FORECASTING THE DEMAND FOR INTER-URBAN RAILWAY TRAVEL IN THE REPUBLIC OF IRELAND

By H. McGeehan*

Models of inter-urban passenger demand are usually variations of gravity models, which attempt to explain travel between city pairs from cross-sectional data. In general these models are suitable for long-term forecasting, since the main explanatory variables (for example, population) do not vary significantly in the short term.

Rail transport operators require a continuous flow of information on the demand for their product. Policy issues relating to such matters as pricing, revenue, and government support, which are dependent upon the level of rail usage, are of immediate importance. Decisions regarding them must therefore be based, as far as possible, on accurate short-term forecasts of demand.

This paper describes a simple aggregate model which has been developed to forecast short-run changes in inter-urban rail passenger travel on the network which is managed and operated by Coras Iompair Eireann (CIE). The modelling framework has been constructed so that forecasts can be made quickly and cheaply, to facilitate the examination of alternative planning strategies.

The need for this type of modelling framework has been recognised in other areas of transport. For example, similar ideas were applied to estimating short-run demand functions for rail freight by Rao (1978) and for urban bus travel by Mullen (1975). Applications to the inter-city passenger market have been confined mainly to air transport (see Lave, 1972), though some results have been published recently for traffic flows from London to seventeen other urban centres on the British Rail network (Jones and Nichols, 1983).

The results reported in this paper are based on analyses of quarterly data covering the period from the beginning of 1970 to the end of 1982. The estimation technique used was that of ordinary least squares.

The paper is divided into six parts. Part 1 examines the identification problem. The choice of variables in the demand equation is discussed in Part 2. In Part 3 empirical results are presented, and in Part 4 the forecasting ability of the model is examined. Further results are reported in Part 5, and some conclusions are drawn in Part 6.

* Economist, Coras Iompair Eireann. The author is indebted to Chris Nash (Leeds University) and several colleagues in CIE for comments and suggestions on an earlier version of the paper, and to Stephen Hirsch for excellent research assistance. The views and opinions given are the sole responsibility of the author and are not necessarily accepted by the Board and Management of CIE.
1. IDENTIFICATION PROBLEM

The use of an OLS estimation technique requires some justification in the light of the discussion in Jones and Nichols (1983) on the identification problem. Without pre-empting later discussion, we may state that the demand for interurban rail travel in Ireland has been characterised, over the study period, by a burgeoning trend in passenger carryings and a low price elasticity. This may mean that real fares and service levels are endogenous, as they may be adjusted by the operator in response to changes in passenger demand. Thus passenger miles, service level and real fare will be jointly determined, and use of a single equation estimation technique will yield biased and inconsistent estimates.

However, the framework within which CIE operates suggests that variations in fares and service levels are exogenously determined.

Increases in fares are granted through the National Prices Commission (NPC), but CIE must first obtain ministerial approval for the application. In general applications to increase fares are made once a year. In the past there have been considerable delays in processing the applications, and increases have been deferred in many instances. For example, increases were deferred completely in 1972/73. In 1975 they were delayed by four months, in 1979 by six months, and in 1981 by eight months. It was planned to increase fares in 1982 on a two-phased basis, by 15 per cent in January and by 10 per cent in July. The second phase was not implemented because of a government embargo. In fact, in only two years of the study period (1978 and 1980) have fare increases been implemented on time and in full.

Because of the considerable lags between increases in passenger demand and increases in fares, it can be presumed that real fare changes are exogenous and are rather a product of government intervention than a reaction to changes in demand.

Throughout the study period, CIE main line operations have been constrained by a shortage of rolling stock, particularly on Friday evenings, when some rolling stock has been diverted from suburban to main line operation.

In 1973, a high frequency rail timetable was introduced, but after a year's operation it was found difficult to maintain schedules and punctuality, because the rolling stock was old and required considerable maintenance.

For this reason the timetable was revised, because it was accepted by management that it was better to limit the passenger's choice and guarantee accurate timekeeping than to maintain a wide choice with subsequent delays. In essence there was a conflict between maintenance capacity and operational demand, which was resolved in favour of the former.

Since that downward revision in 1974, the level of service (in train miles) has remained virtually constant. With the delivery of eighteen new locomotives in 1977 there has been a slight increase in scheduled train miles, but capacity is still constrained.

Therefore output can be treated as having been exogenously determined during the study period. Consequently the use of a single equation OLS technique in the estimation procedure seems justified.
2. CHOICE OF VARIABLES IN THE DEMAND EQUATION

The demand for inter-urban rail travel is hypothesised to depend upon the money cost of travel by rail, disposable income, car ownership, quality of service and seasonality.

Dependent variable

CIE have a record, in the form of a mileage matrix, of all distances between stations. From a combination of these mileage figures and information on ticket sales, data on passenger miles can be extracted. Return tickets count as two journeys and an appropriate factor, based on CIE surveys, is applied to season tickets.

Independent variables

Rail fare

A serious problem in this type of aggregate analysis is the determination of a suitable fares variable. Usually revenue per passenger mile is taken as a proxy, but it has the defect that most railways apply strong distance tapers, so that the fare charged per mile falls with distance travelled. If there is a change in the passenger mix, for example from longer to shorter journeys, the revenue per passenger mile ratio will increase, not because of an actual change in price, but because of a change in the pattern of travel.

A fares variable was therefore constructed to reflect only "pure price" effects. The measure is a weighted average of fare paid with fixed weights \( W_{ij} \) reflecting the proportion of travellers using each of twenty main ticket types, spread over different origin/destination pairs, to give an average fare for the whole network.

The average fare \( F_t \) in period \( t \) is defined as:

\[
F_t = \sum_{i,j} W_{ij} F_{ijt}
\]

where \( \Sigma W_{ij} = 1 \)

and

\( i = \) origin/destination pair
\( j = \) ticket type.

The base year was 1974, and in that year there were 7,148 origin/destination pairs. A representative sample of 21 origin/destination pairs was used to construct the fares variable. (See Appendix.)

One criticism of this measure is that the proportions travelling in each origin/destination pair and the proportions using each ticket type may not remain fixed over the duration of the study period. A test of the validity of this criticism is given in a later section, where results are reported of using the fares variable to predict revenue.
FIGURE 1

Real Industrial Earnings and Retail Sales Volume
1970 – 1982
FORECASTING THE DEMAND FOR INTER-URBAN RAIL TRAVEL

H. McGeehan

Income

It is hypothesised that the demand for inter-urban rail trips depends on real disposable income. Growth in real disposable income should raise the level of business and leisure travel as a whole. However, higher incomes are associated with higher car ownership, so competition with the railway increases as income increases. The overall effect on rail journeys of changes in real income is uncertain, but on balance it would be positive.

The only readily available measure of income is the Index of Average Industrial Earnings; but some characteristics of the index detract from its suitability for forecasting passenger demand.

Firstly, the index refers to gross income (that is, before deduction of income tax), and probably overstates the purchasing power of consumers in the economy.

Secondly, the index is based on earnings in the manufacturing sector only, and thus excludes the very sizeable services sector (which accounted for more than 50 per cent of GDP in 1980). Services employment is heavily concentrated in the eastern region of the country — particularly around Dublin. The inter-urban network has experienced a burgeoning trend in its weekend market over the study period, and out-migration from Dublin has provided the primary impetus. The omission of the services sector could cause a serious distortion in the earnings index, and so undermine the predictive ability of any demand model.

Therefore a second index, measuring the volume of retail sales, was tested in the regression analysis. The index records actual expenditure (or actual income disposed of) and is not subject to the criticisms inherent in the industrial earnings index. The two indexes are shown in Figure 1. The correlation between them is quite high (r = 0.88), but there are some periods when there is a wide divergence (e.g. between 1975 and 1978).\(^1\)

Car Ownership

Car ownership is expected to have an impact on rail carryings. Rising car ownership increases the competition against the railway, and consequently should have a negative impact on rail demand. However, internal CIE studies suggest that car ownership alone may not be a meaningful measure of competition.

In 1977/78, 51 per cent of rail travellers owned or had the regular use of a car. This suggests that their choice of mode is rationalised not on the basis of ownership, but on the basis of some positive aspect of rail travel (CIE, 1981).

Quality of Service

Positive aspects of rail travel may include speed, comfort and suitability; they are all parameters of the rather imprecise term "quality of service". There are several possible measures which could be used to measure service quality: these include frequency of train departures and journey times.

\(^1\) Jones and Nichols (1983) use GDP as a measure of economic activity in their modelling framework. Unfortunately, that procedure could not be followed in this study, because statistics on GDP are only available on an annual basis and are not up to date.
However, these measures of service quality are difficult to calibrate in the type of model being analysed. For example, journey times (the average time of the fastest trains and the average time for all services) show little variation over a period of ten years. At certain points within that ten-year period there may be substantial changes, but overall, changes tend to be small.

It was considered that train miles operated would give a meaningful measure of service quality. Over the study period the network remained virtually unchanged in terms of route length and route coverage. Consequently, changes in train miles have resulted directly from changes in the frequency of train departures. Increased frequencies reduce waiting time, so that the attractiveness of rail travel is increased, and vice versa. But, as noted earlier, the shortage of rolling stock and its age profile have severely constrained the level of service provided. Over the period train miles have generally changed by small amounts, and this lack of variation makes estimation problematic. In preliminary regressions the variable was found to be insignificant (though positive), and it was dropped from the rest of the analysis.

All this, however, is not to suggest that service quality is an unimportant element in rail travel decisions. As mentioned earlier, a large percentage of rail travellers rationalised their choice of mode on the basis of a positive aspect of rail travel. A cross-section study of the inter-urban network (CIE, 1976) found that speed and frequency were significant explanatory variables. The data were selected at points within the period 1973 to 1976 corresponding to major changes in services and fare.

A deficiency of the type of time series modelling used in the present analysis is that it does not "pick up" service factors in the way that cross-section modelling does. However, since the objective of the study is short-term prediction, this is not a serious problem, because service changes are likely to be minor in the short run.

Seasonality

Inspection of the quarterly data on passenger miles suggested that demand was subject to seasonal variation. In order to allow for this, three dummy variables representing the second, third and fourth quarters have been included in the analysis.

Variable Definitions

The definitions of the variables used in the regression analysis are listed below.

Thus:

\[ PM_t \] = Passenger miles in quarter \( t \);
\[ RF \] = Nominal rail fare deflated by the Consumer Price Index;
\[ Y \] = Index of retail sales (volume);
\[ E \] = Index of real industrial earnings;
\[ C \] = Car ownership;
\[ D_2, D_3, D_4 \] = Seasonal dummies representing the second, third and fourth quarters, respectively.
3. EMPIRICAL RESULTS

The model was originally estimated from data covering the period 1970 to 1979. Only those variables which were significant at the 5 per cent probability level were included in the equations.

Several functional forms of the model were tested (for example, double-log and semi-log), but the best results were obtained from a linear specification. The best equation in terms of fit is shown in equation (1) below; the standard errors of the coefficients are given in parentheses.

\[
PM = 37.55 - 42.93 \quad RF + 0.32Y + 26.01 \quad D_2 + 58.92 \quad D_3
\]

\[
\begin{array}{llll}
(9.68) & (0.04) & (3.76) & (3.75)
\end{array}
\]

\[
R^2 = 0.922 \quad D.W. = 1.56 \quad F = 94.4
\]

Equation (1) explains over 92 per cent of the variation in passenger miles. The \(F\) value indicates that the equation is significant (at the 5 per cent level).

The Durbin-Watson statistic suggests some tendency towards positive autocorrelation. However, the value is located in the 'no decision' range; that is, it is greater than the value for which we accept positive serial correlation, but less than the value for which we accept no serial correlation.

Before analysing the coefficients of equation (1), it is worth noting that the industrial earnings variable (\(E\)) was significant and positive in all regressions, but had less explanatory power than the retail sales variable (\(Y\)). Moreover, and this is relevant to later discussion, the predictive ability of equation (1) was greatly impaired when the retail sales variable (\(Y\)) was replaced by the earnings variable (\(E\)).

The results of including the earnings variable in our preferred equation — to the exclusion of the retail sales variable — are reported in equation (2). In a linear formulation there were considerable problems with autocorrelation, and inspection of the residuals suggested that a non-linear specification would be appropriate. The most satisfactory results were obtained from the double-log function in equation (2).

\[
\log PM = 0.59 - 0.40 \log RF + 0.82 \log E + 0.23 \quad D_2 + 0.46 \quad D_3
\]

\[
\begin{array}{llll}
(0.13) & (0.14) & (0.04) & (0.04)
\end{array}
\]

\[
R^2 = 0.864 \quad D.W. = 1.34 \quad F = 55.59
\]

As mentioned above, the predictive power of equation (2) was much less satisfactory than that of equation (1). Therefore, in the analysis which follows the main emphasis will be on the coefficients obtained from equation (1), but comparisons with equation (2) will be made where appropriate.

Fare

As expected, the coefficient of the real fare variable was negative and highly significant. The coefficient implies a point estimate of the fare elasticity of demand of \(-0.4\) at the mean value of real fare and passenger miles. The demand for inter-
urban travel is therefore price-inelastic; this suggests that total expenditure on rail travel will increase with increases with real fare.

The fare coefficient in equation (2) confirms the inelastic nature of passenger demand. Because equation (2) is in double-log form, the coefficients represent elasticities. Thus the fare elasticity from equation (2) is $-0.40$, which increases our confidence in the fact that demand is price-inelastic.

The elasticity is somewhat lower than that determined for British Rail by Jones and Nichols (1983). Comparisons between different railway networks are less than definitive, but certain factors would indicate that the lower elasticity for CIE was to be expected.

For example, in a recent comparison of selected European railways for 1977 (BR/Leeds University, 1979), BR was found to have the highest mean real fare of 2.026 pence. The equivalent CIE fare was 1.469 pence.\(^2\)

No strategic motorway exists in Ireland, and expenditure allocated to roads is mainly spent on maintenance and minor improvements. Taxation of private cars in Britain is relatively low (BR/Leeds University, 1979, p. 33), and the competition BR faces from private cars is more severe than it is in Ireland, where car taxes are quite high.

CIE is the sole operator of scheduled inter-urban coach services, and therefore the competition facing the railway from road passenger transport is quite small.\(^3\) The licensing system for coach services in Britain, even before deregulation in 1980, affords less protection to BR.

Paradoxically, the mean trip length for CIE (59.5 km) is higher than that for BR (42 km). It is generally believed that price elasticities increase with increases in trip length. That somewhat weakens our findings, but at the same time underlines the vagaries of international comparisons!

**Consumer Expenditure**

The coefficient of the retail sales variable \((Y)\) was found to be positive in all regressions and significantly different from zero at the 5 per cent level. The coefficient implies a point estimate of the income elasticity of \(+0.91\), at the mean value. The elasticity is almost unitary; this suggests constancy in the percentage outlay on rail travel as consumer expenditure increases. The coefficient of the industrial earnings variable \((E)\) in equation 2 confirms this high elasticity.

During the analysis it was found, as expected, that there was a high correlation between the retail sales and car ownership variables \((r = 0.87)\). One way of avoiding this collinearity problem is to omit the car ownership variable from the equation. Two possible consequences can arise. Firstly, specification error may be introduced into the model, since the omission of variables will affect the values of the parameters of the retained variables. Secondly, the expenditure elasticity

\(^2\) The figures relate to overall railway functions and are not disaggregated specifically into the inter-urban function. See also note 4.

\(^3\) There has been, since 1980, an upsurge in illegal private operations, but the extent of their involvement is difficult to gauge.
would refer to a “gross” elasticity which includes the effect of increasing car ownership with expenditure, rather than to the “partial” elasticity derived from our preferred equation.

Equation (3) shows the effect of introducing the car ownership variable into our preferred equation.

\[
PM = 38.3 - 45.21 RF + 0.35 Y + 26.22 D_2 + 59.28 D_3 - 0.03 C
\]

\[
(10.47) \quad (0.08) \quad (3.81) \quad (3.83) \quad (0.05)
\]

\[R^2 = 0.916 \quad D.W. = 1.69 \quad F = 74.3\]

As expected, the standard error of the expenditure variable increased considerably with the addition of the car ownership variable. However, the variable retained its significance and its coefficient is virtually unchanged – as are the coefficients of the other variables from our preferred equation. Consequently, the coefficients in equation (1) are extremely stable, and this increases our confidence in the estimates of the expenditure and price elasticities. The car ownership variable was therefore omitted from the analysis.

The estimated income elasticity is higher than the limited number of comparable estimates that have been published for other railway networks (BR/Leeds University, 1979). They range from 0.3 to 0.7 – although a composite elasticity for inter-city and commuting for Netherlands railways (NS) ranges from 0.59 to 0.93, and the upper figure is quite similar to that for CIE.

It is significant that Ireland has the lowest income per head and the lowest level of car ownership per head of all the countries included in the BR/Leeds University study. In 1977 real gross domestic product per head (in PPP sterling)\(^4\) stood at £1,097, and car ownership per head at 0.175.

With such a low level of car ownership, increases in income are likely to increase expenditure on transport in general, and in public transport in particular. During the 1970s real domestic output and real income grew at fairly high average rates of 4.3 per cent and 5.5 per cent respectively, providing a substantial base for growth in rail passenger patronage.

The age distribution of the population in Ireland has shifted significantly in the last decade towards a larger percentage of younger people, who are less likely to have private cars available for trip-making. This, coupled with the primacy of the capital city Dublin as the destination for and origin of those moving within the country, probably explains the burgeoning trend in peak week-end travel, particularly out of the eastern region (Dublin and county) on Friday evenings.

In 1979, approximately 29 per cent of CIE passengers travelled on weekend return tickets – making that fare the most popular in the range of tickets sold. This type of leisure-oriented market is likely to be highly sensitive to changes in disposable income; it may partially explain the high income elasticity derived in the study.

\(^4\) All financial data in the BR/Leeds study is in sterling at purchasing power parity (PPP) exchange rates. The PPP exchange rate for Ireland in 1977 was 0.973 Irish pound equal to one pound sterling.
Car Ownership

The coefficient of the car ownership variable proved to be negative but insignificant in the regressions. This is not very surprising in view of the observation made earlier that a substantial number of rail travellers rationalised their choice of mode on the basis of some positive aspect of rail travel, rather than on car ownership alone.

It was also considered that car owners might have based their decision to travel by rail on the perception that rail travel was cheaper than the car mode. In order to test for this, a variable measuring the cost of motoring relative to rail fare was included in the demand equation. The coefficient, though negative, proved to be not significantly different from zero at even the 10 per cent level.

The lack of significance of the relative cost coefficient may be due to several factors. Firstly, over the period analysed the nominal motoring cost index increased in line with nominal rail fares. Thus the relative index did not change significantly over the period. Trip-makers would have found it difficult to perceive any change in the relative cost attractions of the two modes, so this factor would not have influenced them significantly in their modal choice. The virtual constancy of the relative cost index prevailed during a period when rail passenger miles almost doubled and car occupancy rates remained static.

Secondly, business travellers who may have a car available may be immune to any cost considerations, since travel is funded by the employer. Hence qualitative factors such as comfort and convenience would be decisive in their modal choice rationale.

Thirdly, daily commuters using the inter-urban main line network, particularly into the Dublin metropolis, would consider factors such as congestion and parking charges as important (neither of which can be taken into account here).

Finally, of course, a large percentage of rail travellers do not have the alternative car mode available, so the relative cost variables are irrelevant to them.

Seasonality

The coefficients for the seasonal dummies representing the second and third quarters were found to be positive and highly significant. The model confirms that the demand for rail travel in the period April to September is substantially in excess of that for the remainder of the year, primarily because of the impact of leisure travel.

The seasonal dummy for the fourth quarter was found to be negative but not significant.

---

5 The index of motoring costs was obtained from the Central Statistics Office. The index includes items such as the cost of petrol and motor oil, sparking plugs and miscellaneous parts, and represents a reasonably accurate reflection of the variable costs involved in using a car.

6 If car occupancy rates had increased, the cost of travelling by car would have decreased — assuming cost sharing among the occupants. Given the constancy of car occupancy rates, our motoring cost index reflects more accurately the incidence of price changes.
TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>(1) Actual</th>
<th>(2) Predicted (ex-ante)</th>
<th>(3) Predicted (ex-post)</th>
<th>(4) Error (1 minus 3)</th>
<th>(5) 4 as % of 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger miles</td>
<td>570</td>
<td>567</td>
<td>565</td>
<td>5</td>
<td>0.88</td>
</tr>
<tr>
<td>Revenue (£M)</td>
<td>20.64</td>
<td>20.68</td>
<td>20.61</td>
<td>0.03</td>
<td>0.15</td>
</tr>
</tbody>
</table>

4. FORECASTING ABILITY OF THE MODEL

The primary aim of the analysis was to derive a model capable of predicting passenger demand in the short run. An important test, therefore, of our preferred equation is its ability to forecast outside the sample period.

Obviously the accuracy of any predictions will be conditional upon the accuracy with which the independent variables are predicted. Accordingly Table 1 shows two sets of predictions for 1980. The ex-ante predictions were made on the basis of forecasts of the independent variables. These were made in early 1980. The ex-post predictions, on the other hand, were made in early 1981, when the actual values of the independent variables were known. Any divergence between the two sets of predictions merely reflects the inaccuracy of the forecasts of the independent variables.

Predictions for total revenue are also shown in Table 1; they were determined by multiplying the predicted passenger miles by the average fare determined from our stratified sample. The accuracy of the revenue predictions, therefore, can be considered as a test of the validity of our measure of average fare.

The model performs exceptionally well in predicting values for 1980. The error between the ex-post prediction of passenger miles and the actual total is less than 1 per cent. The error for the corresponding revenue totals is only slightly greater than zero.

The ex-ante and ex-post predictions are also very similar; this indicates that the forecasts of the independent variables were close to their actual values. The accuracy of the revenue prediction tends to suggest that our measure of fares is a good one.

The forecasting ability of equation (2) was less than satisfactory. The ex-post prediction was that demand would total 530 million passenger miles in 1980. The forecasting error was of the order of −7 per cent. Inspection of Figure 1 shows that there was a divergence between industrial earnings and retail sales throughout most of 1980. The difference between the indexes is primarily due to the fact that the earnings index relates to gross pay; but it is also due in part to fluctuations in savings rates. In 1980 there was a fall in the savings rate of about 3 per cent, which would not have been picked up by the earnings index. This suggests that any forecasting model based on disposable income may be unsatisfactory because of unforeseen fluctuations in the savings rate.
### TABLE 2

*Estimates of Demand for Inter-Urban Travel*

<table>
<thead>
<tr>
<th>Equation</th>
<th>Period</th>
<th>Constant</th>
<th>Real Fare</th>
<th>Real Income</th>
<th>2nd Qtr. Dummy</th>
<th>3rd Qtr. Dummy</th>
<th>Strike Dummy</th>
<th>$R^2$</th>
<th>F</th>
<th>DW</th>
<th>Price Elasticity*</th>
<th>Expenditure Elasticity*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) 1970–79</td>
<td>37.55</td>
<td>−42.93</td>
<td>0.32</td>
<td>26.01</td>
<td>58.92</td>
<td>—</td>
<td>0.92</td>
<td>94.4</td>
<td>1.56</td>
<td>0.40</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>(4) 1970–80</td>
<td>33.71</td>
<td>−40.94</td>
<td>0.32</td>
<td>25.40</td>
<td>59.21</td>
<td>—</td>
<td>0.92</td>
<td>108.2</td>
<td>1.56</td>
<td>0.38</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>(5) 1970–81</td>
<td>36.07</td>
<td>−40.91</td>
<td>0.31</td>
<td>24.52</td>
<td>58.68</td>
<td>—</td>
<td>0.92</td>
<td>125.3</td>
<td>1.58</td>
<td>0.37</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>(6) 1970–82</td>
<td>36.73</td>
<td>−41.03</td>
<td>0.31</td>
<td>24.43</td>
<td>57.96</td>
<td>−12.97</td>
<td>0.92</td>
<td>109.2</td>
<td>1.60</td>
<td>0.38</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>

Figures in parentheses are the standard errors of the coefficients. Significance was taken at the 5% level of probability. *Point elasticities at the mean value.*
TABLE 3
Passenger Miles and Revenue, Actual and Predicted

<table>
<thead>
<tr>
<th></th>
<th>(1) Actual</th>
<th>(2) Predicted (ex ante)</th>
<th>(3) Predicted (ex post)</th>
<th>(4) Error (1 minus 3)</th>
<th>(5) 4 as % of 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passenger miles</td>
<td>556</td>
<td>569</td>
<td>563</td>
<td>-7</td>
<td>1.26</td>
</tr>
<tr>
<td>Revenue (£M)</td>
<td>21.92</td>
<td>22.5</td>
<td>22.42</td>
<td>-0.5</td>
<td>2.3</td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passenger miles</td>
<td>500</td>
<td>535</td>
<td>506</td>
<td>-6</td>
<td>1.2</td>
</tr>
<tr>
<td>Revenue (£M)</td>
<td>24.58</td>
<td>26.09</td>
<td>25.0</td>
<td>-0.42</td>
<td>1.7</td>
</tr>
</tbody>
</table>

5. FURTHER RESULTS

The forecasting power of the model was thus established for one year. It has since been used to predict demand and revenue for 1981 and 1982.

The modelling procedure has been repeated each year to incorporate the new values of the independent variables. Equations (4) and (5) associated with this procedure are shown in Table 2; for purposes of comparison equation (1) is also shown. The coefficients are remarkably stable over all the equations; this increases our confidence in the structure of the model. The price and expenditure elasticities are also consistent with those found in equation (1).

In 1982 a strike of six weeks' duration disrupted services on part of the network. A dummy variable taking the value of 1 for the strike period (zero otherwise) was included in the preferred equation, and was found to be significant and of negative sign (see equation (6) in Table 2).

The predictions of passenger miles and revenue for 1981 and 1982, obtained from equations (4) and (5) respectively, are shown in Table 3.

There is a considerable divergence between ex-ante and ex-post predictions for 1982. This was primarily due to the "unforeseen" strike which occurred after the ex-ante predictions were made; but a more important source of error is the under-prediction of the fall in consumer expenditure. The ex-ante forecast was that consumer expenditure would fall by 3½ per cent, whereas the actual fall was 6 per cent. This led to an error in predicting demand of just over 12 million passenger miles. The predictive power of the model is obviously sensitive to the accuracy of the forecasts of the independent variables. This applies particularly to the forecast of consumer expenditure, because of its high elasticity and because of the

---

7 Different values of the independent variables are used for planning purposes, so that a sensitivity analysis can be applied to the various permutations.
FIGURE 2

Actual and Fitted Passenger Miles Per Quarter
fairly large fall in its value in 1982 brought about by a deepening in the recession. The real fare variable is less subject to wide variation, since the nominal fare has normally increased once a year and predictions for inflation are generally accurate within a given range.\footnote{1}

As a measure of the accuracy of the set of predictions given in Table 1 and Table 2 we have computed the Theil $U$ statistic, which is based on a comparison of the predicted and actual changes in the dependent variable (Theil, 1966). It is defined as:

\[
U = + \sqrt{\frac{\sum (P_i - A_i)^2 / n}{\sum A_i^2 / n}}
\]

where
\[
\begin{align*}
P_i &= \text{predicted change in the dependent variable;} \\
A_i &= \text{actual change in the dependent variable;} \\
n &= \text{number of periods being compared.}
\end{align*}
\]

The smaller the value of the inequality coefficient, the better is the forecasting performance of the model. Thus when $U$ is zero the forecast is perfect, since the forecast value ($P_i$) is equal to the actual value ($A_i$).

The computed values for our set of predictions are 0.14 for passenger miles and 0.15 for revenue. Thus the prediction errors are 14 per cent and 15 per cent of the error that would have been observed if the forecasts had been zero-change extrapolated — that is, if it had been assumed that there was no change in the dependent variable.

The computed values of the inequality coefficient are close to zero; this indicates that the forecasting performance of our demand model is quite good. The actual and fitted values of passenger miles are shown in Figure 2, where it can be seen that the model reproduces the pattern of actual movements extremely well.

6. CONCLUSIONS

This study has been concerned with the demand for inter-urban rail passenger travel in the Republic of Ireland, for the period 1970 to 1982. The primary objective of the paper has been to derive a model capable of predicting passenger demand in the short run. Certain variables which $a$ priori were expected to be of importance in determining demand were found to be not significant in regression analysis. These included car ownership, the nominal rail fare relative to motoring costs, scheduled train miles operated and a seasonal dummy (fourth quarter).

An industrial earnings variable was found to be significant and positive in all regressions, but was weak in its forecasting ability, primarily because of fluctuations in savings rates.

Demand was found to be significantly affected by rail fares (deflated by the Consumer Price Index), real consumer expenditure, and seasonal dummies representing the second and third quarters.
The coefficient on the real fare variable suggested that the demand for inter-
urban travel was highly inelastic. From a policy point of view this means that
increases in real fare will reduce passenger miles but at the same time increase
total revenue.

The expenditure elasticity was found to be not significantly different from
unity, implying a proportional relationship in the short run between passenger
travel and consumer expenditure. The fall in passenger miles over the last three
years was primarily due to the effects of the recession, which reduced consumer
expenditure considerably. When the recession ends, and growth returns to the
economy, passenger miles should move upward again.

An encouraging aspect of the study has been the forecasting ability of the
model. Future forecasts can be made over longer periods than one year, with
some degree of confidence. Also, sensitivity analysis can be applied to the fore-
casts of the independent variables, so that different planning scenarios can be
examined.

APPENDIX

Derivation of representative sample size for construction of railway fares index

Table 4 contains data for the number of passenger journeys per origin/destination
in a population consisting of 7,148 origins/destinations for the year ended 31
March 1974. The origins/destinations are arranged in seven strata according to the
number of passenger journeys. It can be seen that the distribution of journeys is
highly skewed. The ten origins/destinations in the largest stratum represent 0.1
per cent of the total number of O/D pairs and 26 per cent of the passenger jour-
neys, whereas the 6,388 O/Ds in the lowest three strata represent 89 per cent of
the total number of O/D pairs, but only 5 per cent of passenger journeys. With
such an asymmetrical distribution, disproportionate sampling is recommended,
and the method applied here is that of optimal allocation (Snedecor and

The basic principle underlying this method is that larger samples are taken in
strata with larger standard deviations (assuming that sampling costs per unit are
the same within all strata).

With the skewed population shown in Table 4, the sample size increases with
size of stratum; that is, the larger strata have larger standard deviations.

The large standard deviation of the largest stratum suggests that we take 28
per cent of the sample from this stratum (column 6, Table 4), whereas we take
only 12 per cent from the second highest stratum.

Only those O/D pairs with a passenger movement of 401 or more were
sampled. This means that O/D pairs below this point are not represented in the
sample. This is not a serious omission, as only 4.5 per cent of the journeys were
made between these points, and in a manageable sample size the second and third
strata would be represented by only one O/D pair each.

By combining columns (2) and (6) of Table 4 it was possible to calculate rela-
tive sampling rates (defined as sample size per stratum divided by number of O/D
TABLE 4

Data for Obtaining the Optimum Sample Sizes in Individual Strata

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>217</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>1–50</td>
<td>5,050</td>
<td>47</td>
<td>0.0</td>
<td>55</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>51–400</td>
<td>1,121</td>
<td>175</td>
<td>0.1</td>
<td>109</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>401–4,000</td>
<td>601</td>
<td>756</td>
<td>0.9</td>
<td>528</td>
<td>0.27 (0.29)*</td>
<td>6</td>
</tr>
<tr>
<td>4,001–20,000</td>
<td>115</td>
<td>1,175</td>
<td>0.9</td>
<td>483</td>
<td>0.24 (0.27)*</td>
<td>6</td>
</tr>
<tr>
<td>20,001–50,000</td>
<td>34</td>
<td>1,038</td>
<td>0.9</td>
<td>235</td>
<td>0.12 (0.13)*</td>
<td>3</td>
</tr>
<tr>
<td>&gt; 50,000</td>
<td>10</td>
<td>1,132</td>
<td>53.8</td>
<td>538</td>
<td>0.28 (0.30)*</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>7,148</td>
<td>4,325</td>
<td></td>
<td>1,948</td>
<td></td>
<td>21</td>
</tr>
</tbody>
</table>

*Adjusted to reflect the non-sampling of the lower strata (i.e. below 401 passenger journeys).

pairs), and it was clear that the lowest sampling rate was to be applied to the 401 to 4,000 stratum. By fixing its rate at 1 per cent — considered to be the minimum desirable sampling rate — a total of 6 O/D pairs were chosen from this stratum. From this base the sample sizes of the other strata were obtained. The overall sample size amounted to 21 O/D pairs.

REFERENCES

British Rail Board/University of Leeds (1979): *A Comparative Study of European Rail Performance*.


