DEMAND FORECASTING FOR NEW LOCAL RAIL STATIONS AND SERVICES

By Jonathan Preston*

1. INTRODUCTION

Passenger rail services in Britain are undergoing something of a revival. Since 1970 about 200 new stations have been opened on publicly owned railways in Britain, and British Rail (BR) have developed a number of new services. Before a new service or a new station can be evaluated a key question has to be posed: how many people will use it?

An immediate problem is that BR’s usual approach to demand forecasting is not suitable for predicting the effects of non-marginal changes in levels of service at a local level. This is discussed in section 2. So the Institute for Transport Studies at Leeds University has been working since 1982 on the development of suitable forecasting approaches. In the course of this work a variety of approaches have been developed and tested. These include a simple Trip Rate Model (TRM), a direct demand model that we have entitled an Aggregate Simultaneous Model (ASM), and a disaggregate Mode Choice Model (MCM). These models are all based on actual behaviour/choices (that is, Revealed Preference (RP) data); they are discussed in section 3. An alternative approach is based on asking people directly how often they would use a new facility. This is termed the Stated Intentions (SI) approach. It is well known that this, if applied without adjustment, is likely to lead to overestimates of usage. A Stated Preference (SP) experiment has been devised in order to correct for this bias. This SI/SP approach is discussed in section 4.

* Institute for Transport Studies, University of Leeds. This work was initially undertaken as a Collaborative Award in Science and Engineering research studentship, funded by the Science and Engineering Research Council and BR (Preston, 1987). Subsequent work has included major research contracts for Leicestershire (Preston and Wardman, 1988) and Nottinghamshire (Preston, 1989) County Councils. The work is now continuing with funding from the Economic and Social Research Council. I am grateful for the comments of Chris Nash, Mark Wardman and an anonymous referee on an earlier version of this paper.
The TRM, ASM and MCM models have been compared with actual usage at six new stations in West Yorkshire. The results are discussed in section 5. Unfortunately, the SI/SP approach has not yet (with one partial exception) been applied to a situation where the new facility has subsequently been opened. However, in a study for Leicestershire County Council it was possible to compare the forecasts with those produced from a TRM and the ASM. These results are discussed in section 6.

In conclusion, it is found that there is a trade-off between the cost/complexity of models and their accuracy. It is felt that models need to be tailored to particular situations, and that major investments require more complex approaches.

2. BACKGROUND

Since 1970 almost 200 new stations have been opened on publicly owned passenger rail services in Britain (including Docklands Light Rail and Tyne and Wear Metro). On the BR network, in the period of rationalisation initiated by Beeching, the number of stations decreased from over 5,000 in 1958 to a low of 2,358 in January 1978. However, since the mid 1970s the number of station openings has begun to exceed the number of closures, so that by the end of March 1989 there were 2,435 passenger stations (BRB, 1989).

The stations that have been opened may be classified into a number of groups. This paper concentrates on unmanned halts serving residential areas on existing services, though the SI/SP approach has been used to study stations on completely new services and to consider destinations as well as origins. This paper does not consider stations related to light rail systems, InterCity stations, Parkway stations, stations giving improved central area access, or stations related to major destinations. Stations of these types will require demand forecasting techniques that are specifically designed to focus on their key characteristics.

BR’s current demand forecasting model, entitled MOIRA, predicts changes in passenger flows as a result of changes to the timetable expressed in terms of Q, an index of level of service quality (Whitehead, 1981). This approach is based on elasticities derived from time series. It was initially developed for InterCity flows only, but has since been extended to Provincial services. However, this incremental approach cannot be applied where existing rail flows are low or zero. This is unlikely for InterCity services with a coarse zoning system, but is likely for local services with a fine zoning system. Even if zones are defined so that some relevant rail trips are made in each zone and elasticities are used to study changes in service, it is likely that changes (for example, in access time) will be so great that the use of constant point elasticities is inappropriate. Hence, the mainstream modelling approach adopted by BR is not really suitable for studying the new services and stations that this paper is interested in.

The problems are magnified because demand forecasting for new rail stations and services is subject to some particular difficulties:

(i) Investments are relatively small, so cheap forecasting methods are required. For example, an unmanned station in West Yorkshire typically costs only around £120,000 (1987 prices). Non-capital costs, such as increased train
operating costs, station maintenance and administration, are often assumed to be minimal (around £2,000 per annum), though consideration is also required of the loss of revenue from existing passengers because of longer journey times. For a new rail service, both capital and non-capital costs are much more substantial.

(ii) Where rail does exist it is often the minor mode, so that large (and hence costly) samples are required if Revealed Preference (RP) mode choice models are to be developed. Moreover, locally calibrated RP models cannot be developed for a mode that does not already exist in that particular area.

(iii) New demand comes from a number of different sources. Experience from West Yorkshire indicates that around 75 per cent of demand will be abstracted from other modes (especially bus), and that the remainder will be split evenly between re-assigned rail trips and generated trips (including a small proportion of re-distributed trips). The TRM and ASM implicitly consider all these sources of demand; so does the SI/SP approach in a more explicit manner. The MCM we have developed with RP data does not include generated trips; but, as the model was only developed for work trips, this omission should not be too important, at least in the short term. Similarly, use of an MCM developed with SP data on an existing O/D matrix could not include generated trips.

3. REVEALED PREFERENCE (RP) MODELS

In this section three model types will be developed. They are (in order of increasing complexity): a Trip Rate Model, an Aggregate Simultaneous Model and a disaggregate Mode Choice Model.

3.1. Trip Rate Model (TRM)

The TRM simply expresses station usage as a function of the population within station catchment areas. Surveys of six new stations in West Yorkshire identified two main catchment areas: the 0–800 metres zone, accounting for 62 per cent of users, and the 800 metres to 2 kilometres zone, accounting for 25 per cent of users. Three examples of simple TRMs are given in Table 1. This table indicates that

<table>
<thead>
<tr>
<th>Trip Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trips per 1000 population</strong></td>
</tr>
<tr>
<td>West Yorkshire</td>
</tr>
<tr>
<td>South Wigston</td>
</tr>
<tr>
<td>Langley Mill</td>
</tr>
</tbody>
</table>
the TRM lacks spatial transferability. This is because it fails to take into account factors such as the socio-economic characteristics of the population of the catchment area, the attractiveness of destinations, the level of rail service, and competition from other modes.

3.2. Aggregate Simultaneous Model (ASM)

These additional factors were taken into account through the development of a direct demand model, which was termed the ASM. This model was calibrated with data on 99 flows estimated from the 1981/82 passenger train surveys for 39 small town, suburban and rural stations in West Yorkshire. Two preferred model forms were developed: a log-linear and a semi-log model. These are represented by Equations (1) and (2) respectively (t-statistics in brackets):

\[
LFLOW = 5.496 + 0.380LOPOP + 0.164LOPOP2 + 0.246LRSOC \\
+ 0.269LDRX - 1.341LGCOTH - 1.239LGCRA \\
R^2 = 0.539 \\
(3.025) (2.617) (1.733) (2.034)
\]

\[
LFLOW = -1.468 + 0.423LOPOP + 0.147LOPOP2 + 0.248LRSOC \\
+ 0.291LDRX + 0.507LGCOTH - 0.007GCRA \\
R^2 = 0.532 \\
(-0.790) (2.868) (1.542) (2.047)
\]

The variables used in these models are defined in Appendix 1. The formulations were chosen partly because they reduced problems of multicollinearity and heteroscedasticity. The models only have a moderate goodness of fit; almost half the variation is unexplained. In particular, the ASM underpredicted flows from long-established commuter stations; this may be partly because these stations draw users from beyond two kilometres. If six outliers of this type were excluded, the \( R^2 \) of Equation (1) would be increased to 0.66.

Apart from some particular specification and measurement problems, the ASM is affected by problems common to cross-sectional models. In particular, it lacks a dynamic structure (which is important when patronage growth over time is considered) and is subject to simultaneity problems. The model was also shown to lack temporal transferability. When Equations (1) and (2) were re-calibrated with 1984 data (based on self-completion questionnaires), five out of 14 parameter values were shown to have significantly changed, at the 5 per cent significance level. In forecasting, the ASM is only used to predict principal flows. Information from adjacent existing stations is then used to factor up these flows to produce a total station usage figure.

The ASM has been applied to over 70 sites in 12 different counties. It has become evident that its spatial transferability is limited, and that it is relatively insensitive to changes in service level. These issues will be further examined in section 6.

A possible improvement might be to develop separate models for work and non-work journeys (or, alternatively, for peak and off-peak trips). It was necessary
DEMAND FORECASTING FOR NEW LOCAL RAIL STATIONS

J. Preston

to develop a non-work trip model for West Yorkshire, to be used in conjunction with the disaggregate models of work trips described in the next section. This model was calibrated for 64 flows based on 1984 data and is described by Equation (3) (with variable definitions in Appendix 1):

\[
L_{FLOW} = -3.580 + 0.562L_{LOPOP} + 0.252L_{REMP} + 0.574L_{RS} - 0.250L_{BS}
\]

\[
= (2.231) (3.230) (4.051) (3.315) (-2.408)
\]

\[
+ 0.966IC - 1.247INTOPP \quad R^2 = 0.709
\]

\[
= (4.634) (-8.077) \quad R^2 = 0.678
\]

(3)

3.3. Disaggregate Mode Choice Models (MCM)

As far as this study is concerned, an aggregate approach has a number of weaknesses. It normally requires catchment areas (=zoning system) to be pre-defined. It fails to establish the importance of factors that exhibit greater intra-zonal than inter-zonal variation: this is particularly true of walk and wait time, which may be critical in the choice of public transport mode. It fails to take into account micro-level information on economic activity, which will clearly affect travel demand. These shortcomings may be overcome by making use of individual data on the times and costs of the mode actually used and at least one alternative (or preferably a full choice set of alternatives) in order to calibrate an MCM.

A data set of this type was provided by the 1981 West Yorkshire Corridor Study, which collected information on the journey to work as part of a study into the value of time (MVA et al., 1987). The model form chosen was the hierarchical logit (HL). This form was chosen because it overcomes the property of independence from irrelevant alternatives which is found in the more widely used multinomial logit model (MNL), and which precludes the possibility of differential substitutability and complementarity. It was hypothesised that rail users, all other things being equal, are more likely to be drawn from bus than from car. This was confirmed by a generalised likelihood ratio test (McFadden et al., 1976). HL models were estimated indirectly (that is, in two stages) using the BLOGIT package (Crittle and Johnson, 1980); the composite cost term (or expected maximum utility (EMU)) was calculated with FORTRAN programs. It is acknowledged that direct estimation (or full information Maximum Likelihood) is preferable to indirect estimation (Daly, 1987), but the requisite software was not available.

At the calibration stage there were difficulties in including socio-economic variables. The preferred model in this study was thus market segmented, and consisted of an MNL model for non-car-owning households and an HL model for car-owning households. The structure of the models is given in Table 2. A model of this form proved very data intensive, and existing data were sufficient only to validate the model for five new stations and to make predictions for a further three potential sites.

A simpler formulation is provided by a single market HL model, as shown in Table 3. The spatial transferability of the HL/MNL model was tested by applying the model to a different data set. A likelihood ratio test showed that the model was not transferable, but this was partly because of problems with the quality of the validation data set.
### TABLE 2

**Market Segmented HL and MNL Models**

<table>
<thead>
<tr>
<th>(A) Non Car Owners</th>
<th>(B) Car Owners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus</td>
<td>Car driver</td>
</tr>
<tr>
<td>Train</td>
<td>Car passenger</td>
</tr>
<tr>
<td></td>
<td>Public Transport</td>
</tr>
<tr>
<td></td>
<td>Bus</td>
</tr>
<tr>
<td></td>
<td>Train</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
<th>(t-stat)</th>
<th>Value</th>
<th>(t-stat)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASC-passenger</td>
<td>-0.844 (-1.305)</td>
<td>ASC-passenger</td>
<td>-0.339 (-0.596)</td>
</tr>
<tr>
<td>ASC-bus</td>
<td>0.427 (1.004)</td>
<td>ASC-driver</td>
<td>1.597 (2.789)</td>
</tr>
<tr>
<td>Wait time</td>
<td>-0.090 (-2.630)</td>
<td>IVT</td>
<td>-0.064 (-3.178)</td>
</tr>
<tr>
<td>Walk time</td>
<td>-0.071 (-2.335)</td>
<td>OVT</td>
<td>-0.059 (-1.481)</td>
</tr>
<tr>
<td>IVT</td>
<td>-0.029 (-1.339)</td>
<td>Total cost</td>
<td>-0.013 (-4.176)</td>
</tr>
<tr>
<td>Availability</td>
<td>-3.012 (-4.643)</td>
<td>EMU</td>
<td>0.377 (4.996)</td>
</tr>
<tr>
<td>Adjusted rho squared</td>
<td>0.500</td>
<td>Adjusted rho squared</td>
<td>0.803</td>
</tr>
<tr>
<td>Number of observations</td>
<td>173</td>
<td>Number of observations</td>
<td>721</td>
</tr>
</tbody>
</table>

#### (i) Upper split

#### (ii) Lower split

| IVT-train | -0.111 (-1.785) |
| IVT-bus | -0.118 (-2.605) |
| Walk time | -0.191 (-3.998) |
| Wait time | -0.276 (-2.565) |
| Total cost | -0.067 (-2.196) |
| Adjusted rho squared | 0.574 |
| Number of observations | 97 |

---

ASC = Alternative specific constant

IVT = In-vehicle time

OVT = Out-of-vehicle time

EMU = Expected maximum utility

\[ EMU = \ln \sum \exp(U_j) \]

where \( U_j \) = utility of mode \( j \).

It is not intended in this paper to comment in detail on the models in Tables 2 and 3, other than to say that their quality was limited by the quality of the RP data set used for calibration. It should also be noted that our HL models have no attributes that are common to the upper and lower splits. For example, many HL models assume that bus and rail fares are the same, and therefore cost is only entered in the upper split as a common attribute.

#### 3.4. The “aggregation issue”

A problem in the use of disaggregate models in forecasting is that the individual choice estimates have to be expanded over the population of interest (=catchment
area) in order to obtain a reliable, unbiased forecast of group behaviour. The difficulty is that, for non-linear functions such as the logit model, the function of averages of variables is not the same as the average of functions. Therefore use of aggregate data with an MCM calibrated with disaggregate data will lead to systematic biases (Westin, 1974).

A method of using aggregate data that is claimed to reduce this problem is the incremental logit model (Kumar, 1980). This model has been extended so as to incorporate an HL structure and accommodate new modes (Koppelman, 1983; Bates et al., 1987); it is then termed the Extended Incremental Logit model (EIL). A form of such a model, where rail has a zero share in the “before” situation, is given by Equations (4) to (7).

Public transport's share in the upper split is given by

\[
P'_{PT} = \frac{P_{PT} \cdot [\exp(U'_{NT} - U_{XT}) + \exp(U'_{XT} - U_{XT})]^{\Phi}}{P_{PT} \cdot [\exp(U_{XT} - U_{XT}) + \exp(U_{XT} - U_{XT})]^{\Phi} + [1 - P_{PT}]} \tag{4}
\]

where

\(P'_{PT}(P_{PT}) = \) proportion choosing public transport in the after (before) situation

\(U'(U) = \) utility measure in the after (before) situation

\(XT = \) existing public transport mode (bus)

\(NT = \) new public transport mode (rail) and

\(\Phi = \) EMU parameter

The lower split shares are then:

\[
P'_{NT} = \frac{\exp(U'_{NT} - U_{XT})}{\exp(U'_{NT} - U_{XT}) + \exp(U_{XT} - U_{XT})} \cdot P'_{PT} \tag{5}
\]

and

\[
P'_{XT} = \frac{\exp(U'_{XT} - U_{XT})}{\exp(U'_{NT} - U_{XT}) + \exp(U_{XT} - U_{XT})} \cdot P'_{PT} \tag{6}
\]
If it is assumed that there is no change in the utility of the existing public transport mode, \( \exp(U_{iTR} - U_{STR}) \) in Equations (4) to (6) simplifies to 1.

The share for all other modes \( P' \) is then estimated as:

\[
P' = P - \frac{1 - P_{PR}}{1 - P_{PT}}
\]  

(7)

The main advantage of the EIL is that it reduces the data requirements of an MCM so that the only information required is the modal shares \( P_M \) and \( P_{XTR} \) and the change in utility \( (U_{iTR} - U_{STR}) \). Given a suitable zoning system, this information may be obtained, for the journey to work, from the Census and from engineering times and costs, respectively.

4. STATED INTENTION/PREFERENCE APPROACH

An alternative approach to forecasting the demand at a new station is simply use market research to ask the question: "If a new station were opened at \( i \) level of service \( Q \), how often, and for what journeys, would you use it?" This approach, which we have termed the Stated Intentions (SI) approach, was, in the essence, used in other studies to forecast usage at new stations in Scotland and Somerset. But it is likely, unless adjusted, to lead to an overestimate of demand, as it is prone to a number of biases:

(i) Self-selectivity bias. In a self-completion survey, potential rail users are more likely to make a response than non-users. To adjust for this bias, Heggie and Papoulas (1976) propose that non-respondents should be treated as non-users of the new facility.

(ii) Non-commitment bias. Respondents are not committed to behaving in the way they have said they will. Although respondents genuinely believe they will use the new service, in actuality they will not. This may be compounded by misinformation and misperceptions. For example, when respondents actually come to use the service they may find the timings inconvenient, or the trains overcrowded or unreliable.

(iii) Policy response bias. Respondents may answer strategically in order to achieve the desired policy response (for example, to get the new station opened).

Thus work by Couture and Dooley (1981) showed that in the case of a new transit system in Danville, Illinois, this simplistic approach resulted in a ratio of intended to actual users of three to one. In a study of South Wigston station (Leicestershire), assuming non-respondents to be non-users reduced the bias, but the predicted usage was still about 56 per cent higher than actual usage.

A more sophisticated approach is based on Stated Preference (SP) approaches. In fact, SP is a generic term which has been applied to a number of methods of analysing hypothetical choices (see Benjamin and Sen (1982) for a comparison of four techniques). A number of studies have compared SP and RP methods (see, for example, Louviere et al., 1981). The main advantages of SP methods, as far as this paper is concerned, are that they can assess new facilities and that they are
more data-efficient than RP methods (that is, they may provide a cheaper way of calibrating a model). Their main disadvantage is that their use in forecasting is still in its infancy.

Typically, SP forecasting experiments develop models of mode choice and then apply these models to O/D matrices of bus and car trips to determine how many trips will divert to rail. This was essentially the approach used in unpublished studies that forecast the patronage of new rail services in Edinburgh and Staffordshire. The main problem with this approach is that, like the MCMs discussed in section 3.3, it does not accommodate generated trips; it will also be affected by the aggregation biases discussed in section 3.4 and the scale factor problem that will be discussed in this section.

A similar approach has been developed in the study of a new rail service between Leicester and Burton-on-Trent (Preston and Wardman, 1988), and has been refined in a study of a service between Nottingham and Worksop. The work consists of two stages. In the first stage an SI survey is undertaken, which collects information on socio-economic characteristics and existing trip patterns as well as in intended use of rail. In the Leicestershire study the SI survey involved distribution of questionnaires to all households within 800 metres and to one in four households within secondary catchment areas (normally 800 metres to 2 kilometres, but this was extended in two cases where it was thought the station could serve communities just beyond the 2 kilometres threshold). Altogether, 29,873 questionnaires were distributed and 4,820 were returned, representing a response rate of 16 per cent.

However, it is known that the SI survey will overestimate usage, unless adjusted. So in a second stage an SP experiment was undertaken to determine the appropriate adjustment factors. This survey focussed on a sub-sample of existing bus and car travellers to central Leicester from the outer Leicester suburbs and the Ashby/Coilville area. Altogether, 1,254 individuals were recontacted, of whom 638 (51 per cent) returned questionnaires. Individuals were presented with 16 paired comparisons of hypothetical times and costs for train and for their existing mode, and were asked to indicate which mode they would use. Binary logit models were calibrated by maximum likelihood, and the BLOGIT package was again used. The resultant models of mode choice are given in Table 4. In comparison with the RP mode choice models that have been developed in this work, the parameter values of the SP models have greater statistical significance (though this is, in part, because there are repeat observations for each individual respondent); socio-economic variables can be explicitly included; and, because of the orthogonal design, these models avoid the problems of collinearity between times and cost which appear to have affected the RP models in Tables 2 and 3.

However, there is a problem with SP models that becomes apparent in the forecasting stage. The models are based on binary logit, the coefficients of which are estimated in units of residual deviation: that is, they are estimated as $\Omega a_i$, where $a_i$ is an unscaled parameter, $\Omega$ is a scalar calculated as $\pi / (\sqrt{3} \sigma)$ and $\sigma$ is the standard deviation of the error differences between modes. A problem arises where $\sigma$ is not what it would be in the actual choices being made. This is likely to happen in SP experiments, as uncertainties/difficulties that respondents have with hypothetical choices are likely to be greater than with actual choices. This in
### Table 4

**SP Models of Mode Choice**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Bus-Train Model</th>
<th>Car-Train Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>(t-stat)</td>
</tr>
<tr>
<td>ASC (Train)</td>
<td>-0.086</td>
<td>(10.21)</td>
</tr>
<tr>
<td>IVT</td>
<td>-0.067</td>
<td>(9.84)</td>
</tr>
<tr>
<td>OVT,</td>
<td>-0.067</td>
<td>(9.84)</td>
</tr>
<tr>
<td>OVT,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COST</td>
<td>-0.056</td>
<td>(19.61)</td>
</tr>
<tr>
<td>FREQ,</td>
<td>+1.327</td>
<td>(15.73)</td>
</tr>
<tr>
<td>FREQ,b</td>
<td>-0.863</td>
<td>(3.34)</td>
</tr>
<tr>
<td>MALE</td>
<td>+0.339</td>
<td>(3.70)</td>
</tr>
<tr>
<td>INCOME &gt; £9,999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE &gt; 39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LEISURE</td>
<td>-0.189</td>
<td>(1.88)</td>
</tr>
<tr>
<td>LEIC SUBURBS</td>
<td>-1.022</td>
<td>(7.64)</td>
</tr>
<tr>
<td>Adjusted rho squared</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>No. of observations</td>
<td>2549</td>
<td></td>
</tr>
</tbody>
</table>

OVT, denotes OVT train

OVT, denotes OVT car

FREQ, and FREQ,b represent the number of trains and buses per hour.

- The deterministic method assigns an individual to the mode with the highest utility, given the estimated weights and the costs and times which would prevail for train and bus/car in the situation to be forecast.
- The probabilistic method calculates the probability of choosing train for each individual, given the estimated utility differences for the situation to be forecast. Aggregate shares are simply the weighted sum of individual shares.

The two methods give different results. In binary choice, because of the shape of the logit function, if rail is the minor mode (that is, if its share is less than 0.5, as in most choice situations) the probabilistic forecast will normally be greater than the deterministic. Where rail is the major mode, the reverse will normally be true. The deterministic method has the advantage that, because the scale factor applies equally to all coefficients and so does not affect relative utilities, the scale factor problem is avoided. The disadvantage is that the deterministic method, by definition, does not include the stochastic component of random utility (that is, the error term). Clearly, this problem requires further theoretical and practical investi-
DEMAND FORECASTING FOR NEW LOCAL RAIL STATIONS

J. Preston

gation. Given that the probabilistic forecasts overstate the market share of rail and that the deterministic forecasts understate it, it may be assumed that these two forecasts bound the actual value.

5. COMPARISON OF RP METHODS — WEST YORKSHIRE CASE STUDY

In this section some validation tests are carried out, based on information on the actual usage of six new stations that were opened in West Yorkshire between 1982 and 1984. Of these stations, Bramley and Deighton serve suburban estates; Crossflatts, Saltaire and Slaithwaite serve mixed communities on the fringes of the conurbation; and Fitzwilliam serves a free-standing mining village.

5.1. Comparison of disaggregate forecasts

For five of the six station sites, three disaggregate approaches may be compared, both with each other and with some actual observations:

(i) Use of the market segmented MNL/HL model (Table 2) with disaggregate data, aggregated by Sample Enumeration (SE). SE simply averages the choice probabilities of a random sample of individuals in the prediction group. A data set was provided by a household survey carried out by the former West Yorkshire County Council in 1981. The forecasts are given in column A of Table 5.

(ii) Use of the market segmented MNL/HL model with aggregate data by using the EIL formulation. The data were the same as in (i), aggregated to form zonal averages. This is given in column B of Table 5.

(iii) Use of the single market HL model (Table 3) with aggregate data. The data were based on census estimates of model shares and engineering estimates of zonal times and costs. This is given in column C of Table 5.

It should be stressed that all three methods only forecast the market share for work journeys to/from the area within 800 metres of new stations.

The first method, given by column A, is accurate in what it does (but this falls far short of producing a total figure of station usage); its forecast shares are, on average, within 17 per cent of actual shares. The second method, given by column B, leads to a deterioration in accuracy because it uses aggregate data. Forecast shares are now, on average, only within 24 per cent of actual shares. This estimate of a 7 per cent worsening due to aggregation is in line with the findings of Talvitie et al. (1980) but, because of the EIL formulation, is less than that in many other studies (Westin, 1974; Reid, 1979). The third method, given by column C, leads to a large deterioration in accuracy, partly because the model that is used is less powerful, but mainly because of the low quality of the data set used. Forecast shares are now only within 69 per cent of actual shares, on average. Comparison with column A might suggest that, in very broad terms, around 25 per cent of this error is due to specification error. Comparison with columns A and B might
TABLE 5

Market Share of Mechanised Work Journeys to/from 800 m of New Stations to/from Existing Rail Station Catchment Areas

<table>
<thead>
<tr>
<th></th>
<th>A MNL/HL Model</th>
<th>B HL Model</th>
<th>C EIL Engineering data</th>
<th>D “Actual”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reported data</td>
<td>EIL data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deighton</td>
<td>0.144</td>
<td>0.185</td>
<td>0.343</td>
<td>0.131</td>
</tr>
<tr>
<td>Crossflatts</td>
<td>0.139</td>
<td>0.156</td>
<td>0.297</td>
<td>0.137</td>
</tr>
<tr>
<td>Slaithwaite</td>
<td>0.153</td>
<td>0.142</td>
<td>0.121</td>
<td>0.147</td>
</tr>
<tr>
<td>Saltaire</td>
<td>0.199</td>
<td>0.201</td>
<td>0.212</td>
<td>0.329</td>
</tr>
<tr>
<td>Bramley</td>
<td>0.219</td>
<td>0.209</td>
<td>0.383</td>
<td>0.236</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.059</td>
<td>0.064</td>
<td>0.148</td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>0.171</td>
<td>0.239</td>
<td>0.686</td>
<td></td>
</tr>
</tbody>
</table>

For definition of RMSE and AD see Appendix 2.

suggest a further 10 per cent is due to aggregation error. The remaining 65 per cent is probably due to measurement error.

5.2. Comparison of aggregate and disaggregate forecasts

The accuracy of four different forecasting methods may be tested for all six new stations:

(i) The West Yorkshire TRM, as given by Table 1. This is given by column A.
(ii) The ASM as given by Equation (1). This is given by column B.
(iii) The HL/MNL model, aggregated by SE. This method produces forecasts for work trips only. Non-work trips are estimated by Equation (3). This is given by column C.
(iv) The HL model, aggregated by EIL. Non-work trips are again estimated by Equation (3). This is given by column D.

Table 6 shows that the TRM produces a forecast that is, on average, within 42 per cent of initial usage, with an RMSE of 71 trips. It is, however, only a very simplistic approach, and is only presented here as a counterpoint to the more sophisticated approaches that have been developed.

Of the three remaining approaches the most accurate, at least initially, is the HL/MNL model, which gives predictions, on average, within around 34 per cent of initial usage.

The predictions of the HL model were estimated to be within 54 per cent of initial usage, with an RMSE of around 97 trips. This approach is only slightly more accurate than the ASM, which gave predictions some 63 per cent above initial usage, with an RMSE of around 108 trips.
TABLE 6

Forecast Weekday Usage of New Stations (Initial Actual Usage Indexed as 1.0)

<table>
<thead>
<tr>
<th></th>
<th>2nd Year Usage</th>
<th>3rd Year Usage</th>
<th>Aggregate</th>
<th>Disaggregate or hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A TRM</td>
<td>B ASM</td>
<td>C HL/MNL</td>
<td>D HL with EIL</td>
</tr>
<tr>
<td>Fitzwilliam</td>
<td>1.56</td>
<td>1.75</td>
<td>1.39</td>
<td>1.31</td>
</tr>
<tr>
<td>Deighton</td>
<td>1.98</td>
<td>1.91</td>
<td>1.76</td>
<td>2.72</td>
</tr>
<tr>
<td>Crossflatts</td>
<td>1.52</td>
<td>1.63</td>
<td>0.80</td>
<td>2.32</td>
</tr>
<tr>
<td>Slackwaite</td>
<td>0.73</td>
<td>1.11</td>
<td>0.67</td>
<td>1.22</td>
</tr>
<tr>
<td>Bramley</td>
<td>0.89</td>
<td>0.86</td>
<td>1.47</td>
<td>1.35</td>
</tr>
<tr>
<td>Slaterdale</td>
<td>1.25</td>
<td>1.48</td>
<td>0.58</td>
<td>1.49</td>
</tr>
<tr>
<td>RMSE — Initial usage</td>
<td>71.3</td>
<td>108.3</td>
<td>78.6</td>
<td>96.9</td>
</tr>
<tr>
<td>— Second year usage</td>
<td>73.5</td>
<td></td>
<td>93.9</td>
<td>85.8</td>
</tr>
<tr>
<td>— Third year usage</td>
<td>64.9</td>
<td></td>
<td>105.8</td>
<td>93.6</td>
</tr>
<tr>
<td>AD — Initial usage</td>
<td>0.422</td>
<td>0.630</td>
<td>0.343</td>
<td>0.540</td>
</tr>
<tr>
<td>— Second year usage</td>
<td>0.377</td>
<td></td>
<td>0.371</td>
<td>0.312</td>
</tr>
<tr>
<td>— Third year usage</td>
<td>0.259</td>
<td></td>
<td>0.370</td>
<td>0.316</td>
</tr>
</tbody>
</table>

With the exception of the TRM (which was calibrated on figures of first-year usage), the models examined in Table 6 are forms of equilibrium models. From count data, it is apparent that the use of new stations has been growing in absolute terms. However, this is against a background of increasing use of rail in West Yorkshire; between 1982 and 1986, demand at 38 existing local stations increased by 48 per cent. In Table 6, second and third year usage figures at six new stations are expressed in relation to the overall increase in demand for rail services as a whole. It can be seen that, except at one station, demand has grown over the first three years at a faster rate than in the network as a whole. Initially, this trend was extrapolated over five years, with the result that demand in year 5 was estimated to be 75 per cent greater than in year 1. Later work suggests that real growth at new stations occurs only in the first three years, so that usage in year 3 is 35 per cent higher than in year 1. Table 6 shows that if these dynamics are taken into account the accuracy of the three equilibrium models is broadly comparable in year 2, but by year 3 the ASM appears to be the most accurate, and that the forecasts are then, on average, within 26 per cent of actual usage. The comparable figures for the MNL/HL and the HL models are 37 and 32 per cent respectively. The ASM’s better performance over time is probably related to its ability to make better forecasts of generated trips, particularly for work journeys.

6. COMPARISON OF RP AND SP METHODS — LEICESTERSHIRE CASE STUDY

A new rail service between Leicester and Burton-on-Trent has been proposed, serving Leicester’s western suburbs and the coal mining area of North West
Leicestershire centred on Ashby and Coalville. For assessing potential demand a number of different forecasting approaches have been compared, though it is not possible to assess their accuracy because the service has not yet been opened. The approaches compared were:

(i) The ASM, again based on Equation (1), but adjusted for findings in South Wigston, where it was found that the model underpredicted demand by 58 per cent.
(ii) The TRM for South Wigston, as given by Table 1.
(iii) The results of the SI survey, assuming that non-respondents would not use the service.
(iv) The results of the SI survey, amended in the light of the SP experiment. This was done by using the SP models in Table 4 with reported time and cost data to predict individuals’ mode choice. The result was then compared with what individuals said they would do in the SI survey.

The methods are compared in Table 7. It can be seen that the lowest demand forecast is provided by the ASM. In part, this reflects problems with transferability, particularly on the measure of attractiveness of destinations, even though the forecasts were adjusted in the light of findings at South Wigston. It may also indicate that a model calibrated for existing rail services may not be appropriate for a new service. The main problem is that the model was calibrated for stations serving small communities. Thus catchment areas are specified as being highly localised. In reality, their size will not be uniform but will vary from site to site. The assumption of localised catchment areas is appropriate for new stations on existing services, as all medium/large communities can be expected to have stations already. Yet new stations on new services may well serve medium/large communities. Experience in validating the ASM in West Yorkshire suggests that the model will vastly underestimate use of the line from such sites.

Usage of the service predicted by the South Wigston TRM is 49 per cent higher than that forecast by the ASM. As expected, the SI survey indicated that usage would be high — over three times the level predicted by the ASM. However, the SP models did suggest that demand would be lower, between 61 and 94 per cent of that given by the SI survey, with the probabilistic forecasts much higher than the deterministic.

Table 8 illustrates that a further way in which the different RP and SP approaches may be compared is through a study of their implied elasticities and

### TABLE 7

<table>
<thead>
<tr>
<th>ASM S. Wigston</th>
<th>TRM S. Wigston</th>
<th>SI Survey</th>
<th>SI Survey Adjusted by SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.49</td>
<td>3.07</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.38</td>
</tr>
</tbody>
</table>

DF = Deterministic forecast.
PF = Probabilistic forecast.
values of time. It should, though, be realised that it is not always possible to compare like with like. This is particularly true for the price elasticity measure.

Deriving price elasticities from the ASM was relatively straightforward, resulting in an elasticity of −0.83 for the log linear formulation (Equation (1)) and −0.65 for the semi-log formulation (Equation (2)). These are within the range used by BR for its Provincial services. The derivation of elasticities from HL models is more complicated but is still feasible (see, for example, McFadden, 1979, pp. 313—316). The single market HL model gives a price elasticity of −0.34; but this is for work journeys only, and is typical of the price elasticities found for commuter rail journeys. The derivation of elasticities from an SP model is more problematical, partly because of the scale factor problem. Elasticity values were derived by examining the effect of 10 per cent increases in fares on the deterministic and probabilistic forecasts. It can be seen from Table 8 that, in this case, the deterministic elasticities had greater absolute values than the probabilistic elasticities. The mean of the two elasticity measures was used for further analysis. This may again be justified, since, given that rail is the minor alternative, the scale factor problem is likely to mean that absolute point elasticities based on the probabilistic forecasts are underestimates and those based on deterministic values are overestimates. This mean value, −1.75, may seem unrealistically elastic, but it is affected by the low values of time (see below) and by the fact that all users have an alternative to rail that they are now using. It may also be related to problems inherent in an SP approach.

A value of in-vehicle time could not be determined from the ASM, so instead the Department of Transport’s standard behavioural value at the time of the study (2.0 pence per minute at 1987 prices) was used in the generalised cost formulations. The derivation of values of time from an HL model is problematical, especially as the parameter value of car in-vehicle time was insignificant in the upper split model. In Table 8 the only value presented is that of rail in-vehicle time in the lower split of the model. This is valued at 2.4 pence per minute. The derivation
of values of time from the binary logit used for the SP model is straightforward, and results in values of 1.5 pence per minute for bus users and 1.8 pence per minute for car users (1987 prices). By contrast, the Department of Transport value has been revised upwards to around 3.2 pence per minute (mid 1987 values). So, the SP derived values in Table 8 appear to be more in line with the old rather than the current Department of Transport values.

7. CONCLUSIONS

This study has identified and developed a wide range of techniques that can be used to forecast the demand for new rail stations and services. In comparing these techniques, it has not been possible to replicate perfect laboratory conditions in which external effects can be controlled for. Comparisons can only be made at a practical level. It has not been the aim of this paper to discuss in detail the models developed, but it should be clear that the approaches we have compared are based on different model formulations (with differing degrees of specification error) and are calibrated with different data sets (with differing degrees of measurement error).

Despite these caveats, a trade-off between accuracy and complexity (or cost) has been identified. All other things being equal, the simplest (or cheapest) and least accurate method is likely to be the TRM, and the most accurate and complicated (or expensive) methods are likely to be provided by disaggregate MCMs. Unless they have already been collected for some other purpose, RP data are likely to be more expensive than SP data. The ASM is somewhere in between the TRM and the disaggregate MCMs, though it is more akin to the former.

Disaggregate methods do appear to be accurate in what they do; but, in this study, this was limited. For example, the MNL/HL model did seem to predict accurately the initial share of rail in the journey-to-work market at six new stations. However, owing to data constraints these forecasts were only produced for the 0–800 m catchment area, and it was not possible to develop comparable models for non-work journeys. Moreover, because the model was based only on mode split and did not cover trip generation and distribution, the forecasts appear to have become less accurate over time. Generally, it appears that the practicality of disaggregate approaches at the forecasting stage is limited by availability of data, by aggregation problems and by difficulties in interpreting parameters. In particular, the econometric implications of the use of HL formulations and of SP data require further investigation.

The conclusion of this study is that forecasting models need to be tailored to the situation they are forecasting. The TRM may be adequate for a “sketch” planning assessment of a very cheap investment (for example, a single new station). The ASM appears to be adequate for the assessment of relatively cheap investments (such as a series of new stations), particularly if it can be calibrated locally (and computerised ticketing and geographic information systems should make this increasingly possible). Disaggregate techniques should be used for major investments. Choice of method will depend crucially on availability of data. SP methods have to be used in situations where suitable RP data do not exist: for example,
where there are no existing rail services or where a completely new mode, such as light rapid transit, is being introduced.

RP and SP approaches should not, however, be thought of as mutually exclusive; given the scale factor problem, it may be sensible in some situations to combine the two approaches. Similarly, the different approaches that we have developed may, in some instances, be applied in a complementary manner. For example, for a major scheme, the TRM or ASM might be used to examine, at a broad level, a wide range of options. A disaggregate modelling approach (probably involving fresh survey work) might then be developed to examine selected options in detail.

Lastly, it should be emphasised that this study has concentrated on the issue of forecasting the usage of a new facility. What in fact is normally required is an evaluation of this facility. For a financial evaluation, the ASM can predict revenue and so is adequate. However, if a social evaluation is to be carried out and user and non-user benefits need to be calculated, some form of disaggregate modelling system would be more appropriate.

APPENDIX 1

The Aggregate Simultaneous Model (ASM): definition of variables

(A) All purposes model

$L$ = denotes a logarithm has been taken.

$FLOW$ = Number of trips from $i$ to $j$ and $j$ to $i$ per average autumn weekday.

$OPOP$ = Population usually resident within a straight line distance of 800 metres of the station.

$OPOP2$ = Population usually resident within a straight line distance between 800 metres and 2 kilometres of the station.

$RSOC$ = Number of residents within social classes 1 and 2 within 800 metres of the station divided by $OPOP$.

$DRX$ = Number of work places within 800 metres of the destination station divided by the economically active population.

$GCOTH$ = Index of competition, expressed as:

$$GCRA/(GCRA + GCRB + GCCA) \quad \text{(Equation (1))}$$

$$GCRB + GCCA \quad \text{(Equation (2))}$$

where

$GCRA$ = Generalised time for rail = $2 \times (\text{walk} + \text{wait time}) + \text{in-vehicle time} + \text{fare} / VOT$.

where

$Walk$ = Access and egress time.
Wait = Calculated as a function of headway = 3.0 + 0.185 service interval.

Fare = Half standard return.

VOT = Department of Transport value of behavioural non-working in-vehicle time.

GCBU = Generalised time for bus = 2 × (walk + wait time) + in-vehicle time + fare/VOT.

where

Walk = Calculated as rail walk time divided by the number of bus stop pairs on competing bus routes within 800 metres of a station.

Wait = Calculated as a function of headway = 1.46 + 0.26 service interval.

GCCA = Generalised time for car = in-vehicle time + operating costs/VOT + parking charge/VOT.

where operating costs are taken as fuel costs only, assuming fuel consumption of 44 km per gallon in urban conditions and 62 km per gallon in rural conditions. In-vehicle time is based on link flow speeds of 46 km per hour in urban conditions and 80 kph in rural conditions.

(B) Non-work trips model

FLOW = Number of non-work trips (excluding education) from i to j and j to i per average weekday.

OPOP = As above.

REMP = Retail employment within the central area shopping zone.

RS = Rail service frequency during off-peak periods (0930–1500 hours and 1800 hours and beyond).

BS = Bus service frequency during off-peak periods.

IC = Dummy variable; = 1 for stations serving medium sized towns, with services timetabled to connect with inter city services. Otherwise = 0.

INTOPP = Proxy variable to take into account the number of competing or intervening variables.

APPENDIX 2

Definition of goodness of fit measures

RMSE (Root Mean Square Error) = \[ \sqrt{\frac{\sum (F - A)^2}{n}} \]

where F = Forecast rail share of journey to work market (Table 5)
Forecast new station daily usage (ons plus offs) (Table 6).

200
DEMAND FORECASTING FOR NEW LOCAL RAIL STATIONS

J. Preston

\[ A = \text{Actual rail share of journey to work market (Table 5)} \]
\[ n = \text{Actual new station daily usage (Table 6).} \]
\[ AD \ (\text{Absolute Deviation}) = \sum |F - A| / \sum A. \]

Date of receipt of final typescript: June 1990.

REFERENCES


