two figures.

The results of such an analysis are given by Table 10.3. From this Table it can be seen that, assuming a project life of 30 years, the total new station package's NPV is positive for all assumptions except no growth and 10% interest rate (although even this is positive if the European Regional Development Fund (ERDF) grant for Saltaire is deducted from the costs). However, the NPV for Bramley is negative for all three no growth assumptions whilst it is also negative for Saltaire (if we exclude the ERDF grant) and Slaithwaite under the no growth and 10% interest rate assumptions.

These results can be compared with those of the PTE as we have not included estimates of revenue lost from existing passengers. It can be seen that the PTE's estimate of an NPV for five stations (Saltaire was excluded) (£1,272,000) is only slightly higher than our estimate for the same five stations assuming high growth (£1,094,195 i.e. a difference of 16%) but is much higher than the value given by our no growth assumption (£308,365 i.e. a difference of 312%).

When a project life of 15 years is considered, the total station package's NPV is again generally positive except for at the 7 and 10% levels under the no growth assumption. The NPVs for Bramley, Slaithwaite and Saltaire are negative under all three no growth assumptions, whilst Bramley has a negative NPV even under the high growth assumption, given a 10% interest rate. This suggests that our results are more sensitive to changes in assumptions concerning project life and interest rate than are those of the PTE.

10.2.2 Modelled Financial Costs and Benefits at Six New Stations

A modelled evaluation of the six new stations was based on the forecasts provided by the "disaggregate" market segmented approach, except at Fitzwilliam where the single market model had to be used. Because we had only been able to model weekday demand, Saturday usage was assumed to be the same as that on an average weekday, even though in reality for the six stations overall it was initially 8.8% higher, although there was considerable variation from station to station.

In order to assess gross revenue use was made of the destination

	5%		7%		10%		PTE
	High Growth	No Growth	High Growth	No Growth	High Growth	No Growth	
Fitzwilliam	164523	61100	113280	32742	61932	4207	80000
Deighton	256366	80184	187457	49883	118481	19391	186000
Crossflatts	336340	124901	249649	85400	162962	45652	257000
Slaithwaite	219103	46789	151048	17332	83023	-12309	265000
Saltaire	202375	46702	136341	14951	70268	-16999	NA
*	(238090)	(82416)	(171389)	(49998)	(104358)	(17091)	
Bramley	117863	-4609	67420	-27721	16975	-50978	484000
TOTAL	1296570	355067	905195	172587	513641	-11036	1271000
*	(1332285)	(390780)	(940243)	(207634)	(547731)	(23054)	

(A) 30 YEARS PROJECT LIFE

Fitzwilliam	76959	13324	55385	1154	29998	-13217
Deighton	138675	29133	109643	16129	75560	772
Crossflatts	187872	58351	151487	41400	108818	21382
Slaithwaite	102440	-2847	73914	-15485	40477	-30410
Saltaire	89447	-6792	61676	-20417	29048	-36508
*	(125161)	(28923)	(96723)	(14629)	(63175)	(-2417)
Bramley	31501	-43548	10321	-53467	-14520	-65180
TOTAL	626894	47621	462426	-30686	269417	-123161
*	(662608)	(83336)	(497473)	(4360)	(303508)	(-89070)

(B) 15 YEARS PROJECT LIFE

TABLE 10.3 FINANCIAL NPV FOR SIX NEW STATIONS (1984 PRICES)

* ERDF grant of £37500 for Saltaire included as a windfall gain

choices implied by the forecasting models, combined with the 1984 fare scale. Using evidence from the County Council's 1984 rail survey, it was assumed that 52% of travellers made use of season tickets or saverstrips, 43% used ordinary tickets, 3% used Day Rover and 2% used staff/privilege tickets. In addition, using our own market research, it was assumed that 18% of travellers made use of half price travel. It was also assumed that, on a weekday, all work journeys were made during the peak and all non-work journeys were made during the off-peak. This resulted, on average, in 44% of trips being forecasted in the off-peak. Journeys to other destinations (non-work journeys to non-modelled destinations) were assumed to cost the system wide mean fare (60 pence). The modelled gross revenue is shown by column 3 of Table 10.2. There is a suggetion that we have underestimated revenue in that the total gross revenue is only 0.812 that of Table 10.1.

For journeys to/from work the percentage of trips abstracted from other modes can be modelled directly, but as with destination choice this may be inaccurate due to the small samples used. To determine the percentage of non-work journeys abstracted and generated we had to rely on our market research which suggested that 40% of non-work trips were abstracted from Public Transport, 20% were abstracted from car and 40% were generated. This allowed us to produce an abstracted revenue figure shown by column 5 of Table 10.2.

Comparing the FYRR given by Table 10.2, column 6 to that of Table 10.1 it can be seen that it is higher in absolute terms by 13%. Although the results at Bramley in Table 10.2 are broadly similar to those in Table 10.1, there are big differences for the other sites, particularly Deighton and Saltaire, and the absolute deviation between the "actual" and modelled FYRRs is 42.6% (RMSE = 0.055).

10.2.3 Modelled Financial Costs and Benefits at 28 Potential Sites

The 28 potential sites identified in 9.1.1 were evaluated. For three of these sites (Osmondthorpe, Gamble Hill and Wrenthorpe) use was made of the market segmented model's prediction of the number of work journeys, whilst for the remaining 25 sites use was made of the single market model.

	1	2	3	4	5	6	7	8	9	10
	Capital costs	Recurrent costs	Mean fare	Gross revenue	Net revenue	Balance (5-2)	FYRR	Existing revenue lost	Balance (6-8)	FYRR
	(£)	(£)	(£)	(£)	(£)	(£)		(£)	(£)	
Osmondthorpe	100000	1700	0.400	59040	14424	12724	0.127	7980	4744	0.047
Gamble Hill	100000	1700	0.390	50544	15116	13416	0.134	3410	10006	0.100
Wrenthorpe	130000*	1700	0.418	13292	4514	2814	0.022	1188	1626	0.013
Steeton	100000	1700	0.700	18690	7770	6070	0.061	1560	4510	0.045
Utley	120000	1700	0.659	16014	6920	5220	0.044	1560	3660	0.031
Newlay	120000	1700	0.392	16783	7178	5478	0.046	4416	1062	0.009
Laisterdyke	150000	1700	0.337	17996	6470	4770	0.032	3047	1723	0.011
Stanningley	120000	1700	0.358	30179	10095	8395	0.070	3324	5071	0.042
Armley	100000	1700	0.338	40357	10241	8541	0.085	3410	5131	0.051
Frizinghall	100000	1700	0.379	25294	6934	5234	0.052	1872	3362	0.034
Bowling	90000	1500	0.286	28301	4959	3259	0.036	2208	1051	0.012
Low Moor	60000	1500	0.343	16877	3705	2305	0.038	2208	97	0.002
Luddendenfoot	120000	1700	0.443	8107	4386	2686	0.022	2160	526	0.004
Lightcliffe	60000	1500	0.332	10558	4681	3181	0.053	2208	973	0.016
Salter Hebble	50000	1500	0.427	21265	8839	7339	0.147	2160	5179	0.104
Wortley	120000	1700	0.318	12593	3530	1830	0.015	1920	-90	-0.001
Cottingley	100000	1700	0.386	27445	5674	3974	0.040	1920	2054	0.021
Milnsbridge	150000	1700	0.426	19681	7023	5323	0.035	602	4721	0.031
Beeston	150000	1700	0.371	25154	5342	3642	0.024	1188	2454	0.016
East Ardsley	130000	1700	0.440	10692	4224	2524	0.019	1188	1336	0.010
Outwood	120000	1700	0.354	8071	2973	1273	0.011	1188	85	0.001
Sandal	100000	1700	0.426	16614	4984	3284	0.033	873	2411	0.024
Hemsworth	100000	1700	0.581	14990	6100	4400	0.044	821	3579	0.036
Burley	100000	1700	0.291	23484	5936	4236	0.042	4464	-228	-0.002
Hawksworth	150000	1700	0.399	28429	10534	8834	0.059	4392	4442	0.027
E. Garforth	100000	1700	0.509	58331	16339	14639	0.146	4620	10019	0.100
E.Knottingley	100000	1700	0.613	15448	9269	7569	0.076	288	7281	0.073
Kirkstall	120000	1700	0.309	27439	7694	5994	0.050	4416	1578	0.013

TABLE 10.4 FINANCIAL COSTS AND BENEFITS AT 28 POTENTIAL SITES (1984 PRICES)

* Assumes PTE responsible for 30% subway provision. If this becomes 100% capital cost increases to £180,000.

Data on construction costs for certain sites were provided by the PTE (1984A). Where such costs did not exist, following consultations with BR and the PTE, an average cost of £100,000 for a double platform station was assumed, rising to £150,000 where there were local access difficulties (for example if the site is located in a cutting) and falling to £60,000 for a single platform station. Recurrent costs were assumed to be £1,700 per annum for double platform stations (£1,500 for single platform stations). These costs are shown by column 1 and 2 of Table 10.4.

Estimates of gross revenues and net revenues were made as in the previous section, by utilising information supplied by the models on destination choice and abstraction (shown by columns 4 and 5 of Table 10.4), whilst existing revenue lost (column 8 of Table 10.4) was calculated as in 10.2.1.

From column 7 of Table 10.4 it can be seen that, using a FYRR measure comparable to that of the PTE only seven stations have a rate of 7% or more (12 have a rate of 5% or more). However, if existing revenue lost is taken into account then only four stations have a FYRR (column 10, Table 10.4) of 7% or more, with this figure only increasing to five if a 5% rate is used.

The data presented in Table 10.4 can be converted into NPVs as before although the definitions of the growth assumptions are modified:

(i) High initial usage. Predicted patronage achieved by year 1 and sustained throughout the project life.

(ii) Low initial usage. Patronage grows additively by 24% pa to achieve predicted patronage in year 5 which is then sustained throughout the project life.

Table 10.5 shows the financial NPVs assuming a project life of 30 years. Although at the 5% interest rate and assuming high growth seven sites give positive returns, only four of these (Osmondthorpe, Gamble Hill, Salter Hebble and East Garforth) are insensitive to further changes.

In the last column of Table 10.5 we relax the assumption that a single Public Transport authority (i.e. the PTE) is responsible for financing both bus and rail and hence switches of revenue from bus to rail can not be considered as a net financial gain. This might be justified for at least two scenarios:

	5%		7%		10%		
	High growth	Low Growth	High Growth	Low Growth	High Growth	Low Growth	Different Financial Scenario
Osmondthorpe	92803	78925	55635	42257	5545	18233	451414
Gamble Hill	106238	91510	66480	52282	13008	26472	496486
Wrenthorpe	-86741	-91141	-95080	-99321	-107494	-103472	-7972
Steeton	-6688	-14259	-24762	-31975	-49401	-42778	12697
Utley	-39755	-46497	-55244	-61724	-76954	-70790	-55588
Newlay	-35789	-42782	-52022	-58764	-74751	-68358	-71868
Laisterdyke	-76672	-82976	-90807	-96844	-110796	-105032	-39471
Stanningley	9053	-783	-15825	-25308	-49852	-40860	89113
Armley	31295	21318	5986	-3634	-28606	-19484	345956
Frizinghall	-19540	-26297	-35051	-41565	-56836	-50659	188220
Bowling	-36826	-41657	-47077	-51735	-61809	-57392	225675
Low Moor	-26104	-29713	-32638	-36118	-42514	-39214	83487
Luddendenfoot	-43437	-47710	-64420	-68540	-98585	-94678	-113727
Lightcliffe	-11100	-15660	-20256	-24923	-34182	-30013	-18048
Salter Hebble	62819	54206	41070	32767	11311	19184	150689
Wortley	-91868	-95308	-97290	-100607	-105893	-102748	-32822
Cottingley	-38909	-44438	-50685	-56016	-67591	-62537	212755
Milnsbridge	-6817	-75013	-83945	-90542	-106075	-99819	-90781
Beeston	-94012	-99218	-104805	-109824	-120425	-115666	155622
East Ardsley	-91199	-95314	-98679	-102646	-109968	-106206	-45527
Outwood	-100430	-103327	-104202	-106995	-110647	-107999	-74650
Sandal	-49516	-54372	-59248	-63929	-73481	-69041	51420
Hemsworth	-32360	-38304	-45400	-51129	-63954	-58521	61366
Burley	-34881	-40665	-47435	-53011	-65354	-60067	52020
Hawksworth	-14198	-24462	-40377	-50272	-76104	-66721	23403
East Garforth	125039	109119	81656	66309	23448	38001	95524
East Knottingley	16355	7323	-6076	-14782	-36903	-26847	76477
Kirkstall	-27856	-35354	-45620	-52847	-70348	-63494	30559
NUMBERS POSITIVE	7	6	5	4	4	4	19

TABLE 10.5 FINANCIAL NPVs FOR 28 POTENTIAL SITES - 30 YEARS PROJECT LIFE (1984 PRICES)

[1] See text - based on 5% interest rate and high growth assumption

(i) One operator might finance rail services (i.e. BR) whilst another operator(s) might finance bus. This situation currently exists in most Shire counties. Net financial gain to BR should then include abstracted bus revenue.

(ii) Following bus deregulation in October 1986 around 71% of West Yorkshire's pre-deregulation bus network was operated as a commercial service. Abstraction of revenue from these commercial services to a social service (rail) might be considered by the PTA as a net financial gain. Most of the potential new station catchment areas studied are served by commercial bus services for most parts of the day, although an exception is provided by East Garforth. Of course bus deregulation is likely to have other effects, especially due to changes in bus networks, frequency and fare, which, however, we have not had time to investigate.

From the last column in Table 10.5 it can be seen that changing the financial organisation assumption results in 19 of the sites returning a positive NPV (compared to seven under previous assumptions and assuming a 5% interest rate and high growth), even though the effect of lost revenue from existing passengers has been taken into account.

The effects of assuming a 15 year project life were also tested. At the 5% and 7% interest rates Osmondthorpe, Gamble Hill, Salter Hebble and East Garforth continue to return positive NPVs. However at the 10% interest rate, even assuming high growth, Osmondthorpe's NPV becomes negative as do the NPVs for the other three sites assuming a low growth rate. Changing the assumptions concerning financial organisation results in 11 sites returning positive NPVs at the 5% interest rate and high growth assumption. Compared to a 30 year project life the NPVs of eight sites have become negative (Burley, Hawksworth, Hemsworth, Kirkstall, Milnsbridge, Sandal, Stanningley and Steeton). It should be noted that throughout the above we have assumed that there is existing spare capacity to carry the additional traffic.

10.3 SOCIAL COST-BENEFIT ANALYSIS

10.3.1 Determination of Benefit to New Station Users from Market Research

In section 5.5.2 it was shown how the data collected by our market research could be expanded so as to allow estimation of benefit to new station users. As generalised time savings are expected to be the main social benefit in our analysis we shall attempt to evaluate them in this section. It is likely that different evaluation measures will give different results and because of this at least five benefit measures were tested. These are shown by Table 10.6 and include:

A. Calculating the differences between standard Department of Transport generalised cost measures for each individual new station user and then determining the mean difference. To determine the mean benefit use was made of the rule of half for those trips that were generated, whilst those trips that were abstracted were considered to be existing travellers and their benefit was given a weighting of one. Thus:

$$B_i = \sum_n \{B(2\Delta Wk + 2\Delta Wt + \Delta IVT + \Delta CO/VOT + \Delta PA/VOT)\}/n \quad \text{Equation 10.1}$$

$$= \sum_n \{B(GCp - GCr)\}/n$$

where B_i = mean benefit measure at station i, Wk = Walk time (minutes),
Wt = Wait time (minutes), IVT = In Vehicle Time (minutes).
CO = Public Transport fare or car operating cost (pence)
PA = Parking cost (pence), VOT = Department of Transport Value of time (pence per minute), n = number of observations
GCp = Generalised time of previous mode,
GCr = Generalised time of rail, and
B = 1 if trip abstracted, 0.5 if generated.

A problem with this measure was apparent in that a large percentage of travellers (29.8% of those surveyed) were accrueing negative benefits. This might be due to:

(i) Measurement error, in particular use of Bus IVT from published timetables may have understated actual IVT, particularly in the peak.

(ii) Specification error, as the generalised cost formulation has failed to take into account a number of attributes that may be perceived as being important, for example convenience, comfort and reliability.

(iii) Variability of the value of time.

		A	B (i)	B (ii)	C	D	E
Saltaire	(W)	6.08	21.93	25.55	12.66	4.10	14.81
Saltaire	(S)	9.10	16.37	18.45	8.32	2.40	7.68
Bramley	(W)	6.11	11.30	12.33	5.05	0.46	3.62
Bramley	(S)	2.68	9.27	16.95	2.41	-1.82	1.21
Slaithwaite	(W)	3.39	14.57	13.04	7.25	3.78	13.40
Slaithwaite	(S)	8.44	16.75	55.50	9.53	7.52	11.54
Crossflatts	(H)	1.07	15.81	15.72	0.80	-1.57	6.75
Crossflatts	(W)	8.72	16.14	18.36	9.45	5.29	8.17
Crossflatts	(S)	-1.03	16.83	32.98	6.72	2.54	8.76
Deighton	(S)	20.33	24.11	36.96	21.84	14.13	19.10
Deighton	(W)	15.05	23.67	25.13	12.17	6.04	8.64
Fitzwilliam	(W)	37.88	60.04	49.86	41.05	32.43	35.85
Fitzwilliam	(S)	34.15	50.23	66.93	24.76	18.54	26.59
Weighted Mean Weekday		8.62	20.31	21.07	10.34	4.74	10.96
Weighted Mean Saturday		9.81	19.00	33.68	9.99	5.23	10.15
Mean All		9.10	19.73	26.14	10.20	4.94	10.63

TABLE 10.6 GENERALISED TIME SAVINGS (IN MINUTES) FOR NEW STATION USERS

W = Weekday S = Saturday H = Holiday (See text for further explanation)

Explanations (i) and (ii) above imply adding a fixed term to the mean benefits (or scaling up the generalised costs proportionately) hence in

B(i) ignoring the negative benefits might be a minimum adjustment to make i.e.

$$B_i = \{ \sum_n B \lambda (GCp - GCr) \} / z \qquad \text{Equation 10.2}$$

where λ = 1 if GCp - GCr $\geq$ 0, = 0 otherwise
z = number of observations where λ = 1.

or

B(ii) alternatively it might be argued that a constant term should be added to all observations so that the entire distribution shifts to the right. This constant term might be the mean negative benefit at each station i.e.

$$N_i = \{ \sum_n B \delta (GCp - GCr) \} / y \qquad \text{Equation 10.3}$$

where δ = 1 if GCp - GCr < 0, = 0 otherwise
y = number of observations where δ = 1

and thus mean benefit may be calculated as

$$B_i = \{ \sum_n B (GCp - GCr) \} / n + | N_i | \qquad \text{Equation 10.4}$$

C. Method A might be reworked in aggregate as

$$B_i = \bar{B} (\bar{GCp} - \bar{GCr}) \qquad \text{Equation 10.5}$$

This type of approach should be used where only aggregate data, particularly concerning the proportions abstracted and generated, are available.

D. A fourth approach might be to make use of the disaggregate parameters from the models developed in Chapter Seven. In order to keep the analysis simple, use was made of the MNL model given by column 1 of Table 7.7. It should be noted that although the out of vehicle time parameters are roughly double those of in-vehicle time, this formulation implies a much higher value of in-vehicle time (£2.72 per hour) than previously used. Although unmeasured attributes are taken into account by the ASCs, these ASCs were calibrated with data on all trip makers and not just those who choose rail. As a result the inclusion of the ASCs is unlikely to eradicate the problem of negative benefits.

E. A possible way of updating ASCs has been put forward by Anas (1982, page 148).

$$a_i^* = \hat{a}_i - \ln(\tilde{H}_i / H_i) \qquad \text{Equation 10.6}$$

where a_i^* = unbiased ASC, a_i = biased ASC
$\tilde{H}_i$ = proportion choosing alternative i in the sample (market share)
H_i = proportion choosing alternative i in the total population.

(A recent study in the West Midlands (WMCC, 1984) uses a slightly different variation of this formula).

Equation 10.6 was thus used with the data in Table 10.7 to modify the ASCs of the disaggregate model as shown by Table 10.8 (it should be noted that train is maintained as the base alternative i.e. ASC-Train = 0). The disaggregate utility function of Table 7.7, column 1 was thus modified to include these ASCs and rerun with the new station users data set.

The results of the five benefit measures are shown by Table 10.6 (although it should be noted that the value of time used in D and E is 3.3 times greater than that used in A, B and C). In terms of generalised time saved in minutes the highest benefits are given by method B, with method B(ii) (inclusion of constant term based on the mean negative benefit) giving a slightly higher benefit, especially on Saturdays, than B(i) (exclusion of negative benefits). The second highest benefits, in terms of time savings, are given by method E, which demonstrates the effect of modifying ASCs. Comparison with method D suggests that "unmeasured attributes" account for around 6.22 minutes of benefits on a weekday (5.85 minutes on a Saturday). The third highest benefits are given by method C, as aggregated data seems to result in an increase in overall benefits of 12% compared to method A. The lowest generalised time savings are given by method D, as they are around half those in A, C and E. This is due to the higher value of time used and the use of

	User popn (1)	Sample popn (2)	Sample popn (3)
	(H_i)	(H_i)	
Bus	0.53	0.18	0.46
Train	0.28	0.09	0.01
Car Driver	0.13	0.56	0.40
Car Passenger	0.06	0.17	0.13
Numbers	197	952	2768

TABLE 10.7 MODAL SHARES

Notes

(1) Weekday work travellers recorded in surveys 1982-84. Generated trips excluded

(2) 1981 Corridor Study

(3) Journey to work travellers living within 800m of a station and travelling to a rail served destination (From the 1981 Census, Special Workplace Statistics)

	"Biased"	"Unbiased"
ASC - Bus	0.448	0.393
ASC - Car Driver	1.824	-0.771
ASC - Car Passenger	-0.092	-2.678

TABLE 10.8 ADJUSTED ALTERNATIVE SPECIFIC CONSTANTS

See text for explanation

ASCs that have not been calibrated for the user population.

It should be noted that the mean generalised time savings of A, C and E are broadly consistent with the results determined, with different data, by the market share recovery rate method shown by Table 8.7. Given this our preference is to use the values implied by method A, although acknowledging that these values may be an underestimate. As a result we shall also make use of methods B(ii) and E in later analysis, as they both consistently give positive returns.

It should be noted that the total level of benefit (for all five measures) is much less than the mean net social benefit given by the PTE which was calculated as being equivalent to around 54 minutes per new station user (at April, 1985 Department of Transport VOT) (Cottham, 1985). This included a time penalty of two minutes for each passenger on the train prior to the new station. We were unable to ascertain precisely how the PTE calculated these time savings, but possible causes for the discrepancy might include that, compared to our study, the PTE underestimated out of vehicle time for rail travellers or may have overestimated car operating costs.

10.3.2 Actual Social Costs and Benefits at Six New Stations

In our analysis of social costs and benefits we shall include the following.

(i) Capital and recurrent costs, defined as before.

(ii) Operator's surplus, which for the six new stations is given by column 8 of Table 10.1 (i.e. gross revenue minus revenue abstracted from other Public Transport operators, revenue lost from existing Public Transport users switching to private modes or not travelling).

(iii) Time savings to new station users. Initially we used the estimates given by column A of Table 10.6, to calculate the value of generalised time savings per annum at each station. This is shown by Column B of Table 10.9. In order to test the sensitivity of the results to time saving estimates, two other methods were tested:

(a) Inclusion of a constant term equivalent to the mean negative benefit at each station (method B(ii) Table 10.6). This is shown by Column C of Table 10.9.

	A Financial balance[1]	B Time savings[2]	C Time savings[3]	D Time savings[4]	E Time losses[5]	F Time losses[6]	G Balance (A+B-E)	H Balance (A+C-E)	I Balance (A+D-E)
Bramley	3833	4930	11448	9377	10442	34040	-1679	4839	-20830
							(-0.013)	(0.039)	(-0.167)
Crossflatts	11054	3692	10506	13599	7162	23347	7584	14398	1306
							(0.095)	(0.180)	(0.016)
Deighton	7681	6407	10975	14234	3842	12526	10246	14814	9389
							(0.133)	(0.192)	(0.122)
Fitzwilliam	8760	12195	17638	36344	2181	7111	18774	24217	37933
							(0.218)	(0.282)	(0.441)
Saltaire	8415	6889	25641	46846	8952	29184	6352	25104	26077
							(0.054)	(0.213)	(0.221)
Slaithwaite	9458	2578	12780	24425	1058	3450	10978	21180	30433
							(0.104)	(0.200)	(0.287)
TOTAL	49201	36691	88988	144815	33637	109658	52255	104552	84358
							(0.083)	(0.177)	(0.143)

TABLE 10.9 "ACTUAL" SOCIAL FYRR FOR SIX NEW STATIONS (1984 PRICES) (FYRR in brackets)

[1] Based on column 9, Table 10.1 [2] Method A, Table 10.6 [3] Method B(ii) [4] Method E [5] Dept of Transport VOT - 83.3 pence per hour [6] Disaggregate VOT - 272 pence per hour

	A	B	C	D	E	G	H	I
Bramley	3507	14728	5786	4965	10442	7793	-1149	-1970
						(0.062)	(-0.009)	(-0.016)
Crossflatts	12909	4681	11588	7774	7162	10428	17335	13521
						(0.130)	(0.217)	(0.169)
Deighton	2141	10248	9607	4033	3842	8547	7906	2332
						(0.110)	(0.103)	(0.030)
Fitzwilliam	4840	N/A	10790	15802	2181	N/A	13449	17741
							(0.156)	(0.206)
Saltaire	19166	21554	16819	10943	8952	31768	27033	21157
						(0.269)	(0.229)	(0.179)
Slaithwaite	14280	4254	7790	2104	1058	17476	21012	15326
						(0.165)	(0.198)	(0.145)
TOTAL	56843	55465	62380	44901	33637	76012	85586	68107
						(0.150)	(0.145)	(0.115)

TABLE 10.10 "MODELLED" SOCIAL FYRR FOR SIX NEW STATIONS (1984 PRICES)

(b) The use of disaggregate parameters, including ASCs updated to take into account the effect of "unmeasurable" attributes (method E of Table 10.6). This is shown by Column D of Table 10.9.

(iv) Time losses to existing users. Assuming a one minute penalty and given the information on passengers on the train prior to a new station (already used in the calculation of existing revenue lost) this can be easily calculated. Column E of Table 10.9 assumes a value of time of 83.3 pence per hour, whilst Column F assumes a value of time of £2.72 per hour.

Column G of Table 10.9 shows the FYRR, given the time savings estimated from method A of Table 10.6, for all six stations. It can be seen that although the new station package overall exceeds a 7% return, the returns of Saltaire and Bramley are below 7%. If method B(ii) of Table 10.6 is used then the net social benefits almost double, as shown by column H of Table 10.9, although the return at Bramley is still below 7%. If method E of Table 10.6 is used then column I of Table 10.9 suggests that, compared to Column G, net social benefits increase by over 60%, although the returns at Bramley and also Crossflatts are below 7%. These results clearly illustrate that our results are sensitive to the type of benefit measure used, although the effect on policy implications is rather less drastic. For all three measures the value of net social benefits is considerably less than the £352,000 per annum estimated by the PTE (Cottham, 1985, Appendix pviii).

Using the same assumptions as in 10.2.1 the information in Table 10.9 can easily be converted into NPVs, as Table 10.11 shows. This table shows that, assuming a 30 year project life, all stations give positive returns under the high growth assumption, although Bramley and Saltaire consistently give negative returns under the low growth assumption and are joined by Slaithwaite and Crossflatts when there is a 10% interest rate. Overall the new station package returns a positive NPV except under the no growth assumption at a 10% interest rate.

Assuming a 15 year project life it can be seen that Bramley consistently produces a negative NPV, as does Saltaire under the low growth assumption at all three interest rates, whilst Slaithwaite and Crossflatts give

	5%		7%		10%	
	High Growth	No Growth	High Growth	No Growth	High Growth	No Growth
Fitzwilliam	419354	202604	315505	146967	211488	90982
Deighton	367027	89915	274101	57737	181171	25358
Crossflatts	299133	36586	218063	14110	137161	-8506
Slaithwaite	432799	91297	170649	29303	97133	-3415
Saltaire	210459	-28454	140605	-45717	70866	-63087
*	(246174)	(7620)	(175653)	(-10670)	(104976)	(-28997)
Bramley	57147	-139749	16328	-136809	-24259	-133849
TOTAL	1,785,919	252,181	1,135,251	65,591	673580	-92517
*	(1,821,634)	(287,895)	(1,170,299)	(100,638)	(707,670)	(-58427)

(A) 30 YEARS PROJECT LIFE

Fitzwilliam	241994	108869	198239	84992	146807	56797
Deighton	207941	35703	168917	21894	123153	5587
Crossflatts	159590	-1280	125800	-10925	86270	-22315
Slaithwaite	118003	6996	87199	-6497	51104	-23409
Saltaire	90221	-57538	61108	-64946	27036	-73694
*	(125935)	(-21823)	(96155)	(-29900)	(61127)	(-39603)
Bramley	-13772	-134796	-30561	-133535	-50121	-132045
TOTAL *	803,977	-42,046	610,702	-109,467	384,249	-189,079
	(839,691)	(-6,331)	(645,749)	(-74,421)	(418,340)	(-154,988)

(B) 15 YEARS PROJECT LIFE

TABLE 10.11 SOCIAL NPV FOR SIX NEW STATIONS (1984 PRICES)

* ERDF grant of £37500 for Saltaire included as windfall gain

negative NPVs assuming low growth at the 7% and 10% interest rates. Overall it can be seen that the new station package consistently gives positive NPVs under the high growth assumption and negative returns under the low growth assumption, although the mid point between the two values is always positive. These results again show that our evaluation method is sensitive to values concerning project life, interest rate and demand growth over time.

10.3.3 Modelled Social Costs and Benefits at Six New Stations

In assessing the time losses to existing users the same procedure as in the previous section was used, as shown by Column E of Table 10.10. However assessment of time savings to new station users is much more problematic. A number of different approaches were tested and these are shown in Table 10.10.

(i) Firstly time savings may be assessed directly from the disaggregate model provided a sample enumeration approach is adopted. As column B of Table 10.10 shows this can be done for five of the six new stations. However we have had to assume that the extent of time savings for work and non-work journeys are the same (which is likely to be unrealistic given differences in destination choice, congestion, fare and service levels), whilst for generated journeys the rule of half was applied. For Fitzwilliam (and 25 of the 28 potential sites) disaggregate data on times and costs do not exist, and hence the sample enumeration technique can not be used. Estimation of time savings from the incremental logit approach is complicated by the use of zonal data (as for many zonal pairs rail does not have a generalised cost advantage over bus) and the lack of times and costs for private transport. An evaluation measure might, however, be based on changes in the composite Public Transport utility.

Column G of Table 10.10 shows the effect of using the disaggregate model's estimate of time savings, as compared to Column G of Table 10.9 it can be seen that the FYRR of all stations, except Deighton, increases. This suggests that time savings for work/peak journeys are greater than for non-work/off-peak journeys. Moreover, compared to the actual FYRR figures of Table 10.9, our modelled results are only within ± 88.7% (as measured by our AD figure).

(ii) An alternative approach might make use of the model's predictions of the proportions abstracted from bus, train and private transport and the proportion generated, combined with zonal information on modal times and costs (as used with the naive IL model and with the aggregate models) for each destination choice. Again, however, for many zonal pairs the generalised cost of travelling via the new station was higher than the generalised cost of other modes. This was because, due to the coarseness of the zoning systems used, walk time for rail was much greater than walk time for bus, although it is clear that this variable will exhibit a great degree of intra-zonal variation. From our market research for example, it can be shown that overall walk time for rail travellers is only estimated to be 14.9% higher than the walk time of the best alternative mode (although there is considerable variation between stations and between days of the week). Assuming such a relationship exists across all new stations it was possible to calculate the time savings due to a new station with respect to each mode (it should be noted that for zonal pairs with good road links, for example dominantly motorway, these savings were still negative). The rule of half was again applied for generated journeys and the annual value of all time savings is given by column C of Table 10.10, with the resultant FYRRs given by column H. Again it can be seen that, except at Deighton and Fitzwilliam, there is a tendency to overestimate returns compared to column G of Table 10.9, although the results do tend to be in line with the adjusted results given by columns H and I of Table 10.9. Overall, again using our AD measure, it may be shown that the modelled FYRRs differ from the "actual" returns by around 71.9%.

(iii) A third, and final, approach was based on our market research, which suggested that a relationship existed between time savings to new station users and the generalised cost of Public Transport (in most cases bus) prior to the new station opening. This relationship was confirmed by a simple regression based on 13 observations and shown by Table 10.12.

This simple relationship suggests that the benefits of a new station are related to the level of accessibility of existing Public Transport and has advantages in that it can be simply applied to potential sites. The main disadvantage of such an approach is that it assumes a constant relationship between the generalised cost of travelling via a new station and the generalised cost of existing Public Transport, although it is likely that rail will have a relative advantage in some corridors (due to road conditions, rail speed and service levels) compared to others. Nonetheless, the annual value of time savings implied by such an approach are shown by Column E of Table 10.10 with the resultant FYRRs given by Column I. It can be seen that although compared to column G of Table 10.9 there is still a tendency to overestimate returns, overall "actual" and modelled FYRRs differ by only about 52.7%, as measured by the AD measure, representing a deterioration of only 10% compared to the modelled financial FYRR given in 10.2.2. This suggests that this approach may be helpful in providing a "sketch-planning" estimate of the level of benefits and thus is used in the next section to estimate possible time savings at 28 sites.

	Intercept	Public Transport Generalised Cost	R^2	$\bar{R}^2$
Disaggregate parameters (based on the single market model)	-18.797	0.418	0.780	0.760
	(-4.177)	(6.240)		
Aggregate parameters	-17.060	0.349	0.685	0.656
	(-2.253)	(4.893)		

TABLE 10.12 THE RELATIONSHIP BETWEEN TIME SAVINGS AND PUBLIC TRANSPORT GENERALISED COST IN THE BEFORE SITUATION (t-statistics in brackets).

10.3.4 Modelled Social Costs and Benefits at 28 Potential Sites

Using the approach given by (iii) above, it was possible to estimate time savings to new station users and time losses suffered by existing users at the 28 potential sites, and a social FYRR may then be derived (Table 10.12). It can be seen that there are nine sites where the social FYRR exceeds 7%, whilst there are also seven sites with negative social FYRRs, being concentrated in the inner city areas.

The information presented in Table 10.13 can be used to develop NPVs by using the same procedure as in 10.2.3. Table 10.14 shows the results for a 30 year project life, whilst similar results were also produced for a 15 year project life. Again it was apparent that our results are sensitive to assumptions concerning project life, discount rates and demand growth. Hence

	Time savings	Time increases	Balance	(FYRR)
Osmondthorpe	19473	12941	11277	(0.113)
Gamble Hill	19943	12847	17102	(0.171)
Wrenthorpe	11827	2624	10829	(0.079)
Steeton	16159	4850	15819	(0.158)
Utley	13542	4850	12352	(0.103)
Newlay	8356	12885	-3464	(-0.029)
Laisterdyke	3204	11479	-6552	(-0.044)
Stanningley	8679	12523	1227	(0.010)
Armley	5551	12847	-2165	(-0.022)
Frizinghall	5256	5318	3300	(0.033)
Bowling	8114	6761	2404	(0.027)
Low Moor	7205	6761	541	(0.009)
Luddendenfoot	3464	5893	-1903	(-0.016)
Lightcliffe	5058	6761	-730	(-0.012)
Salter Hebble	1538	5906	811	(0.016)
Wortley	726	4568	-3932	(-0.033)
Cottingley	6994	4568	4480	(0.045)
Milnsbridge	6692	1570	9843	(0.066)
Beeston	9611	2624	9441	(0.063)
East Ardsley	6959	2624	5671	(0.044)
Outwood	2820	2624	281	(0.022)
Sandal	9499	2373	9537	(0.095)
Hemsworth	9846	2262	11163	(0.112)
Burley	1120	8919	-9027	(-0.080)
Hawksworth	8657	8896	4203	(0.028)
East Garforth	36718	7440	39297	(0.393)
East Knottingley	4860	562	11566	(0.116)
Kirkstall	7560	12849	-3711	(-0.031)

TABLE 10.13 SOCIAL FYRRS FOR 28 POTENTIAL SITES (1984 PRICES)

at the 5% interest rate, assuming high growth and a 30 year project life, 10 sites return positive NPVs, whilst at the other extreme assuming low growth, a 15 year project life and a 10% interest rate only one site (East Garforth) returns a positive NPV.

In order to be consistent with the PTE and because we suspect estimation of time savings based on method A of Table 10.6 to be under estimated, our preferred results may be those of the high growth, 30 year project life and 5% interest rate assumption. This would suggest that the following sites have potential for development.

(i) In the Leeds District: East Garforth, Gamble Hill and Osmondthorpe.

(ii) In the Wakefield District: East Knottingley, Hemsworth and/or Sandal and Wrenthorpe.

(iii) In the Bradford District: Steeton and/or Utley

(iv) In the Kirklees District: Milnsbridge.

(N.B. None of the three sites evaluated in Calderdale gave a positive social return).

It should be noted that all work in this Chapter, so far, has been based on 1984 fare levels. However in January 1985 a major change took place in that Metrocard was extended to rail, effectively ensuring for regular travellers a maximum fare of 50 pence for journeys made within West Yorkshire. This clearly has had a dramatic effect on usage. Using the elasticity from our disaggregate approach we have calculated that Metrocard will have led to a 6.4% increase in overall rail patronage in West Yorkshire. This is likely to be an underestimate as we have been unable to measure the effect on non-work trips. In addition it should be noted that for certain flows the changes in fare are non marginal and thus our elasticity measures will be inappropriate. (Indeed Dunne (1984) has suggested that the use of disaggregate point elasticities is only justifiable where the relative change is 5% or less).

The last column of Table 10.14 illustrates the impact of Metrocard on the social NPVs of 28 sites, given high growth and a 5% interest rate. (It is clear that given an elasticity of less than unity the financial effect will be negative). All stations incur an increase in patronage of between 1 and 10%, although it should be noted that the majority of forecasted trips were short distanced and unlikely to be affected by Metrocard. Metrocard increases the

	5%		7%		10%		
	High growth	Low Growth	High Growth	Low Growth	High Growth	Low Growth	Metrocard[1]
Osmondthorpe	70544	37961	37667	5997	4582	25451	78630
Gamble Hill	162901	128740	112220	79228	61220	29991	160657
Wrenthorpe	36470	20546	4378	-10972	-27915	-42471	40190
Steeton	143178	119863	96303	73823	49125	28109	151095
Utley	69882	49943	33527	14056	-3556	-21784	81642
Newlay	-173296	-188431	-163021	-177611	-147682	-166518	-176124
Laisterdyke	-250720	-260147	-231302	-240349	-211763	-220382	-257053
Stanningley	-101137	-119430	-104773	-122408	-108532	-125155	-106379
Armley	-136801	-152818	-130215	-145048	-122954	-130719	-140567
Frizinghall	-49271	-61148	-59050	-70501	-68891	-79749	-50776
Bowling	-49970	-62707	-57687	-69967	-65452	-77097	-49723
Low Moor	-53221	-63851	-54528	-64775	-55843	-65561	-53912
Luddendenfoot	-75679	-83328	-97207	-104579	-137938	-144914	-151251
Lightcliffe	-71222	-80711	-68787	-78207	-66882	-75557	-72882
Salter Hebble	-37533	-47646	-39936	-49685	-42355	-51599	-43113
Wortley	-180444	-184592	-168791	-172790	-157066	-160858	-182150
Cottingley	-31131	-43474	-44406	-56306	-57767	-69050	-30331
Milnsbridge	1313	-12050	-27856	-40739	-57209	-69426	1943
Beeston	-4867	-19347	-32845	-46892	-60999	-74319	-2930
East Ardsley	-42822	-53717	-59628	-70132	-76539	-86500	-40500
Outwood	-115679	-121324	-116511	-121954	-117350	-122510	-115926
Sandal	46608	32479	18345	4832	-10094	-22995	51820
Hemsworth	71604	56066	38522	23445	5233	-8970	76431
Burley	-223394	-230270	-199607	-206236	-175669	-181955	-225240
Hawksworth	-85388	-104087	-97843	-115870	-110377	-127471	-86019
East Garforth	504095	452398	387638	337802	270450	223192	537562
East Knottingley	77999	64231	43684	30413	9155	-3430	77569
Kirkstall	-225378	-237175	-205063	-216435	-184620	-195404	-180905
NUMBERS POSITIVE	10	9	9	8	6	4	10

TABLE 10.14 SOCIAL NPVs FOR 28 POTENTIAL SITES - 30 YEARS PROJECT LIFE (1984 PRICES)

1 5% interest rate, high growth

total benefits to new station users, but causes a slight decrease in net revenue. Conversely existing revenue lost due to increased journey time will decrease slightly but the number of "existing" passengers experiencing time losses will increase. It can be seen that, to some extent, these changes cancel out, as comparing the first and last columns of Table 10.14 no station's NPV changes sign. In addition for 13 stations the NPV increases but for 15 it decreases. Thus, according to our analysis, Metrocard has little effect on the social NPV of new stations.

10.4 FURTHER EVALUATION ISSUES

In this section we shall relax some of the assumptions that have been made so far in this Chapter and we shall also attempt to take into account a number of additional variables. It should be stressed that the results in this section are meant to be illustrative rather than definitive.

10.4.1 Financial Arrangements

In 10.2 our main concern was the financial effect on the operators of both bus and rail, although in the last column of Tables 10.5 this was relaxed, so as to look at the effect on the profitability of rail alone. Another scenario might be to look at the effect on the finances of WYPTE, which is only allocated the revenue attributable to passenger mileage on supported sections. Our market research shows that, overall, only 80.5% of revenue will be allocated to the PTE, with the rest going to BR and adjacent authorities that support rail services (e.g. Greater Manchester, South Yorkshire). Table 10.15 shows that the effect of reallocating revenue reduces the financial benefit to the PTE. Typically the FYRR to the PTE is reduced by to 2 to 3 percentage points, whilst the other Public Transport operators receive a windfall gain (assuming stopping time penalties of one minute have little effect on revenue). However, the information in Table 10.15 was obtained from our own market research and may be subject to sampling error; for example we suspect that Manchester bound journeys from Slaithwaite have been under represented and that seasonal variation has not been fully taken into account (a survey at Crossflatts suggested that contributory revenue is higher during holiday periods).

	Gross[1] revenue	Gross[2] revenue	FYRR[3]	FYRR[2]
Fitzwilliam	22946	16797	0.111	0.074
Deighton	33286	25131	0.128	0.092
Crossflatts	41199	37773	0.167	0.150
Slaithwaite	19634	19375	0.094	0.090
Bramley	37988	26623	0.057	0.036
Saltaire	40731	31858	0.095	0.079
TOTAL	195784	157557	0.103	0.079

TABLE 10.15 FINANCIAL BENEFIT TO THE PTE (1984 Prices)

1 As column 3, Table 10.1

2 Only revenue on supported sections allocated to the PTE

3 As column 6, Table 10.1

Section	Miles	Section 20 Cost (£000 pa)	FYRR 1	FYRR 2	FYRR 3
Knottingley to Knottingley East	1	24	0.076	-0.161	0.126
Keighley to Utley	1.2	65	0.044	-0.499	0.072
Utley to Steeton	1.9	103	0.061	-0.971	0.119

TABLE 10.16 EFFECT OF EXTENSIONS OF SECTION 20 PAYMENTS (1984 Prices)

1 PTE responsible for construction costs, no extension to Section 20 payments.

2 PTE responsible for construction costs, Section 20 payments extended

3 BR responsible for construction costs.

Three out of the 11 sites favoured in 10.3.3 (East Knottingley, Steeton and Utley), are on sections of line not subject to financial support by the PTE under Section 20 of the 1968 Transport Act. Although this is only a transfer payment it can have a dramatic effect on financial returns to the PTE and hence it is desirable to estimate the costs of such extensions. This was crudely done by estimating the total cost per mile of the relevant services (from the PTE, 1979, updated to 1984 prices, although real costs are likely to have risen since) and calculating the cost of extending supported mileage as the ratio of extended mileage over existing mileage times existing cost. This method clearly fails to take into account the effect of indivisibilities as, for example, the extensions to Steeton and East Knottingley would include level crossings. As a result the costs in Table 10.16 might be seen as underestimates. However this Table clearly shows that extensions to Section 20 payments lead to the financial return to the PTE becoming highly negative. The last column of Table 10.16 shows that if one operator was responsible for financing rail and others for financing other Public Transport then (excluding the effect of revenue lost from existing passengers) all three sites exhibit returns in excess of 7%. This suggests that if different financial arrangements were adopted for these boundary stations, benefits for both the PTE and BR may result.

10.4.2 Operating Arrangements

(i) Rail Operating Costs. Previously we have assumed that there were no significant increases in rail operating costs other than those included in the recurrent cost figure. Suppose it was decided to open four stations on the Leeds-Bradford line (at Laisterdyke, Stanningley, Gamble Hill and Armley) as examination of column 6, Table 10.4 shows that this would give a financial FYRR to the PTE of 7.4%. However, this would inevitably lead to increases in operating costs because:

(i) Examination of train working diagrams shows that although the mean turn round time is 21 minutes for Leeds-Bradford workings, the minimum turn round time is only 3 minutes, which could no longer be achieved if four more stops were introduced.

(ii) Given existing pricing policies the Leeds-Bradford service is already overcrowded during peak periods and would be unable to accommodate the extra 400 or so commuters that we predict would want to use the service.

Thus if four more stations were opened on the Leeds-Bradford line it is likely that the service would need to be upgraded. One way this might be done would be to introduce a new two car unit on the line, operating a shuttle service between 0700 and 1800 hours Monday to Saturday thus allowing frequency to be increased from every 20 to every 15 minutes. This would require rescheduling of other services which we assume can be done at minimal costs. Given daily train mileage of around 220 miles the additional costs might be as in Table 10.17. Increases in net PTE revenue were estimated using a service elasticity of 0.26 as used in the METS model (Department of Transport, 1984), although this may be on the low side (Nash and Whiteing, 1984). Despite such revenue increases it can be seen from Table 10.18 that upgrading the Leeds-Bradford rail service, on its own, results in a negative FYRR to the PTE.

(ii) Bus Operating Costs.

Upgrading a rail service might be accompanied, in a regulated Public Transport market, by downgrading the competing bus service, which in this case study would be the service 72 (Leeds Central Bus Station to Bradford Interchange). At the time of our study this service operated for most of the day on 15 minute headways, timetabled journey time was 48 minutes which, along with a 10 minute turn round time, meant that the service accounted for eight buses most of the day. If we suppose that the service is reduced to half hourly intervals this would allow for the withdrawal of four buses. The Bradford Bus Study (Travers, Morgan and Partners, 1974) estimated the cost of all day one man operated double decker bus operation as (updated to 1984 prices) £44,537 pa (again though we suspect that real costs may have increased during the intervening period). If the costs of bus withdrawal are taken into account Table 10.18 shows that our improvement package may achieve a financial FYRR to the PTE of 17%, although we have assumed that all revenue from the withdrawn bus services has switched either to the remaining bus service, replacement rail services or alternative bus services, which may be over optimistic. The social implications of this kind of improvement package are wide ranging but are, unfortunately, beyond the scope of this study.

In the case of certain stations the introduction of bus feeders might be desirable, for example at Hemsworth and Steeton. Costs from the Peat, Marwick, Mitchell (1975) study of seven potential bus feeders in West

	Price (£)
Capital Cost of Unit (Class 141)	365180
Time related maintenance pa	13853
Mileage related maintenance pa	22858
Fuel pa	13839
Cleaning etc. pa	5145
Crew pa	48542
Total Operating Expenses pa	104237

TABLE 10.17 POSSIBLE COSTS OF UPGRADING LEEDS-BRADFORD SERVICE

Based on Figures given by the PTE (1981)

	Net Financial Benefit	Capital Cost	FYRR
Stations opened, rail service unchanged	35122	470000	0.074
Stations opened, rail service[1] upgraded	-38.667	835180	-0.046
Stations opened, rail service[1] Upgraded, bus service down-graded	139481	835180	0.167

TABLE 10.18 POSSIBLE FINANCIAL EFFECTS OF UPGRADING THE LEEDS-BRADFORD SERVICE

1 Service elasticity of 0.26 assumed

	Previous Balance[1]	Net revenue	Existing revenue cost	Time savings	Time gains	New balance
Laisterdyke	-6552	5915	1030	2929	3840	2274
Stanningley	1227	9229	1117	7934	4208	10138
Gamble Hill	17102	13812	1204	18232	4535	26312
Armley	-2165	9362	1204	5075	4535	8698
TOTAL	9612					47422
(FYRR)	(0.020)					(0.101)

TABLE 10.19 EFFECT OF CHANGES IN STOPPING PATTERNS

1 from last column of Table 10.4

Yorkshire indicate that, at 1984 prices, supplying a bus feeder would involve capital costs of around £28,000 and operating costs per route mile of about £8,000 pa. It should, though,be noted that bus deregulation resulting from the 1985 Transport Act seems likely to seriously reduce the prospects of bus and rail integration (Kilvington, 1985).

(ii) Changes in stopping patterns

So far we have assumed that all local passenger trains stop at new stations. This assumption was relaxed for the four potential stations on the Leeds-Bradford line by assuming the Caldervale trains (to Halifax and beyond) did not stop at these stations. This reduces service frequency from three to two trains an hour, leading to a decrease in patronage at affected stations of 8.7% (assuming a service elasticity of 0.26) and a subsequent loss of gross revenue. However, as around two-thirds of passengers already travelling on this line used the Caldervale trains, there were large reductions in existing revenue lost and time increases suffered by existing passengers due to the additional four stops. As a result Table 10.19 shows that, for the four stations, changes in stopping patterns may increase the overall social FYRR from 2 to 10%. It should, though, be noted that the Leeds-Bradford service was the only service in West Yorkshire, with the possible exception of the Airedale services, where frequencies were high enough to allow experimentation with stopping patterns.

10.4.3 Extending the Cost-Benefit Equation

In this sub-section we shall attempt to extend our benefit measure by considering a number of additional variables:

(i) Time savings due to reduced congestion

In a number of social cost benefit studies of railways the reduction in road congestion has provided the main benefits. For example Foster's (1974) study of two suburban lines in Greater Manchester showed that road congestion costs accounted for 60 to 65% of gross benefits. However, we would argue that the effect of reduced congestion is likely to be of a lower magnitude for new stations as abstraction from car is limited. Indeed market research indicates that the six new stations have led to a reduction of only 78 car driver trips per day.

Where there is a congested junction such small reductions in car usage might still lead to substantial benefits. To examine this effect we studied the potential effect of East Garforth and Osmondthorpe stations on congestion in the York Road Corridor and in particular at the Woodpecker junction, which has been identified by WYTCONSULT (1977A) as being severely congested in the peak. By making crude assumptions about route assignment and departure time our forecasts indicate that these two new stations might lead to a reduction of up to 26 cars passing through Woodpecker junction in the peak hour.

Since the withdrawal of the Appendix J method (Cmnd 3686, 1968) in 1977 there has been no standard method for evaluating congestion effects at saturated junctions. A possible approach is given by the series of equations presented by Akcelik (1981) and shown by Appendix 4. Data for these equations were provided from the Urban Congestion Monitoring project (May et al. 1983) and are also presented in Appendix 4). This method estimates that such a small reduction in peak hour flow would lead to 1394 driver minutes of time savings. Assuming car occupancy of 1.35 (from the 1981 WYTS) and a symmetrical congestion pattern (which is unlikely given the particular layout of this junction) this would result in savings of 62.7 hours per weekday, which given a value of time of 83.3 pence per hour, might imply a saving of £13,057 per annum, although we have taken no account of rerouteing, change in destination and time of travel due to the reduced congestion.

	Net revenue	Time saving to new station users	Time saving to road users	Gross benefit
Osmondthorpe } East Garforth }	30763	56191	13057	100011

TABLE 10.20 POSSIBLE SCALE OF TIME SAVINGS FOR ROAD USERS

Table 10.20 shows that at these two sites reduction in road congestion accounts for almost 13% of gross benefits. Other possible congested junctions where new stations might reduce delay include Armley Gyratory (for stations on the Leeds-Bradford line) and Fox Corner (for stations west of Shipley on the Airedale line). However, it is clear that even in the extreme cases of Osmondthorpe and East Garforth time savings to road users, although significant, form a much smaller proportion of gross benefits than many other social cost-benefit studies of passenger railways. Thus for most sites we may be justified in omitting such benefits from detailed analysis. This conclusion is compounded by the problems in accurately measuring time savings to road users.

(ii) Accident savings

In road investment schemes accident costs typically form 20% of benefits, although in Foster's (1974) study of Manchester suburban railways they only formed 1.5 to 2% of gross benefits. Again we might argue that because the level of abstraction by new stations from private transport is limited accident reductions will only be of a small scale.

For example Gregg (1983) estimated that there were 2 fatalities per 10^9 passenger km for rail compared to 10.1 for car and 3.65 for bus and coach. Clearly there will be little change when, due to the six new stations only an estimated 794,850 passenger km are abstracted from car and 3,187,800 passenger km from bus per annum. In addition examination of "Transport Statistics: Great Britain" (Department of Transport, 1983) indicates that there are around 975 non fatal casualties per 10^9 passenger km for road travel (with the rate for private car being some 2.5 times greater than for bus) compared to 136 per 10^9 passenger km for rail travel. Using such rates and the values used by COBA (Department of Transport, 1981) we were able to estimate accident savings due to the six new stations as being worth around £19,784 pa. From Table 10.21 it can be seen that these savings, although small in absolute terms, represent between 11 and 16% of gross benefits for the new stations that have already been opened.

Net Revenue	Time savings to new station users	Accident savings	Gross benefits
71123	36691[1]	19784	127594
	88988[2]		179895

TABLE 10.21 POSSIBLE ACCIDENT SAVINGS

[1]column B, Table 10.9 [2] column C, Table 10.9

(iii) Fuel savings

It can be shown, for example from Department of Energy (1976) statistics, that rail is an energy efficient form of transport. In the case of new stations the main saving will be due to abstraction from car drivers. Given that around 23,350 car driver trips per annum are abstracted by the six new stations, the mean trip length is 21 km and average fuel consumption is 44 km per gallon then the annual petrol saving is 11,144 gallons. If we assume that the market price reflects the opportunity cost then this effect will already have been taken into account by the assessment of car operating cost savings in 10.3.1. As we were unable to show that the opportunity cost is significantly higher than the market price, we must conclude that the value of fuel savings is minimal. We have assumed that the additional fuel costs due to stopping trains have been included in the recurrent cost figure.

(iv) Changes in Taxation

It has been argued, for example by Gwilliam and Mackie (1975, p203) that where there is substantial generated traffic a non resource correction should be made in order to take into account the tax benefit issue. Thus we shall take into account changes in petrol tax (estimated at 55% of pump price) and we shall estimate losses in indirect taxation as 15% of the net change in Public Transport revenue. In total this is estimated as £20,103 per annum for the six new stations.

10.4.4 Effect of Extending the Cost-Benefit Equation

Given points (i) to (iv) above we can estimate an extended benefit measure as:

$$B = V_1 + V_2 + F + A + FS + S_1 + S_2 + \Delta T - C \qquad \text{Equation 10.7}$$

where V_1 = New station user time savings,
V_2 = Existing rail user time savings
F = Road user time savings, A = Accident savings,
FS = Fuel savings, S_1 = Revenue gained due to new station,
S_2 = Revenue lost due to increased journey time
ΔT = Change in indirect taxation and C = recurrent costs.

= 36691[1] - 33637 + 8058[2] + 19784 + 0 + 71123 - 11882 - 20103-10500

= 59534

[1] column B, Table 10.9 [2] Estimated for Bramley, Crossflatts and Saltaire only

It can thus be shown that by extending our benefit equation we get a FYRR for the six new stations of 0.101 compared to 0.088 from column G Table 10.9. This suggests that if we extend our benefit algorithm to include F, A and ΔT the net benefit of new stations increases by around 15%. This is a significant proportion but, due to lack of reliable data, we were unable to extend the benefit algorithm for the 28 potential sites. These results indicate that using our social NPV measure may well underestimate total benefits.

10.4.5 Other External Effects

(i) Environmental Effects

The main adverse environmental effect of introducing an additional stop to a rail service will be in terms of additional noise due to stopping and starting. Other effects might include increased vibration, fumes, dust/dirt and visual intrusion. It should be noted that the effect of visual intrusion has been minimised by choice of site and building materials but has still caused public concern. For example residents at East Garforth complained that a new station would allow passengers to look into their homes. Disbenefits due to noise may be greatest for those living close to the new station, but research by Fields and Walker (1980) showed that rail noise causes less annoyance to people living within 100 metres of a station than to other people living close to a railway line because:

a) the "convenience factor" of the station outweighs the environmental disadvantage, and,

b) the "convenience factor" influences the perception of the environmental disadvantage of rail.

The main environmental benefit is likely to be the reduction in road noise and pollution due to abstraction from car drivers. Most of this benefit is likely to accrue away from the new station site, although road traffic (and the accompanying environmental disbenefits) might increase at new station sites where park and ride or kiss and ride are extensively used.

As the number of car drivers who have switched to train is limited it is debateable whether any environmental benefits will be perceived. As we also have doubts whether any environmental disbenefits will be perceived we must conclude that environmental changes due to new stations are highly marginal in magnitude and effect, although they are likely to be positive.

(ii) Secondary and Tertiary Effects

Our evaluation framework has concentrated on studying primary effects, i.e. changes in travel patterns. In section 2.3.2 it was shown that even for major rail investments, such as those in Glasgow and Tyne and Wear, secondary effects (i.e. changes in accessibility) and tertiary effects (i.e. changes in land use) are limited and the benefits of such effects should not be overstressed (Dodgson, 1984). Moreover, due to the small scale of the new station investment programme we would expect secondary and tertiary effects to be even more limited. It might also be argued that much of the benefit of improved accessibility has already been taken into account by estimating the time savings accrueing to new station users. In addition tertiary effects might be merely a redistribution of benefits; a relative increase in house prices near to a new station might be accompanied by a relative decrease in another (non rail served) area.

For the new stations that have been already opened the following specific comments might be made concerning secondary and tertiary effects:

a) In terms of accessibility.

The new station at Fitzwilliam greatly improved Public Transport accessibility to Leeds, particularly for work and shopping and resulted in large benefits for new station users. Evidence from the PTS on/off counts suggests that Deighton and Saltaire have attracted small numbers of work journeys and this may be due to improved accessibility to ICI Dalton Green and SALTS mills respectively. It should be noted that our disaggregate mode split models have not fully taken into account such attracted work journeys and may explain why such models underpredict usage compared to our aggregate all-trips model.

b) In terms of land use.

There are unlikely to be any commercial and/or industrial developments. For example retail facilities in the new station catchment zones are of a low order and are unlikely to be affected by improved rail access to high order centres.

There may, of course, be changes in destination choice amongst higher order centres due to the introduction of a rail link e.g. shopping journeys to Leeds from Fitzwilliam replacing those to Barnsley. The most likely effects will be in terms of residential land use but none are currently apparent. Such effects are likely to take a long time to develop and are likely to take the form of increased house prices and/or private residential development similar to that which has taken place close to Baildon, Burley in Wharfedale, Garforth and Horsforth stations in recent years (although not necessarily because of the railway). However, private residential development is unlikely to occur at Saltaire (due to the conservation area status) or at Bramley, Deighton and Fitzwilliam (where much of the existing housing is in public ownership) whilst developments at Crossflatts and Slaithwaite may be limited by green belt policy. Hence it seems evident that tertiary effects have been, and will be, limited.

Similar comments might be made for the 28 potential sites:

a) In terms of accessibility.

WYTCONSULT (1977A) identified a number of areas with problems in terms of accessibility to work, education, shopping, health and welfare facilities. Table 10.22 shows that 12 of our potential new station catchment areas were identified as having problems for at least one of the five variables.

Purpose	
Work	Sandal, East Garforth, Wrenthorpe.
Education	Hawksworth, Beeston, East Garforth, Sandal, Gamble Hill, Luddendenfoot, East Knottingley, Low Moor.
Shopping	Bowling, East Knottingley, Gamble Hill, Sandal, Low Moor, East Garforth, Hemsworth
Health	East Knottingley, Gamble Hill, Hemsworth, Osmondthorpe
Welfare	Bowling, East Garforth, Low Moor.

TABLE 10.22 AREAS WITH ACCESSIBILITY PROBLEMS (in descending order of magnitude) Source: WYTCONSULT, 1977A.

A new station is most likely to reduce accessibility problems to work and high order shopping centres. It is less likely to improve accessibility to education, health and welfare where the problems are normally in terms of low order provision.

In addition it might be noted that Gamble Hill, Salter Hebble and Steeton could serve hospitals, Frizinghall would attract traffic from Bradford Grammar School, whilst Armley, Beeston, Bowling, Cottingley, East Knottingley, Laisterdyke, Low Moor, Luddendenfoot, Outwood, Stanningley and Wortley serve industrial areas of varying size and traffic potential. Also Beeston or Cottingley might double up as serving Elland Road football ground (although this would involve additional access time compared to a site outside the ground).

b) In terms of land use.

Again we would expect our potential stations to have little effect, although stations at, for example, East Garforth and Wrenthorpe might increase residential development pressure there whilst stations at, for example, Gamble Hill and Osmondthorpe might reduce the number of vacant council properties. Certain sites might be related to development projects, although as an efect rather than a cause. Thus Low Moor might be related to the proposed West Yorkshire Transport Museum and the Spen Valley railway (Thornhill on the Huddersfield to Wakefield line is another potential site related to this project which we have not evaluated due to uncertainties regarding the service's future). Similarly, Kirkstall station could be related to developments around the Abbey, which include a proposed Brewery Museum and light railway. Stanningley, or a resited New Pudsey, could be related to a proposed major out-of-town shopping centre development. Lastly, a new station at Leeds Parish Church, (which because of its special nature we have not evaluated) might be linked to the proposed market redevelopment and revamped Central Bus Station. This is the one site that might have some development effects of its own.

10.5 CONCLUSIONS

In this Chapter we have carried out actual and modelled financial and social evaluation of the six new stations already opened in West Yorkshire. This work has shown that in addition to usage forecasts, data is also required on the mean fare paid, previous mode used and the generalised time savings (for which a number of measures might be used) for new station users as well the number of passengers on trains prior to reaching the new station. It was shown that, for a number of reasons (in particular due to different definitions of what constitutes a generated trip and different calculations of time savings to new station users), our actual financial and social evaluation results tended to be lower than those of the PTE. These data requirements ensure that any modelled evaluation will have broad confidence limits as it was shown that the modelled financial FYRRs deviate from the "actual" FYRRs by 43%, whilst for social FYRRs the respective figure is at least 53%.

We have also been able to model the financial and social returns for 28 potential situations. In terms of financial NPVs, it was shown that if net revenue to Public Transport operators was considered only a maximum of seven sites give positive returns, but if net revenue to the rail operator only was considered this figure increases to 19. In terms of social NPVs our preferred measure (5% interest rate, 30 year project life, high growth rate - given by the first column of Table 10.13), resulted in 10 sites giving positive values. The policy implications of these findings are discussed in more detail in section 11.4.

It was noted that our findings were sensitive to the assumptions made concerning project life, interest rates and demand growth over time. Our results were also sensitive to changes in rail or bus operating costs and changes in stopping patterns. It was also shown that if the cost-benefit equation was extended to include time savings to road users, accident reductions and changes in indirect taxation then first year net benefits for the six new stations already opened might increase by around 15% indicating that the case for new stations has been strengthened.

Lastly, it should be noted that although we were unable to explicitly measure environmental, secondary and tertiary effects it is suspected that these will be minimal, although likely to be positive.

CHAPTER ELEVEN CONCLUSIONS

11.1 INTRODUCTION

This Chapter constitutes the last stage of our research framework, whereby this study's aims, methods and results are brought together to form conclusions at a number of levels. In the next section some theoretical findings, particularly from our modelling work, will be summarised. In a third section we shall examine the empirical findings of this research, with respect to the results of our market research, forecasting methods and evaluation framework. In a fourth section the policy implications of these findings will be discussed and the goals and objectives of the West Yorkshire new station programme re-examined. Lastly, prior to drawing our final conclusions, we shall, in line with convention, make some recommendations for further research.

11.2 THEORETICAL FINDINGS

Two main approaches have been developed to produce forecasts of new station usage:

(i) Aggregate Simultaneous Models (ASMs) based on zonal data.

(ii) Disaggregate approaches (DAs) based primarily on mode split models for the journey to work, calibrated using individual (and household) data.

Trip rate and regression models, previously used to forecast usage at new local rail stations (see 3.2), may be seen as simple applications of an ASM-type approach, whilst in 3.3 it was demonstrated how both the ASM and DA represented theoretical improvements over such simple approaches. In the rest of this section some aspects of the ASM and DA, as developed, will be compared.

11.2.1 Aggregate Appraoch

The ASM for all trips that was used in further analysis was the double log model given by Equation 6.8. The main advantage of the ASM, compared to the DA, was its relative simplicity. Given access to good quality rail OD data (which this study did not have) and Census data (via SASPAC) such a model can easily be calibrated. Moreover, from a practitioner's viewpoint, the data needed is more likely to be available or more easily acquired than that required by a DA. In addition our study found that the development of an ASM required less than a third of the research effort devoted to developing a DA, quite apart from the less demanding requirements in terms of knowledge and experience.

11.2.2 Disaggregate Approach - Some Advantages

Two main DAs were developed and used in further analysis

(i) The Market Segmented model. For members from car owning households this was based on an HL model of mode choice for work journeys, given by model WBA9F-1 for the lower split and initially by WBA9F-3 for the upper split (see Table 7.12), although the ASCs for the upper split model were subsequently enriched (see model TRAN-2, Table 8.3). For members from non-car owning households an MNL model was found to be appropriate, as given by WBA9F-5, Table 7.12. This approach was used to produce forecasts for five of the six new stations already opened but was used for only three of the 28 potential sites.

(ii) The Single Market model. This was based on an HL model of mode choice for work journeys, with the lower split model given by WBA9C-4 (Table 7.8) and the upper split model given by WBA9D-1 (Table 7.9). Using aggregate data, based on the Census Special Workplace Statistics, this approach could produce forecasts for all existing and potential new stations.

In order to produce total station usage figures both sets of models were used in conjunction with an aggregate non-work model (NWM), given by column 3 of Table 6.11. Thus our approach to forecasting total station usage should more correctly be referred to as hybrid rather than disaggregate. Nonetheless, we shall continue to use the acronym DA to refer to this approach.

The main advantage of the DA relates to the firm behavioural basis of the work mode split models (although the ASM also has links with utility theory). This allowed us to determine the parameter values of each mode's utility function, whereas the ASM had used a generalised cost formulation based on existing values, as determined by the Deprtment of Transport. It should be noted that the DA did not always give valuations in line with those of the Department of Transport. The DA's behavioural basis also meant that it was able to provide information on abstraction from existing modes for work journeys as well as benefit measures for new station users.

Our main theoretical finding in calibrating mode split models was to show that, given our data on mode choice between car driver, car passenger, bus and train, a MNL model structure is inappropriate as the IIA axiom is violated because bus and train are not perceived as independent alternatives. It was thus shown that an HL model structure was more appropriate particularly as the value of the Expected Maximum Utility (EMU) parameter (ϕ) was significant and within the prescribed range.

11.2.3 Disaggregate Approach - Some Disadvantages

Despite the above much of our work in Chapters Seven and Eight has shown that logit modelling in general, and hierarchical logit modelling in particular, is something of a "minefield". While calibrating disaggregate mode split models a number of problems emerged:

(i) As shown in 8.1.1, there are problems in developing appropriate goodness of fit measures. It was evident that the rho-squared measure, even if it was adjusted to take into account restricted choice sets, does not possess the intuitive qualities of R^2 in OLS. As a result it was difficult to compare directly the goodness of fit of disaggregate mode split models with the ASM, although, based on informal tests, we have reason to believe that the disaggregate models, as developed, are superior in this respect.

(ii) It was evident that the adoption of an HL structure increases the complexity of interpreting parameter values and in determining elasticity values (see 8.1.2). Moreover, we have seen that the inclusion of socio-economic variables may lead to further complications due to the number of ways such variables can enter the utility function and an apparent need to keep the utility functions in the upper and lower nests consistent. Market segmentation, for example on the basis of car ownership, adds yet further complications, although it has some theoretical and practical advantages.

(iii) The use of HL models means that a potentially large number of decision tree structures can be tested as Sobel (1980) has demonstrated and Figure 3.7 illustrates. Furthermore, common attributes need to be determined for alternatives in the lower nest(s). Ortuzar (1980B) argued that, for bus and rail in West Yorkshire, fare may be considered a common attribute. However, we have found the parameter value of fare in the lower split to be significant and of a plausible magnitude, whilst the assumption of no common attributes simplifies model structure.

(iv) Table 7.7 has shown that the parameter values of logit models are sensitive to choice set specification. It has been shown that such problems are, in part, related to data availability and quality as in an ideal situation (from the modeller's view point) all individual's would have all the relevant choices available to them.

(v) It has been noted that certain parameter values, in particular cost, are unstable. This might be due to specification error, although there are indications that this is more likely to be due to measurement error caused by deficiencies in the data set used. In particular we suspect that time and cost attributes may be correlated.

Points (iv) and (v) indicate that the practical extent of our disaggregate modelling calibration stage was limited by data availability (provided by the Corridor data set described in 4.2.4) and quality, particularly if it is remembered that previous users of the Corridor data set had formed "an overall impression that the data was very variable in quality and coverage" (MVA, 1984).

The problems with the disaggregate mode split models in the calibration stage were further compounded in the validation (section 8.2) and prediction (section 9.2) stages. It was shown that, although such models may possess a certain degree of temporal transferability between 1981 and 1984 (in marked contrast to the ASM) we were unable to demonstrate that these models were spatially transferable. Again it was suspected that this lack of transferability may be partly related to deficiencies in the data set used (provided by the Countywide data set described in 4.2.5), especially as we were forced to make assumptions about choice set specification and had to combine reported and engineering data.

In section 8.3 we noted the effect of the aggregation problem and in 9.3.1 we were able to show how the naive use of disaggregate parameters with aggregate data led to errors in line with those detected by the literature. Thus the use of a sample enumeration technique to produce aggregate forecasts was preferred. However for 25 of our 28 potential new station sites data at a suitably disaggregated level were not available and, moreover, we were unable to apply a market segmented approach. We were thus restricted to modifying the Extended Incremental Logit (EIL) model of Koppelman (1983) so as to be consistent with our single market model and thus make use of aggregate data. This approach was seen to greatly increase inaccuracy, although this was mainly due to the use of engineering data but we were still able to show that these models may be useful in assessing planning priorities.

11.3 EMPIRICAL FINDINGS

This section will consist of three sub-sections, covering the empirical findings of our market research, statistical modelling and evaluation framework.

11.3.1 Findings from Market Research

The results of our market research exercise have been discussed in detail in Chapter Five. Some of the main findings are listed below:

(i) As Figures 5.3 to 5.8 showed, new station catchment areas are clearly defined and highly localised, with Table 5.18 showing that 63% of first year weekday usage originated within a straight line distance of 800m from the station (67% on Saturdays). As Table 5.19 shows, the catchments are localised because walk is the dominant access mode, accounting for 83% of users on a weekday (86% on a Saturday). Two sub groups were identified:

(a) Urban sites (Bramley, Deighton and Saltaire) where the mean access distance is between 400 and 750 metres and walk accounts for between 80 to 100% of travellers' access mode.

(b) Semi-urban or rural sites (Crossflatts, Fitzwilliam and Slaithwaite) where the mean access distance tends to be higher (in excess of 750 metres) with between 15 and 45% of users making use of mechanised access modes.

As Figure 5.2 shows there is evidence that destination catchment areas are similarly well defined and concentrated. In addition it was apparent that trips were short distanced and local in nature, with 90% being made to destinations within West Yorkshire.

(ii) Important information has been produced on the proportion of travellers abstracted from other modes and the proportion generated. Generated trips were strictly defined as trips generated due to the new station. Trips that were generated due to exogeneous reasons were redefined as being abstracted from existing modes in the same proportions as for all non generated trips (in contrast to the PTE who defined such trips as being generated). The results of our approach were given by Table 5.16. Although there are a number of differences between new station sites and days of the week, it was estimated that 56% of first year weekly (Monday to Saturday) demand was abstracted from bus, 13% from train, 16% from car and only 13% generated. Abstraction from existing Public Transport was thus much greater than that anticipated by the PTE following the Airedale and Wharfedale surveys in 1978. Our results showed, at least in the short term, that generated trips were almost entirely restricted to non-work journeys made either during off-peak periods on weekdays or on Saturdays.

(iii) Leading on from this point, it was evident that there were marked differences between weekday and Saturday usage with work being the main purpose for the former (55%) and shopping for the latter (63%), whilst it was also noticeable that the propensity to make long distance trips, party size and the number of accompanied children were greater on a Saturday (and during school holidays) than on a weekday.

(iv) As has already been noted there were important differences between station users. In particular, as Table 5.23 indicated, users at Crossflatts were typical of the traditional view of rail commuters (i.e. dominantly in white collar occupations with high car ownership rates) whilst at the other extreme usage at Deighton was typified by members of households with low car ownership, low economic activity rates and an emphasis on blue collar occupations.

(v) As Table 5.24 showed our second year surveys indicated that up to 75% of weekly travellers would have used the new station one year ago, although 16% were abstracted from bus and 8% generated. The second year surveys also indicated that there has been some overall growth in patronage and that the people for whom travelling via a new station provides the most marginal benefits are those who live on the edge of new station catchment areas.

11.3.2 Findings from Statistical Modelling

A comparison of the elasticities implied by the ASM all-trips and the single market disaggregate mode split model was given by Table 9.8, whilst the values of time implied by the disaggregate models were discussed in Chapter Seven. Important results have been found for the EMU parameter (ϕ) which, depending on model formulation, may take a value of between 0.15 and 0.44 (see Table 7.9). The parameter ϕ might be interpretated as an index of similarity, with values positive but close to zero indicating that higher nest decisions are not very dependent on lower nest decisions (i.e. bus and train are perceived to be similar), whilst values closer to one imply a virtually simultaneous structure, akin to the MNL (i.e. bus and train are perceived to be independent). It is interesting to note that the value of ϕ determined by the single market model (0.205) is lower than the values determined by Ortuzar (1980A) and Matzoros (1982) who found values for the Garforth corridor of 0.329 and 0.360 respectively, whilst Sobel (1980) found a value of 0.4 in applications with Dutch data.

Our main empirical findings with regards to the statistical models developed were related to the examination of their forecasting accuracy, as summarised by Table 9.6. Table 8.5 showed that a market segmented disaggregate mode split model can be highly accurate in what it does, but this was, from the point of view of estimating new station demand, limited as we were only able to predict the proportion of work journeys to rail served destinations within pre-defined new station catchment areas. In order to produce an estimate of total station usage, knowledge was required of other journey purposes and origin distance breakdowns at new stations. This was provided by the aggregate NWM and market research respectively. As a result of this process the forecasting inaccuracy of the market segmented approach, as measured by the Absolute Deviation (AD) measure, more than doubles so that forecasts are only within ± 34% of actual first year usage. Moreover, the market segment model can only be applied to five existing sites. If a single market approach, based on aggregate, engineering data used in conjunction with the EIL, is used to forecast usage at the six existing new stations, inaccuracy, as measured by the AD, increases further to ± 54%. This is only slightly better than the results from the all-trips ASM (AD = + 63%). All three modelling approaches can, however, be shown to be superior to the PTE's initial regression model where the AD = + 193%, but only the market segment approach is more accurate than the trip rate model (AD = ± 42%). This suggests that although a trip rate approach lacks a theoretical basis, is unlikely to be transferable and fails to replicate rankings correctly, it may provide the easiest and cheapest way of determining possible usage for one-off, low cost new stations (Preston and Nash, 1986).

It should be noted that, given patronage growth over time, the accuracy of the DAs compared to the ASM diminishes. Indeed a tendency for the DAs to underpredict usage in subsequent years (following patronage growth) is apparent and is related to problems with incorporating long term generated work trips and difficulties in measuring education trips and attracted work trips. In section 9.4 it was noted that evidence for patronage growth at new stations is not completely consistent, partly due to the difficulty of excluding exogeneous factors. However, analysis of trends in West Yorkshire's new station patronage suggests that real growth of around 75% over the first five years might be expected, although evidence from elsewhere suggests that this ought to be considered as an upper limit.

11.3.3 Findings from the Evaluation Framework

In Chapter Ten we developed financial and social FYRRs and NPVs for the six new stations that have been opened and for 28 potential sites. It was shown that the NPV results were sensitive to assumptions concerning interest rates, project life and patronage growth. Moreover, it was implicit in our comparison of forecasting approaches that evaluation results will be sensitive to the type of model estimates used. This is explicitly illustrated by Table 11.1 for financial FYRRs. As measured by the AD, the evaluation provided by the ASM is the most accurate, although this makes use of information from our market research on the average proportion of users abstracted from other Public Transport modes, hence introducing a degree of circularity into the evaluation process. Use of the DA to determine financial FYRRs leads to an increase in inaccuracy of almost 20 percentage points, as measured by the AD, although such an approach has advantages in that, at least for work journeys, estimates of net revenue are supplied by the modelling process. Use of the trip rate model leads to a further increase in inaccuracy of around 20 percentage points. This indicates that the use of a system wide mean fare is inappropriate. Table 11.1 indicates that the results of our evaluation framework have very broad confidence limits due to inaccuracies inherent in the usage forecasts and difficulties in determining net revenue from these forecasts. Furthermore when determining social costs and benefits, as shown in section 10.3.1, there were problems in measuring the extent of benefits accrueing to new station users, with it being apparent that a number of unmeasured attributes, such as comfort, should have been taken into account.

Such difficulties are further illustrated by comparing the ranking of new stations. For example one new station site is ranked bottom (or near the bottom) in terms of initial usage and forecasted usage, but is ranked in the middle in terms of financial returns and at the top in terms of social returns. This illustrates that in evaluating new stations, gaining information on total usage is only the beginning, information is also required on the fare paid, the previous mode used and the generalised cost of the previous mode vis-a-vis rail.

At the request of the PTE our evaluation results were based on a 5% interest rate and a 30 year project life and hence represent our most optimistic estimates. Some of our main results are summarised by Table 11.2 and are as follows:

(i) The package of six new stations already opened return positive financial and social NPVs under both the high growth and no growth assumptions. The results of the financial NPV under the high growth assumption for six stations are similar to the results for five stations

	"Actual"	Modelled		
	(1)	DA (2)	ASM (3)	Trip Rate Model (4)
Bramley	5.7	5.5	7.5	13.6
Crossflatts	16.7	19.0	16.6	3.9
Deighton	12.8	5.6	14.4	9.1
Fitzwilliam	11.1	6.6	5.7	4.8
Saltaire	9.5	18.6	14.6	5.5
Slaithwaite	9.4	13.9	7.5	3.0
TOTAL	10.3	11.6	10.8	6.9
RMSE		5.5	3.3	7.5
AD		0.43	0.24	0.63

TABLE 11.1 COMPARISON OF "ACTUAL" AND "MODELLED" FINANCIAL FYRRs (EXPRESSED AS PERCENTAGES, 1984 PRICES)

(1) Column 6, Table 10.1.

(2) Column 6, Table 10.2.

(3) Usage forecasts from Table 9.6. Mean fare based on modelled destinations (assuming 39% of weekday traffic off peak). Net revenue estimated as 31% of gross revenue.

(4) Usage forecasts from Table 9.6. Average fare paid 60 pence. Net revenue estimated as 31% of gross revenue.

	Financial		Social	
	High growth	Low/No growth	High growth	Low/No growth
6 opened stations (PTE estimate	1296570 1272000)	355067	1785919	252187
28 potential stations (Net rev. to rail	-582450 2433791)	-782849	-723359	-1164296
10 possible openings (Net rev. to rail	57204 1272605)	-32709	1184594	950177

TABLE 11.2 SUMMARY OF NPV RESULTS (1984 prices, 5% interest rate, 30 year project life). ERDF grant at Saltaire excluded.

produced by WYPTE (1984), although our social NPVs are lower than those produced by the PTE.

(ii) A package of a further 28 new stations would result in negative financial and social returns under most assumptions. If, however, a financial evaluation was based on net revenue to the rail operator rather than net revenue to Public Transport operators then this package may return a positive financial NPV, although paradoxically society as a whole would still suffer a loss, mainly due to the time penalties incurred by existing users.

(iii) If out of these 28 potential sites, a subset of 10 stations (as listed in 10.3.4) were opened then positive social NPVs would be achieved, although the financial NPV would only be positive to a global Public Transport operator under the high growth assumption. It is noticeable that the returns from this package are lower than the returns for the six new stations already opened.

Lastly, it should be noted that our evaluation results may be sensitive to assumptions concerning the definition and extent of the benefit algorithm. In sections 10.4.1 and 10.4.2 it was shown how changes in rail or bus operating costs or procedures (especially stopping patterns) and changes in financial arrangements (including consideration of net revenue to the rail operator only rather than Public Transport operators as a whole) affected our results, although it was noted that the introduction of Metrocard (in January 1985) seemed to have little effect on our social evaluation results. Furthermore, it was shown, in 10.4.3, that attempts to take into account reductions in road congestion and accidents as well as changes in indirect taxation led to a significant increase (around 15%) in net social benefits for the six new stations already opened, although these effects will be reduced where abstraction from car is limited and there is spare road capacity.

11.4 POLICY IMPLICATIONS

Despite the caveats given above the results of our evaluation framework have a number of policy implications. Firstly, it may be considered that the opening of six new stations in West Yorkshire between 1982 and 1984 has been a success in both financial and social terms, although there is evidence that Bramley consistently performs worse than the other stations.

This success has been achieved despite the fact that initial patronage and also benefits (in particular net revenue and benefits to new station users) do not appear to have been as high as the PTE had hoped. This success is however insufficient to radically improve the finances of the local rail network (which had been seen by some observers (for example Sully, 1981) as being an objective of the programme) as financial benefits of between £0.3m and £1.3m over a 30 year period are limited compared to a system wide operating deficit of £8.66m per annum in 1984/5 (WYPTE, 1985). The new station programme may be seen as being more consistent with a policy of maximising passenger mileage within a budget constraint. By November 1985 the six new stations had led to an additional 1300 trips per average weekday on the local rail system, whilst around 34,000 people living within one km of a new station had gained access to the rail system (representing around 1.7% of the County's population).

Secondly, our results suggest that there are a number of sites that represent good investments as measured by the social NPVs given by Table 10.13 and may be seen as being consistent with a policy of maximising social benefits within a budget constraint. These sites are, in descending order of priority:

1. East Garforth
2. Gamble Hill
3. Steeton
4. East Knottingley
5. Hemsworth
6. Osmondthorpe
7. Utley
8. Sandal
9. Wrenthorpe
10. Milnsbridge

Of these sites three (Steeton, East Knottingley and Utley) are affected by problems related to Section 20 payments (discussed in 10.4.1) whilst two sites (Utley and Wrenthorpe) give negative financial NPVs even on a net revenue to the rail operator basis (Table 10.5). The PTE have plans to open East Garforth and Frizinghall stations (with the latter not being in our list partly due to problems in assessing the level of school traffic related to Bradford Grammar School) in the near future.

Thirdly, our results allow us to refine the search procedure discussed in 9.1.1. In most circumstances we would exclude sections of the passenger rail network that are within about one mile ($\approx$1.6 km) of an existing station, sections in non built up areas and sections affected by engineering and/or operational difficulties. The definition of what constitutes a built up area might be determined by considering a number of different scenarios:

(i) A single platform station costing £60,000, with recurrent costs of £1,500 pa. In order for a financial NPV to the rail operator to be positive this would require net revenue to BR of £6,339 pa

(ii) A double platform station costing £100,000 with recurrent costs of £1,700 pa. In order for a financial NPV to be positive this would require net revenue to BR of around £9,675 pa.

(iii) A double platform station costing £150,000 (due to problems with access etc.), with recurrent costs of £1,700 pa. For a positive financial NPV this would require net revenue to BR of around £13,798 pa.

For each of these scenarios a low fare (30 pence - typical of inner city sites), a medium fare (60 pence - which is the West Yorkshire average fare) and a high fare (£1 - typical of sites on the periphery of the West Yorkshire network) were considered. In order to determine daily patronage required it was assumed, from our own market research, that only 13.4% of revenue was abstracted from existing BR services. No consideration was made of the effect of a one minute time penalty on existing revenue. From Table 11.3 it can be seen that between 25 and 177 weekday ons and offs are required in order to break even. For a typical West Yorkshire station (capital cost £100,000, mean fare 60 pence) this figure would be 63. The population within 800 metres of a station that would be required in order to achieve the break even number of ons/offs is also shown by Table 11.3. This was calculated by assuming 63% of usage originates within 800 metres of a station and that the mean weekday trip rate in the 0 to 800 m band was 20.17 per thousand population (from Table 5.27). The figures in Table 11.3 suggest a minimum definition of a built-up area might be based on a population of 800 within 800 metres of a railway line. For a typical West Yorkshire station this figure might be close to 2000 population. These results assume that the dominant access mode is walk and may be lower where feeder bus, kiss and ride or park and ride are extensively used (and hence catchment areas are larger).

Mean fare (pence)	Scenario	Number of weekday on/offs required	Population within 800 m required (to nearest 50)
30	1	81	2550
	2	125	3900
	3	177	5550
	1	41	1300
60	2	63	1950
	3	89	2800
	1	25	800
100	2	38	1200
	3	54	1700

TABLE 11.3 SOME EVALUATION GUIDELINES Scenarios 1: Capital cost £60,000 2: Capital cost £100,000 3: Capital cost £150,000. Assumes 7% interest rate, 30 year project life and population determined by the trip rate model (no growth assumed).

This method is obviously based on a crude approach. Our results have shown that patronage will be higher, the larger the existing number of commuters to nearby rail served towns from the new station area. In addition it is evident that social benefits will be increased if roads and bus services are poor and the train prior to reaching a new station is lightly loaded.

11.5 RECOMMENDATIONS FOR FURTHER RESEARCH

It is inevitable that any research will leave a number of questions unanswered as well as answering others. These might include:

(i) Could our forecasting methods be further improved? For example the trip rate model might be segmented to take into account car ownership or social class, whilst the ASM might be recalibrated given more reliable OD data, preferably disaggregated by journey purpose, as well as more reliable measures of the attraction of destinations and the generalised cost of modes.

However, as we saw in Chapter Seven, it was with the DA that potential refinements are most obvious. In particular we would wish to model journey purposes other than work and to develop generation and distribution stages. Such a model system would, however, be very data intensive and it is debateable whether the increased forecasting accuracy compared to simpler approaches would compensate for the increased research effort. In addition it might also be worthwhile to make use of other software rather than BLOGIT, particularly if HL models could be estimated directly and it was possible to test other functional forms, for example multinomial probit if taste variation is believed to exist. Furthermore in section 3.2.4 it was shown that it would be interesting to develop a disaggregate forecasting approach based on stated, rather than revealed, preference data, although this could only be achieved by major new survey work. An opportunity to develop such an approach might be provided by a research contract with Leicestershire County Council to evaluate the effect of reopening the Leicester to Burton rail line to passenger traffic.

Lastly, as was shown in section 9.4, the lack of detailed time-series information has made the assessment of patronage change over time difficult. Demand at new stations ought to continue being monitored so that a longitudinal data set might be developed and suitable analyses undertaken.

(ii) Could the evaluation framework be improved? Aside from making use of more reliable forecasts, the main way in which our evaluation framework might be improved would be to extend the benefit algorithm so as to include the effects of changes in road congestion, accident rates and indirect taxation for all stations. This would require reliable predictions of the journeys abstracted from car, as well as detailed information on road traffic flows and saturated junctions in order to assess congestion effects.

(iii) Could our methods be applied to other areas? To an extent this has been done in that trip rate models have been used to assess sites in Derbyshire and the Selby and Ryedale districts of North Yorkshire, whilst the ASM has been used in Tyne and Wear, Leicestershire and the Craven district of North Yorkshire. However, further tests of the spatial transferability of all our forecasting approaches would provide useful insights.

(iv) To what extent are our findings affected by recent changes in policy? In section 1.4.1 two Acts of Parliament were seen as having particularly important implications:

(a) The Local Government Act, 1985. This included the replacement of the PTA with a Joint Board of District Councils, in which it would be permissible for Districts to opt out (which would have serious implications were this to happen, although this currently seems unlikely in West Yorkshire) and the separation of responsibility for Public Transport and for highways and traffic management (which might restrict the amount of integration between public and private transport). However, more important effects may be caused by:

(b) The 1985 Transport Act. As we have seen in 10.4.1 this legislation has important implications in terms of the financing of Public Transport. In addition the abolition of road service licensing for stage carriage bus services is likely to lead to competition between bus and rail, if the experience of the 1980 Act on express coaches (Kilvington and Cross, 1986, Douglas, 1985) and commuter coaches (Dyer et al. 1985) is anything to go by. In addition theoretical work by Evans (1986) has indicated that, particularly in terms of frequency, bus competition will be most intense on urban and short inter-urban corridors. As West Yorkshire's passenger railways mainly serve short inter urban routes it seems likely that rail usage will decrease, particularly if competition between the former NBC and PTE bus fleets were to emerge.

Rail usage is also likely to be reduced by the reduction in integration with bus (Kilvington, 1985). Rail's response to competition might include predatory pricing, exploitation of advantages of speed (which would have implications for stopping patterns) and comfort and improved marketing (in particular to exploit the uncertainties that bus users might face due to changing bus services). In addition it should be noted that, in contrast to most bus services, rail services in West Yorkshire remain under the control of the restructured PTE/PTA and this may have important, and possibly unexpected, implications in terms of policy and objectives compared to commercial bus services.

Further to these two pieces of legislation, reduced Central Government support for Public Transport expenditure in West Yorkshire could mean the withdrawal of up to four passenger rail services (WYPTE, 1986).

Given such important policy changes it would clearly be of interest to determine their effects on both the existing new stations and on potential new stations. From our previous financial analysis it does appear that in certain cases, somewhat perversely, the 1985 Transport Act might strengthen the case for new stations, particularly in metropolitan areas.

11.6 FINAL CONCLUSIONS

If we re-examine the research aims presented in section 1.2 it can be seen that we have:

(i) established the volume and characteristics of patronage at six existing and 28 potential new station sites.

(ii) established the extent to which traffic is generated and abstracted through our market research.

(iii) measured the costs and benefits of new stations to new station users, rail and bus operators and the public at large. It has been shown that the West Yorkshire new station programme may be considered a success in both financial and social terms and up to 10 further sites have been identified as worthy of development.

(iv) made brief studies (in section 10.4.4) of the limited effects on accessibility to jobs, shopping and other facilities (although the main way these effects were taken into account were through changes in user's generalised cost).

In terms of our subsidiary aims our main conclusions relate to analysis of the trade-off between complexity and accuracy in evaluating new local rail stations. Our findings are as follows:

(i) The simplest approach we have developed is the trip rate model, which performs reasonably accurately in terms of replicating initial usage, but less well in producing revenue forecasts and assessing changes in patronage over time. Such an approach may be adequate for evaluating low cost, one-off stations particularly if trip rate information at similar nearby stations is available.

(ii) Our next simplest approach was provided by the ASM. Although this does not replicate first year patronage very well, it performs more accurately where patronage growth occurs. Such an approach might be suitable for evaluating packages of new stations, including the introduction of new services, particularly if calibrated for nearby existing services.

(iii) Our most complex approach was the DA. Although our disaggregate mode split models were accurate in what they did, due to resource constraints this was limited to forecasting journey to work shares within predefined areas. In applying these models to produce total station usage forecasts use had to be made of an aggregate NWM and, in some cases, aggregate time and cost data. Thus our DA was not, in application, totally disaggregate by any means, and this may explain, particularly when examining post first year usage, why the expected gains in accuracy did not materialise. Given the complexity of disaggregate modelling approaches it is apparent that their main role is in evaluating major projects in large conurbations (hence their application in Amsterdam, San Francisco and Santiago, for example), although we have illustrated how the EIL may be used to assess the impact of a much more modest project on journey to work flows.

Thus in searching for the most appropriate technique for evaluating new local rail stations we must conclude that the increased complexity (especially in terms of information requirements) of the DA does not appear to be compensated by similar increases in accuracy and hence a simpler approach, as provided by the ASM, may be more appropriate.

APPENDIX 1 NEW STATION SURVEY FORM

EVALUATION OF NEW LOCAL RAILWAY STATIONS PROJECT
NEW STATION SURVEY 3

Form number [ ] Accompanied by [ ] [ ] [ ]

Station name ______ Weather ______

Date & time [ ] [ ] [ ] [ ] [ ]

---

1. WHERE ARE YOU GOING TO ? Could you please give:
a) The name of the railway station ______

b) The full address of the place where you will finish this journey including: Street number (or name of shop, factory etc) ______

Street ______

Town ______

2. HOW WILL YOU GET THERE FROM THE STATION WHERE YOU LEAVE THE THE TRAIN ? Will it be:

On Foot [ ]

By Bus [ ]

By Taxi [ ]

Other ______

As Car Driver [ ] if YES where is car parked? in station car park [ ] elsewhere [ ] did not park [ ]

As Car Passenger [ ] if YES will the driver also arrive by train ? YES [ ] NO [ ]

3. WHAT IS THE MAIN PURPOSE OF YOUR JOURNEY ? Is it:

To or from work [ ] Visiting friends [ ]

To or from school/college [ ] Entertainment [ ]

On Firm's Business [ ] On Personal Business [ ]

Shopping [ ] Other/multi-purpose ______

4. HOW WILL YOU RETURN FROM (Insert information from Question 1 b)) ? ______

5a) EXCLUDING TODAY, HOW MANY TIMES IN THE PAST FOUR WEEKS HAVE YOU BEEN TO (Insert information from Question 1b)) ? Was it:

[ ] None [ ] Once [ ] 2-4 times [ ] 5-10 times [ ] 11-15 times [ ] More than 15 times

5b) HOW MANY OF THESE JOURNEYS TO (Insert information from Question 1b)) INVOLVED TRAVEL FROM THIS STATION ? Was it:

None of them	About a quarter	About a half	About threequarters	All of them
☐	☐	☐	☐	☐

5c). HOW ELSE DID YOU GET TO (Insert information from Question 1 1b)) ? ____________________

---

6. HOW MANY TIMES A MONTH DID YOU GO TO (Insert information from Question 1b)) BEFORE THIS STATION WAS OPENED ? Was it:

None	Once	2-4 times	5-10 times	11-15 times	More than 15 times
☐	☐	☐	☐	☐	☐

None: Go to Question 8a) or 8b) (See Interviewer's Note 8)

Others: Go to Question 7 then go to 8a) or 8b)

7. HOW DID YOU GET TO (Insert information from Question 1b)) BEFORE THIS STATION WAS OPENED ? Was it:

As Car Driver ☐ | By Taxi ☐

As Car Passenger ☐ | By Motorcycle, Moped or Scooter ☐

By Bus ☐ | If Bus please give details of service name and number ____________________

Other ____________________

8a) I NOTICE THAT YOU ARE TRAVELLING LESS/MORE OFTEN TO (Insert information from Question 1b)) THAN BEFORE THIS STATION WAS OPENED. IS THIS DUE TO THE OPENING OF THIS STATION OR DUE TO SOME OTHER REASON ?

Opening of Station ☐ | Other reasons (specify) ____________________

8b) WHICH OF THE FOLLOWING REASONS WOULD YOU SAY WAS MOST IMPORTANT IN YOUR DECISION TO TRAVEL BY TRAIN TODAY ? Was it:

Speed of Journey ☐ | Low Cost ☐

Service Reliability ☐ | Nearness of Station ☐

Comfort ☐ | Other/details ____________________

---

9. WHERE HAVE YOU COME FROM TO GET TO THIS STATION ? Please give the full address of the place where you started this journey including: Street number (or name of shop, factory etc)

____

Street ____________________

Town ____________________

10. HOW DID YOU GET TO THIS STATION ? Was it:

On Foot [ ]
By Motorcycle, Moped or Scooter [ ]
By Bus [ ]
Other ____________

As Car Driver [ ] — if YES where is car parked? in station car park [ ] / elsewhere [ ] / did not park [ ]

As Car Passenger [ ] — if YES is the driver also getting on the train ? YES [ ] / NO [ ]

11. WHAT TICKET WILL YOU (AND ANY CHILDREN TRAVELLING WITH YOU) USE TO TRAVEL ON THIS TRAIN ? Will it be:

Ordinary Single (or Return) Fare [ ]
Off Peak Single (or Return) Fare [ ]
Child's Fare [ ]
Multirider/Season Ticket [ ]
British Rail Senior Citizen Card [ ]
British Rail Under 24/ Student Card [ ]
MetroPass/Permit [ ]
Other ____________

---

12. HOW MANY CARS AND VANS ARE AVAILABLE FOR USE BY YOUR HOUSEHOLD ? Are there:

None [ ] One [ ] Two or more [ ]

Were any of them available to you for this journey ? YES [ ] NO [ ]

13. HOW MANY PEOPLE, INCLUDING YOURSELF, ARE IN YOUR HOUSEHOLD ? Are there:

1 [ ] 2 [ ] 3 [ ] 4 [ ] 5 [ ] 6+ [ ]

14. HOW MANY ARE THERE IN EACH OF THE FOLLOWING CATEGORIES ? How many are:

A Full-time Employed [ ]
B Full-time Students/ Schoolchildren [ ]
C Part-time Employed [ ]
D Fully Retired [ ]
E Children of Pre-School Age [ ]
F Other (unless already mentioned) e.g. Housewife,Unemployed etc [ ]

15. I SEE THERE ARE (Insert information from Question 14) EMPLOYED PEOPLE IN YOUR HOUSEHOLD. CAN YOU TELL ME WHAT THEY DO ?

1. ____________ 3. ____________

2. ____________ 4. ____________

THANK YOU FOR YOUR HELP

M [ ] F [ ] No. of accompanied children 0 [ ] 1 [ ] 2 [ ] 3+ [ ]

APPENDIX 2 CODING INVENTORY

Variable	Variable Label	and Code			Columns
CASENO	PASSENGERNO	As X02 for originating station then form number code.			1 - 4
X01	TIME	Use 24 hour clock.			5 - 8
X02	DESTINATION STATION	Use 2 digit-local code			10 - 11
X03	DESTINATION TOWN	Use 6 column Post Code			12 - 14
X04	DESTINATION STREET	"			15 - 17
X05	EGRESS MODE				18 - 19
	On Foot	01	Car Driver	04	
	Bus	02	Taxi	05	
	Car Passenger	03	Other	09	
X06	JOURNEY PURPOSE				20 - 21
	Work	01	Personal Business	05	
	Education	02	Firm's Business	06	
	Shopping	03	Visiting Friends	07	
	Leisure	04	Other	09	
X07	RETURN MODE	As X05.			22 - 23
X08	REGULARITY				24 - 25
	None	01	5 - 10 Times	04	
	Once	02	11 - 15 Times	05	
	2 - 4 Times	03	More than 15 Times	06	
X09	PROPORTION BY RAIL				26 - 27
	None	01	Half	03	
	Quarter	02	Three Quarters	04	
			All	05	
X10	ALTERNATIVE MODE	As X05.			28 - 29
X11	PREVIOUS REGULARITY	As X08.			30 - 31
X12	PREVIOUS MODE				32 - 33
	Bus	02	Taxi	05	
	Car Passenger	03	Train	06	
	Car Driver	04	Motorcycle	07	
			Other	09	
X13	REASON FOR CHANGE				34 - 35
	New Station	01	Personal	04	
	Employment	02	Other Rail	05	
	Residence	03	Other	09	
X14	ATTITUDE TO RAIL				36 - 37
	Speed	01	Safety	05	
	Cost	02	Comfort	06	
	Reliability	03	Convenience	07	
	Nearness	04	Other	09	
X15	ORIGIN TOWN	As X03, X04			38 - 40
X16	ORIGIN ADDRESS	"			41 - 43
X17	ORIGIN DISTANCE	Straight line distance from origin to station in 10s of metres.			44 - 46
X18	ACCESS MODE	As X05.			47 - 48
X19	PARKING				49 - 50
	Station Car Park	01	Passenger - Driver on Train	04	
	Elsewhere	02	Passenger - Driver not on train	05	
	Did not park	03			

X20	TICKET PRICE	Revenue in Pounds and Pence.				51 - 54
X21	TICKET TYPE					55 - 56
	Ordinary	01		B.R.Senior Citizen	06	
	Off Peak	02		Multi-Rider	07	
	Child	03		Family	08	
	Metropass	04		Conference	09	
	B.R. Under 24	05		Day Rover	11	
	Privilege	12		Duty Pass	15	
X22	CAR OWNERSHIP					57 - 58
	None	00		Two	02	
	One	01		etc.		
X23	CAR AVAILABILITY					59 - 60
	Unavailable	01		Available	02	
X24	HOUSEHOLD SIZE		As X22.			61 - 62
X25	FT EMPLOYED		"			63 - 64
X26	FT STUDENT		"			65 - 66
X27	PT EMPLOYED		"			67 - 68
X28	RETIRED		"			69 - 70
X29	PRESCHOOL CHILDREN		"			71 - 72
X30	OTHER		"			73 - 74
X31	OCCUPATION					75 - 76
	Unskilled Manual	01		Professional	04	
	Skilled Manual	02		Managerial	05	
	Clerical	03		Other	09	
X32	SEX					77 - 78
	Female	01		Male	02	
X33	ACCOMPANIED CHILDREN		As X22.			79 - 80
X34	PARTY SIZE		"			1 - 2

APPENDIX 3 DERIVATION OF THE EXTENDED INCREMENTAL LOGIT MODEL (EIL) WITH A HIERARCHICAL STRUCTURE

The upper split of a hierarchical model, for our purposes, might be expressed as:

$$P_T = \frac{\exp(U_{PT})}{\exp(U_{PT}) + \exp(U_{CD}) + \exp(U_{CP})} \qquad \text{Equation 1}$$

where P_T = probability of choosing Public Transport

U_{PT} = utility measure of Public Transport,

U_{CD} = utility measure of Car Driver and

U_{CP} = utility measure of Car Passenger.

Changes to Equation 1 could be expressed incrementally as follows:

$$P'_T = \frac{\exp(\Delta U_{PT} + U_{PT})}{\exp(\Delta U_{PT} + U_{PT}) + \exp(\Delta U_{CD} + U_{CD}) + \exp(\Delta U_{CP} + U_{CP})}$$

Equation 2

where P'_T = probability of choosing Public Transport in after situation.

If we divide through by $\sum_{\alpha'} \exp U_\alpha$, Equation 2 can be shown to be equivalent to:

$$P'_T = \frac{P_T \cdot \exp(\Delta U_{PT})}{P_T \exp(\Delta U_{PT}) + P_{CD} \cdot \exp(\Delta U_{CD}) + P_{CP} \cdot \exp(\Delta U_{CP})}$$

where P_{CD} = probability of choosing Car Driver and

P_{CP} = probability of choosing Car Passenger.

Assuming no change in Private Transport utilities Equation 3 simplifies to:

$$P'_T = \frac{P_T \cdot \exp(U'_{PT} - U_{PT})}{P_T \cdot \exp(U'_{PT} - U_{PT}) + (1 - P_T)} \qquad \text{Equation 4}$$

Prior to a new station opening we can represent Public Transport utility as:

$$U_{PT} = \phi\, U_{XT} \qquad \text{Equation 5}$$

In the after situation we have two Public Transport modes available and hence:

$$U'_{PT} = \phi \ln(\exp U'_{NT} + \exp U'_{XT}) \qquad \text{Equation 6}$$

Substituting Equations 5 and 6 into 4 and rearranging gives:

$$P'_T = \frac{P_T \cdot \{(\exp U'_{NT} + \exp U'_{XT})^{\phi} / (\exp U_{XT})^{\phi}\}}{P_T \cdot \{(\exp U'_{NT} + \exp U'_{XT})^{\phi} / (\exp U_{XT})^{\phi}\} + (1 - P_T)} \qquad \text{Equation 7}$$

This can be shown to be equivalent to Equation 8:

$$P'_T = \frac{P_T \cdot \{(\exp(\Delta U_{NT}) + \exp(\Delta U_{XT})\}^{\phi}}{P_T \cdot \{\exp(\Delta U_{NT}) = \exp(\Delta U_{XT})\}^{\phi} + (1 - P_T)} \qquad \text{Equation 8}$$

where $\Delta U_{XT} = U'_{XT} - U_{XT}$ (in our case normally equals 0) and

$\Delta U_{NT} = U'_{NT} - U_{XT}$.

This approach would need to be modified for situations where both rail and bus are available in the before situation.

Where rail has a significant market share in the before situation ($P_{NT} > 0$) then Equation 7 may be rewritten as

$$P'_T = \frac{P_T \cdot \exp\{\phi \ln(\exp U'_{NT} + \exp U'_{XT}) - \phi \ln(\exp U_{NT} + \exp U_{XT})\}}{P_T \cdot \exp\{\phi \ln(\exp U'_{NT} + \exp U'_{XT}) - \phi \ln(\exp U_{NT} + \exp U_{XT})\} + \{1 - P_T\}}$$

Equation 9

which simplifies to

$$P'_T = \frac{P_T \cdot (\{\exp U'_{NT} + \exp U'_{XT}\}^{\phi} / \{\exp U_{NT} + \exp U_{XT}\}^{\phi})}{P_T \cdot (\{\exp U'_{NT} + \exp U'_{XT}\}^{\phi} / \{\exp U_{NT} + \exp U_{XT}\}^{\phi}) + \{1 - P_T\}}$$

Equation 10

APPENDIX 4 FORMULAE FOR THE CALCULATIONS OF JUNCTION DELAY FOR OVERSATURATED LINKS

$$D = \frac{qc\,(1-v)^2}{2(1-y)} + N_o\,x \qquad (1)$$

where

D = Delay in vehicle hours per hour

q = Flow in vehicles per second

c = Cycle time in seconds

v = Green time ratio (g/c)

y = Flow rate (Flow/Saturation Flow (q/s)

and

$$N_o = QT_f\left(Z + \sqrt{z^2 + \frac{12\,(x - x_o)}{QT_f}}\right) \quad x > x_o \qquad (2)$$

$$= \text{zero} \qquad \text{otherwise}$$

where

N_o = Average overflow queue

Q = Capacity, in vehicles per hour

T_f = Flow period i.e. the time in hours during which an average arrival (demand) flow rate q persists

x = Degree of saturation (q/Q)

x_o = Degree of saturation below which the average overflow queue is approximately 0 (given by $x_o = 0.67 + sg/600$ where s = saturation flow in vehicles per second and g = effective green time in seconds.

$$z = 1 - x \qquad (3)$$

$$d = D/q$$

where

d = delay in seconds.

For the Woodpecker junction (0800 - 0900 hours):

q = 0.437 (before) 0.403 (after)

c = 110

v = 0.364

x = 0.924 (before) 0.909 (after)

Q = 1702

x_o = 0.702 and hence:

D = (from 1)) 131.82 (before) 108.20 (after)

as N_o = (from (2)) 4.26 (before) 3.40 (after)

thus d = 301.65 (before) 247.60 (after)

N.B. Peak hour flow (from Urban Congestion Monitoring project) = 1573 (before) 1547 (after)

REFERENCES

Adler, T.J. and M.E. Ben-Akiva. 1976. "Joint-Choice Model for Frequency, Destination and Travel Mode for Shopping Trips". Transportation Research Record. No. 569, pp136-150.

Akcelik, R. 1981. "Traffic Signals: Capacity and Timing Analysis". Australian Road Research Board. Research Report ARR123.

Aldcroft, D.H. 1981. "Rail Transport". SSRC/RSS Review of U.K. Statistical Sources. Vol. XIV, No.24. Pergamon Press, Oxford.

Alderson, R. 1984. "Forecasting Usage of a New Station at Dunston, Tyne and Wear". Undergraduate Dissertation, Department of Civil Engineering, University of Newcastle upon Tyne. (unpublished).

Alonso, W. 1968. "The Quality of Data and the Choice and Design of Predictive Models". Highway Research Record. Special Report 97, pp178-192.

AMV. 1972 "London Typology Study". Presented to Passenger Systems Officer, British Rail. February. (Unpublished).

Anas, A. 1982. "Residential Location Markets and Urban Transportation. Economic Theory, Econometrics and Policy Analysis with Discrete Choice Models". Studies in Urban Economics. Academic Press, New York.

Atherton, T.J. and M.E. Ben-Akiva. 1976. "Transferability and Updating of Disaggregate Travel Demand Models". Transportation Research Record. No. 610, pp12-18.

Automobile Association. 1981. "Drive and Trail Magazine". October.

Baker, S.K. 1984. "Rail Atlas of Britain and Ireland". Fourth Edition. Oxford Publishing Company, Oxford.

Bates, J.J. and M. Roberts 1983. "Recent Experiences with Models Fitted to Stated Preference Data". PTRC Summer Annual Meeting. Seminar N.

Baumol, W.J. and R.E. Quandt. 1968. "Demand for Abstract Transport Modes: Theory and Measurement". Journal of Regional Science. Vol. 6, No. 2, pp13-16.

Beckmann, M.J. and T.F. Golob. 1971. "A Utility Model for Travel Forecasting". Transportation Science. Vol. 5, pp79-91.

Beesley, M.E. and C.D. Foster. 1965. "The Victoria Line: Social Benefits and Finances". Journal of the Royal Statistical Society. Series A. Vol. 128, pp67-88.

Bellamy, P.H. 1978. "Seasonal Variations in Traffic Flows". Supplementary Report 194. TRRL, Crowthorne.

Ben-Akiva, M.E. 1975. "Structure of Passenger Travel Demand Models". Transportation Research Record. No. 526, pp.26-41.

Benjamin, J. and L. Sen. 1982. "Comparison of the Predictive Ability of Four

Multi-attribute Approaches to Attitudinal Measurement". Transportation Research Record. No. 890, pp1-6.

Black, J. 1966. "Some Retail Models". Paper for the Urban Studies Conference. September.

Black, W.C. and T.A. Black. 1982. "Deriving Service Areas for Mass Transit Systems". Transportation Research A. Vol. 16, No. 3, pp185-197.

Bonsall, P.W. and A.F. Champernowe. 1979. "The Simulation of Organised Car Sharing : 1 : The Synthesis of a Population Base". Working Paper 107. Institute for Transport Studies. University of Leeds. March.

Bovy, P.H.L. and M.A. Bradley. 1985. "Route Choice Analysed with Stated-Preference Approaches". Transportation Research Record. No. 1037, pp11-20.

Bradley, M.A., P.A. Marks and M. Wardman. 1986. "A Summary of Four Studies into the Value of Travel Time Savings". PTRC Annual Summer Meeting, Seminar M.

Bristow, M.R. and F. Rodriguez. 1973. "A Cost-Benefit Study of the Manchester to Buxton Railway". Centre for Urban and Regional Research, University of Manchester.

British Railways Board. 1963. "The Reshaping of British Railways".

British Railways Board. 1986. "Passenger Demand Forecasting Handbook-Technical Report". Issued by Director, Policy Unit on Behalf of Passenger Demand Analysis and Forecasting Group. (Unpublished).

Broom, D., S.R. Lowe, H.F. Gunn and P.M. Jones, 1983. "Estimating Values of Time: An Experimental Comparison of Transfer Price Methods with the Revealed Preference Approach". PTRC Annual Summer Meeting,. Seminar N.

Buchanan, C.D. and Partners. 1970. "Canterbury Traffic Study". October.

Burnett, K.P. and D.C.L. Prestwood. 1976. "The Development of Disaggregate Trip Distribution Models". In Stopher, P.R. and Meyburg, A.H. (Eds.). "Behavioural Travel Demand Models". Heath, Lexington, Mass.

Burnett, K.P. and S. Hanson. 1979. "A Rationale for an Alternative Mathematical Approach to Movement as Complex Human Behaviour". PTRC Summer Annual Meeting, Seminar M.

Button, R.J. and A.D. Pearman. 1983. "The Practice of Urban Transport Appraisal". Gower, Aldershot.

Cardell, N.S. and B.J. Reddy. 1977. "A Multinomial Logit Model which permits Variations in Taste Across Individuals". Charles River Associates Inc. (Unpublished).

Central Regional Council. 1980. "An Estimate of Revenue that might be Generated by Re-opening Bridge of Allan Station". TPP Team. August. (Unpublished).

Chan, D.P.1982. "A Study of the Role of Rail in the British P.T.E. Areas". MSc.

Dissertation, Department of Civil Engineering, University of Leeds. (Unpublished).

Clarke, W.R. and A.R. Cotton. 1983. "The Choice of a Light Transit System for London's Docklands". Presented at "Electrifying Urban Public Transport Conference". Blackpool. November.

Clayton, G. and J.H. Rees, 1965. "The Economic Problems of Rural Transport in Wales". Welsh Economic Studies, No. 5.

Cmnd 3686. 1968. "Transport in London". HMSO, London.

Cmnd 7131. 1978. "Nationalised Industries". White Paper. H.M.S.O. London.

Cook, R.D. 1977. "Detection of Influential Observations in Linear Regression". Technometrics. Vol. 19, pp15-18.

Cottham, G.W. 1985. "The Cost-Effectiveness of Integrated Public Transport. West Yorkshire - A Case Study". Presented at "Passenger Transport : Planning for Radical Change". University of Nottingham. March.

Coslett, S.R. 1981. "Efficient Estimation of Discrete Choice Models". In Manski, C.F. and D. McFadden (Eds.) "Structural Analysis of Discrete Data with Econometric Applications". MIT Press, Cambridge (Mass.).

Cottrell, R. 1980. "A Metro for Avon". Modern Railways, pp18-23. January.

Couture, M.R. and T. Dooley. 1981. "Analyzing Traveler Attitudes to Resolve Intended and Actual Use of a New Transit Service". Transportation Research Record. No. 794, pp27-33.

Coventry, P.S. 1977. "Sheffield and Rotherham Land Use/Transportation Study 3. Trams for Sheffield?" Traffic Engineering and Control, pp470-474. October.

Crittle, F.J. and L.W. Johnson. 1980. "Basic Logit (BLOGIT) Technical Manual". Australian Road Research Board. ATM No. 9.

Crow, R.T., K.H. Young and T. Cooley, 1973. "Alternative Demand Functions for Abstract Transportation Modes". Transportation Research. Vol. 7, No. 4, pp335-354.

Crumlish, J.D. 1966. "Notes on the State-of-the-Art of Benefit-Cost Analysis as Related to Transportation Systems". Clearinghouse for Federal Science and Technical Information. Springfield, Va.

Daganzo, C.F. 1980. "Multinomial Probit - The Theory and its Applications to Demand Forecasting". Academic Press, New York.

Daly, A.J. 1978. "Issues in the Estimation of Journey Attribute Values of Travel Choice". In Hensher, D.A. and M.Q. Dalvi (Eds.) "Determinants of Travel Choice". Saxon House, Farnborough.

Daly, A.J. 1982. "Binary and Multinomial Choice Models". PTRC Autumn Course. "New Methods for Analysing Travel Demand". November.

Daly, A.J. and S. Zachary. 1977. "The Effect of Free Public Transport on the

Journey to Work". Supplementary Report 338. TRRL. Crowthorne.

Daly, A.J. and S. Zachary. 1978. "Improved Multiple Choice Models". In Hensher, D.A. and M.Q. Dalvi (Eds.) "Determinants of Travel Choice". Saxon House, Farnborough.

Dansie, B.R. 1985. "Parameter Estimability in the Multinomial Probit Model". Transportation Research B. Vol. 19, No. 4, pp526-529.

Davies, P. and J.J. Gribble. 1978. "Bus Passenger O-D Surveys using the Seat Matrix Method". Traffic Engineering and Control, pp559-562, December.

Demetsky, M.J. and J.L. Korf. 1979. "Modelling Park'n'Ride and Kiss'n'Ride as Submodal choices". Transportation, Vol. 8, pp409-426.

de Neufville, R. and J.H. Stafford. 1971. "Systems Analysis for Engineers and Managers". McGraw-Hill, New York.

Department of Energy. 1976. Advisory Council on Energy Conservation. Paper No. 2. "Passenger Transport : Short and Medium Term Considerations". Energy Paper 10.

Department of Environment. 1985. "Commercial and Industrial Floorspace Statistics. England. 1981-84. No. 13".

Department of Industry. 1975. "Report on the Census of Distribution and Other Services, 1971. Part 5 - Area Tables - Yorkshire and Humberside Region". Business Statistics Office. SD14. Business Monitor.

Department of Transport. 1980. "Value of Time and Vehicle Operating Costs for 1979". Highway Economic Note 2.

Department of Transport. 1981. "COBA9 Manual". Policy and Methods Division. May.

Department of Transport. 1983. "Transport Statistics: Great Britain. 1972-82". HMSO, London.

Department of Transport. 1984. "Model for Evaluating Transport Subsidies: User Manual".

Destor, G. 1975. "Binary Station Choice Models for a Rail Rapid Transit Line". Transportation Research. Vol. 9, pp31-41.

Dickinson, G.C. 1967. "Passenger Transport Developments". In Beresford, M.W. and G.R.J. Jones (Eds.) "Leeds and Its Region". Leeds Local Executive Committee of the British Association for the Advancement of Science.

Dobson, P.R. 1979. "Market Segmentation: A Tool for Transport Decision Making". In Hensher, D.A. and P.R. Stopher (Eds.) "Behavioural Travel Modelling". Croom-Helm, London.

Dobson, R. and W.E. McGarvey. 1977. "An Empirical Comparison of Disaggregate Category and Regression Trip Analysis Techniques". Transportation. Vol. 6, pp287-307.

Dodgson, J.S 1977. "Cost Benefit Analysis, Government Policy and the British Railway Network". Transportation. Vol. 6, pp149-170.

Dodgson, J.S. 1984. "Railway Costs and Closures". Journal of Transport Economics and Policy. Vol. 18, No. 2, pp219-235.

Domencich, T.A. and D. McFadden. 1975. "Urban Travel Demand: A Behavioural Analysis" (A Charles Rivers Associates Research Study). North Holland/American, Elsevier.

Douglas, N.J. 1985. "A Welfare Assessment of the Deregulation of the Express Coach Market in 1980". PhD Thesis, School of Economic Studies, University of Leeds.

Dunne, J.P. 1984. "Elasticity Measures and Disaggregate Choice Models". Journal of Transport Economics and Policy. Vol. 18, No.2, pp189-197.

Dunne, J.P. 1985. "Disaggregate Mode Choice Models and the Aggregation Issue: Some Empirical Results". Transportation Research A. Volume 19, No. 4, pp315-324.

Dyer, I., D.K. Robbins and P.R. White. 1985. "Development of Commuter Coach Services into London". UTSG Annual Conference, Birmingham. January.

Else, P.K. and M. Howe, 1969. "Cost Benefit Analysis and Railway Services". Journal of Transport Economics and Policy. Vol. 3, No. 2, pp178-194.

Emerson, D. 1979. "A Study of the Market for and Feasibility of Joint Bus/Rail Tickets in West Yorkshire". MSc Dissertation, Department of Civil Engineering, University of Leeds. (Unpublished).

European Conference of Ministers of Transport (ECMT). 1980. "Scope for Railway Transport in Urban Areas". Round Table 47.

Evans, A.W. 1969. "Inter City Travel and the LM Electrification". Journal of Transport Economics and Policy. Vol. 3, No. 1, pp69-94.

Evans, A. 1986. "A Theoretical Comparison of Competition with Other Economic Regimes for Bus Services". UTSG Annual Conference, Cardiff. January.

Farrington, J.H. 1986. "Rail Traffic Potential-Newtonhill". Transport Research Group, Aberdeen University. February. (Unpublished).

Fields, J.M. and J.G., Walker, 1980. "Reactions to Railway Noise: A Survey Near Railway Lines in Great Britain: Volume 1". University of Southampton, Institute of Sound and Vibration Research, Technical Report 102.

Fleet, C.R. and S.R. Robertson. 1969. "Trip Generation in the Transport Planning Process". Highway Research Record. No. 240, pp11-27.

Foot, D.H.S. and D.N.M. Starkie. 1970. "Ashford-Hastings Railway Line". University of Reading. Geography Paper No. 3.

Foster, C.D. 1974. "Social Cost Benefit Study of Two Suburban Surface Rail Passenger Services". Report to British Railways Board. (Unpublished).

Fowkes, A.S. and M. Wardman. 1985. "The Logit Model and the Consequences of Inter Personal Taste Variations". UTSG Annual Conference, Birmingham.

Fowkes, A.S. 1983. "Report on West Yorkshire RP Re-Analysis". In "Value of Time, Phase 2". Progress Report for Meeting on 22nd November. (Unpublished).

Frankenna, M.W. 1978. "The Demand for Urban Bus Transit in Canada". Journal of Transport Economics and Policy. Vol. 12, No. 3, pp280-303.

Fukuyama, M. 1983. "Optimal Station Location for a Two Hierarchy Transit System". In Hurdle, V.F., E. Hauer and G.N. Stevart (Eds.) "Proceedings of the 8th International Symposium on Transportation and Traffic Theory". University of Toronto Press, Toronto.

Gaudry, M.J.I. and M.G. Dagenais. 1979. "The DOGIT Model". Transportation Research. B. Vol. 13, pp105-111.

Glaister, S. 1983. "Some Characteristics of Rail Commuter Demand". Journal of Transport Economics and Policy. Vol. 17, No. 2, pp115-132.

Goodwin, P.B. 1978. "On Grey's Critique of Generalised Cost". Transportation. Vol. 7, pp281-295.

GPO (General Post Office) 1982. "Using Post Codes in Business Systems".

Granger, C.W.J. and P. Newbold. 1976. "The Use of R^2 to Determine the Appropriate Transformation of Regression Variables". Journal of Econometrics. Vol. 4, pp205-210.

Gregg. S.J. 1983. "Are Roads Safer than Railways?" Modern Railways, pp200-201, April.

Grey, A. 1978. "The Generalised Cost Dilemma". Transportation. Volume 7, pp261-280.

Gunn, H.F. 1980. "Hierarchical Logit". Unpublished Note. "Advanced Modelling Seminar". University of Leeds.

Gunn, H.F. 1982. "Introduction to Individual Choice Models". Technical Note 94. Institute for Transport Studies, University of Leeds. (Unpublished).

Gunn, H.F. 1984. "An Analysis of Transfer Price Data". PTRC Annual Summer Meeting, Seminar H.

Gunn, H.F. and J.J. Bates 1980. "Some Statistical Considerations in Fitting Travel Demand Models". In "International Conference on Research and Applications of Travel Demand Models", University of Leeds.

Gunn, H.F., P.J. Mackie and J.D. Ortuzar. 1980. "Assessing the Value of Travel Time Savings - A Feasibility Study". Working Paper 137, Institute for Transport Studies, University of Leeds.

Gunn, H.F., M.E. Ben-Akiva and M.A. Bradley. 1985. "Tests of the Scaling Approach to Transferring Disaggregate Travel Demand Models". Transportation Research Record. No. 1037, pp21-30.

Gwilliam, K.M. and P.J. Mackie. 1975. "Economics and Transport Policy". George Allen and Unwin, London.

Gwilliam, K.M., A.D. May and P.W. Bonsall. 1984. "Transport in the Metropolitan Counties: Current Performance and Future Prospects. Volume 1 : Conclusions". Institute for Transport Studies, University of Leeds.

Haigh, A. 1978. "Railways in West Yorkshire". Second Edition. Dalesman Books, Clapham (North Yorkshire).

Hall, P. and C. Hass-Klau. 1985. "Can Rail Save the City? - The Impacts of Rail Rapid Transit and Pedestrianisation on British and German Cities". Gower, Aldershot.

Harata, N. and K. Ohta. 1986. "Some Findings on the Application of Disaggregate Nested Logit Models to Railway Station and Access Mode Choice". Presented at Fourth WCTR, Vancouver, January.

Harrison, A.J. 1974. "The Economics of Transport Appraisal". Croom-Helm, London.

Hartgen, D.T. and P.S. Liou. 1976. "Issues for Implementing Disaggregate Travel Demand Models". In Stopher, P.R. and A.H. Meyburg. "Behavioural Travel Demand Models". Heath, Lexington, Mass.

Hartley, T.M. 1979A. "Rail Passenger Service Demand Data for West Yorkshire". Technical Note 19. Institute for Transport Studies, University of Leeds.

Hartley, T.M. 1979B. "Aggregate Modal Split Models - Refining the Goodness of Fit". Technical Note 27. Institute for Transport Studies. University of Leeds.

Hartley, T.M. and C.A. Nash. 1980. "Management Objectives for Local Rail Services". Working Paper 132. Institute for Transport Studies, University of Leeds.

Hartley, T.M. and J.D. Ortuzar. 1979. "Modal Split Models for the Journey to Work in West Yorkshire by Car, Bus and Rail: Some Comments". Technical Note 9. Institute for Transport Studies. University of Leeds.

Harvey, G. 1985. "Research Directions in Travel Demand Analysis". Transportation Research. A. Vol. 19, No. 5/6, pp455-459.

Hauser, J.R. 1978. "Testing the Accuracy, Usefulness and Significance of Probabilistic Choice Models. An Information - Theoretical Approach". Operations Research. Vol. 26, No. 3, pp406-421.

Hausman, J.A. and D.A. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions recognizing Interdependence and Heterogeneous Preferences". Econometrica. Vol. 46, No. 2, pp403-426.

Haywood, P.J. and D.A. Blackledge. 1982. "West Midlands Rail Improvements". TRRL Symposium. "Effect of Rail Improvements on Glasgow and Other Cities". November.

Heggie, I.G. and D. Papoulias. 1976. "A Comparative Evaluation of Forecast and Use of Park and Ride in Oxford". Traffic Engineering and Control, pp144-149, April.

Hellewell, D.S. 1977. "Light Rapid Transit". Traffic Engineering and Control, pp532-537. November.

Helwig, J.T. (Editor). 1983. "SAS Introductory Guide". Revised Edition. SAS Institute Inc., Cary, N. C.

Hensher, D.A. 1977. "The Origins and Recent History of Disaggregate Demand Analysis". In Wigan, M.R. (Editor). "New Techniques for Transport Systems Analysis". Australian Road Research Board and the Bureau of Transport Economics, Vermont, Victoria.

Hensher, D.A. 1979. "Individual Choice Modelling with Discrete Commodities: Theory and Application to the Tasman Bridge Re-opening". The Economic Records, pp243-260. September.

Hensher, D.A. and L.W. Johnson. 1977. "A Two Period Analysis of Commuter Mode Choice : The Predictive Capability of Individual Choice Models". Logistic and Transportation Review. Vol. 13, No. 4, pp361-375.

Hensher, D.A. and R.G. Bullock. 1979. "Price Elasticity of Commuter Mode Choice: Effect of a 20 per cent Fare Reduction". Transportation Research. A. Vol. 13, No.3, pp193-202.

Hensher, D.A. and P.R. Stopher (Eds.). 1979. "Behavioural Travel Modelling". Croom-Helm, London.

Hensher, D.A. and L.W. Johnson. 1981. "Applied Discrete Choice Modelling". Croom-Helm, London.

Hepburn, D.R.C. 1977. "Analysis of Changes in Rail Commuting to Central London. 1966-1971". Supplementary Report 268. TRRL, Crowthorne.

Heydecker, B.G. 1985. "Development of a Travel Model which Includes Car Parks". UTSG Annual Conference, Birmingham.

Hill, D.M. and H.G. Von Cube. 1963. "Development of a Model for Forecasting Travel Mode Choice in Urban Areas". Highway Research Record. Vol. 38.

Hockenhull, G. 1984. "Investigation of Demand for New Local Rail Stations in Leicestershire examining Five Potential Sites". M.Sc. Dissertation, Department of Civil Engineering, University of Leeds. (Unpublished).

Holder, D.R. and G.C. Wood. 1980. "Adjustment and Validation of Disaggregate Models for Use in Aggregate Forecasting". In "International Conference on Research and Applications of Disaggregate Travel Demand Models". University of Leeds.

Horowitz, J.L. 1985A. "Travel and Location Behaviour: State of the Art and Research Opportunities". Transportation Research. A. Vol. 19, No. 516, pp441-453.

Horowitz, J.L. 1985B. "The Usefulness of Prediction Success Tables for Discriminating Among Random Utility Models". Transportation Research Record No. 1037, pp440-443.

Huff, D.L. 1963. "A Probabilistic Analysis of Shopping Centre Trade Areas". Land Economics. February.

Hughes, R.V. 1971. "Urban Railways Resurgent". Modern Railways, pp12-15. July.

Hunt, E.K. and J. Schwartz, 1972. "A Critique of Economic Theory". Penguin Books, Harmandsworth.

Hurdle, V.F. and S.C. Wirasinghe. 1980. "Location of Rail Stations for Many to One Travel Demand and Several Feeder Modes". Journal of Advanced Transportation. Vol. 14, No. 1, pp29-45.

Jeanes, C. and L. Lesley. 1984. "The Effect of New Stations upon the Wrexham-Bidston Railway Line". Wrexham-Bidston Railway Line Research Project. Report No. 5. Department of Town and Country Planning. Liverpool Polytechnic. May.

Johnson, I. and C.A. Nash. 1981. "Preliminary Results from a Survey of Present and Past Rail Commuters in the Hertfordshire Area". Working Paper 153, Institute for Transport Studies, University of Leeds.

Jones, G. 1984. "Capital Investment in Urban Railways - A Study of the Proejct Development Process". UTSG Annual Conference. Loughborough.

Jones, I.S. and A.J. Nichols. 1983. "The Demand for Inter City Rail Travel in the United Kingdom: Some Evidence". Journal of Transport Economics and Policy. Vol. 17, No. 2, pp133-154.

Joy, D. 1978. "A Regional History of the Railways of Great Britain. Volume 8: South and West Yorkshire". David and Charles, Newton Abbot.

Kanafani, A. 1983. "Transportation Demand Analysis". McGraw-Hill, New York.

Kemp, M.A. 1973. "Some Evidence of Transit Demand Elasticities". Transportation. Vol. 2, No. 1, pp25-52.

Kennedy, P. 1979. "A Guide to Econometrics" Martin Robertson, Oxford.

Kilvington, R.P. 1985. "Can Integrated Transport Planning Authorities be Justified?" In Harrison, A. and J. Gretton (Eds.). "Transport UK 1985: An Economic, Social and Policy Audit". Policy Journals, Newbury.

Kilvington, R.P. and A.K. Cross. 1986. "Deregulation of Express Coach Services in Britain". Oxford Studies in Transport. Gower, Aldershot.

Kingham, R.I. 1976. "An Application of Mode-Choice Methodologies to Infrequent Commuter Rail Services". Transportation Research Record. No. 610, pp30-36.

Knight, R.L. and L.L. Trygg. 1977. "Evidence of Land Use Impacts of Rapid Transit Systems". Transportation. Vol. 6, pp231-247.

Kocur, G., W. Hyman and B. Aunet. 1982. "Wisconsin Work Mode-Choice Models based on Functional Measurement and Disaggregate Behavioural Data". Transportation Research Record. No. 895, pp25-31.

Kompfner, P. 1979. "Notes on Light Rail Transit in Great Britain". Supplementary Report, 482. TRRL, Crowthorne.

Koppelman, F.S. 1975. "Prediction with Disaggregate Models, the Aggregation Issue". Transportation Research Record. No. 527, pp73-80.

Koppelman, F.S. 1976A. "Methodology for Analysing Errors in Prediction with Disaggregate Choice Models". Transportation Research Record. No. 592, pp17-23.

Koppelman, F.S. 1976B. "Guidelines for Aggregate Travel Prediction Using Disaggregate Choice Models". Transportation Research Record. No. 610, pp19-24.

Koppelman, F.S. 1983. "Predicting Transit Ridership in Response to Transit Service Changes". Journal of Transportation Engineering. Vol. 109, pp548 to 564.

Koppelman, F.S. and M.E. Ben-Akiva. 1977. "Aggregate Forecasting with Disaggregate Travel Demand Models Using Normally Available Data". Presented at World Conference on Transport Research, Rotterdam.

Koppelman, F.S. and C.G. Wilmot. 1982. "Transferability Analysis of Disaggregate Choice Models". Transportation Research Record. No. 895, pp18-31.

Koppelman, F.S., G.K. Kuah and G.G. Wilmot, 1985. "Transfer Model Up-dating with Disaggregate Data". Transportation Research Record. No. 1037, pp102-107.

Kumar, A. 1980. "Use of Incremental Form of Logit Models in Demand Analysis". Transportation Research Record. No. 775, pp21-27.

Lamb, G.M. and G. Havers, 1970. "Introduction to Transportation Planning. 2. Treatment of Networks". Traffic Engineering and Control, pp486-489. November.

Lancaster, K.J. 1966. "A New Approach to Consumer Theory". Journal of Political Economy. LLXIV, pp132-137.

Landau, U. 1978. "Aggregate Prediction with Disaggregate Models: Behaviour of the Aggregation Bias". Transportation Research Record. No. 673, pp100-105.

Lane, R., T.J. Powell and P. Prestwood Smith. 1971. "Analytical Transport Planning". Duckworth, London.

Langdon, M.G. 1984. "Improved Algorithms for Estimating Choice Probabilities in the Multinomial Probit Model". Transportation Science. Vol. 18, pp267-299.

Layard, R. (Ed.) 1972. "Cost Benefit Analysis - Selected Readings". Penguin, Harmondsworth.

Leake, G.R. 1971. "Inter-City Modal Split in Great Britain: Air Versus Rail". Centre for Transport Studies, University of Leeds.

Le Clercq, F. 1980. "Experiences with the Application of Disaggregate Techniques in Aggregate Models". In "International Conference on Research and Applications of Disaggregate Travel Demand Models", University of Leeds.

Leeds City Council and the Ministry of Transport, Housing and Local Government. 1969. "The Leeds Approach".

Leitch, Sir G. (Chair) 1977. "Report of the Advisory Committee on Trunk Road Assessment". HMSO, London.

Lerman, S.R. and C.F. Manski. 1979. "Sample Design for Discrete Choice Analysis of Travel Behaviour: The State of the Art". Transportation Research. A. Vol. 13, pp29-44.

LGORU (Local Government Operational Research Unit) 1973. Daly, A.J., G.W. Phillips, K.G. Rogers and P.J. Smith. "Planning Urban Bus Services". Report C149.

LGORU. 1976. "Huddersfield Bus Study 1: The Context". Report C230.

Litchfield, N. 1971. "Cost Benefit Analysis in Planning: A Critique of the Roskill Commission". Regional Studies. Vol. 5, pp157-183.

Liou, P.S. and A.P. Talvitie. 1975. "Disaggregate Access Mode and Station Choice Models for Rail Trips". Transportation Research Record. No. 526, pp42-52.

Liou, P.S., G.S. Cohen and D.T. Hartgen. 1975. "Application of Disaggregate Modal Choice Models to Travel Demand Forecasting for Urban Transit Systems". Transportation Research Record. No. 534, pp52-62.

Los, M. 1977. "Determining Modal Choice and Station Choice for a Rail Transit Line". Department des Sciences Economiques. Centre de Recherche sur les Transports, Universite de Montreal. October.

Louviere, J.J., D.H. Henley, G. Woodworth, R.J. Meyer, I.P. Levin, J.W. Stoner, D. Curry and D.A. Anderson. 1981. "Laboratory-Simulation versus Revealed-Preference Methods for Estimating Travel Demand Models". Transportation Research Record. No. 794, pp42-51.

Louviere, J.J. and D.A. Hensher. 1982. "Design and Analysis of Simulated Choice or Allocation Experiments in Travel Choice Modelling". Transportation Research Record. No. 890, pp11-17.

Lowe, S. 1981. "Elasticities in Logit Models". Unpublished Note.

Luce, R.D. and P. Suppes. 1965. "Preference, Utility and Subjective Probability". In Luce, R.D., R.R. Bush and E. Galanter (Eds.). "Handbook of Mathematical Psychology. Volume 3". John Wiley, New York.

Lupton, D.R. 1976. "Free Standing Town Studies: Public Transport Accessibility Model". WYTS Study Document 405.

McFadden, D. 1974. "Conditional Logit Analysis of Qualitative Choice Behaviour". In Zarembka, P. (Editor) "Frontiers in Econometrics". Academic Press, New York.

McFadden, D. 1976. "Quantal Choice Analysis: A Survey". Annals of Economic and Social Measurement. Vol. 5, pp363-390.

McFadden, D. 1979. "Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments". In Hensher, D.A. and P.R. Stopher. "Behavioural Travel Modelling". Croom-Helm, London.

McFadden, D. 1981. "Econometric Models of Probabilistic Choice". In Manski, C.F. and D. McFadden (Eds.). "Structural Analysis of Discrete Data with Econometric Applications". MIT Press, Cambridge, Mass.

McFadden, D. and F. Reid. 1975. "Aggregate Travel Demand Forecasting from Disaggregated Behavioural Models". Transportation Research Record. No. 534, pp24-37.

McFadden, D., W. Tye and K. Train. 1976. "Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model". WP7616, Institute of Transportation Studies, University of California.

McFadden, D. et al. 1979. "Overview and Summary: Urban Travel Demand Forecasting Project". Research Report 796, Institute of Transportation Studies, University of California.

McGillivray, R.G. 1972. "Mode Split and the Value of Travel Time". Transportation Research. Vol. 6, pp309-316.

McLeod, G.M., J.T. Everest and N.J. Paulley. 1980. "Analysis of Rail and Air Passenger Flows between London and Glasgow using Box-Jenkins Method". Supplementary Report 524. TRRL, Crowthorne.

McLynn, J. 1973. "Some Consideration of Modal Split Models for the 1972 Transportation Needs Study". National Bureau of Standards.

McLynn, J.M. and R.H. Watkins. 1967. "Multi-Mode Assignment Model in Approaches to the Modal Split for Inter City Transport". NE Corridor Project, Technical Paper 7.

Mackie, P.J. 1977. "Cost of Operating Bus and Rail Services in Urban Areas". In "Integrating Public Transport Symposium", University of Newcastle, March.

Manchester Rapid Transit Study. 1967. "Volume 1 : Report of the Working Party". September.

Manski, C.F. and S.R. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples". Econometrica. Vol. 45, No. 8, pp1977-1988.

Martin, K.M. and M.Q. Dalvi, 1976. "The Comparison of Accessibility by Public and Private Transport". Traffic Engineering and Control, pp509-513.

Martin and Voorhees Associates, Scottish Development Department and TRRL. 1980. "The Rail Survey". Working Paper 6. February.

Martin and Voorhees Associates, Scottish Development Deprtment and TRRL. 1982. "Glasgow Rail Impact Study. Final Report". September.

Matzaros, A. 1982. "The Estimation of Nested Logit Choice Models". MSc Dissertation, Department of Civil Engineering, University of Leeds (Unpublished).

May, A.D., A.S. Fowkes and F.O. Montgomery. 1983. "Travel Time Monitoring in Urban Areas". Working Paper 166, Institute for Transport Studies, University of Leeds.

Merseyside P.T.E. and B.R. 1981. "The Merseyrail Projects: A Monitoring Report".

Meyer, R.J., I.P. Levin and J.J. Louviere. 1978. "Functional Analysis of Mode Choice". Transportation Research Record. No. 673, pp1-7.

Mills, T.C. 1983. "The Functional Form of the U.K. Demand for Money". The Journal of the Royal Statistical Society. Series C. Vol. 27, No. 1, pp52-57.

Ministry of Transport. 1969. "The Cambrian Coast Line. A Cost-Benefit Analysis of the Retention of Railway Services on the Cambrian Coast Line (Machynlleth-Pwllheli). HMSO.

Modern Railways. 1985. "More Stations for Manchester". August.

Moser, C.A. and G. Kalton. 1971. "Survey Methods in Social Investigation". Heinemann, London.

Moss, A.J. and G.R. Leake. 1977. "Some Passenger Access Characteristics to Certain Inter-City Rail Termini". Traffic Engineering and Control, pp315-318. June.

Mufti, R.K., L.S. Golfin, and C.D. Dougherty. 1977. "A Procedure to Estimate Demand for Regional Fringe Parking Facilities". 56th Annual Meeting of the Transportation Research Board. Washington D.C. January.

Munby, D.L. 1978. "Inland Transport Statistics. Great Britain. 1900-70. Volume 1. Railway, Public Road Passenger Transport, London Transport". Clarendon Press, Oxford.

MVA Consultancy. 1984. "Evaluation of Transfer Pricing Methods - Final Report". Unpublished Report to the Department of Transport.

MVA Consultancy, Institute for Transport Studies, Leeds University and Transport Studies Unit, Oxford University. 1986. "Research into the Value of Time - Draft Final Report". Prepared for Department of Transport. February.

Nash, C.A. 1982. "Economics of Public Transport" Longman, London.

Nash, C.A. and A.E. Whiteing. 1984. "A Note on the Application of Mets to Rail Services". Technical Note 141. Institute for Transport Studies, University of Leeds.

Neuberger, H. 1971. "User Benefit in the Evaluation of Transport and Land Use Plans". Journal of Transport Economics and Policy. Vol. 5, pp52-57.

Nie, N.H., C. Hadlai Hull, J.G. Jenkins, K. Steinbrenner and D.H. Bent. 1975. "Statistical Package for the Social Sciences". Second Edition. Mc-Graw-Hill, New York.

Nitta, Y. and M. Mori. 1986. "Station Choice Models Incorporating Generalised Time". Presented at Fourth WCTR, Vancouver. January.

Oi, W.Y. and P.W. Shuldiner. 1962. "An Analysis of Urban Travel Demand". North Western University Press, Evanston, Ilinois.

OPCS (Office of Population Censuses and Surveys). 1979. "Census, 1971. England and Wales. General Report. Part 1. Definitions".

OPCS. 1983 "1981 Census - Special Workplace Statistics". User Guide 87. October.

Ortuzar, J.D. 1980A. "Modal Choice Modelling with Correlated Alterntives". PhD Thesis. Department of Civil Engineering, University of Leeds.

Ortuzar, J.D. 1980B. "Multimodal Choice Modelling - Some Relevant Issues". Working Paper 138. Institute for Transport Studies, University of Leeds.

Ortuzar, J.D., F.J. Achondo and A. Espinosa. 1986. "On the Stability of Logit Mode Choice Models". PTRC Annual Summer Meeting. Seminar M.

Ou, F-L. and J.C. Yu. 1983. "Transferability of Urban Travel Demand Models". Journal of Transportation Engineering. Vol. 109, No. 4, pp579-595.

Peacock, A. 1973. "Cost Benefit Analysis and the Control of Public Investments". In J.N. Wolfe (Editor). "Cost-Benefit and Cost Effectiveness". Allen and Unwin, London.

Peacock, A. and C. Rowley. 1975. "Welfare Economics: A Liberal Restatement". Martin Robertson, Oxford.

Pearce, D.W. and C.A. Nash. 1981. "The Social Appraisal of Projects: A Text in Cost Benefit Analysis". MacMillan, London.

Peat, Marwick, Mitchell & Co. 1975. "Draft Final Report on Local Rail Interchanges in West Yorkshire". April.

Pickup, R. 1984. "Increasing Patronage in West Yorkshire". Presented at Annual Conference on Public Transport Operations. University of Leeds. July.

Pindyck, R.S. and D.L. Rubinfeld. 1981. "Econometric Models and Economic Forecasts". McGraw-Hill, New York.

Pownall, C.E. and A.G. Wilson. 1976. "A New Representation of the Urban System for Modelling and for the Study of Micro Level Interdependence". Area. Vol. 8, No. 4, pp246-258.

Prest, A.R. and R. Turvey. 1965. "Cost Benefit Analysis: A Survey". Economic Journal. Vol. 78, pp638-731.

Preston, J.M. 1983. "Main Report of Pilot Survey". Evaluation of New Local Rail Stations Project. Work Paper 1. (Unpublished).

Preston, J.M. 1984A. "Rail Demand in West Yorkshire". Evaluation of New Local Rail Stations Project. Work Paper 4. (Unpublished).

Preston, J.M. 1984B. "Evaluation of New Local Rail Stations". UTSG Annual Conference, Loughborough.

Preston, J.M. 1987. "A Review of Recent Developments in Urban Railways". Technical Note 208, Institute for Transport Studies, University of Leeds. (Forthcoming).

Preston, J.M. and C.A. Nash. 1986. "Some Guidelines for Evaluating New Local Rail Stations". Working Paper 228. Institute for Transport Studies, University of Leeds.

Prideaux, J.D.C.A. 1983. "Transport Interchanges". Paper to Institution of Highway Engineers National Conference. "Transportation in Conurbations". December.

Projectbureau, 1977. "The SIGMO Study". Four Reports to the Ministry of Transport of the Netherlands, The Hague.

Quarmby, D.A. 196. "Journey to Work - Some Findings". Journal of Transport Economics and Policy. Vol. 1, No. 3, pp273-314.

Railway Development Society. 1984. "Bring Back the Trains - the Case for Railway Reopenings". Second Edition. Cambridge.

Rao, P. and R.L. Miller. 1971. "Applied Econometrics". Wadsworth, Belmont, California.

Rassam, P.R., R.H. Ellis and J.C. Bennett. 1971. "The N-Dimensional Logit Model: Development and Application". Highway Research Record. No. 369, pp135-147.

Read, M. 1977. "Assessment of Bus Feeder Services to Rapid Transit Stations". PhD Thesis. Department of Civil Engineering. University of Leeds.

Reaume, D. 1974. "Cost-Benefit Techniques and Consumer Surplus: A Classificatory Analysis". Public Finance. Vol. 29, No. 2.

Reid, F.A. 1978. "Minimizing Error in Aggregate Predictions from Disaggregate Models". Transportation Research Record. No. 673, pp59-65.

Reid, F.A. 1979A. "Roles of Behavioural Travel Models in Urban Policy Analysis". In Hensher, D.A. and P.R. Stopher (Eds). "Behavioural Travel Modelling". Croom-Helm, London.

Reid, F.A. 1979B. "Aggregation Methods and Tests". In McFadden, D. et al. "Overview and Summary: Urban Travel Demand Forecasting Project".

Richards, M.G. 1980. "Can Disaggregate Models Improve our Forecasting Approach?" In "International Conference on Research and Applications of Disaggregate Travel Demand Models", University of Leeds.

Richards, M.G. and M.E. Ben-Akiva. 1975. "A Disaggregate Travel Demand Model". Saxon House, Farnborough.

Roberts, J. 1985A. "Manifestations - Part 1 - Metropolitan Counties". Transport Retort, Vol. 8, No. 5.

Roberts, J. 1985B. "Manifestations - Part 2 - Southern England Shire Counties". Transport Retort, Vol. 8, No. 8.

Roberts, J. 1986A. "Manifestations - Part 3 - Scotland and Wales". Transport Retort. Vol. 9, No. 1.

Roberts, J. 1986B. "Manifestations - Part 4 - Northern English Shires". Transport Retort. Vol. 9, No. 4.

Roberts, J. 1986C. "New Life in the Hills". Report to the Settle-Carlisle Joint Action Committee.

Ruhl, A., A. Daly and G.C. Dobson. 1979. "The Use of Disaggregate Models for Railway Investment Decisions in Holland". PTRC Annual Summer Meeting. Seminar M.

Savage, I.P. 1983. "A Probability Method for Bus Revenue Estimation". Traffic Engineering and Control, pp32-33. January.

Schneider, M. 1959. "Gravity Models and Trip Distribution Theory". Papers and Proceedings of the Regional Science Association.

Schuman, H. and S. Presser. 1981. "Questions and Answers in Attitude Surveys". Academic Press, New York.

Scott, A.J. 1971. "Dynamic Location - Allocation Systems: Some Basic Planning Strategies". Environment and Planning. Vol. 3, pp73-82.

Self, P. 1972. "Econocrats and the Policy Process". Macmillan, London.

Sherret, A. 1979. "The Role of Disaggregate Travel Models in Transport Policy Analysis". In Hensher, D.A. and P.R. Stopher (Eds) "Behavioural Travel Modelling". Croom-Helm, London.

Silman, L.A. 1979. "The Time Stability of a Disaggregate Modal Split Model". Working Paper. Israeli Institute of Transportation Planning and Research. Tel Aviv.

Small, K. and D. Brownstone. 1982. "Efficient Estimation of Nested Logit Models: An Application to Trip Timing". Working Paper, Princeton University. March.

Sobel, K.L. 1980. "Travel Demand Forecasting by Using the Nested Multinomial Logit Model". Transportation Research Record. No. 775, pp48-55.

Spear, B.D. 1977. "Applications of New Travel Demand Forecasting Techniques to Transportation Planning - A Study of Individual Choice Models". U.S. Department of Transportation. Federal Highway Administration. Office of Highway Planning. Urban Planning Division.

Stark, D.C. 1981. "Time Series Analysis of Glasgow Suburban Rail Patronage". Supplementary Report 649, TRRL. Crowthorne.

Steer, Davies and Gleave Ltd. 1984. "Methodology Review: Forecasting the Effects of Changes in Access on BR". Unpublished Report to BR. September.

Stewart, M.B. and K.F. Wallis. 1981. "Introductory Econometrics". Second Edition. Basil Blackwell, Oxford.

Stopher, P.R. 1975. "Goodness-of-Fit Measures for Probabilistic Travel Demand Models". Transportation. Vol. 4, pp67-83.

Stopher, P.R. and S.O. Lavender. 1972. "Disaggregate Behavioural Travel Demand Model: Empirical Tests of Three Hypotheses". Transport Research Forum Proceedings.

Stopher, P.R. and A.H. Meyburg. 1975. "Urban Transportation Modelling and Planning". Heath, Lexington, Mass.

Stopher, P.R., A.H. Meyburg and W. Brog (Eds.). 1981. "New Horizons in Travel Behaviour Research". Lexington Books, Lexington, Mass.

Sugden, R. 1972. "Cost Benefit Analysis and the Withdrawal of Railway Services". Bulletin of Economic Research. Vol. 1, No. 24, pp23-32.

Sully, J. 1981. "New Stations in West Yorkshire". Modern Railways, p289. June.

Talvitie, A.P. 1973. "Aggregate Travel Demand Analysis with Disaggregate or Aggregate Travel Demand Models". Proceedings-Transportation Research Forum. Vol. 14, No. 1.

Talvitie, A.P. and D. Kirshner. 1978. "Specification, Transferability and the Effect of Data Outliers in Modelling the Choice of Mode in Urban Travel". Transportation. Vol. 7, pp311-331.

Talvitie, A.P., Y. Dehgahni and M. Anderson. 1980. "An Investigation of Prediction Errors in Mode Choice Models". In "International Conference on Research and Applications of Disaggregate Mode Choice Models", University of Leeds.

Traffic Research Corporation. 1969. "West Yorkshire Transportation Study. Technical Report." July.

Train, K. 1978. "A Validation Test of a Disaggregate Mode Choice Model". Transportation Research. Vol. 12, pp167-174.

Transecon International. 1986. "Estimating the Demand Effects of Significant Changes in Access/Egress Times - A Case Study of Bristol Parkway". Unpublished Report to BR. January.

Transport 2000. 1983. "Investing in British Rail". TEST Report 49. London.

Travers, R., Morgan and Partners. (In Conjunction with West Yorkshire M.C.C. and West Yorkshire P.T.E.). 1974. "Buses in Bradford: An Interim Report of the Bradford Bus Study".

TRRL, University of Newcastle upon Tyne, Tyne and Wear County Council and Tyne and Wear PTE. 1986. "The Metro Report. Summary Findings of the Impact of Metro and Public Transport Integration in Tyne and Wear". Metro Monitoring and Development Study.

Truong, T.P. and D.A. Hensher. 1980. "Valuation of Commodity Time Savings and Budget Time". School of Economic and Financial Studies, Macquarie University, Sydney.

Tyler, J. and R. Hassard. 1973. "Gravity/Elasticity Models for the Planning of the Inter-Urban Rail Passenger Business". (Unpublished paper, British Railways Board).

UMRCC (University of Manchester Regional Computing Centre). 1982. "SASPAC User's Guide".

University of Leeds. 1980. "International Conference on Research and Application of Disaggregate Travel Demand Models".

Van Zwam, H.H.P. and A.J. Daly. 1981. "Travel Demand Models for the Zuidvleugel Study". PTRC Annual Summer Meeting, Seminar N.

Von Ehrenrook, W.H. 1975. "Arlington Park Commuter Station". Traffic Engineering, pp19-24. January.

Vuchic, V.R. 1966. "Interstation Spacing for Line-Haul Passenger Transportation". PhD Dissertation - Graduate Report. The Institute of Transportation and Traffic Engineering, University of California, Berkeley.

Vuchic, V.R. 1969. "Rapid Transit Interstation Spacings for Maximum Number of Passengers". Transportation Science. Vol. 3, No. 3, pp214-232.

Vuchic, V.R. and G.F. Newell. 1968. "Rapid Transit Interstation Spacings for Minimum Travel Time". Transportation Science. Vol. 2, No. 4, pp303-309.

Wabe, J.S. 1969. "Commuter Travel to Central London". Journal of Transport Economics and Policy. Vol. 3, No. 1, pp48-68.

Wagon, D.J. and A.F. Hawkins. 1970. "The Calibration of the Distribution Model for the SELNEC Study". Transportation Research. Vol. 4, pp103-112.

Walmsley, D.A. 1982. "Putting it all together: General Conclusions from the Impact Studies". Presented at TRRL Symposium. "Effects of Rail Improvements in Glasgow and Other Cities". Glasgow, November.

Wardman, M. 1986. "Route Choice and the Value of Motorists' Travel Time: Empirical Findings". Working Paper 224, Institute for Transport Studies, University of Leeds.

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

Watson, P.L. 1975. "Comparison of the Model Structure and Predictive Power of Aggregate and Disaggregate Models of Inter City Mode Choice". Transportation Research Record No. 527, pp59-65.

Webb, M. 1980. "British Rail's Ticket Sales Data System". Technical Note 42, Institute for Transport Studies, University of Leeds.

Webster, F.V. 1977. "Urban Passenger Transport: Some Trends and Prospects". Laboratory Report 771. TRRL, Crowthorne.

Westin, R.B. 1974. "Predictions from Binary Choice Models". Journal of Econometrics. Vol. 2, No. 1, pp1-16.

White, P.R. 1976. "Planning for Public Transport". Hutchinson, London.

White, P.R. and S.R. Williams. 1976. "Modelling of Cross Country Rail Passenger Trips". PTRC Summer Annual Meeting, Seminar N.

Whitehead, P.M. 1981. "Estimating the Effects on Revenue of Rail Service Changes". PTRC Summer Annual Meeting, Seminar K.

Whiteing, A.E. 1977. "Some Aspects of the Cost of Access to the Bus System: An Investigation of Walking Distances to Bus Services in Leeds". MA. Dissertation, School of Economic Studies, University of Leeds.

Wilks, S.S. 1962. "Mathematical Statistics". John Wiley and Sons, New York.

Williams, H.C.W.L. 1977A. "On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit". Environment and Planning. A. Vol. 9, pp285-344.

Williams, H.C.W.L. 1977B. "Generation of Consistent Travel-Demand Models and User-Benefit Measures". In Hills, P.J., M.Q. Dalvi and P.W. Bonsall (Eds). "Urban Transportation Planning: Current Themes and Future Prospects". Abacus Press, Tunbridge Wells.

Williams, H.C.W.L. and J.D. Ortuzar. 1979. "Behavioural Travel Theories, Model Specification and the Response Error Problem". PTRC Annual Summer Meeting, Seminar M.

Williams, H.C.W.L. and M.L. Senior. 1977A. "Model Based Transport Policy Assessment. 1. The Use of Alternative Forecasting Models". Traffic Engineering and Control, pp402-406. September.

Williams, H.C.W.L. and M.L. Senior. 1977B. "Model Based Transport Policy Assessment. 2. Removing Fundamental Inconsistencies from the Models". Traffic Engineering and Control, pp464-469. October.

Williams, H.C.W.L. and M.L. Senior. 1977C. "Accessibility, Spatial Interaction and the Evaluation of Land-Use Transportation Plans". Working Paper 203, School of Geography, University of Leeds.

Williams, H.C.W.L. and M.L. Senior. 1978. "Accessiblity, Spatial Interaction and the Spatial Benefit of Land-Use Transportation Plans". In Karlquist, A, L. Lundqvist, F. Snickars and J.W. Weibull (Eds.). "Spatial Interaction Theory and Planning Methods". North Holland Press, Elsevier.

Wilson, A.G. 1970. "Advances and Problems in Distribution Modelling". Transportation Research. Vol. 4, pp1-18.

Winch, R.M. and S.A. Smith. 1980. "Short Term Impacts of a Suburban Rapid Rail Station: Case Study Results for Silver Spring, Maryland". Transportation Research Board. January.

WMCC (West Midlands County Council). 1984. "West Midlands Rail Mode Choice Models". Joint Transport Planning Unit (with the assistance of the MVA Consutancy). May.

Wohl, M. and C. Hendrickson. 1984. "Transportation Investment and Pricing Principles". John Wiley and Sons, New York.

Wootton, H.J. 1980. "The Nature of Problems and the Type of Model". In "International Conference on Research and Applications of Disaggregate Travel Demand Models", University of Leeds.

WYMCC. 1978A. "Structure Plan - Written Statement for Submission to the Secretary of State". November.

WYMCC. 1978B. "Structure Plan - Statement on Participation". December.

WYMCC. 1980. "Structure Plan - As Approved by the Secretary of State for Environment". July.

WYMCC. 1982. "Transport Policies and Programmes". July.

WYMCC. 1985. "Transportation Trends. A Review of Trends affecting Transport in West Yorkshire".

WYPTE. 1979. "Public Transport Plan for West Yorkshire".

WYPTE. 1983. "Annual Report and Accounts 1982-3".

WYPTE. 1984A. "New Rail Station Programme". Public Transport Sub Group. Report of Meeting on 20th March. (Unpublished).

WYPTE, 1984B. "New Rail Station, Project Review". Directorate of Finance and Administration. February. (Unpublished).

WYPTE, 1985. "West Yorkshire Transport Plan 3. 1986/7-1988/8". July.

WYPTE and BR. 1982. "A New Concept in Local Transport: Introduction of Lightweight Diesel Multiple Units on Local Services in and around West Yorkshire". (Unpublished).

WYTCONSULT. 1975A. "Rail Studies Station Inventory". WYTS Study Document 703.

WYTCONSULT. 1975B. "Rail Network for Study". WYTS Study Document 701.

WYTCONSULT. 1975C. "Speed Flow Relationships on Road Networks". WYTS Study Document 504.

WYTCONSULT. 1976A. "The Railway System". WYTS Interim Report 2. August.

WYTCONSULT. 1976B. "Rail Passenger Survey". WYTS Study Document 507.

WYTCONSULT. 1976C. "Leeds 1981 Short Term Economic Base and Public Transport Tests". WYTS Study Document 312.

WYTCONSULT. 1976D. "Trip Generation Model. Household Category Analysis". WYTS Study Document 515.

WYTCONSULT. 1976E. "Free Standing Towns Studies: Public Transport Accessibility Model". WYTS Study Document 405.

WYTCONSULT. 1977A. "The Extent and Impact of Travel in 1975". WYTS Final Report. Volume 1. October.

WYTCONSULT. 1977B. "Development, Testing and Evaluation of the Medium Term Provisional Plan and Highway Sub Tests for Leeds". WYTS Study Document 318.

WYTCONSULT. 1977C. "Light Rail Transit for Leeds: An Initial Appraisal". WYTS Study Document 316.

WYTCONSULT. 1977D. "The Development of Medium Term Alternatives for Bradford". WYTS Study Document 310.

WYTCONSULT. 1977E. "Development, Testing and Evaluation of the Provisional Medium Term Plan for Bradford". WYTS Study Document 322.

WYTCONSULT. 1977F. "The Dimensions of Future Change". WYTS Final Report. Volume 2. October.

WYTCONSULT. 1977G. "Recommended Policies and Plans". WYTS Final Report. Volume 3. October.