Monograph:
Preston, J.M and Aldridge, D.M. (1991) Greater Manchester PTE New Railway Station Demand Prediction Model. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 326

Institute of Transport Studies
University of Leeds

Published paper

Preston, J.M., Aldridge, D.M. (1991) *Greater Manchester PTE New Railway Station Demand Prediction Model.* Institute of Transport Studies, University of Leeds. Working Paper 326

Working Paper 326

July 1991

GREATER MANCHESTER PTE NEW RAILWAY STATION DEMAND PREDICTION MODEL

J M Preston and D M Aldridge

This work was sponsored by the Economic and Social Research Council

Contents

1. INTRODUCTION

This paper reports on work that was initiated in February 1989 to develop a simple model that can accurately predict the usage of new stations in the Greater Manchester area. The starting point for this study was the work carried out by Moss in 1988 in which an attempt was made to develop a new station demand model based on patronage data for 9 of the 10 new stations that had been opened in Greater Manchester since 1984. In his research the main explanatory variable was the population within 1,000m and the sub-divisions 0-300m, 300-600m, 600-800m and 800-1,000m were also examined. Arbitrary dummy variables were studied to assess the effect of service frequency, car ownership, alternative routes and park and ride. The main finding was that population, on its own, did not appear to be an adequate explanatory variable. For example, Mills Hill was found to attract 7.9 times as many passengers per 1,000 households as Hag Fold. In other words, there was very large variation in the trip rates at these nine new stations, with a mean of 36.2 and a standard deviation of 22.4 daily trips per thousand households.

An alternative approach is that developed in West Yorkshire based on multiple regression techniques (Preston, 1987). This model predicts the number of rail trips between two stations as a function of:

(i) the population within 800 metres of the origin station,
(ii) the proportion of that population in social classes I and II,
(iii) the population between 800 metres and 2 kilometres of the origin station,
(iv) the number of jobs within 800 metres of the destination station,
(v) the generalised cost of rail,
(vi) the generalised cost of competing modes (bus and car).

This model is a form of direct demand model in that it forecasts the number of trips (T) between origin i and destination j by mode k (ie. Tijk). In this paper, we shall develop a simpler version of this model which will simply predict the number of trips from origin i by mode k (ie. Tik). We shall call this a trip end model.

The West Yorkshire model (called the Aggregate Simultaneous Model - ASM) was calibrated for 39 existing stations based on patronage data collected in the early 1980s. In this work, we shall attempt to calibrate a similar model for 36 existing stations in Greater Manchester, based on patronage data collected in 1987/8. These stations are listed in Appendix 1. Of these stations 16 are on what we have termed the Oldham Loop, 9 are on the Bury Line, 9 are on the Altrincham Line and the remaining 2 are on the Buxton Line. It was felt that this sample was reasonably representative of Greater Manchester new stations although well used commuter stations may be unavoidably over represented due to the availability of patronage data which dictated at data set.

In developing a new station model for Greater Manchester, we have borne in mind comments made about Moss's earlier work by Greater Manchester PTE, Greater Manchester Transportation Unit, Manchester City Council and British Rail, Provincial (Midland) and in particular the consideration of existing trip patterns, especially to central Manchester. Whilst we encountered difficulties in obtaining relevant data, we were able to explicitly incorporate a number of other points. Our work will be based on multiple regression and will be able to examine the effect of distance, frequency and car ownership directly. Use will be made of 0-800m and 800m - 2km populations (rather than households within 1km), adjusted to take into account overlapping catchment areas. Figure 1 shows the zoning scheme used.

Having dealt with the background to this study, the rest of the report will be as follows:

- in section 2 we outline the data sought and made available for our study;

- in section 3 we describe the process of calibrating a simple trip end model;

- in section 4 our simple trip end model is developed further and in greater depth with the aim of maximising the goodness of fit;

- in section 5 we develop a more generalised framework;

- in section 6 we calibrate a simple trip end model for walk access patrons;

- in section 7 we comment on the criteria of model choice, discuss statistical problems related to cross-sectional data and summarise our findings;

2. DATA REQUIREMENTS AND AVAILABILITY

In this section we shall discuss the data that was required for this study and the data made available. Firstly, we shall describe the data that was available on rail demand, which formed the basis for our dependent variable, secondly the Census data used in the study and thirdly the level of service data required for rail, car and bus. The Census data and the level of service variables provided the basis for our independent variables.

2.1 Rail Demand Data

The rail demand data was provided from origin and destination surveys that were primarily designed to assess the impact of Light Rapid Transit. These included:

(i) Loadings by time periods (between 07.00 and 24.00 hours) for weekdays, Saturdays and Sundays for the Altrincham and Bury Lines. This was based on fieldwork carried out by the Harris Research Centre between 5th and 26th September 1987. Data was also made available on ticket type (standard single/return concession, cheap day return, cheap day return concession, BR seasons and GMPTE saver tickets) and access and egress modes.

(ii) Loadings for Saturday and the weekday peak and off-peak periods for the Rochdale Line and Oldham Loop. This was based on fieldwork carried out by the Harris Research Centre between 29th October and 12th November, 1988. Information was made available on journey purpose (work/education and other), ticket type (as before, but also including BR Pass/Scholars Pass), car availability (driver, passenger, not available) and access/egress modes.

(iii) Data on services from Piccadilly to Alderley Edge, Prestbury and Buxton (which we shall call the Buxton Line). This data was based on 1984 data expanded to 1988 levels. In addition to origin/destinations, it included information on access mode, ticket type (single, return, saver seven, saver monthly, weekly season, monthly season, employee season, concessionary, saver annual and other), time of day (07.00-09.30, 09.30-14.00), trip purpose (home-work, home-education, homeshop/personal business, home-social/recreational, home-employer's business, other home, non-home employer's business and other non home) and egress mode.

We encountered two problems with this data. Firstly, it was supplied as hard copy, often in a very disaggregate form necessitating the need for manually produced summary statistics. These were prepared for boarding passengers but we were not able to repeat this process for alighting passengers and our analysis was therefore based on boarding passengers only. Clearly, for the system as a whole the number of boarding and alighting passengers are the same. However, the work carried out in West Yorkshire indicated this does not hold true for individual stations due to factors such as topography, location of stations and competing bus stops, variations in levels of rail and bus services by time of day.

Secondly, although the three main data sources were broadly similar, their formats were not identical. In particular, the data on the Buxton Line only covered the period 07.00-14.00 hours (ie. roughly half a day). The data for the Altrincham, Bury and Rochdale Lines and Oldham Loop covered the period up to 24.00 hours and was used to provide a suitable factor to transform the Stockport Line data into an all day usage estimate. There were also differences in the level of detail of time of day, journey purpose and day of week information. In this paper we shall only study weekday traffic and shall distinguish between peak (defined as 07.00-09.30 and 15.00-18.00 hours) and off-peak periods.

2.2 Census Data

The 1981 Census, up-dated wherever possible, provided a large number of potentially useful variables based on the zones given in Figure 1. This was done by use of Small Area Statistics. The following information was requested, for both 0-800m and 800m-2km zones:

(i) usually resident population,
(ii) social class,
(iii) number of cars,
(iv) number of households owning none, one, two or three and more cars,
(v) the number of economically active residents,
(vi) usual mode for the journey to work. Provided by the Greater Manchester Research and Information Planning Unit (GMRIPU) although there were some problems with additional data,
(vii) the number of jobs and other employment characteristics within the destination station zones can only be determined from the Census of Employment at Ward level. GMRIPU did not make this information available,
(viii) tables that relate origins and destinations for the journey to work. Neither GMRIPU nor the Greater Manchester Transportation Unit (GMTU) had this information readily at hand.

The problems and delays we encountered in obtaining Census information was particularly disappointing given recent progress in the handling of geo-demographic data, through information technology applications and the development of geographic information systems.

2.3 Level of Service Variables

Information was required on the level of service offered by the three main modes that we wished to study.

For rail this information involved:

(i)the number of trains, each way, per weekday (peak and off peak). This was obtained from a full set of timetables supplied by GMPTE,
(ii)the journey time to central Manchester and other destinations. This again was obtained from the timetables,
(iii)estimates of the mean fare to Manchester and other stations. The PTE gave us information on the mean fare paid at each station, but this was not disaggregated by destination,
(iv)indicators of reliability, over-crowding, quality of rolling stock, etc. if these were believed to vary from line to line.

For bus this information involved:

(i)the number of buses (each way) in direct competition with rail. Appendix 2 lists the main bus services that are in direct competition with rail for journeys to central Manchester. Timetables were made available by GMPTE for all these services,
(ii)the journey time to central Manchester and other main destinations. This was obtained from the timetables,
(iii)estimates of mean fares to central Manchester and other destinations. This was obtained from fare tables supplied by GMPTE.

For car this information involved:

(i)estimates of the peak and off peak journey times to central Manchester and other destinations,
(ii)estimates of petrol costs for journeys to central Manchester and other destinations,
(iii)estimates of the mean parking charge in central Manchester and other main destinations.
No car related information was made available.

3. <u>CALIBRATION OF A TRIP END MODEL</u>

In this section, we shall describe how we calibrated a simple trip end model. This will be described in four stages with regards to the choice of explanatory variables, the choice of functional form, the analysis of residuals and statistical tests. Definitions of all variables presented in this paper are given in Appendix 3 in the order in which they appear.

3.1 <u>Choice of Explanatory Variables</u>

In our initial exploratory analysis, models were developed for six dependent variables:

(i)weekday boardings - all trips,
(ii)weekday peak boardings - all trips,
(iii) weekday off-peak boardings - all trips,
(iv) weekday boardings - central Manchester trips only,

(v) weekday peak boardings - central Manchester trips only,
(vi) weekday off peak boardings - central Manchester trips only.

(NB: central Manchester defined as Deansgate, Oxford Road, Piccadilly and Victoria stations).

For each dependent variable, there were 53 possible explanatory variables, of which 50 were variables based on Census data and the other three were rail journey time, fare and frequency.

Regression models were developed using the Statistical Analysis Package (SAS) (SAS Institute Inc., 1982). Use was made of the stepwise procedure's forward selection option which detects the independent variable with the highest F value. Once a variable is in a model, it stays. The procedure goes on, sequentially, to add more variables until no F-statistic has a value greater than a pre-set value (the default is 0.5). In our example, the degrees of freedom provided a more important constraint, ie. variables were added until there were no remaining degrees of freedom.

The ten most powerful explanatory variables appeared to be, in descending order (with the sign of the relationship in brackets):

(i) rail frequency (+),
(ii) number of households owning three or more cars 800m-2km (+),
(iii) number of people in social class II 800m-2km (+),
(iv) number of people who work at home 0-800m (+),
(v) number of people who travel to work by pedal cycle 800m-2km (-),
(vi) number of households owning two cars 800m-2km (+),
(vii) number of people who travel to work by car pool 800m-2km (-),
(viii) number of people who travel to work on foot 0-800m (+),
(ix) number of people in social class IIIM 0-800m (-),
(x) number of people in social class I 0-800m (+).

Although the signs of the relationships were not always easy to explain, this exploratory analysis gave a number of interesting insights. As might be expected, rail frequency was a powerful explanatory variable (though simultaneity problems complicate matters), with other variables showing rather surprising features. It appeared that rail demand was strongly related to groups with high car ownership and in social classes I and II (professional and managerial). These variables are likely to be correlated and may be viewed as a proxy for income. Rail demand appeared negatively related to social class IIIM (skilled manual) which might be related to location of workplace; skilled manual workers are generally employed in factories which have inner city and suburban locations which are poorly served by rail. The reverse is true of professional and managerial groups whose workplaces are concentrated in central areas (particularly in Manchester itself) and are well served by rail. The relationship with the journey to work variables was more difficult to explain and is examined later in this section.

Surprisingly, population did not appear to be an important explanatory variable and nor did rail fare or journey time; these variables are considered later in this section. It is also interesting to note that the 800m-2km variables appeared as powerful as the 0-800m variables.

In Table 1 some of these issues are examined in more detail, with all boardings on an average weekday being the dependent variable. From model 1 it can be seen that neither the 0-800m population (POP1) nor the 800m-2km population (POP2) parameter values are significant at the 10% level, whilst the POP1 parameter is the wrong sign. Some of these problems are overcome if the intercept is dropped (model 2) as the POP1 parameter value becomes significant. This model might be thought of as a simple trip rate model which estimates that there will be 46 boardings per thousand population within 800m of each station and 18 boardings per thousand population between 800m and 2km of each station. However, it should be noted that in instances where the intercept is dropped the R^2 is no longer comparable with that of a model with an intercept. In model 3 the effect of using a variable based on the population between 0 and 2km (POP = POP1 + POP2) was tested, but the parameter value was again found to be insignificant. If the intercept is dropped the POP parameter value becomes highly significant, implying 27 boardings per thousand population (model 4).

The parameter values for the numbers in social class I/II (SOC1) and IIIM (SOC3) are both significant at the 10% level in model 5, although SOC3 is not quite significant at the 5% level, whilst the intercept value is insignificant. However, we might expect SOC1 and SOC3 to be correlated with POP. Thus, in model 6 they are expressed as ratios with respect to POP. It can be seen that all four parameter values (including the intercept) are significant at the 10% level and the R^2 indicates that over 50% of all variation is explained by this model.

In model 7 the parameter value for the number of households owning two or more cars (CAR2) is highly significant, although the intercept value is not. The CAR2 parameter is expressed as a ratio in model 8, in this case with respect to the number of households (HHOLD). All three variables are significant at the 10% level and almost half of all variation is explained. A further variable (CPOP, defined as the number of cars 0-2km divided by POP) was tested, although the results are not shown in Table 1. The parameter value of this variable was negative but not significant. What this result suggests is that moving from owning no car to one car has a negative effect on rail usage, but this is largely counter balanced by the positive effects of moving from one to two or more cars. However, we postulate that this is due to income/social effects rather than car ownership per se.

Model 9 attempts to assess the effect of the number of people who are unlikely to switch to rail for travelling to work because they work locally, either at home or by travelling to work by foot or by pedal cycle (WLOCAL). However, the parameter value for this variable was insignificant. In model 10 we expressed this variable as a ratio with respect to the number of economically active residents (EACT). However, neither parameter value is significant, although the RLOCAL value is of the right sign (ie. if the proportion of people working locally increases, the numbers using rail decrease).

In Table 2 we examine the influence of level of service variables. Model 11 clearly illustrates the importance of frequency, with an average of over 9 people estimated as boarding each additional train. If the intercept is dropped (model 12), it can be seen that this figure reduces to less than 8.

In model 13, it appears that the effect of rail journey time (RJT) is insignificant and moreover is of an implausible sign. The greater the journey time (and hence distance from Manchester), according to this model, the greater the number of trips by rail. Model 14 investigates a polynomial function for journey time, which gives much better results, with the RJT and RJT2 parameter values being significant at the 5% level. Moreover, the signs of the coefficients are plausible. This model suggests that rail demand initially increases with journey time (and hence distance) from central Manchester, but then eventually decreases. This is probably because, for short distance trips, bus is very competitive with rail but this competitiveness decreases with distance from central Manchester. In later work we attempted to model this directly.

Fare effects are examined in model 15,in which the parameter value is negative, yet is insignificant,and model 16 where a new variablewas created by dividing fare by journey time (as a proxy for distance - in later work we include distance as a separate variable). This variable (FRJT) has a parameter value significant at the 10% level, although the goodness of fit (as measured by R^2) is relatively low.

Lastly, model 17 highlights the possible effect of the central Manchester stations to which each origin station has a direct rail link. It can be seen that the parameter value of this variable (CMS) is significant, although the intercept is not. It implies that a station with a service to Deansgate, Oxford Road and Piccadilly will attract, on average, 456 more boarding passengers than a station with a service to Victoria.

Given the findings of Tables 1 and 2, Table 3 summarises our initial development of a multiple regression trip end model. In model 18 a high R^2 is achieved (0.766) but the parameter values for RCAR2, RJT and RJT2 are insignificant, whilst the values for RLOCAL and CMS do not display the expected sign. However, it maybe that the RLOCAL parameter sign indicates a strong local employment base such that rail trips would be attracted. The variables RJT and RJT2 are dropped and replaced by FRJT in model 19. The values for the intercept, RCAR2 and FRJT are not significant at the 10% level (although the latter is significant at the 13% level) whilst RLOCAL and CMS are still of the wrong sign. The variable RCAR2 is dropped in model 20, as this is likely to be correlated to RSOC1. We also dropped CMS as we have strong reason to believe this should have a positive value as it reflects the degree of accessibility to Manchester city centre. It can be seen that, with the exception of the intercept, all variables in model 20 are significant at the 10% level and that although the R^2 is not as high as model 18, the value adjusted for degrees of freedom (R^2) is higher.

3.2 Choice of Functional Form

Given the findings in 3.1, we decided to examine model 20 in further detail. So far we have only considered (with one minor exception), linear additive functional forms. Model 20 is of this type and is reproduced in Table 4 with two models having alternative functional forms.

(i)a semi-log formulation, as represented by model 21 (this might alternatively be described as an exponential model);
(ii)a double-log formulation, as represented by model 22 (this might alternatively be described as a log-linear model).

In comparing the goodness of fit in models 21 and 22, it is sufficient to compare their R^2 values as they share the same dependent variable (LDEP). Model 21 has a small advantage over model 22 yet comparing these models with 20 is more problematical.

However, a log likelihood ratio test has been devised by Mills (1978) to overcome the problem and is as follows:

$$\text{LLR} = \frac{-N}{2} \text{ Log MSE} + (\lambda-1)(\text{LogY}).N$$

where:

LLR = Log Likelihood Ratio
N = Number of observations
MSE = Mean Square Error
λ = 1 if linear, 0 if log
LogY = Mean value of dependent variable

For model 20:

$$\text{LLR} = \frac{-36}{2} \text{ Log } (117128.4) = -210.1$$

For model 21:

$$\text{LLR} = \frac{-36}{2} \text{ Log } (0.260076) - 6.137 .36$$

$$= 24.2 - 220.9$$

$$= -196.7$$

For model 22:

$$\text{LLR} = \frac{-36}{2} \text{ Log } (0.267327) - 220.9$$

$$= -197.2$$

This test indicates that both models 21 and 22 are slightly superior to model 20 and again indicates that 21 has a better fit than 22.

3.3 Analysis of Residuals

In Table 5 (contained in a confidential annexe) we examine the implications of model 21 in more detail. This is done by examining the difference between the actual number of boarding passengers and the model's predicted number, which is called the residual. In absolute terms, the five largest residuals are for (in descending order):

(i) Sale (underestimates by 900 boardings),
(ii) Altrincham (overestimates by 837 boardings),
(iii) Rochdale (underestimates by 774 boardings),

(iv) Davenport (overestimates by 659 boardings),
(v) Radcliffe (underestimates by 407 boardings).

The first three are medium-sized free standing towns which might be thought to have characteristics different to those of the typical new station site. However, one would have to also include Bury and Oldham in this category. The overprediction of demand at Davenport may be due to incorrect specification of the catchment area; this site is likely to be dominated by nearby Stockport station which has a better level of rail service. This situation may also be replicated between Brooklands and Sale, Oldham Mumps and Werneth, Crumpsall and Bowker Vale and Radcliffe and Whitefield, although the last two pairs show no readily apparent difference in level of service.

It is disturbing to note that model 21 underpredicts demand at the three new stations included in this sample (Derker, Mills Hill and Smithy Bridge) by 39% (combined usage of 1203, combined prediction of 738). This error is particularly acute at Mills Hill.

As a result of the residual analysis, model 21 was re-calibrated with a data set that excluded the six stations that serve free-standing towns. The result of this re-calibration is model 23, shown in Table 6. The exclusion of these six observations led to a slight improvement in the overall goodness of fit, as theR^2 measure increasing from 0.719 to 0.724. Moreover, a z-test at the 5% level shows there are no statistically significant differences in the parameter values of models 21 and 23. This is encouraging as it indicates that our model form is relatively stable.

The residuals of model 23 are examined in Table 7 (again in the confidential annexe). The largest residuals (at Davenport, Radcliffe, Crumpsall and Bowker Vale) seem likely to be due to incorrect specification of catchment areas. Model 23 continues to underpredict usage at the three new stations by 38% (actual usage 1163, predicted usage 726) only a small improvement on model 21.

3.4 Statistical tests

Model 23 was tested for two statistical problems that commonly afflict regression models based on cross-sectional data, namely collinearity (correlation between the independent variables) and heteroscedasticity (non constant variance of the error term).

A collinearity problem is said to occur where a component associated with a high condition index contributes strongly to the variance of two or more variables. The collinearity diagnostics for model 23 are presented in Table 8 and, although there is some evidence that (from component 6) RSOC1 and RSOC3 are correlated, the correlation matrix presented in Table 9 suggests it is not severe enough to be a major concern.

Visual inspection of a scatterplot of the dependent variable (LDEP) and the residuals suggested that heteroscedasticity was not a significant problem. This was confirmed by a Park-Glesjer test, regressing the absolute values of the residuals against the independent variables, in which none of the parameter values were significant (indeed only two parameters had t-statistics with absolute values greater than 1) and the error term thus more likely to have a constant variance.

Model 21, explaining over threequarters of all variance in the logarithms, was used for further analysis having the best functional form and being based on the full data set. Model 23, whilst giving marginal improvement in terms of fit and predictive capability did not lend itself to further development being based on a smaller data set. All parameters in model 21 have the correct sign and are significant at the 10% level, with the exception of RSOC3 and RLOCAL. However, the parameter values of both these variables have absolute t-statistic values greater than 1, indicating that they do just contribute to the goodness of fit as measured byR^2 (ie. having taken into account the loss of two degrees of freedom).

Mean elasticities may be produced from model 21 by multiplying the parameter value by the mean value of its variable. From this model, it is estimated the elasticity of rail with respect to POP is 0.7, with respect to FRJT is -0.85 and with respect to FREQ is 1.1. However, these values are critically affected by the size of the mean variable. It is interesting, therefore, to comment on the values of the parameters in the Log linear model (model 22). The LPOP parameter implies that a 10% increase in population would lead to a 9.9% increase in demand ie. a constant population elasticity value close to unity. This LFREQ parameter implies a constant rail frequency elasticity of 0.685, whilst the LFRJT parameter implies a constant fare elasticity (with respect to journey time) of -0.411. All these values seem eminently plausible.

4. DEVELOPMENT OF THE MODEL

The semi-log model of section 3 (model 21) was used as a starting point for further model development. Two variables were however dropped from this model, FRJT, because the parameter sign gave implausible results on speeding or slowing services, and RLOCAL because it lacks an intuitive interpretation and in many cases is not statistically significant anyway. The data set was expanded by the addition of a number of new variables based on the injection of data from the following sources:

(i)The distance by rail (in kilometres) from each station to either Victoria or Piccadilly, derived from the BR timetable.

(ii)The number of buses (both ways) on a weekday travelling to/from the station catchment area to/from central Manchester, obtained from GMPTE supplied timetables. This was calculated for both peak (defined as 05.00 up to 09.30 and 15.00 to 18.00 hours) and off-peak periods.

(iii)The journey time (in minutes) by bus from the station catchment area to central Manchester, for both peaks and off-peak periods, obtained from timetables provided by GMPTE. These timetables did generally indicate that bus services were slower (journey times longer) in the peak period.

(iv)The journey time (in minutes) by car from the station catchment zone (as defined by GMTU) to Piccadilly Gardens. For the Bury and Altrincham lines this was based on 1986 run times. For the Oldham and Rochdale it was based on 1989 run times (see Appendix 2). Again a distinction has been made between the peak (defined in this instance as 08.00 - 09.00) and the off-peak (defined as 10.00 - 12.00 hours).

In addition, Davenport's 800m - 2km catchment population was adjusted to take into account an overlap with Cheadle Hulme station.

Using this enlarged and amended data set a number of new model versions were developed and these are presented in Table 10.

4.1 The Models

Model 24 of Table 10 splits POP into two variables; POP1 is the 0 - 800m population and POP2 the 800m - 2km population. A feature of our early runs in Table 10 was that the POP1 parameter value was never statistically significant and, in several instances, implausibly had a lower value than the POP2 parameter value.

Model 24 also included 4 of the 5 new variables listed in section 2; namely DIST, BFREQ, BJT and CJT. In addition the variable FRJT is replaced by RJT and FARE is not included as a variable. Leaving aside problems of collinearity for the moment, it can be seen that in addition to POP1, the values for RSOC3, RJT and BJT were not significant at the 10% level.

In models 25 and 26 DIST and RJT were examined to determine which of these 2 variables is statistically most significant. As model 25 clearly shows, the RJT variable can be dropped without affecting the R^2 measure (andR^2 increases from 0.728 to 0.734). However, the BJT parameter value remains insignificant.

In models 27 to 30 the effect of SPEED (defined as DIST divided by RJT) is examined. In all cases the parameter value is insignificant and, more importantly, of the wrong sign. This may be related to problems of simultaneity;well used commuter lines (such as Bury and Altrincham) have closely spaced stations and hence relatively slow services, whilst less well used lines (such as the Rochdale line) have more widely spaced stations and hence faster services. In model 27, the parameter values for BJT and CJT are both insignificant, and for BJT is of the wrong sign. In model 28 these variables were re-defined by dividing through by RJT to produce RBJT and RCJT but the same pattern of insignificance and wrong signs emerged. The same is true in model 29 where the variables were divided through by DIST to produce BSPEED and CSPEED. Lastly, in model 30 a new variable, BGT, was tested which attempts to combine BFREQ and BJT into a generalised time measure. This is done by estimating the mean headway of bus services. Waiting time for frequent services with random passenger arrivals will be half this headway. However, waiting time is typically valued at double the value of in-vehicle time, thus BGT was estimated (in minutes) as in-vehicle time (BJT) plus the service interval. This amendment failed to produce a parameter of bus competition that was significant or of the right sign.

In model 31, a generalised time measure was developed for rail (RGT) in a similar manner to that for bus. Although the parameter value for this variable had the correct sign, it was not quite significant at the 10% level. However, it is interesting to note that in both models 30 and 31 the parameter value for CJT (a variable CGT could not be developed as we had no measure of out-of-vehicle time) was of the right sign and significant.

At this stage, at GMPTE's insistance, the models developed had not included a rail fare variable. This was changed by model 32, in which the FARE parameter value is significant and of the right sign. Bus competition is reflected by the BFREQ variable and car competition by the CJT variable. The reason that BJT (or its variants) has not proved to be a plausible variable may be due to the fact that real bus journey times are quite different from those published in timetables. At first glance, model 32 provided a good representation of the data in that all parameters were significant at the 10% level (except RSOC3, which is just insignificant) and of the right sign. Model 33 includes the additional DIST variable and R^2 increased noticeably (up from 0.717 to 0.774) but the POP1 parameter value became insignificant and smaller than the value for POP2. In model 34, the mean rail frequency elasticity is estimated to be around 1.1 and, because of the semi-log formulation, is estimated to increase as frequency increases. In most situations, this is counter-intuitive since it is at low frequencies that we might expect rail demand to be most elastic and this high frequency elasticity may be largely due to simultaneity problems.

In model 35, POP1 and POP2 were re-combined to form the POP variable. Rail frequency was entered by taking natural logarithms (LFREQ). A constant elasticity of around 0.8 is implied. In this model all parameter values were significant and of the right sign, whilst the R^2 measure indicated that over 80% of variation in the logarithms was explained. However, the model is plagued by statistical problems and in particular (and not surprisingly) the strong correlation between FARE, DIST and CJT. Moreover, a Park-Glesjer test indicated that the BFREQ variable is heteroscedastic. Model 36 has FARE divided by DIST, to produce FDIST, and CJT replaced by CSPEED. This removed most of the problems related to multicollinearity and reduced heteroscedasticity. Moreover, compared to model 35, it lead to a marked improvement in goodness of fit, with R^2 increasing from 0.752 to 0.841. However, residual analysis indicated poor model performance in some respects with usage at Altrincham still grossly over-estimated, in absolute terms.

A possible cause of this might be the sensitivity of the model to RSOC1, which is the proportion of the population in social classes 1 and 11 (professional and managerial). It should be noted though that SOC1 was from the 10% sample, whilst POP was from a 100% sample. However, if separate measures are developed for social class 1 (RSOC11) and social class ll (RSOC12), as in model 37, very different values emerge. The parameter value for RSOC11 is neither statistically significant nor of the expected sign but that for RSOC12 is both significant and of the expected sign (and highly positive). It is the professional rather than the managerial group which is the main user of rail services, which might be expected "a priori". In model 37 the Altrincham residual reduced to -219 (still an overestimate). However, if, as in model 38, RSOC11 is dropped, this residual increases again to -752. Despite this, the model is well specified and explained almost 88% of logarithmic variation.

As a result model 38 was re-estimated for peak and off-peak trips to give models 39 and 40 respectively. The parameter values for these models were compared with those of model 38 by means of a z-test, the results of which are also shown in Table 10. For the peak period model, three parameter values were significantly different (at the 5% level); RSOC3, FREQ and BFREQ. It is evident that people in social class 111M are even less likely to use rail in the peak than they are for the day as a whole. Peak users are also more sensitive to rail and bus frequency. For the off-peak model, three parameter values were again insignificantly different; this time they were POP, FREQ, CSPEED. The POP value indicated that, as expected, the propensity to travel in the off-peak is reduced. Somewhat perversely, the FREQ variable indicates that demand is more sensitive to frequency than for the all day model, but less so than the peak. This somewhat strange result indicates that multicollinearity may still be having some effects. By contrast, CSPEED indicated demand to be less sensitive to car speeds during the off-peak.

Model 41 was based on trips to Manchester only and the RSOC3 variable was no longer found to be significant and was therefore dropped. Compared to model 38 theR^2 measure fell from 0.846 to 0.736 and the BFREQ parameter value was not quite significant at the 10% level, but a statistically significant distinction can be made between POP1 and POP2. However, it should be noted that the fare variable included here was MFARE, which was divided by DIST to produce MFDIST (this variable had a mean value of 9.88, implying a mean fare elasticity in excess of -2). Nonetheless, the model did provide some plausible results and is examined in more detail in the next section.

Model 41 was re-estimated for peak and off-peak trips and labelled 42 and 43 respectively. For model 42, four parameter values were significantly different at the 5% level; namely, POP2, FREQ, BFREQ and CSPEED. Demand in the peak period is more sensitive to rail frequency, bus frequency and car speed than during the day as a whole, and likewise the propensity to travel by rail from the 800m to 2km zone is greater during the peak period.

For model 43, 5 parameter values are signficantly different from the values for model 41. These are the intercept, RSOC12, FREQ, BFREQ and CSPEED. Demand from social class 11 is much less in the off-peak period than for the day as a whole. Once again, perversely, demand in the off-peak is more sensitive to rail frequency but less sensitive to bus frequency (indeed the value for this parameter becomes positive) and car speed.

4.2 Residual Analysis

In this section model 38 (for all trips) and model 41 (for Manchester trips only) are examined in further detail with particular attention given to the predictions inferred (see Tables 11 and 12 contained in the confidential annexe).

For model 38, the 5 largest (in absolute terms) residuals are for:

(i) Sale (underestimates by 853 boardings),
(ii) Altrincham (overestimates by 752 boardings),
(iii) Brooklands (overestimates by 699 boardings),
(iv)Timperley (overestimates by 579 boardings),
(v) Crumpsall (underestimates by 360 boardings).

These indicate the model poorly predicts demand for stations on the outer section of the Altrincham line (although, because these stations have high demand, the performance in relative terms is not so bad). In particular, the model fails to pick up the rail heading that seems to occur at Sale station. However, the 3 new stations opened on the Oldham Loop and Rochdale Line fare better with total predicted demand being 1018 compared to actual demand of 1203. Most of this error is due to Mills Hill, with a 372 prediction compared to 551 actual boarders, an underestimate of 32%.

Two statistical measures have been computed to assess model accuracy; a Root Mean Square Error (RMSE) measure and an Absolute Deviation (AD) measure (for definitions, see Preston, 1987). For model 38 the RMSE measure indicates that, on average, forecasts are within 276 of actual values and the AD measure that, on average, forecasts are within 25% of actual values (this is a marked improvement on the West Yorkshire figure of around 40%).

However, the model implies high elasticity values; for example, a mean rail frequency elasticity of 1.22 and a mean rail fare elasticity of -1.01. In addition, the model implies a rail cross elasticity with respect to bus frequency of -0.28 and a cross elasticity with respect to car speed of -2.25.

For model 41, the 5 largest (in absolute terms) residuals are:

(i) Sale (underestimates by 464 boardings),
(ii)Bury (underestimates by 408 boardings),
(iii)Brooklands (overestimates by 360 boardings),
(iv)Rochdale (underestimates by 319 boardings),
(v)Crumpsall (underestimates by 249 boardings).

The dependent variable here is trips to Manchester which are estimated as accounting for 51% of all trips. Thus, although Manchester is clearly the dominant destination, there does appear to be substantial travel elsewhere. However, it should be noted that the lines we have studied have important secondary destinations (e.g. Oldham, Rochdale, Bury, Stockport, Altrincham) whilst on other lines (e.g. Styal, Hadfield/Glossop, Rose Hill/Marple) Manchester might be expected to be more dominant. In addition trips to many destinations involve changing trains, and possibly stations, in central Manchester.

Analysis of the residuals suggest the model may have problems accommodating rail heading, although it should be noted that usage at Altrincham is overestimated. A possible amendment would be to exclude stations serving main towns (e.g. Oldham (Mumps and Werneth), Rochdale, Bury, Altrincham, Sale) for two reasons. Firstly, they are likely to be contributing to the simultaneity problems that result in a high frequency elasticity. Secondly, it is at these stations that railheading occurs. As population within 2 kilometres only is included in the models, this is likely to lead to underpredictions. However, exclusion of these major stations did not lead to significant improvements to the models.

A more correctable feature is the consistent underpredicted usage of the Bury Line. This may be due to the very good reliability record of this line which might be accounted for by a reliability dummy variable that takes the value 1 if the station is on the Bury line and 0 otherwise (in lieu of any better data on reliability). This was done in model 44.

Model 41 more accurately predicts usage for the three new stations. Predicted boardings to Manchester are 607, actual boardings 519, with most of the error now being due to Smithy Bridge. The Mills Hill forecast is accurate but the data suggests that only 35% of demand of this station is to Manchester which casts some doubt on the accuracy of the Manchester flow data. As would be expected from the goodness of fit measures, model 41 is relatively less accurate than model 38. The RMSE and AD measures indicate that forecasts are, on average, within 160 and ±30% (i.e. more in line with the accuracy of the West Yorkshire model) of the actual values.

5.A MORE GENERALISED MODEL OF RAIL DEMAND

The approach adopted so far is exploratory in nature and has been primarily driven by ad-hoc attempts to improve the model formulation and goodness of fit. This can have undesirable consequences. For example redefining variables may lead to other variables picking up their effects e.g. variable CSPEED has a high value in model 38 perhaps because it is picking up the effects of DIST and as a result is probably biased upwards. In this section we attempt to develop a more generalised framework, which will allow redefinitions on a more consistent approach.

5.1 Developing a framework

In line with work previously carried out by ITS in West Yorkshire this development is based on a log-linear form, implying constant elasticities, consistent with most variable definitions, and the additional attraction of a possible reduction in multicollinearity.

For example, a model

$$T = a\,(P1)^{b}\,(P2)^{c}\,(GCR)^{d}\,(GCB)^{e} \qquad (1)$$

can be re-expressed as:

$$T = a\,(P1)^{b'}\,(P2/P1)^{c'}\,(GCR)^{d'}\,(GCB/GCR)^{e'} \qquad (2)$$

where

T = Number of trips
P1 = Population variable 1
P2 = Population variable 2
GCR = Generalised Cost of Rail
GCB = Generalised Cost of Bus

a, b, c, d, e, a', b', c', d', e' = parameters

For example, in equation 1 rail's generalised cost elasticity is d. In equation 2 it is d' - e' (it is assumed that e' > 0).

A model of the form of equation 2 has been developed for all trips and trips to central Manchester only, models 45 and 46 respectively in Table 10. Three new variables have been introduced:

(i)GCR. This is the generalised cost (in pence) of travel by rail. This consists of in-vehicle time (RJT), wait time (estimated as a function of FREQ using a relationship used in West Yorkshire) and MFARE. The value of in-vehicle time was based on the Department of Transport's recommended values (DTp, 1987, Table 2A) and estimated to be 5.0 pence per minute at mid 1988 prices. Wait times value was estimated to be twice that of in-vehicle time.

(ii)GCB. This is the generalised cost (in pence) of travel by bus. This consists of in-vehicle time (BJT), wait time (estimated as a function of BFREQ using a relationship used in West Yorkshire) and MFARE. The value of in-vehicle time was again based on the Department of Transport's recommended values (this time being 2.9 pence per minute at 1988 prices) and the value of wait time assumed to be double that value.

(iii)GCC. This is the generalised cost (in pence) of travel by car. This consists of in-vehicle time (CJT) and cost. The cost of travelling by car was found by dividing DIST by the average fuel consumption in urban driving of a family saloon multiplied by the price of petrol. An amount was also added for parking costs. These calculations were carried out so as to be consistent with previous work we have carried out in West Yorkshire (Preston, 1987). The value of in-vehicle time was estimated as 4.4 pence per minute at mid 1988 prices.

It should be clear that, out of necessity, our generalised cost measures are rather crude and hence the results in this section should only be treated as illustrative. Moreover, it should be evident that our generalised cost expressions are incomplete. We have been unable to measure the differences in access and egress time for the three modes we are considering. This is likely to be an important barrier to estimating the relationship between bus and rail as access/egress time is likely to be a prime influence on choice. In addition, the generalised cost measure fails to take into account attributes such as comfort and reliability.

5.2 Model development

In model 45, both GCB and GCC are divided by GCR in order to produce the variables GCBR and GCCR respectively. In order to produce a log linear formulation, the natural logarithms of all variables are taken. It is evident that this model represents a better model than those based on generalised time (e.g. models 30 and 31) but, in terms of goodness of fit, represents a deterioration compared to model 38.

In particular, the intercept and LGCBR parameters are insignificant at the 10% level, whilst the cross elasticity implied by LGCCR seems to be implausibly large, suggesting that if car generalised cost increases by 10%, rail usage increases by almost 18%. By contrast, a corresponding increase in bus generalised cost is only estimated to result in a 1% increase in rail usage. However, the direct elasticity measure for rail generalised cost is more plausible (-2.259). Given the breakdown of mean RGC, this implies a fares elasticity of -0.66, a journey time elasticity of -0.77 and a wait time elasticity of -0.82 (and because of the headway function used a frequency elasticity of only -0.52). These estimates of rail elasticities are more in line with previous studies (including those by BR) than those elasticities derived from earlier models in this work.

However, model 45 has statistical problems of its own as a Park-Glesjer test indicated both the LPOP and LGCCR variables are affected by heteroscedasticity, and attempts to remove this problem, through the use of weighted least squares, failed to produce sensible results. Moreover, as Table 14 (contained in the confidential annexe) indicates, a number of very large residuals result from the model, and in particular at Sale and Davenport. The RMSE measure is 339 and the AD measure 0.343, a marked deterioration compared to model 38. Attempts to improve on model 45 by entering the rail generalised cost variables separately failed to prove successful.

The generalised model of rail demand to Manchester only (model 46) is much less successful. Not only are the values for the intercept and LGCBR parameter insignificant, but so are the values for LGCR and LRSOC3. Moreover, the value of LGCBR is of the wrong sign. As a result, both the implied direct and cross elasticities are implausible.

6. WALK MODE ACCESS RAIL DEMAND

Greater Manchester new stations are predominantly targeted at patrons who walk to the station and so we amended the patronage counts accordingly to study this demand. Appendix 5 shows the contribution that patrons with a walk mode access make to total rail demand. Overall, 67% of users walk to the station but this is partly due to the effect of stations such as Bury, Oldham Mumps and Rochdale. Excluding the stations that serve the five largest towns increases the figure to 74%.

To facilitate our analysis, the quality of the data set was improved further by removing three stations that had caused particular concern. Derker was dropped because 1981 Census data was no longer considered appropriate given changes to the catchment area around the station. The removal of Davenport and Hazel Grove partially eliminated data incompatibility problems commented on previously. All fare related variables were also removed.

6.1 Developing a basic model for walk mode access patronage

Three journey time related and eight population based variables were modelled together with rail and bus frequency variables and a distance measure. However, collinearity problems involving the distance and the three journey time variables led us to replace these variables by a car journey time and rail speed variable. We concentrated on population measures in the 800m-2km catchment and our initial results are shown in Table 17. In particular, it should be noted that the rail speed variable is of the wrong sign (and this feature re-occurred whenever this variable was included in the model).

Our preliminary searches, highlighted the number of people 800m-2km (POP2), percentage of people in social classes I or II (managerial and professional) within 800m-2km of the origin station (FSOC2), and rail frequency (FREQ), to be the three most significant variables. The 800m-2km catchment variables were more significant than 0-800m catchment ones and although we used the above three variables as our foundation for model development we also researched a trial model based around the variables FREQ, the number of people 0-800m (POP 1) and the percentage of people in social classes I or II 0-800m of the origin station (NSOC2). At this stage the FREQ-POP2-FSOC2 model gave anR^2 value of 74% and all four parameter values (an intercept was included) were of the right sign and significant at the 90% level. The FREQ-POP1-NSOC2 model fared considerably worse with anR^2 value of 59% and the POP1 variable insignificant at the 90% level. We nevertheless persevered with this model for comparative purposes. To these models we added and eliminated further variables (both singularly and in combinations) and searched for the best functional form in each case, logarithm and inverse functions being adopted where considered appropriate. This approach led to the following three models each including an intercept.

(i) Model 48

This model is of an exponential form (only the dependent variable is logged), and contains the variables FREQ, POP1, POP2, FSOC2 and previously undefined variables BFREQSQ (the square of bus frequency) and FSOCIIIN (the percentage of people in social class IIIN 800m-2km).

(ii) Model 49

An exponential model with variables FREQ, POP1, BFREQSQ and CJT (car journey time).

(iii) Model 50

This model is loglinear (both dependent and independent variables are in logarithmic form). The explanatory variables are FREQ, POP1, CJT and NSOC2.

Table 18 compares the three models for each of the boarding periods and destinations. All estimated parameters are significant at the 90% level (unless stated otherwise) and the sign of the estimated parameters are correct where the relationship between explanatory and dependent variable is known. The table entries for models 48 and 49 indicate that the trend is for improved goodness of fit in the off peak, peak, all trips order of ascendancy for `all destination' trips, though whether it is the nature of each particular demand or the rising patronage count that gives rise to this is unclear. However, for central Manchester bound journeys this trend breaks down with peak boardings showing an improved fit relative to the aggregated boardings. Model 50 table entries show the reverse situation with an apparent trend for central Manchester bound trips and the absence of any pattern for `all destination' trips.

Peak trips comprise approximately 64% of the demand, irrespective of trip destination, and trips made to central Manchester comprise 44% of the total daily patronage although it should be noted that 10% of trips on the Altrincham line were categorised as having the destination `Other'. A more refined definition of trips with destinations classed as `Other' is dependent on the survey technique and classification method employed and we cannot comment further on their composition. Appendix 6 shows the percentage of central Manchester destined trips from each station.

The variables rail speed and distance from central Manchester were added to the model and substituted for each variable at each stage of analysis and at no time did an improved fitted model result. Indeed the speed parameter estimate was negative on numerous occasions. Speeds, frequencies and journey time for all forms of transport were tested along with population magnitude and social class percentages. The best fitting model for each variable could be provided yet we feel the limited value of these models precludes a detailed analysis.

The poor fit of each model for off peak period boardings is noted, with central Manchester bound trips poorly reproduced by model 49 in particular.

6.2 Walk mode access: towards a best fitting model

Model 48 produces the highestR^2 values and at the same time comprises a large number of explanatory variables relative to the two other models under consideration. Parameter estimates and their associated t-statistics are shown in Tables 19, 20 and 21. The estimates for peak period demand are similar to those for all period demand for both central Manchester bound and all destination travel. However there is a marked change for the off peak period with both bus and rail frequency parameters increasing in magnitude, and population within the 0-800m catchment area becoming more significant to the detriment of the social class percentages further from the station. The implication here is that off peak period demand is very time sensitive with frequency and access time becoming more dominant.

Whilst the population variable for the 800m-2km distance is more significant than the 0-800m population variable in all cases, its parameter estimate is of a smaller magnitude. All other things being equal our results suggest people living within 800m of the station make 66% more trips than residents in the 800m to 2km catchment area. Rail frequency is the most significant variable for all periods and destinations. Other than the higher R^2 values for the all destination trips in comparison with the central Manchester destined ones, the only point to note is the higher frequency parameter estimates for the former. It is interesting to note the goodness of fit pattern for central Manchester destined trips. The lower R^2 value for off peak trips indicate a poorer fit and their inclusion has a detrimental effect on the fit of the all trips data.

The mean rail frequency elasticity implied by Table 19 is 1.3, much higher than that which we would expect and suggests our models are affected by simultaneity. This figure is perhaps best interpreted as being confounded with a feedback effect so that the impact of policy change on frequency would be expected to be much lower.

Table 21 summarises our analysis of non-central Manchester bound trips. Unlike Tables 19 and 20 some parameter estimates are highly insignificant, although the overall goodness of fit is generally good. The variable BFREQSQ is derived from bus frequency data for central Manchester destined trips and is therefore not particularly relevant to the modelling of trips not destined for central Manchester. Dropping it from the above model results in the POP1 variable becoming even less significant and the R^2 value falling.

The predictive capability of model 48 is examined by comparing the differences between the actual numbers of alighting passengers at each station, as given by the dependent variable, and the anticipated patronage as generated by the model. These residuals are summarised in Table 21 contained in the confidential annexe. In descending order the largest five residuals, in absolute terms are:

(i)Crumpsall (underpredicts by 521 boardings)
(ii)Altrincham (underpredicts by 334 boardings)
(iii)Bury (underpredicts by 270 boardings)
(iv)Prestwich (underpredicts by 258 boardings)
(v)Stretford (overpredicts by 243 boardings).

Three of these stations are on the Bury line, the other two on the Altrincham line. Now that Derker has been omitted from our data set, only two new stations remain, Smithy Bridge and Mills Hill, and in comparison with our earlier work the performance of the model has improved dramatically with a 14% underestimate of actual usage predicted by our model. The two statistical measures, RMSE and AD, indicate that, on average, forecasts are within 153 and 23% of actual values respectively.

A brief survey of the residuals' data reveals that the Oldham loop rail demand is particularly well predicted with usage of the Bury line stations tending to be underestimated and Altrincham line overestimated. Usage of the two new stations has been particularly well predicted as 86% of the actual patronage yet both are on the group of services we have termed the Oldham Loop and hence bias the new station demand predictive capability of the model.

Table 23 shows that no two variables are correlated to the extent that we should be concerned with collinearity problems. Visual inspection of a scatter plot of the dependent variable and the residuals suggest heteroscedasticity is not a significant problem either, confirmed by a Park Glesjer test in which the absolute residual was regressed against the independent variables.

6.3 Walk mode access: redefining our population variables

Given the low significance of POP1 (only significant at the 5% level in two of the nine models in tables 19, 20 and 21), we adopted a more aggregated set of population variables by combining the 0-800m and 800m-2km catchment areas. Model 38 from section 4 was considered to have a favourable composition with the variables CJT and DIST prominant. A negative correlation exists between the POP1 variable and the CJT and DIST variables and we have focused attention on the former which has prevented the introduction of the latter into a significant model. However the use of the POP variable reduced the correlation allowing the inclusion of all three variables in model 51 shown in Table 24.

The Bury line provided four of the five largest residuals, in absolute terms for this model. The usage of Mills Hill and Smithy Bridge, the two new stations, have been underpredicted by 12%. On average forecasts are within 204 of actual values, as indicated by the RMSE measure, and within 31% of actual values, using the AD measure. This model does however suffer from high collinearity between the distance and car journey time variables, the correlation coefficient being 0.970.

Model 52 gives the highestR^2 value for walk mode access passenger rail demand. In this model, shown in Table 25, the main point of note is the significant contribution made by the percentage of population residing in the 0-2km catchment who are classed as unskilled workers. This variable was the most significant of all the social class defined variables tested. In particular the social classes represented by FSOC2 and FSOCIIIN became more insignificant with the catchment area expanded to cover 0-2km.

The differences between the actual numbers of alighting passengers at each station, and the model 52 generated patronage are shown in Table 26 (contained in the confidential annexe). In descending order the largest five residuals, in absolute terms are:

(i) Altrincham (underpredicts by 518 boardings),
(ii) Crumpsall (underpredicts by 489 boardings),
(iii) Brooklands (overpredicts by 289 boardings),
(iv) Stretford (overpredicts by 237 boardings),
(v)Navigation Road (overpredicts by 221 boardings).

The Oldham Loop demand is fairly well predicted yet the Altrincham Line provides four of the five stations listed above. This line is better predicted by model 48 which incorporates the social class I and II percentages whereas this model features a variable representing the less affluent populace. The implication is that the stations that comprise the Altrincham Line have rail demand best modelled with variables more suited to its more affluent population composition, furthering the arguments, for increasing homogeneity by developing a more disaggregate modelling approach.

New station usage has been predicted at 97% of the actual patronage, though the Oldham Loop stations have again been well predicted and bias the model capability. The RMSE measure indicates that, on average, forecasts are within 169 of actual values and the AD measure that they are within 25% on average.

The variables that comprise model 52 are not found to be correlated to an extent that concerns us and application of a Park Glesjer test substantiates our belief that the model is free from heterostedastic problems, borne out through the inspection of a scatter plot of the dependent variable and residuals.

6.4 Combining demand specific models

To increase the homogeneity of trip destination we developed model 53 to be a combination of a central Manchester destined trip model and the model giving the best predictions for all other trips.

For the central Manchester destined trips we used a model based around the car journey time and distance variables and included a social class variable despite its slight insignificance at the 90% level of confidence. The model chosen, shown in table 27, compares slightly unfavourably in terms of fit (as measured by theR^2 figure), on comparison with model 48, yet was adopted due to the favourable explanatory variables. The model used for trip destinations other than central Manchester was the best fitting model developed in our searches.

The predictive capability of model 53, indicated by the comparison of actual and model generated patronages, is shown in table 28 within the confidential annexe. In descending order the largest five residuals, in absolute terms are:

(i) Crumpsall (underpredicts by 495 boardings),
(ii) Altrincham (underpredicts by 431 boardings),
(iii) Brooklands (overpredicts by 311 boardings),
(iv) Prestwich (underpredicts by 264 boardings),
(v) Bury (underpredicts by 260 boardings).

The same batch of stations continue to be the most poorly predicted though the greatest absolute error is marginally reduced from previous residual analysis. Of the five stations highlighted two are on the Altrincham line and the other three on the Bury line. The RMSE measure indicates that on average forecasts are within 163 of actual values. The AD measure shows the forecasts to be within 24% of actual values on average. As with our best `trip end' model (model 52) new station usage has been predicted at 97% of the actual patronage and, like model 51, the joint existance of the car journey time and distance explanatory variables introduce a collinearity problem.

Whilst we feel that the general development of such a model may lead to slight improvements in fit, in this case we are precluded by the continuing failure to accurately predict central Manchester destined journeys.

7 CONCLUSION

The final choice of model is necessarily judgemental. For all access mode demand the generalised model 45 provided the best theoretical framework and the most realistic estimates of rail elasticity. However, it performed less well in terms of predictive accuracy. Models 38 and 41 both perform better in this respect. It has, though, become clear that our explanation of total boardings is currently better than for boardings to Manchester only. This may, in part, reflect measurement error in the dependent variable. It was also found that modelling peak and off-peak periods separately did not lead to significant improvements.

Model 38 may have some features that are unattractive from GMPTE's point of view, the specification of the FREQ variable may be unrealistic, whilst for predictive reasons the use of the average FARE variable is not ideal. Amendments include replacing FREQ with LFREQ, or its reciprocal IFREQ, and replacing FDIST with MFDIST or DIST. However, Table 13 shows that these changes lead to a marked reduction in goodness of fit, as measured by R^2. If there is concern about the specification of FDIST and the fact that we can not identify greater propensity to travel from within 800m of the station, model 41 may be considered superior. The goodness of fit of this model is improved if a dummy variable for the Bury line (BDV) is included (R^2 up from 0.736 to 0.780 - see model 44 in table 10). Amendments to the FREQ function fail to improve on this; LFREQ causes R^2 to decrease to 0.717, IFREQ causes a decrease to 0.644.

Our preference, therefore would be based on a model of the form of model 38 but adjusted so as to reduce statistical problems related to simultaneity, multicollinearity and heteroscedasticity. A number of further amendments were tested:

(i) Exclusion of the six major stations (Altrincham, Bury, Oldham, Werneth and Mumps, Rochdale and Sale) from the data set. Only 30 observations remain. This should reduce simultaneity.
(ii) Experiment with different functional forms. It may only be sensible for certain variables to have elasticities that increase with the size of the variable. Other variables may more sensibly have constant elasticities. Moreover, taking logs will help reduce collinearity.
(iii) Develop a new dependent variable LTR (Log Trip Rate) defined as the log of (DEP/POP). This should reduce problems of heteroscedasticity and gain one degree of freedom.
(iv) Replace FARE by MFARE.

These amendments resulted in model 47 which is shown in Table 15. All parameter values are significant at the 5% level, with the exception of FDIST. Moreover, the implied elasticities are believable. In particular, both rail fare and frequency elasticities are around 0.8 (in absolute terms). The cross elasticities are relatively high (possibly reflecting specification error) but less so than in earlier model runs.

The R^2 measure indicates that the model has a good degree of fit, although Table 16, in the confidential annexe, illustrates that some large residuals still occur, particularly at Brooklands and Timperley. The RMSE measure for this model is 264 and the AD measure is 0.307. It is not possible to provide confidence intervals that have a firm statistical basis but we would recommend at the 95% confidence level +50% for individual stations and ±20% for a package of new stations. In order to be applied, the model would require the following information to be collected.

(i) POP The population within 2 kilometres of the station, adjusted for overlapping catchment areas.

(ii) SOC12The 10% population within 2 kilometres in social class II.
(iii)SOC3The 10% population within 2 kilometres in social class IIIM.
(iv)FREQThe number of trains (each way) to/from central Manchester.
(v) MFDISTAdult cash fare to central Manchester (in pence) divided by distance (in kilometres).
(vi)BFREQThe number of buses (each way) to/from central Manchester.
(vii) CSPEEDDistance to central Manchester divided by car journey time (defined as mean of peak and off-peak journey times) to give car speeds in kilometres per minute.

It is believed that a model of this type would provide reasonable forecasts of usage of the type of stations that Greater Manchester PTE are considering opening in the near future. For example, model 47's forecasts of usage at the three new stations in the data set are, on average, within ±25%.

For walk access patrons our best model, in terms of highest R^2 value, was model 52 and it resulted from combining the two catchment areas into one. This model highlighted the negative relationship between unskilled workers and walk mode access rail demand. The R^2 values for these rail demand models were lower than those of our all access mode ones, yet the implication that walk access mode rail demand is more difficult to model should be interpreted cautiously. Indeed, demand at two new stations, Mills Hill and Smithy Bridge, was well predicted, each having a higher than average percentage of users walking to the station.

Catchment analysis suggested that whilst total population in the 0-800m area was a significant variable it could not be broken down by social class, yet the 800m-2km catchment fared more favourably with three variables in our first walk access model (model 48).

The Oldham Loop patronage continually predicted more accurately than Bury or Altrincham Line usage and we experimented with two sets of dummy variables to identify possible Line related factors. The first related to the Line upon which the trip was undertaken and the second separated trips depending on the central Manchester station from which the Line originated, namely Victoria or Piccadilly. Neither of these variables revealed anything of interest with regards to railway demand.

Unfortunately data limitations preclude a more detailed comparison of trip destination other than those to central Manchester. In particular there is a lack of a suitable measure for the relative attractiveness of a location as a business, shopping and recreational centre. This applies to both central Manchester and other important population centres. Some of these towns (eg Bolton, Stockport, Wigan) are not in our data set and thus prevent detailed destination analysis. This may explain our relatively low R^2 measures from fitting central Manchester destined trips and it is our view that further improvements to the models presented would depend on such analysis being undertaken, as compared to the total demand models all central Manchester destined ones fared worse in all cases.

Throughout our work we have been continually beset by statistical problems inherant in aggregate cross-sectional data analysis. One major problem referred to throughout this paper, and surfacing in many guises, is simultaneity. This is the chicken and egg problem in that we are unable to correctly identify demand-side and supply-side effects. For example, is there a high level of rail frequency on the Bury-Altrincham lines because there is a high level of demand or is there a high level of demand because there is a high level of rail frequency? Possible solutions to the simultaneity might include:

(i)the development of a two stage least squares model of demand and supply,
(ii)the development of models from time series or pooled data,
(iii)the dilution of the simultaneity by making frequency part of a generalised cost formula that incorporates in-vehicle time, fare etc.

Solutions (i) and (ii) might be considered beyond the scope of this study because they require additional information with respect to variables that explain variations in supply, whilst solution (iii) has been investigated in earlier work.

The other two main problems, multicollinearity and heteroscedasticity, have been tested for throughout, and appropriate alterations made where discovered at significant levels. Such adjustments were of a simplistic nature and mainly took the form of variable re-definitions and, as above, we considered the neccessary course of action beyond the scope of this study.

REFERENCES

Department of Transport (1987) "Values for Journey Time Savings and Accident Prevention". March

Mills, T. (1978) "The Functional Form of the UK Demand for Money". Journal of the Royal Statistical Society. Series C, 27, 1, 52-57

Moss, P. (1988) "New Stations. Patronage Predictions for Future Proposals". GMPTE Discussion Paper. June

Preston, J. (1987) "The Evaluation of New Local Rail Stations in West Yorkshire". PhD Thesis. School of Economic Studies, University of Leeds

SAS Institute Inc. (1982) "Statistical Analysis System User's Guide: Statistics". Cary, North Carolina

Table 1: Examination of Explanatory Variables

Weekday boarding - all trips - with t-statistics in brackets

Population variables

	1	2	3	4
Intercept	789.6 (2.12)	-	484.6 (1.97	-
Pop 0-800m (POP1)	-0.043 (-0.89)	0.046 (1.73)	-	-
Pop 800m-2km (POP2)	0.016 (1.33)	0.018 (1.44)	-	-
Pop 0-2km (POP)	-	-	0.009 (0.87)	0.027 (6.10)
R^2	0.056	0.5227	0.022	0.515

Social Class variables

	5		6
Intercept	248.0 (1.28)	Intercept	581.1 (1.87)
SOC1	1.701 (5.11)	RSOC1	32869.6 (5.88)
SOC3	-0.518 (-1.94)	RSOC3	-44167.1 (-4.57)
		POP	0.026 (3.39)
R^2	0.442	R^2	0.563

Car Ownership variables

	7		8
Intercept	80.0 (0.51)	Intercept	-713.5 (2.42)
CAR2	0.629 (4.50)	RCAR2	6892.5 (5.07)
		HHOLD	0.066 (3.01)
R^2	0.373	R^2	0.457

Journey to Work variables

	9		10
Intercept	569.1 (2.68)	Intercept	941.2 (2.02)
WLOCAL	0.507 (0.594)	RLOCAL	-2907.8 (-1.26)
		EACT	0.281 (1.26)

R^2	0.010		R^2	0.071	

Table 2: Explanatory Variables: Level of Service Attributes

The Frequency variable

	11			12	
Intercept	-194.4	(-0.934)		-	
FREQ	9.297	(4.569)	FREQ	7.550	(9.424)
R^2	0.381		R^2	0.717	

The Journey Time variable

	13			14	
Intercept	588.9	(2.264)	Intercept	-328.2	(-0.681)
RJT	5.862	(0.379)	RJT	145.5	(2.249)
			RJT2	-4.370	(2.216)
R^2	0.004		R^2	0.133	

The Fare variable

	15			16	
Intercept	1067.7	(3.051)	Intercept	1075.8	(4.353)
FARE	-5.881	(-1.161)	FRJT	-80.373	(-1.754)
R^2	0.038		R^2	0.083	

Central Manchester Stations (CMS) variable

	17	
Intercept	70.789	(0.272)
CMS	456.367	(2.511)

R^2	0.157

Table 3: Multiple Regression - Trip End Models

	18		19		20	
Intercept	-1477.2	(-2.653)	-878.8	(-1.389)	-305.4	(-0.652)
POP	0.027	(3.700)	0.026	(4.104)	0.021	(3.411)
RSOC1	30607.2	(3.319)	31058.1	(3.564)	24901.3	(4.277)
RSOC3	-23810.7	(-2.175)	-26160.5	(-2.508)	-31053.3	(-3.289)
RCAR2	2099.9	(0.876)	1362.8	(0.670)	-	
RLOCAL	4054.3	(2.230)	3984.1	(2.519)	2843.0	(1.859)
FREQ	8.703	(3.919)	8.642	(4.334)	6.781	(3.796)
RJT	25.534	(0.493)	-		-	
RJT2	-0.580	(-0.398)	-		-	
FRJT	-		-51.125	(-1.543)	-77.870	(-2.753)
CMS	-430.5	(-2.263)	-382.235	(-2.113)	-	
R^2	0.766		0.782		0.745	
R^2	0.686		0.718		0.692	

Table 4: Trip End Model - Choice of Functional Form

	20		21			22	
Dependent variable	DEP		LDEP			LDEP	
Intercept	-305.4	(-0.652)	5.455	(7.814)		-3.110	(-1.445)
POP	0.021	(3.411)	0.00003	(4.287)	LPOP	0.986	(5.187)
RSOC1	24901.3	(4.277)	21.867	(2.521)	LRSOC1	1.052	(4.286)
RSOC3	-31053.3	(-3.289)	-19.095	(-1.357)	LRSOC3	-0.392	(-1.147)
RLOCAL	2843.0	(1.859)	-2.342	(-1.028)	LRLOC	0.099	(0.211)
FREQ	6.781	(3.796)	0.012	(4.586)	LFREQ	0.685	(3.476)
FRJT	-77.870	(-2.753)	-0.173	(-4.108)	LFRJT	-0.411	(-1.362)
R^2	0.745		0.767			0.760	
R^2	0.692		0.719			0.711	

Table 6: Re-calibration of model 21 (model 23)

	VALUE	T-STAT	Z-TEST
Intercept	5.665	8.707	1.458
POP	0.000036	3.590	1.154
RSOC1	19.894	2.189	0.896
RSOC3	-16.141	-1.053	0.809
RLOCAL	-3.269	-1.392	1.618
FREQ	0.0118	4.413	0.597
FRJT	-0.170	-4.030	0.291
R^2	0.781		
R^2	0.724		

Table 8: Collinearity Diagnostics

Component	Condition No.	VARIANCE PROPORTIONS						
		Intercept	POP	RSOC1	RSOC3	RLOCAL	FREQ	FRJT
1	1.000	0.0004	0.0030	0.0020	0.0008	0.0008	0.0023	0.0028
2	3.760	0.0000	0.0448	0.1610	0.0040	0.0026	0.0000	0.0353
3	6.056	0.0005	0.0592	0.0022	0.0364	0.0187	0.3229	0.0665
4	7.235	0.0010	0.6160	0.0051	0.0030	0.0322	0.0565	0.2655
5	8.883	0.0127	0.1536	0.1458	0.0023	0.0785	0.2159	0.5981
6	16.880	0.0236	0.0733	0.5948	0.9356	0.3013	0.1769	0.0282
7	22.278	0.9617	0.0502	0.0890	0.0180	0.5659	0.2255	0.0035

Table 9: Correlation Matrix

		POP	RSOC1	RSOC3	RLOCAL	FREQ	FRJT
POP		1.000	-0.370	-0.089	0.152	0.037	0.257
RSOC1	-0.370	1.000	0.578	-0.453	0.128	-0.337	
RSOC3	-0.089	0.578	1.000	0.062	-0.297	-0.287	
RLOCAL		0.152	-0.453	0.062	1.000	-0.319	0.219
FREQ		0.037	0.128	-0.297	-0.319	1.000	0.204
FRJT		0.257	-0.337	-0.287	0.219	0.204	1.000

Table 10: Greater Manchester - Additional Models

	24	25	26	27	28	29
Intercept	3.861	3.866	3.677	4.911	5.484	6.036
	(3.872)	(4.037)	(3.354)	(3.970)	(6.339)	(5.327)
POP1	0.00003*ǂ	0.00003*ǂ	0.00007*	0.00007*	0.00006*ǂ	0.00005*ǂ
	(0.623)	(0.658)	(1.324)	(1.502)	(1.185)	(0.939)
POP2	0.00006	0.00006	0.00004	0.00006	0.00007	0.00007
	(4.091)	(4.307)	(3.173)	(3.863)	(4.311)	(4.650)
RSOC1	33.023	33.143	22.978	34.007	39.764	40.985
	(2.716)	(3.157)	(1.812)	(2.831)	(3.842)	(4.107)
RSOC3	-22.072*	-22.215*	-25.100*	-32.717	-36.700	-38.360
	(-1.303)	(-1.465)	(-1.384)	(-2.006)	(-2.451)	(-2.509)
FREQ	0.012	0.012	0.015	0.013	0.011	0.011
	(3.227)	(3.425)	(3.607)	(3.387)	(3.563)	(3.453)
DIST	-0.162	-0.162	-	SPEED -1.484ǂ	-1.476*ǂ	-1.265*ǂ
	(-2.561)	(-2.647)	-	(-1.664)	(-1.212)	(-1.261)
RJT	-0.0008*	-	-0.018*	-	-	-
	(-0.021)	-	(-0.382)	-	-	-
BFREQ	-0.002	-0.002	-0.002	-0.002	-0.003	-0.003
	(-2.654)	(-2.790)	(-1.902)	(-2.549)	(-3.079)	(3.495)
BJT	0.005*	0.004*	0.005*	-0.011*ǂ	RBJ -0.154*ǂ	BSPEED 1.988*ǂ
	(0.174)	(0.187)	(0.172)	(-0.414)	(-0.547)	(0.865)
CJT	0.102	0.102	0.028*	0.033*	RCJT 0.262*	CSPEED -2.277*
	(2.407)	(2.188)	(0.619)	(0.876)	(0.315)	(-0.979)
R^2	0.805	0.805	0.754	0.777	0.769	0.775
^{2}	0.728	0.734	0.669	0.700	0.689	0.698

* Not significant at the 10% level

ǂ Implausible parameter value

NB: The figures in each row refer to the nearest variable descriptor to the left.

	30	31	32	33	34	MEAN VALUE
Intercept	3.400	5.946	4.314	4.481	3.944	
	(3.173)	(9.764)	(4.594)	(5.331)	(4.665)	
POP1	0.00005*	-0.00002*╪	0.00009	0.00005*	0.00003*	╪7848.0
	(0.924)	(-0.281)	(1.871)	(0.981)	(0.672)	
POP2	0.00004	0.00005	0.00004	0.00005	0.00006	14091.0
	(2.765)	(3.025)	(2.621)	(3.593)	(4.382)	
RSOC1	26.719	39.786	18.100	26.430	33.449	0.023
	(2.149)	(3.493)	(1.654)	(2.558)	(3.284)	
RSOC3	-23.715*	-50.944	-22.675*	-18.103*	-23.385	0.028
	(-1.372)	(-3.176)	(-1.560)	(-1.384)	(-1.723)	
FREQ	0.013	RGT -0.014*	FREQ 0.015	0.012	0.012	93.972
	(3.285)	(-1.550)	(4.337)	(3.984)	(3.663)	
SPEED	-0.697*╪	-	FARE -0.017	-0.016	-	66.056
	(-0.701)	-	(-1.884)	(2.056)	-	
BGT	-0.0005*╪	-0.004*╪	BFREQ-0.002	-0.003	-0.002	262.44
	(-0.215)	(-1.463)	(-2.666)	(-3.552)	(-2.915)	
CJT	0.046	0.044	0.052	0.138	0.107	1.377
	(2.337)	(1.811)	(2.043)	(3.616)	(2.887)	
DIST	-	-	-	-0.160	-0.162	11.00
				(-2.805)	(-2.915)	
R^2	0.721	0.639	0.782	0.832	0.805	
^{2}	0.638	0.548	0.717	0.774	0.747	

	35	36	37	38
Intercept	2.253	7.381	7.235	7.333
	(1.825)	(9.799)	(9.590)	(9.899)
POP	0.00005	0.00005	0.00005	0.00005
	(4.815)	(6.383)	(6.362)	(6.587)
RSOC1	31.768	24.441	RSOC11-47.553	RSOC1234.442
	(3.718)	(3.784)	(-0.827)	(3.962)
RSOC3	-27.580	-21.241	RSOC12 51.521	-
	(-2.155)	(-2.046)	(2.297)	-
LFREQ	0.831	FREQ 0.013	RSOC3 -28.591	RSOC3 -24.309
	(4.050)	(6.502)	(2.419)	(-2.302)
FARE	-0.0169	FDIST -0.140	FREQ 0.014	FREQ 0.013
	(-2.084)	(-5.086)	(6.686)	(6.755)
DIST	-0.179	-	FDIST 0.140	-0.139
	(-3.231)	-	(-5.135)	(-5.141)
BFREQ	-0.003	-0.002	-0.002	-0.002
	(-3.810)	(-3.553)	(-3.501)	(-3.586)
CJT	0.144	CSPEED -4.050	-4.076	-4.069
	(3.587)	(-3.622)	(-3.683)	(3.697)
R^2	0.809	0.873	0.880	0.877
^{2}	0.752	0.841	0.847	0.846

	39		40	
	PEAK		OFF-PEAK	
		|Z-TEST|		|Z-TEST|
Intercept	7.062	0.894	6.279	4.826
	(6.703)		(9.135)	
POP	0.00006	1.656	0.00004	3.340
	(5.004)		(5.841)	
RSOC12	38.406	1.105	29.493	1.173
	(3.081)		(3.588)	
RSOC3	-37.105	2.984	-18.112	1.763
	(-2.517)		(-1.779)	
FDIST	-0.139	0.024	-0.148	0.888
	(-3.687)		(-5.671)	
FREQ	0.033	13.446	0.020	7.458
	(5.111)		(6.173)	
BFREQ	-0.007	32.592	-0.002	1.477
	(-3.354)		(-1.781)	
CSPEED	-4.582	1.073	-3.322	2.194
	(-2.658)		(-3.650)	
R^2	0.810		0.861	
^{2}	0.762		0.827	

	41	42		43	
	MANCHESTER ONLY	PEAK	|Z-TEST|	OFF-PEAK	|Z-TEST|
Intercept	6.955	7.609	1.368	4.479	5.342
	(5.759)	(5.350)		(2.932)	
POP1	0.00008	0.00005*ǂ	1.821	0.00011	0.304
	(1.779)	(0.966)		(1.797)	
POP2	0.00004	0.00006	2.530	0.00001*	0.522
	(3.399)	(3.592)		(1.086)	
RSOC12	21.103	20.070*	0.269	10.190*	2.693
	(2.037)	(1.638)		(0.752)	
MFDIS	-0.234	-0.268	1.789	-0.228	0.296
	(-4.538)	(-4.482)		(-3.363)	
FREQ	0.014	0.038	15.475	0.024	10.601
	(5.773)	(5.679)		(4.379)	
BFREQ	-0.0011*	-0.0053	8.503	0.0019*ǂ	6.903
	(-1.444)	(-2.500)		(1.051)	
CSPEED	-3.721	-5.826	3.442	-0.933*	5.023
	(-2.395)	(-2.836)		(0.542)	
R^2	0.788	0.788		0.613	
^{2}	0.736	0.735		0.517	

	44		45	46
	MANCHESTER ONLY		ALL TRIPS	MANCHESTER ONLY
Intercept	6.411		2.260*	-0.772*
	(5.719)		(0.710)	(-0.232)
POP1	0.00010	LPOP	1.003	1.060
	(2.421)		(4.720)	(4.743)
POP2	0.00004	LRSOC12	1.030	0.994
	(3.323)		(3.999)	(3.737)
RSOC12	20.041	LRSOC3	-0.579	-0.326*
	(2.119)		(1.663)	(-0.879)
MFDIST	-0.247	LGCR	-0.609	-0.028*
	(-5.213)		(-1.839)	(-0.081)
FREQ	0.012	LGCBR	0.134*	-0.566*╪
	(5.308)		(0.193)	(-0.576)
BFREQ	-0.0006*	LGCCR	1.785╪	2.812╪
	(-0.933)		(2.439)	(2.486)
CSPEED	-2.731			
	(-1.862)			
BDV	0.517			
	(2.586)			
R^2	0.830		0.764	0.711
^{2}	0.780		0.716	0.652

Table 13: Effects of Amendments to Model 38 on Goodness of Fit

	R^2
	0.846
Amendments	
Replace FDIST by MFDIST	0.721
Replace FDIST by DIST	0.705
Replace FREQ by LFREQ	0.809
Replace FREQ by IFREQ	0.753
Replace FDIST by MFDIST and FREQ by LFREQ	0.671
Replace FDIST by DIST and FREQ by IFREQ	0.662

Table 15: Model 47 - Dependent Variable = LTR

	parameter value	t-statistic	mean elasticity
Intercept	-5.271	-3.615	
LSOC12	0.916	4.875	0.916
LSOC3	0.752	-4.403	-0.752
LFREQ	0.813	4.793	0.813
MFDIST	-0.0912	-1.273	-0.790
BFREQ	-0.0023	-2.708	-0.551
CSPEED	-2.665	-2.665	-1.325
R^2	0.863		
^{2}	0.827		

Table 17: Basic model for walk mode access patrons

	Parameter Estimate	t Statistic
Intercept	2.536	(1.324)*
POP1	0.00013	(2.546)
NSOC2	5.782	(0.360)*
NSOCIIIN	2.536	(0.063)*
NSOCIIIM	-1.570	(-0.073)*
FREQ	0.0195	(5.778)
BFREQ	-0.0011	(-0.865)*
SPEED	-0.1023	(-0.085)*
CJT	0.0302	(1.593)*
^{2}	61%	

* = insignificant at 90% level of confidence

Table 18: Model comparisons

	Model 48	Model 49	Model 50	Mean number of trips per week day from each station originating from walk access mode
All Destinations				
All Boardings	^{2}=77%	^{2}=70%	^{2}=63%	451
Peak Boardings	^{2}=75%	^{2}=69%	^{2}=54% if NSOC2 dropped	283
Off Peak Boardings	^{2}=73%	^{2}=69%* if FSOCIIIN dropped	^{2}=64%	168
Destination - central Manchester				
All Boardings	^{2}=62%	^{2}=49% if BFREQSQ dropped	^{2}=48%	198
Peak Boardings	^{2}=64%	^{2}=58% if NSOC2 dropped	^{2}=42%	128
Off Peak Boardings	^{2}=54%*	^{2}=45%*	^{2}=41% if NSOC2 dropped	70

* At least one variable is insignificant at the 90% level

Table 19: Model 48 - all destinations

	All Trips	Peak Trips	Off Peak Trips
Intercept	2.427 (4.847)	2.172 (4.462)	1.376 (2.531)
POP1	0.000072 (1.795)	0.000071 (1.813)	0.000113 (2.312)
POP2	0.000044 (5.745)	0.000040 (3.377)	0.000042 (3.032)
FREQ	0.0138 (3.655)	0.0124 (5.320)	0.0191 (7.309)
BFREQSQ	-0.0000025 (-2.310)	-0.0000023 (-2.122)	-0.0000043 (-3.573)
FSOC2	21.206 (2.330)	21.819 (2.466)	20.959 (1.828)
FSOCIIIN	61.651 (1.964)	59.783 (1.959)	-
^{2}	77%	75%	73%

Table 20: Model 48 - destination = central Manchester

	All Trips	Peak Trips	Off Peak Trips
Intercept	2.023 (3.633)	1.826 (3.523)	0.92 (1.446)*
POP1	0.000084 (1.885)	0.000080 (1.915)	0.000124 (2.169)
POP2	0.000039 (2.887)	0.000033 (2.662)	0.000040 (2.489)
FREQ	0.0089 (3.341)	0.0085 (3.411)	0.0125 (4.104)
BFREQSQ	-0.0000021 (-1.740)	-0.0000023 (-2.009)	-0.0000032 (-2.249)
FSOC2	22.189 (2.193)	19.821 (2.104)	27.024 (2.013)
FSOCIIIN	67.386 (1.931)	67.354 (2.073)	-
^{2}	62%	64%	52%

Table 21: Model 48 - destination = non central Manchester

	All Trips	Peak Trips	Off Peak Trips
Intercept	1.441 (2.657)	-352.5 (-3.289)	-0.0861 (-0.144)*
POP1	0.000053 (1.208)*	0.0085 (0.985)*	0.000074 (1.534)*
POP2	0.000050 (3.793)	0.0066 (2.538)	0.000051 (3.512)
FREQ	0.0189 (7.278)	2.130 (4.152)	0.0235 (8.215)
BFREQSQ	-0.0000029 (-2.400)	-0.000092 (-0.394)*	-0.0000041 (-3.159)
FSOC2	19.672 (1.995)	8068.6 (4.143)	15.349 (1.410)*
FSOCIIIN	52.842 (1.554)*	-2517.4 (-0.375)*	57.778 (1.539)*
^{2}	81%	66%	83%

* = insignificant at 90% level of confidence

Table 23: Correlation matrix

	POP1	FREQ	POP2	BFREQSQ	FSOC2	FSOCIIIN
POP1	1.00000	-0.18376	0.37438	0.39250	-0.38794	-0.15556
FREQ	0.18376	1.00000	-0.05117	0.08889	0.35225	0.38414
POP2	0.37438	-0.05117	1.00000	0.25230	-0.48826	-0.43743
BFREQSQ	0.39250	0.08889	0.25230	1.00000	-0.36053	-0.41628
FSOC2	-0.38794	0.35225	-0.48826	-0.36053	1.00000	0.38340
FSOCIIIN	-0.15556	0.38414	-0.43743	-0.41628	0.38340	1.00000

Table 24: Model 51 - CJT/Distance based model for all trips (walk mode station access only)

Intercept	3.339	(7.034)
FREQ	0.0170	(7.993)
CJT	0.1053	(2.494)
BFREQSQ	-0.0000035	(-3.190)
DIST	-0.1492	(-2.258)
POP	0.000026	(2.879)
^{2}	74%	

Table 25: Model 52 - Best `trip end' model (walk mode station access only)

Intercept	4.510	(14.719)
POP	0.000040	(4.860)
FREQ	0.0137	(6.997)
BFREQSQ	-0.0000024	(-2.680)
SOCV	-142.9	(-3.907)
^{2}	80%	

Table 27: Model 53 components

	non-C.Manchester trip		**C.Manchester trips**	
Intercept	3.190	(9.733)	2.996	(5.474)
POP	0.000043	(4.931)	0.000031	(2.521)
FREQ	0.0190	(9.053)	0.0097	(3.329)
BFREQSQ	-0.0000028	(-2.972)	-0.0000030	(-2.351)
CJT	-	0.0883	(1.767)	
DIST	-	-0.1437	(-1.916)	
SOC2	-	10.953	(1.364)*	
SOCV	-130.153	(-3.327)	-	
^{2}		84%		58%

* insignificant at 90% level of confidence

Appendix 1: Greater Manchester New Stations Model Calibration Data Set - Stations

	STATION CODE	STATION NAME
Oldham Loop (Oldham/Rochdale Line)		
	MPL	Miles Platting
	DLA	Dean Lane
	FAI	Failsworth
	MOS	Mostyn
	HOL	Hollinwood
	OWE	Oldham Werneth
	MHI	Mills Hill
	OMU	Oldham Mumps
	DER	Derker
	SHA	Shaw
	NHE	New Hey
	MIL	Milnrow
	SBR	Smithy Bridge
	LIT	Littleborough
	ROC	Rochdale
	CAS	Castleton
Bury Line		
	WRO	Woodlands Road
	CRU	Crumpsall
	BVA	Bowker Vale
	HPA	Heaton Park
	PRE	Prestwich
	BOB	Bess o'th' Barn
	WHI	Whitefield
	RAD	Radcliffe
	BUR	Bury
Altrincham Line		
	OTR	Old Trafford
	WRD	Warwick Road
	STR	Stretford
	DRO	Dane Road
	SAL	Sale
	BRO	Brooklands
	TIM	Timperley
	NRO	Navigation Road
	ALT	Altrincham
Buxton Line		
	DAV	Davenport
	HGR	Hazel Grove

Appendix 2: Bus-Rail Competition

ALTRINCHAM/SALE/MANCHESTER

Bus services 261, 262, 264 (GMB); North Western 343 and 344

BURY/MANCHESTER

RADCLIFFE service 97, plus from centre, services 94 and 98

WHITEFIELD services 90, 91, 94, 97, 98, 135, 137

BOWKER VALE AND CRUMPSALL services 57, 59-61, 147, 160/161

ROCHDALE services 16, 17, 23, 24

LITTLEBOROUGH service 23

CASTLETON services 16/17

MILNROW/NEW HEY/SHAW services 181/182

DERKER/OLDHAM/OLDHAM LOOP services 82, M82 (Ribble), 180, 183, 184

STOCKPORT services 190-2, plus Bee Line City Sprint minibus routes 1 & 2

HAZEL GROVE services 190-2

DAVENPORT peak hour trip only Trent service 252, otherwise no direct link to Manchester, only Stockport

Appendix 3: Variable Definitions

(In the order they appear in the text)

VARIABLE	DEFINITION
POP1	Usually resident population within 800m of origin station
POP2	Usually resident population between 800m and 2km of origin station
POP	Usually resident population within 2km of origin station
SOC1	Population in household headed by a member of Social Class I or II (managerial and professional) and within 2km of origin station
SOC3	Population in household headed by a member of Social Class IIIM (skilled manual) and within 2km of origin station
RSOC1	SOC1/POP
RSOC3	SOC3/POP
CAR2	Number of households within 2km of origin station with 2 or more cars
HHLD	Number of households within 2km of origin station
RCAR2	CAR2/HHLD
WLOCAL	Number of people within 2km of origin station and either working at home or travelling to work by pedal cycle or on foot
EACT	Number of people within 2km of origin station who are economically active
RLOCAL	WLOCAL/EACT
FREQ	Number of trains in both directions between origin station and central Manchester on a weekday
RJT	Journey time in minutes by rail from origin station to central Manchester (either Piccadilly or Victoria stations)
RJT2	RJT squared
FARE	Mean rail fare in pence from station to all destinations
FRJT	FARE/RJT
CMS	Central Manchester Stations, Victoria=1, Deansgate, Oxford Road, Piccadilly=2
DEP	Dependent variable - all boardings
L	This prefix denotes that the natural logarithm has been taken
DIST	Distance in kilometres by rail from origin station to central Manchester (either Piccadilly or Victoria stations)
BFREQ	Number of buses (both ways) between origin station and central Manchester
BJT	Journey time by bus in minutes from origin station to central Manchester
CJT	Journey time by car in minutes from origin station to central Manchester
MFARE	Mean standard rail fare in pence from station to central Manchester, based on GMPTE standard adult cash fare
SPEED	DIST/RJT
BSPEED	DIST/BJT
CSPEED	DIST/CJT
RBJT	BJT/RJT
RCJT	CJT/RJT
BGT	Generalised time of travel in minutes to central Manchester by bus Defined as BJT plus service interval (see text for explanation)
RGT	Generalised time of travel in minutes to central Manchester by rail Defined as RJT plus service interval (see text for explanation)
RSOC11	Population in household headed by a member of social class I (professional) and within 2km of origin station, divided by POP
RSOC12	Population in household headed by a member of social class II (managerial) and

	within 2km of origin station, divided by POP
MFDIST	MFARE/DIST
IFREQ	1/FREQ
BDV	Bury line dummy variable
GCR	Generalised cost of rail (pence, mid 1988 prices)
GCB	Generalised cost of bus (pence, mid 1988 prices)
GCC	Generalised cost of car (pence, mid 1988 prices)
GCBR	GCB/GCR
GCCR	GCC/GCR
MANC	Dependent variable - Manchester boardings only
RESID	Residual value (DEP (or MANC) minus PREDICT)
PREDICT	Predicted value
TR	DEP/POP
SOC12	Population in household headed by a member of social class II(managerial) and within 2km of origin station
NSOCIIIN	Population in household within 800m of origin station headed by a member of social class IIIN (skilled non manual), divided by the population currently resident within 800m of the origin station
NSOCIIIM	Population in household within 800m of origin station headed by a member of social class IIIM (skilled manual), divided by the population currently resident within 800m of the origin station
FSOC2	Population in household between 800m and 2km of origin station headed by a member of social class I or II (managerial and professional), divided by the population currently resident between 800m and 2km of the origin station
NSOC2	Population in household within 800m of origin station headed by a member of social class I or II (managerial and professional), divided by the population currently resident within 800m of the origin station
BFREQSQ	BFREQ squared
FSOCIIIN	Population in household between 800m and 2km of origin station headed by a member of social class IIIN (skilled non manual), divided by the population currently resident between 800m and 2km of the origin station
SOCV	Population in household headed by a member of Social Class V (unskilled manual) and within 2km of the origin station

Appendix 4: Times by Highway to Piccadilly Gardens

1986

STATION NAME	ZONE	PEAK TIME	OFF-PEAK TIME
Bury	260	28.53	23.61
Radcliffe	244	28.31	23.52
Whitefield	251	23.53	19.08
Bess o'th' Barn	250	20.16	16.96
Prestwich	249	16.52	13.57
Heaton Park	248	15.19	12.89
Bowker Vale	7	12.91	10.94
Crumpsall	7	12.91	10.94
Woodlands Road	5	9.35	8.37
Old Trafford	87	7.53	6.87
Warwick Road	71	11.95	10.72
Stretford	72	14.09	12.49
Dane Road	73	20.64	16.88
Sale	73	20.64	16.88
Brooklands	77	23.46	18.97
Timperley	78	27.53	23.02
Navigation Road	83	28.87	22.67
Altrincham	84	29.9	33.63

1989

Dean Lane	19	11.40	9.25
Failsworth	299	14.99	12.19
Hollinwood	300	16.71	13.78
Oldham Werneth	302	23.97	20.07
Oldham Mumps	309	26.79	22.68
Derker	314	28.86	24.33
Shaw	311	31.93	25.84
New Hey	290	33.09	27.13
Milnrow	290	33.09	27.13
Miles Platting	11	6.20	5.73
Moston	21	16.66	13.75
Mills Hill	272	21.80	18.02
Castleton	279	28.12	22.74
Rochdale	283	32.96	26.66
Smithy Bridge	292	39.01	31.59
Littleborough	292	39.01	31.59
Davenport	377	31.15	23.43
Hazel Grove	372	39.17	29.62

Appendix 5: Access mode of boarding passengers

			DAILY RAIL PATRONAGE		
STATION CODE	STATION NAME	MODES	ALL ACCESS MODES	WALK ACCESS %	
MPL	Miles Platting		29	27	93%
DLA	Dean Lane		198	161	81%
FAI	Failsworth		189	158	84%
MOS	Moston		194	160	82%
HOL	Hollinwood		131	118	90%
OWE	Oldham Werneth		146	107	73%
MHI	Mills Hill		391	270	69%
OMU	Oldham Mumps		526	310	59%
SHA	Shaw		360	202	56%
NHE	New Hey		82	72	88%
MIL	Milnrow		197	161	82%
SBR	Smithy Bridge		203	157	77%
LIT	Littleborough		174	103	59%
ROC	Rochdale		1279	440	34%
CAS	Castleton		208	152	73%
WRO	Woodlands Road		272	238	88%
CRU	Crumpsall		932	830	89%
BVA	Bowker Vale		551	416	75%
HPA	Heaton Park		671	592	88%
PRE	Prestwich		984	767	78%
BOB	Bess-O'th'-Barn		590	469	79%
WHI	Whitefield		1008	654	65%
RAD	Radcliffe		922	489	53%
BUR	Bury		2155	1055	49%
OTR	Old Trafford		606	438	72%
WRD	Warwick Road		548	484	88%
STR	Stretford		822	514	63%
DRO	Dane Road		378	333	88%
SAL	Sale		1879	1131	60%
BRO	Brooklands		1075	761	71%
TIM	Timperley		1062	927	87%
NRO	Navigation Road		560	449	80%
ALT	Altrincham		2680	1737	65%
			22002	14882	

Appendix 6: Trip Destination of Boarding Passengers

		DAILY RAIL PATRONAGE WITH A WALK MODE ACCESS		
STATION CODE	STATION NAME	ALL TRIPS	TRIPS DESTINED TO CENTRAL MANCHESTER	%
MPL	Miles Platting	27	11	41
DLA	Dean Lane	161	91	57
FAI	Failsworth	158	101	64
MOS	Moston	160	125	78
HOL	Hollinwood	118	81	69
OWE	Oldham Werneth	107	64	60
MHI	Mills Hill	270	177	66
OMU	Oldham Mumps	310	164	53
SHA	Shaw	202	102	50
NHE	New Hey	72	41	57
MIL	Milnrow	161	104	65
SBR	Smithy Bridge	157	80	51
LIT	Littleborough	103	68	66
ROC	Rochdale	440	242	55
CAS	Castleton	152	96	63
WRO	Woodlands Road	238	151	63
CRU	Crumpsall	830	470	57
BVA	Bowker Vale	416	202	49
HPA	Heaton Park	592	238	40
PRE	Prestwich	767	201	26
BOB	Bess O'th'Barn	469	238	51
WHI	Whitefield	654	358	55
RAD	Radcliffe	489	210	43
BUR	Bury	1055	355	34
OTR	Old Trafford	438	156	36
WRD	Warwick Road	484	186	38
STR	Stretford	514	168	33
DRO	Dane Road	333	124	37
SAL	Sale	1131	597	53
BRO	Brooklands	761	292	38
TIM	Timperley	927	324	35
NRO	Navigation Road	449	173	39
ALT	Altrincham	1737	538	31
		14882	6528	

CONFIDENTIAL ANNEXE

Table 5: Analysis of Residuals (Model 21)

Station Code	Actual Usage	Prediced Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	29	55	-26
DLA	198	347	-149
FAI	189	225	-36
MOS	194	158	36
HOL	131	216	-85
OWE	146	440	-294
MHI*	551	263	288
OMU	526	358	168
DER*	335	279	56
SHA	594	471	123
NHE	82	185	-103
MIL	197	282	-85
SBR*	317	196	121
LIT	174	173	1
ROC	1279	505	774
CAS	208	149	59
Bury Line			
WRO	251	206	45
CRU	914	523	391
BVA	545	936	-391
HPA	659	556	103
PRE	978	659	319
BOB	586	521	65
WHI	973	1137	-164
RAD	903	496	407
BUR	1865	1665	200
Altrincham Line			
OTR	604	549	55
WRD	545	555	-10
STR	809	827	-18
DRO	378	386	-8
SAL	1897	997	900
BRO	1066	1424	-358
TIM	1045	1281	-236
NRO	551	596	-45
ALT	3104	3941	-837

Buxton Line

DAV	546	1205	-659
HGR	1085	957	128

* = New Station

Table 7: Analysis of Residuals (model 23)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	29	51	-22
DLA	198	340	-142
FAI	189	222	-33
MOS	194	164	30
HOL	131	206	-75
MHI*	551	263	288
DER*	335	264	71
SHA	594	465	129
NHE	82	200	-118
MIL	197	290	-93
SBR*	317	199	118
LIT	174	165	9
CAS	208	147	61
Bury Line			
WRO	251	203	48
CRU	914	500	414
BVA	545	959	-414
HPA	659	554	105
PRE	978	660	318
BOB	586	538	48
WHI	973	1143	-170
RAD	903	462	441
Altrincham Line			
OTR	604	515	89
WRD	545	533	12
STR	809	762	47
DRO	378	386	-8
BRO	1066	1331	-265
TIM	1045	1239	-194
NRO	551	543	8
Buxton Line			
DAV	546	1030	-484
HGR	1085	955	130

* = New station

Table 11: Analysis of Residuals (model 38)

Station Code	Actusl Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	29	45	-16
DLA	198	309	-111
FAI	189	195	-6
MOS	194	164	30
HOL	131	138	-7
OWE	146	316	-170
MHI*	551	372	179
OMU	526	229	297
DER*	335	366	-31
SHA	594	450	144
NHE	82	131	-49
MIL	197	185	12
SBR*	317	280	37
LIT	174	214	-40
ROC	1279	1074	205
CAS	208	181	27
Bury Line			
WRO	251	183	68
CRU	914	554	360
BVA	545	585	-40
HPA	659	542	117
PRE	978	650	328
BOB	586	693	-107
WHI	973	921	52
RAD	903	893	10
BUR	1865	1763	102
Altrincham Line			
OTR	604	505	99
WRD	545	632	-87
STR	809	887	-78
DRO	378	378	0
SAL	1897	1044	853
BRO	1066	1765	-699
TIM	1045	1624	-579
NRO	551	732	-181
ALT	3104	3856	-752
Buxton Line			

DAV	546	700	-154
HGR	1085	908	177

* = New station

Table 12: Analysis of Residuals (model 41)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	12	17	-5
DLA	112	136	-24
FAI	111	116	-5
MOS	141	80	61
HOL	94	95	-1
OWE	100	270	-170
MHI*	192	204	-12
OMU	346	165	181
DER*	219	225	-6
SHA	233	277	-44
NHE	45	110	-65
MIL	136	107	29
SBR*	108	178	-70
LIT	117	99	18
ROC	902	583	319
CAS	173	106	67
Bury Line			
WRO	157	142	15
CRV	492	243	249
BVA	209	206	3
HPA	266	181	85
PRE	295	168	127
BOB	318	314	4
WHI	595	479	116
RAD	540	522	18
BUR	1078	670	408
Altrincham Line			
OTR	203	245	42
WRD	192	226	-34
STR	194	351	-157
DRO	152	188	-36
SAL	1007	543	464
BRO	482	841	-360
TIM	512	681	-169
NRO	222	416	-194
ALT	1052	1104	-52
Buxton Line			

DAV	486	497	-11
HGR	914	757	157

* = New station

Table 14: Analysis of Residuals (Model 45)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	29	81	-52
DLA	198	224	-26
FAI	189	238	-49
MOS	194	148	46
HOL	131	172	-41
OWE	146	172	-26
MHI*	551	457	94
OMU	526	342	184
DER*	335	223	112
SHA	594	505	89
NHE	82	93	-11
MIL	197	256	-59
SBR*	317	423	-106
LIT	174	262	-88
ROC	1279	890	389
CAS	208	190	18
Bury Line			
WRO	251	276	-25
CRU	914	701	213
BVA	545	856	-311
HPA	659	408	251
PRE	978	536	442
BOB	586	406	180
WHI	973	994	-21
RAD	903	427	476
BUR	1865	1324	541
Altrincham Line			
OTR	604	351	253
WRD	545	683	-138
STR	809	893	-84
DRO	378	309	69
SAL	1897	911	986
BRO	1066	1409	-343
TIM	1045	1211	-166
NRO	551	516	35
ALT	3104	3885	-781
Buxton Line			
DAV	546	1487	-941
HGR	1085	1394	-309

* = new stations

Table 16: Analysis of Residuals (Model 47)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	29	57	-28
DLA	198	219	-21
FAI	189	179	10
MOS	194	157	37
HOL	131	112	19
MHI*	551	408	143
DER*	335	327	8
SHA	594	583	11
NHE	82	90	-8
MIL	197	214	-17
SBR*	317	463	-146
LIT	174	204	-30
CAS	208	193	15
Bury Line			
WRO	251	319	-68
CRU	914	578	336
BVA	545	597	-52
HPA	659	506	153
PRE	978	541	437
BOB	586	686	-100
WHI	973	997	-24
RAD	903	728	175
Altrincham Line			
OTR	604	355	249
WRD	545	576	-31
STR	809	748	61
DRO	378	256	122
BRO	1066	1982	-916
TIM	1045	1663	-618
NRO	551	676	-125
Buxton Line			
DAV	546	938	-392
HGR	1085	642	443

Table 22: Model 48 - residual analysis of daily patronage (walk mode access only)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	27	61	-34
DLA	161	187	-26
FAI	158	150	8
MOS	160	154	6
HOL	118	162	-44
OWE	107	165	-58
MHI*	270	254	16
OMU	310	104	206
SHA	202	301	-99
NHE	72	145	-73
MIL	161	185	-24
SBR*	157	114	43
LIT	103	78	25
ROC	440	377	63
CAS	152	120	32
Bury Line			
WRO	238	282	-44
CRU	830	309	521
BVA	416	292	124
HPA	592	473	119
PRE	767	509	258
BOB	469	449	20
WHI	654	739	-85
RAD	489	510	-21
BUR	1055	785	270
Altrincham Line			
OTR	438	506	-68
WRD	484	560	-76
STR	514	757	-243
DRO	333	396	-63
SAL	1131	1273	-142
BRO	761	887	-126
TIM	927	925	2
NRO	449	629	-180
ALT	1737	1403	334

* = new stations

Table 26: Model 52 - residual analysis of daily patronage (walk mode access only)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	27	48	-21
DLA	161	171	-9
FAI	158	179	-21
MOS	160	127	33
HOL	118	170	-52
OWE	107	235	-128
MHI*	270	248	22
OMU	310	106	204
SHA	202	298	-96
NHE	72	82	-10
MIL	161	186	-25
SBR*	157	165	-8
LIT	103	114	-11
ROC	440	316	124
CAS	152	95	57
Bury Line			
WRO	238	262	-24
CRU	830	341	489
BVA	416	342	74
HPA	592	516	76
PRE	767	573	194
BOB	469	459	10
WHI	654	522	132
RAD	489	616	-127
BUR	1055	864	191
Altrincham Line			
OTR	438	475	-37
WRD	484	566	-82
STR	514	751	-237
DRO	333	431	-98
SAL	1131	1120	11
BRO	761	1050	-289
TIM	927	816	111
NRO	449	670	-221
ALT	1737	1219	518

* = new stations

Table 28: Model 53 - residual analysis of daily patronage (all trips)

Station Code	Actual Usage	Predicted Usage	Residual
Oldham Loop (Oldham/Rochdale Line)			
MPL	27	61	-34
DLA	161	168	-7
FAI	158	146	12
MOS	160	131	29
HOL	118	140	-22
OWE	107	181	-74
MHI*	270	244	26
OMU	310	114	196
SHA	202	284	-82
NHE	72	119	-47
MIL	161	131	30
SBR*	157	171	-14
LIT	103	102	1
ROC	440	350	90
CAS	152	112	40
Bury Line			
WRO	238	296	-58
CRU	830	335	495
BVA	416	337	79
HPA	592	455	137
PRE	767	503	264
BOB	469	493	-24
WHI	654	544	110
RAD	489	660	-171
BUR	1055	795	260
Altrincham Line			
OTR	438	494	-56
WRD	484	505	-21
STR	514	672	-158
DRO	333	429	-96
SAL	1131	1064	67
BRO	761	1072	-311
TIM	927	914	13
NRO	449	633	-184
ALT	1737	1306	431

* = new stations

jmp\wp326(lma)