THE DEMAND FOR INTER-CITY RAIL TRAVEL IN THE UNITED KINGDOM

Some Evidence

By Ian S. Jones and Alan J. Nichols*

INTRODUCTION

Very little has been published in the United Kingdom on the characteristics of the demand for inter-urban travel, either in total or by individual modes (in contrast to the large volume of work on urban travel demand). Of published studies of the demand for inter-city rail travel in the UK there appear to be only three. The first, Evans (1969), used data from one-day before-and-after on-train surveys to estimate the effects of a major electrification scheme on rail traffic. The second, Leake (1971), examined the modal split of air and rail travellers for a small sample of inter-city journeys. The third, Tyler and Hassard (1971), estimated a gravity-type model to explain variations in the volume of rail travellers between London and a set of provincial centres.

The results reported in the present paper (which originated in a study undertaken as a contribution to a joint forecasting exercise by the Department of Transport and the British Railways Board) are based on the analysis of data (assembled by BR) on four-week ticket sales covering the period from the beginning of 1970 to the end of 1976 for each of the 17 main London-based flows. Demand functions for each of the flows have been estimated by means of ordinary least squares.

The seventeen traffic flows used in the modelling of demand were those from London to: Bath, Birmingham, Bristol, Cardiff, Carlisle, Edinburgh, Glasgow, Leeds, Leicester, Liverpool, Manchester, Newcastle, Norwich, Nottingham, Preston, Swansea and Swindon.

The remainder of the paper is organised as follows. Part 1 examines problems of estimating a demand function. Part 2 examines the choice of variables in the demand equation in the light of the data available to the study. Part 3 presents the results of the estimation, and Part 4 considers what conclusions might be drawn from the results.

* Ian S. Jones is at the National Institute of Economic and Social Research. Alan J. Nichols is Economic Adviser, Department of Energy. The authors are extremely grateful to Andrea Pygall for assistance in carrying out the research described in the paper. The views are those of the authors and do not necessarily reflect those of either the Department of Transport or the British Railways Board.

1 For seven flows (those to Bath, Bristol, Cardiff, Glasgow, Preston, Swansea and Swindon), the period of estimation was extended in order to study the effects of significant changes in service level.
1. ESTIMATION OF THE DEMAND MODEL

We wish to examine the influence of, *inter alia*, a set of own-price and service quality variables on the demand for inter-city rail travel. However, if some of these variables were endogenous (that is, if they were subject to adjustment by the operator in response to changes in the demand for the service), estimates of their effect on demand obtained by OLS methods might be biased.

The nature of the identification problem can be illustrated by the following model of a public transport operator, who faces a (true) demand function of the kind (ignoring stochastic terms):

\[ Q_t = K e^{-at} F_b M_t^c, \quad b = 0 \]  

(1)

That is to say, demand declines exogenously with time at rate \( g \) but is perfectly price-inelastic (\( F \) is the price of a trip, \( M \) is bus miles run and \( b \) and \( c \) are constant elasticities). The operator is constrained to offer a specified level of service represented by bus miles run:

\[ M_t = M_t^* \]  

(2)

Costs (\( TC \)) are a function of bus miles run, and the level of profit (or revenue support) is constrained to equal a specified sum (\( \Pi \)): Thus:

\[ TC_t = a M_t^c = a M^{*c} \]  

(3)

\[ R_t - TC_t = \Pi^* \]  

(4)

where \( R_t = F_t Q_t \) and \( e \) is a constant.

We also assume that the real rate of inflation of unit costs is zero. If we differentiate this system of equations and constraints with respect to time we obtain the following set of conditions:

\[ \frac{1}{Q} \frac{\partial Q}{\partial t} = -g \]  

(5.1)

\[ \frac{\partial M^*}{\partial t} = 0 \]  

(5.2)

\[ \frac{1}{TC} \frac{\partial TC}{\partial t} = \frac{1}{a} \frac{\partial \alpha}{\partial t} + e \left( \frac{1}{M} \frac{\partial M}{\partial t} \right) + \log M \frac{\partial e}{\partial t} = 0 \]  

(5.3)

\[ \frac{\partial \Pi^*}{\partial t} = 0 \]  

(5.4)

\[ \frac{1}{R} \frac{\partial R}{\partial t} = 0 \]  

(5.5)

134
It follows that

\[ \frac{1}{F} \frac{\partial F}{\partial t} + \frac{1}{Q} \frac{\partial Q}{\partial t} = 0 \quad (5.6) \]

and hence

\[ \frac{1}{F} \frac{\partial F}{\partial t} = g \quad (5.7) \]

The behaviour described by this set of relationships would generate a sample of price-quantity observations like the one shown in Figure 1 over a period of time. Suppose we now estimate a demand equation of the form:

\[ Q_t = K F_t^\beta \quad (6) \]

on this set of data. The resulting OLS estimate of \( \beta \) would be a biased estimate of the true value, \( b \), and would in fact reflect the pattern of supply side response. An exactly analogous identification problem would arise if real fares were constrained to be constant, but bus mileage were allowed to decline in response to the exogenous change in demand.

The paper in this journal by Frankena (1978) is one of the few attempts to specify and estimate a demand and supply model for urban bus transport which takes account of these difficulties. But it is, as Frankena recognises, extremely difficult to specify appropriate supply side relationships and operator constraints.

Can simultaneity be a serious problem for the present study? We have approached
this question by postulating certain conditions which are *a priori* conducive to the existence of simultaneity and then assessing to what extent these conditions appear to apply in the inter-city rail market.

We suggest that the problem of simultaneity is likely to be most acute if the market is one where (a) demand is relatively price-inelastic and (b) there is a relatively strong exogenous trend in the demand for the service. A low price elasticity of demand means that changes in real fare levels may be an effective way of altering total revenue in response either to changes in costs or to (unfavourable) changes in exogenous demand. However, the discussion in Part 2 of this paper suggests that BR management may have worked on the assumption that the elasticity of demand for inter-city rail travel with respect to fares was approximately unity during the period of estimation. The real fare level then becomes a variable which may be altered in pursuit of objectives such as filling spare capacity, or in response to government intervention.

The existence of a strong exogenous trend in demand (up or down) makes it likely that service levels will be systematically adjusted in response. However, the market for inter-city rail travel appears not to have exhibited any very strong exogenous trends either up or down during the 1970s. It can therefore be presumed that changes in the provision of capacity reflect railway management's views on their potential profitability and can be treated as exogenous.

For these reasons we believe that variations in fare and output levels can be treated as if they were exogenous for the period over which the models have been estimated. We have therefore felt justified in examining the influence of the variables discussed in Part 2 on the demand for inter-city travel by single equation OLS methods.

Anticipating later discussion, we note that a fundamental problem of identification is present in the estimation of cross-sectional (or gravity) models of the inter-city rail market, such as the one estimated by Tyler and Hassard (1971) to explain variations in journeys between London and a set of provincial centres as a function of population and socio-economic characteristics at the provincial end of the journey, and of distance and service quality. The problem of simultaneity occurs because, *ceteris paribus*, we would expect that the quality of service would itself be jointly determined as a function of expected traffic volumes; that is, we expect that rail management will provide the highest service quality on the heaviest flows. In these terms, it appears to be no accident that the first and, to date, the only large-scale inter-city electrification scheme in the UK was carried out on the West Coast main line, which serves no less than four large provincial conurbations.

Because Tyler and Hassard (1971) estimated their demand equation by ordinary least squares, rather than as one of a pair of simultaneous relationships, we would expect that estimates of the service quality parameter would be biased.

**Lag effects**

The model of demand represented by equation (6) above assumes instantaneous adjustment of demand to a change in the level of fares. Consumers' reaction to certain types of changes may be more complex. For example, major improvements in service may be accompanied by a lot of publicity, which may generate high demand in the period immediately after introduction; once this "halo" effect has gone, the level of demand may decline for a while, before increasing again as relatively infrequent
travellers become aware of the new service. BR themselves have suggested that the full effects of a new service may take up to two years to be felt, though they believe that about 80% of the effect occurs by the end of the first year. We investigated the effect of changes in journey time, using an Almon lag model which requires the prior specification of both the degree of polynomial and the length of time over which the lag structure operates. Lacking any strong *a priori* hypothesis about the form of duration of the lag structure, we experimented with models incorporating first varying degrees of polynomial and then varying durations of lag. We found no evidence of any well-defined lag structure, and, typically, a very high proportion of any effect appeared in the first period after the introduction of the service change; but the estimation of lag effects in the present context is hampered by the relatively small number of “experiments” in the form of service level changes which have occurred on any of the routes in our sample. The results presented in Part 3 are therefore based on models which assume instantaneous adjustment of demand to changes in the level of each of the explanatory variables in the set.

2. CHOICE OF VARIABLE IN THE DEMAND EQUATION

(1) The variable to be explained

The specification of the demand variable used in the research was constrained by the nature of the data available. This is best described with the aid of the network diagram shown in Figure 2.

In the figure, $A$ is the London terminal and $B$ the provincial terminal; $X$ and $Y$ are suburban stations. The data used in the study cover journeys made on tickets printed with the names of $A$ and $B$, so they only cover a sub-set of the total number of journeys between $A$ and $B$; and this, in turn, is a sub-set of the total number of

---

2 For example, journeys between $X$ and $B$ may be made on a single printed (or blank) ticket denominated by $X$ and $B$; or on two tickets (from $X$ to $A$ and from $A$ to $B$). Some journeys using rail mode only between $A$ and $B$ may also be made on blank tickets issued, for example, by travel agents.
journeys made on trains running between A and B. That total would include a potentially large number of flows to and from intermediate stations such as C. Moreover, there is no simple mapping between the ticket sale data and travel between specified origin and destination zones, such as the GLC area and Greater Manchester.

In this paper we use the ticket sale data as a proxy for total rail travel between city centre and city centre.

The analysis is restricted to total ticket sales rather than to sales of particular types of tickets or groupings of ticket types (first and second class, etc.), because disaggregated ticket type data only became available at a late stage in the study, and these disaggregated data covered a significantly shorter period of time than that for which aggregate data were available.

2) Explanatory variables

We hypothesise that the demand for rail travel as measured in the study is a function of rail fare and level of service; of costs and level of service on competing modes; of population and employment at the trip ends; and of economic activity. We now discuss the specification of each of these variables.

Rail fares

Previous studies of the demand for inter-city travel have used one of two alternative measures of fare. The first is the unweighted average revenue per passenger journey (or per passenger mile), defined as:

\[ F_t = \frac{\sum_i (Q_{it} F_{it})}{\sum_i Q_{it}} \]  

(7)

where \( Q_{it} \) = journeys on ticket type \( i \) in period \( t \)
\( F_{it} \) = price in period \( t \) of tickets of type \( i \).

Use of this measure of "price" is open to the objection that \( F_t \) can vary without any proper "price experiment" having occurred. Consider a service on which full and reduced fare tariffs are available. Even if each tariff is constant, the average fare may change as a result of variation in the market share of each type of ticket. Three potential sources of market share variation in inter-city rail travel are: (1) seasonal features (more business travel on first-class full fare tickets at certain times of year); (2) changes in service quality (which may affect certain types of travel more than others); and (3) longer-term changes reflecting differences in income elasticity of demand between journeys for different purposes.

The second measure of fare is a weighted average of fare paid, with fixed weights reflecting the proportion of travellers using each particular type of ticket at some base date. Thus

\[ F_t(W_t) = \sum_i W_{tn} F_{it} \]  

(8)

where \( W_{tn} \) = proportion of travellers using ticket type \( i \) in period \( n \).
DEMAND FOR INTER-CITY RAIL TRAVEL IN UK

Ian S. Jones et al

As we have explained, this option was not in fact open to us until a comparatively late stage in the study; but, even if the data had been available, use of this measure would have required a resolution of the price index problem of incorporating quality changes or new products. "Quality change" in the present context occurs when the availability of a particular type of ticket is altered without any change in its relative price. A new product is simply the introduction of a new type of ticket. In each case, the price of travel perceived by some travellers changes, even if published tariffs are constant or even if the prices remain in some constant proportional relationship with each other.

Some changes both in quality and in the nature of a new product affected many of the flows we were using during the period of the study.

In working with average fare variables of the type given by equation (7), we have attempted to minimise the problem of variation in fare arising from seasonal variations in traffic mix by using a fare variable derived by calculating the average fare paid within a fares "regime", i.e., between fares increases. We have then deflated four-weekly period by four-weekly period, using the Retail Price Index (RPI). We have not attempted to test whether the demand function is homogeneous of degree zero in money fares, as is strictly required by this formulation of the fares variable.3

We were particularly interested in testing what appeared at the time of the study to be a rule of thumb employed by BR management, that the elasticity of demand with respect to fares is not significantly different from unity in the inter-city market (British Railways Board, 1976—see, for example, para. 2.24. On the "received income" on fares elasticity, see also Pryke and Dodgson, 1975).

Rail service levels

The term "quality of service" in the context of inter-city rail passenger services covers the following characteristics:

(1) the point-to-point journey time,
(2) frequency of train departures,
(3) quality of rolling stock.

There may be some degree of "jointness" in these aspects of service quality, but there are examples among the sample of routes which we have studied in which changes along one dimension of service quality have occurred independently of changes along other dimensions. For example, changes in train departure frequency and in quality of rolling stock have occurred without any alteration to journey times. Changes in rail journey times have also occurred without any significant changes in departure frequency. In an ex-ante context, therefore, each of these aspects of rail service should be regarded as independent instruments by which railway management can influence

---

3 With demand function of the forms:

\[ Q_i = B_i P_{i1} P_{i2} Y_{i3} \]

where \( P_i \) = money fare, \( I_i \) = price index, and \( Y_i \) = money income, this condition requires that:

\[ \sum_{i=1}^{i=3} B_i = 0 \]
the level of demand for the service. We have, therefore, specified train departure frequency and journey time as separate arguments in the demand function used in the present study. We have not included a variable measuring quality of rolling stock, because we have had no precise information on when changes in quality have occurred. We would judge (by analogy with the results on train departure frequency discussed later in the paper) that this omission has not led to seriously biased estimates of the co-efficients of the variables included.

An alternative approach used by both Evans (1969) and Tyler and Hassard (1971) is to combine the effects of journey time and frequency into a single quality of service variable, expressed either in journey time or in point-to-point speed units. In order to do this, it is necessary to make assumptions about the distribution of desired arrival or departure times. The composite or generalised time variable for travel between \( i \) and \( j \) \((X^g_{ij})\) is the sum of in-vehicle travel time \((T_{ij})\) plus "expected" waiting time. This latter may be written as some function of the average time interval between train departures \((N_{ij})\). Thus:

\[
X^g_{ij} = T_{ij} + g(N_{ij})
\]  

The equivalent "composite speed" variable may be written:

\[
X^s_{ij} = D_{ij} \left(T_{ij} + g(N_{ij})\right)^{-1}
\]

where \(D_{ij} = \text{distance from } i \text{ to } j\).

Evans assumed that \(g(N)\) is equal to half the average service interval; the Tyler-Hassard study gave \(g(N)\) a value of about a quarter of the average service interval.

One basic objection to the use of any "composite" measure of journey time or speed is that an implicit estimate of responsiveness of demand to the two components of journey time and train departure frequency may be obtained from data in which only one of the components displays any variation. A second is that there is no direct evidence currently available for inter-city travel on the relative weights attached by consumers to in-vehicle and waiting time. Given an estimate of the responsiveness of demand to whichever composite variable is used, the responsiveness of each component is then a function of the relative weights of the two components in the composite function, and their proportions are quite arbitrary.

**Journey time**

Unless journey times are completely uniform for all trains running between points A and B, the specification of the journey time variable introduces difficulties of a similar nature to those encountered in specifying fares variables. Lacking any data on train load factors, we chose from a number of possible measures: for example, the average journey time of all daytime trains; the average journey time of all fast trains where a service includes both fast and semi-fast trains; the fastest journey time; and the average of the fastest \( n \) or \( n\% \) of trains. On a sample of flows which had undergone extensive changes in journey time, the behaviour of the journey time series was found not to be very sensitive to the choice of measure of journey time. The results given later in the paper are based on either the average journey time for all daytime trains or the average journey time of all fast trains.
Train departure frequency

The measure of train departure frequency is based either on the total number of daytime Monday–Friday trains, or on the total number of daytime fast trains. We would expect the demand for inter-city travel to be an increasing function of train departure frequency, for two principal reasons. First, an increase in frequency allows travel behaviour to be more closely matched to the timetable of other activities for which travel is an input. Second, increases in train departure frequency stand proxy for increases in passenger seat miles provided; other things being equal, an increase in passenger seat miles offers an improvement in comfort.

In their response to a consultative document on Transport Policy published in 1976, BR stated that their aim in respect of their inter-city business was to "operate the frequency of service which maximises net revenue so that reductions in frequency aim to fall short of the point where revenue losses outweigh cost savings" (British Railways Board, 1976, para. 2.25). This implies that the marginal revenue earned from an additional train should be no less than the marginal cost of providing the train. By using published data on the relationships between costs and revenues for the inter-city business as a whole, and by proceeding as if output could be varied continuously rather than in discrete steps, this first order condition can be expanded into a required relationship between elasticities of demand and cost with respect to frequency. With revenue per passenger mile held constant at a level $P_o$, the first order condition may be written as:

$$\frac{\partial Q}{\partial F} \frac{F_o}{Q_o} = \frac{C_o}{P_o Q_o} \left( \frac{\partial C}{\partial F} \frac{F_o}{C_o} \right)$$

(11)

where $F =$ train departure frequency.

The first term on the right-hand side is the ratio of costs to revenues, and the second the elasticity of costs with respect to output. The relationship between so-called direct costs and revenues for BR inter-city services as a whole appeared in the Price Commission (1978) report on BR fares. The ratio of direct costs to revenue appears to be of the order of 0.6. If output were continuously variable, and the costs of producing that output arose in a uniform fashion, then we would expect to find that the elasticity of demand with respect to frequency at current output levels was about 60% of the elasticity of direct costs with respect to frequency. Since direct costs are defined as being those which vary closely with output, we might expect this latter elasticity to be close to unity. For relatively long-distance inter-city services, without any strong daily peak in passenger demand of the kind experienced in urban transport, these assumptions might not be unrealistic. In this case, the contribution maximising elasticity of demand with respect to frequency would be of the order of 0.5—0.6. However, on services with more marked daily peaks in passenger demand, which are reflected in differential service levels, the derivation of the required demand elasticity depends on whether or not the change in train departure frequency occurs in the peak. If the change is entirely off-peak, the required elasticity may be as low as about 0.2 (see Appendix A). In the light of BR's statement of policy, we therefore expected that

---

4 See, for example, British Railways Board (1978).
the estimated elasticity of demand to train departure frequency would lie in the range of about 0.2—0.6.

**Costs and service levels on competing modes**

Between 1970 and 1976 competing modes experienced great improvements, both in level of service and in money costs, on a number of the routes analysed in the present study. These changes resulted, first, from the extension of the national motorway network; and second, from the introduction of shuttle air services between London and Scotland. Other things being equal, they should have served to depress the demand for rail travel.

Extension of the national motorway network resulted in quite dramatic reductions in journey times by road between London and a number of provincial centres. This, in turn, encouraged the development of scheduled coach services, so that inter-city rail services faced increased competition both from car and from another public transport mode. Some of the changes in coach journey time between 1970 and 1976 on routes affected by motorway openings are shown in Table 1.

To examine the effect of the improved level of service for both private car and coach on the demand for rail travel on each individual flow, we have used a dummy variable, taking the value 0 before and 1 after the extension or completion of the relevant motorway. The justification for using a single variable is that the time lag between the motorway opening or completion and the subsequent adjustment of coach service schedules has typically been very short. An alternative approach would have been to specify the level of service by road in terms of the published coach journey times. However, changes in coach journey time may be misleading indicators of changes in the potential journey time by private car if the motorway completion has been used as an opportunity for changing the nature of the service offered by coach. The effect on the relevant flows of the introduction of shuttle air services has also been examined by the inclusion of a dummy variable.

The effect of variations in petrol prices on the demand for rail travel has been examined through the inclusion of an index of real petrol prices as an explanatory variable. After declining during the first part of the period, petrol prices increased quite sharply in real terms for a time in the wake of the 1973 crisis.
Population and employment

The choice of explanatory variable introduces certain problems for the choice of explanatory demographic variables. First, the definition of the catchment area is uncertain; for example, should we define the catchment area at the London end of the trip in terms of the Greater London Council’s definition of Central London, or of the GLC administrative area, or of something even wider? The problem is that the behaviour of the population and employment variables over time may differ according to the area chosen. Second, on the basis of cross-sectional evidence on the demand for travel, we would hypothesise that the value of inter-city business trips is related to the volume of certain types of employment in an area. Though mid-year population estimates were available for local authority areas up to 1976, data on the volume and socio-economic characteristics of the labour force in local areas were available only for census years. Unfortunately, the disappearance of the 1976 sample census eliminated any information on even the direction of change in the relevant local labour market variables. Third, relatively robust parameter estimates (in the sense of estimates with standard errors which are small relative to the estimated co-efficient) require significant variation in the explanatory variable. We examined population changes in areas which correspond to the catchment areas at the provincial end of the flows studied, and in nearly all cases the population change over the period 1970–76 was of an order of \( \pm 1–2\% \). We judged that changes in local population at the provincial trip end would not have exerted any strong influence on the demand for travel except in Liverpool and Manchester, where resident population fell by \( 6\frac{1}{2} \) and \( 4\frac{3}{4} \% \) respectively. In each case we included a time trend variable which was very highly correlated with changes in population.

Economic activity

We hypothesise that the demand for inter-city rail travel is a function of the level of economic activity. There are two aspects of the term economic activity which we sought to separate out. First, we wished to examine the relationship between variations in the demand for rail travel and the level of trend variables such as real GDP. We also expected that, irrespective of the existence of any relationship with trend variables, the demand for inter-city rail travel would also be a function of the level of cyclical activity.

Let us assume that the relevant variable is \( Y_t \); then \( Y_t/Y^*_t \) is the ratio to actual \( (Y_t) \) to trend value \( (Y^*_t) \) and is taken as a measure of cyclical activity. Our model may be specified as:

\[
Q_t = a_0 + a_1 Y_t + a_2 (Y_t/Y^*_t)
\]

It follows that:

\[
Q_{t+n} - Q_t = a_1 (Y_{t+n} - Y_t) + a_2 \left\{ (Y_{t+n}/Y^*_{t+n}) - (Y_t/Y^*_t) \right\}
\]

We hypothesise that, irrespective of the sign of \( a_1, a_2 \) is greater than zero.

The implication of the model may be seen as follows. We first note that:

\[
Q_{t+n} = a_0 + a_1 Y_{t+n} + a_2 (Y_{t+n}/Y^*_{t+n})
\]

143
Suppose that \( Y_{t+n} In = Y_t, Y_{t+n}^* > Y_t \). It would follow that for \( a_2 > 0 \) the sign of \((Q_{t+n} - Q_t)\) would be negative.

We expected the demand for inter-city rail services to be positively related to the level of cyclical activity for the following reasons:

1. Business travel accounted for perhaps 30–40% or more of the inter-city rail traffic between London and the main, provincial centres represented in the study. High levels of cyclical activity tend to be associated with a relatively high level of business, and of demand for investment in plant and equipment. We expected that each of these would be associated with increased demand for inter-city rail travel. Conversely, low levels of cyclical activity will be associated with a relatively low volume of new business and investment. Each of these factors would tend to depress the demand for business travel.

2. Variations in the level of economic activity tend to be correlated with "consumer confidence" factors which in turn might influence the demand for certain types of highly optional leisure travel.

We used the GDP output series as the trend economic activity variable. For the period in question, this was very highly correlated with real PDI, which is often specified as the income variable in travel demand studies. The relationship between the demand for inter-city rail travel and this trend variable would be the outcome of two opposing tendencies. On the one hand, high-income people would tend to do more travelling both in the course of their work and for leisure, but they would also tend to own more cars and therefore be less likely to make a given trip by rail. Cross-sectional evidence for the UK contained in Nichols and Jones (1977) suggests that within the period covered by their study, these two tendencies operated rather differently for work-related and for leisure travel. The bulk of longer-distance work-related travel by rail was done by people from relatively high-income, professional/managerial and car-owning households. By contrast, the pattern of longer-distance rail use for leisure purposes was more evenly spread by income, but on average members of non-car-owning households made about 2–2 1/2 times as much use of rail for longer-distance leisure travel as did members of car-owning households. If the national relationship between GDP and growth in car ownership was replicated in the local areas generating and attracting travel in our sample, then, other things being equal, we would have expected the demand for business travel to have increased and the demand for leisure travel to have been at best static or to have declined over the period 1970–76. We have no strong indications on whether the net effect on total traffic might have been positive or negative. In the absence of a car ownership variable (or any other variables relating to the cost of car ownership) in the demand equation we had no hypothesis for the sign of the GDP variable, which should be interpreted as an estimate of the "total" elasticity of demand for rail travel with respect to GDP.

**Specification of the cyclical activity variable**

We considered a number of possible alternative measures of the level of cyclical activity:

1. Measures based on the absolute level of unemployment or of vacancies.

144
DEMAND FOR INTER-CITY RAIL TRAVEL IN UK

Ian S. Jones et al

2. Measures based on estimation of a trend rate of growth of GDP and the level of economic activity, expressed as the ratio of actual and trend levels of the variable.

3. Measures based on time trends modified or corrected for the estimated impact on short output growth of changes in the rate of growth of factors of production.\(^5\)

We ruled out use of labour market variables as a measure of cyclical activity, because we thought that the unemployment series in particular contained an underlying upward trend which might reflect underlying changes in the operation of the labour market.

The approach used in the present study has been based on fitting trend curves to the quarterly GDP output series. The level of economic activity (\(EA_t\)) is then measured in terms of the divergence between the estimated trend value and the actual value of GDP. A number of experiments in curve fitting were carried out over varying periods of time, terminating in the third quarter of 1973. The main cyclical "events" of the period 1970–76 were judged to have been the recession of 1971–72, the cyclical upturn and peak of 1972–73, and the very deep recession from late 1974 onwards. It was found that representation of these features was not very sensitive to the period of time or the specification of the trend curve.

Seasonal variation

We have tested for the existence of seasonal variations in traffic volumes through the inclusion of a set of twelve seasonal dummy variables for the thirteen four-weekly periods.\(^6\)

The demand model

Demand equations of the following kind were therefore estimated for each of the 17 flows:

\[
Q_t = \beta_0 P_t^{81} JT_t^{82} EA_t^{84} GDP_t^{85} PP_t^{86} e^{(\sum_{i=1}^n \beta_i D_i)} + \varepsilon_t
\]  

where \(F_t = \) average revenue per journey (as defined above)

\(DF_t = \) departure frequency by rail

\(JT_t = \) Journey time by rail

\(EA_t = \) index of cyclical activity

\(GDP_t = \) index of real GDP

\(PP_t = \) index of petrol prices

\(S_t = \) service level on non-rail modes

\(D_i = \) seasonal dummy variable

\(\varepsilon_t = \) a random error

\(^5\) For a discussion of alternative measures of cyclical fluctuations, see OECD (1973).

\(^6\) Dummy variables were also used to take account of the effects of events such as strikes.
3. EMPIRICAL RESULTS

Table 2 presents estimates of the ordinary least squares estimates obtained for the demand equations estimated. We now discuss these estimates.

Fares

The estimates of fares elasticities are all negative, and, with one or two exceptions, are all highly significant (see Appendix B).

Almost half the elasticities obtained lie in the range $-0.6$ to $-0.8$, with an unweighted mean for the 17 flows of $-0.64$. In Part 2 we said that we were particularly interested in testing the rule of thumb that the elasticity of demand was equal to unity. Specified in this way, it is appropriate to use a two-tailed test with:

$$H_0: \beta = -1.0$$

The acceptance region is given by

$$-t_{(n-k)\alpha/2} \leq \frac{\hat{\beta} - (-1.0)}{S_\beta} \leq t_{(n-k)\alpha/2}$$

Typically $(n - k)$ is equal to approximately 75. At the 1% level, we find that the estimated elasticities of nine of the seventeen flows are significantly different from unity; at the 5% level, twelve of the seventeen flows are significantly different. If we ask the slightly different question whether the estimated elasticities are significantly less than unity, a one-tailed test is appropriate. Here we find that the estimated elasticities on eleven of the seventeen flows are significantly less than unity at the 1% level. These results appear to cast some doubt on the validity of BR's rule of thumb, but some caution is required in their interpretation. It is possible that the journey purpose mix on the sample of flows we have examined is not representative of the journey purpose mix of all journeys on BR's inter-city trains. Our sample is confined to London-based flows. Data from Long Distance Travel Surveys suggest that work-related travel accounts for a higher proportion of London-based than of non-London-based travel. The sample is also concentrated on longer-distance flows. It is possible, though we have no direct evidence of this, that journey purpose mix differs significantly by length of journey.

In comparison with the results obtained in the present study, Tyler and Hassard (1971) reported an elasticity with respect to average revenue per journey of $-1.2$. However, the average revenue variable was used as a proxy for the distance, or generalised cost, of travel, and was not strictly a co-efficient on fares. An OECD study on the future of European Transport (OECD, 1977) assumed fares elasticities of between $-0.42$ and $-0.69$ for train services over distances of about 300 m at average speeds of 100 m per hour.

---

7 The report indicates that these estimates were derived from an unpublished "parametric" study of inter-city travel in the UK undertaken by the Transport and Road Research Laboratory. The elasticities are not strictly analogous to those obtained in the present study, since they refer to changes in modal shares on the assumption that the total volume of traffic on all competing modes is constant.
### Table 2

**Estimates of the Demand for Inter-City Rail Travel between London and Provincial Centres**

<table>
<thead>
<tr>
<th></th>
<th>Fare</th>
<th>Journey Time</th>
<th>Departures</th>
<th>Economic Activity</th>
<th>GDP</th>
<th>Other Mode</th>
<th>$R^2$</th>
<th>D.W.</th>
<th>Period of Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bath</td>
<td>-0.14</td>
<td>-0.97</td>
<td>ns (2)</td>
<td>1.45</td>
<td>ns (2)</td>
<td>-</td>
<td>0.77</td>
<td>2.07</td>
<td>1972.1—1977.6</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.12)</td>
<td></td>
<td>(0.20)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birmingham</td>
<td>-0.67</td>
<td>-0.97</td>
<td>ns (2)</td>
<td>1.65</td>
<td>ns (2)</td>
<td>-</td>
<td>0.84</td>
<td>1.36</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.12)</td>
<td></td>
<td>(0.24)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bristol</td>
<td>-0.70</td>
<td>-0.64</td>
<td>ns (2)</td>
<td>1.65</td>
<td>0.76</td>
<td>-0.105</td>
<td>0.77</td>
<td>2.19</td>
<td>1970.1—1977.6</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.12)</td>
<td></td>
<td>(0.31)</td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiff</td>
<td>-0.85</td>
<td>-0.48</td>
<td>ns (2)</td>
<td>1.01</td>
<td>ns (2)</td>
<td>-0.061</td>
<td>0.73</td>
<td>1.62</td>
<td>1970.1—1977.6</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.04)</td>
<td></td>
<td>(0.24)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carlisle</td>
<td>-0.34</td>
<td>-0.37</td>
<td>ns (2)</td>
<td>1.59</td>
<td>ns (2)</td>
<td>-</td>
<td>0.69</td>
<td>1.87</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.15)</td>
<td></td>
<td>(0.28)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edinburgh</td>
<td>-0.60</td>
<td>-0.87</td>
<td>ns (2)</td>
<td>1.04</td>
<td>0.28</td>
<td>-0.038</td>
<td>0.87</td>
<td>1.87</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.13)</td>
<td></td>
<td>(0.38)</td>
<td>(0.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glasgow</td>
<td>-1.18</td>
<td>-0.87</td>
<td>ns (1)</td>
<td>ns (1)</td>
<td>ns (2)</td>
<td>-0.068</td>
<td>0.83</td>
<td>1.51</td>
<td>1969.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.13)</td>
<td></td>
<td>(0.13)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leeds</td>
<td>-0.62</td>
<td></td>
<td>ns (2)</td>
<td>1.43</td>
<td>0.34</td>
<td>-</td>
<td>0.69</td>
<td>1.47</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td>(0.25)</td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leicester</td>
<td>-0.67</td>
<td></td>
<td>ns (2)</td>
<td>2.30</td>
<td>-0.41</td>
<td>-</td>
<td>0.82</td>
<td>2.04</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td></td>
<td></td>
<td>(0.28)</td>
<td>(0.18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liverpool</td>
<td>-0.85</td>
<td></td>
<td>ns (2)</td>
<td>1.28</td>
<td>ns (2)</td>
<td>-0.022</td>
<td>0.73</td>
<td>1.90</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td></td>
<td>(0.25)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manchester</td>
<td>-0.65</td>
<td></td>
<td>ns (2)</td>
<td>1.80</td>
<td>ns (2)</td>
<td>-0.038</td>
<td>0.80</td>
<td>1.93</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td></td>
<td></td>
<td>(0.27)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newcastle</td>
<td>-1.03</td>
<td></td>
<td>ns (2)</td>
<td>2.00</td>
<td>0.31</td>
<td>-</td>
<td>0.60</td>
<td>1.91</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td></td>
<td></td>
<td>(0.26)</td>
<td>(0.09)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwich</td>
<td>-0.60</td>
<td></td>
<td>ns (2)</td>
<td>1.82</td>
<td>ns (2)</td>
<td>-</td>
<td>0.59</td>
<td>1.35</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td></td>
<td></td>
<td>(0.27)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nottingham</td>
<td>-0.68</td>
<td></td>
<td>ns (2)</td>
<td>1.20</td>
<td>-0.8</td>
<td>-</td>
<td>0.82</td>
<td>1.87</td>
<td>1970.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td></td>
<td></td>
<td>(0.26)</td>
<td>(0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preston</td>
<td>-0.11</td>
<td>-0.93</td>
<td>ns (1)</td>
<td>1.24</td>
<td>ns (2)</td>
<td>-</td>
<td>0.69</td>
<td>1.57</td>
<td>1969.1—1976.13</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.09)</td>
<td></td>
<td>(0.25)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swansea</td>
<td>-0.87</td>
<td>-0.42</td>
<td>ns (2)</td>
<td>1.61</td>
<td>ns (2)</td>
<td>-0.53*</td>
<td>0.77</td>
<td>1.42</td>
<td>1971.10—1977.6</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.12)</td>
<td></td>
<td>(0.36)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swindon</td>
<td>-0.48</td>
<td>-0.31</td>
<td>ns (2)</td>
<td>0.45</td>
<td>ns (2)</td>
<td>-0.086</td>
<td>0.73</td>
<td>1.84</td>
<td>1971.10—1977.6</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td></td>
<td>(0.23)</td>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses.

* Estimated with respect to coach journey time

ns (1) = Excluded from regression because of $F$ ratio and tolerance
ns (2) = Excluded from regression because of $F$ ratio only
ns (3) = Excluded from regression because of tolerance only
Studies of the demand for domestic air travel in the USA have suggested that the price elasticity of demand in that market is typically less than unity. For example, Verleger (1972) suggests that, on routes on which business travel is predominant, the elasticity of demand is very low indeed. On longer-distance routes, where the share of personal travel is somewhat higher, estimated price elasticities are also somewhat higher. In the present context it may be significant that the only two flows for which the estimated elasticity is above unity (Glasgow and Newcastle) are relatively long-distance ones, which compete with well developed scheduled air services. The air services might in turn attract a relatively high proportion of business traffic, leaving the rail mode with a relatively large share of personal traffic, which is more responsive to price.

**Rail journey time**

The estimated elasticities of demand with respect to rail journey time are all negative and (typically) highly significant. The distribution of outcomes is bimodal in character, with three relatively high estimates (in the range $-0.8$ to $-1.0$) and four rather low ones (in the range $-0.3$ to $-0.5$).

Comparison of our results with those obtained in the studies by Evans (1969) and Tyler and Hassard (1971) referred to above is complicated because, as noted, they each estimated demand response with respect to a composite journey time or speed variable. A further complication with the Tyler and Hassard model is that it does not yield a constant elasticity of demand with respect to their composite speed variable. It would nevertheless appear that, at the levels of departure frequency and speed current on the flows where journey time elasticities have been obtained in the present research, the Tyler and Hassard model yielded estimates of journey time elasticities in the range of about $-1.3$ to $-1.7$, whilst Evans estimated an elasticity with respect to journey time of about $-1.1$. We have indicated that the Tyler–Hassard estimates are certainly biased. By comparison with our own and Evans’s estimates the direction appears to be upward.

**Departure frequency**

It was suggested earlier that, given the relationship between direct costs and revenues for BR’s inter-city services as a whole, the elasticity of demand with respect to departure frequency implied by an objective of maximising contributions would be in the range of about $-0.2$ to $-0.6$. In the present study we have found no evidence that the elasticity of demand with respect to departure frequency is significantly different from zero on any of the flows examined.

This result is puzzling, both in the light of BR’s own stated objectives and by comparison with the estimates of the elasticity of demand with respect to bus miles obtained in some studies of the demand for urban bus travel (see, for example, Frankena, 1978). We offer three comments on this outcome. First, though we are not aware of any obvious source of estimation bias, we would emphasise that many of the

---

8 The model estimates an elasticity with respect to $(1 - 1/x)$. In this form the elasticity varies inversely with $x$.  

---

148
frequency "experiments" observed in the present study have involved increases in frequency from a "base" in which the level of departure frequency was already very high. We have not been able to observe the effects of changes in departure frequency consisting of either a reduction in frequency or an increase from a very low base. Given the very high "base", we would expect that any reduction in the mismatch between the timing of travel and other activities would be very small.

Second, we would expect that an important element in the relatively high elasticity of demand estimated for urban bus travel is the relative ease of substitution between bus and walk as a means of making shorter-distance urban journeys. Finally, very few of the studies of the demand for urban bus travel have addressed the problem of simultaneity discussed earlier. The suspicion remains that, in some of these studies, what is described as the elasticity of demand with respect to bus miles run may be closer to an elasticity of bus miles run with respect to demand.

**Changes in costs and service levels on other modes**

The co-efficients on the dummy variables representing substantial changes in service levels on other modes are negative, and are nearly all significantly different from zero at the 5% level. The size ordering of the co-efficients on the motorway dummy variables fairly closely matches the estimated ordering of the changes in road journey time produced by the motorway completions.

By contrast, the co-efficient on price of petrol did not enter any of the estimated equations at the minimum level of significance specified in the stepwise regression programme. There are two possible explanations for the apparent insignificance of price of petrol. First, changes in journey time directly affect the level of service on both car and coach modes. The effect of variations in petrol prices on the cost of coach travel is relatively small, because petrol prices are only one element in the total cost or provision. Second, a relatively high proportion of all longer distance inter-city journeys are made in course of work. The share of petrol costs in the generalised cost of such journeys is probably very small. This effect is reinforced by the relatively high car occupancy rates which are typical of longer-distance inter-city journeys. By analogy with the earlier discussion of composite journey time and speed, we would therefore expect that the cross elasticity of demand with respect to price of petrol would be small relative to the cross elasticity with respect to road journey time.

**Cyclical activity**

As hypothesised, the estimated co-efficient on the index of economic activity is positive across the sample of flows, and in most cases is significantly different from zero at the 5% level. In order to get some intuitive feel for what the economic activity variable was contributing to the explanation of variations in traffic flows, we have compared the residual plots of equations which do and do not include it. What we have found is that, in general, introduction of the variable reduces the errors of estimation in two particular sub-periods:

1. from late 1972 to mid 1973, when real fares were roughly constant on many flows but when traffic levels increased quite sharply;
(2) from early 1975 to late 1975, when traffic levels turned down in advance of any really significant increase in real fares.

The first sub-period coincided with quite rapid increase in cyclical activity measured by our economic activity variable; the second corresponded with a rapid decline.

**Gross domestic product**

We have found no uniformity in either the sign or the significance of the GDP variable in our sample of flows. Some cases have yielded estimates which are positive and significant at the 5% level; others have yielded estimates which are negative and significant. By comparison, studies of the demand for urban bus travel have suggested that the "total" elasticity of demand for transit with respect to real income is negative, while studies of the US domestic airline market clearly indicate that the real GDP or income elasticity of demand in the market is strongly positive (Verleger, 1972).

**Population**

The time trend variable included in the Liverpool and Manchester equation was negative (as hypothesised) but not significant at the 5% level. Unfortunately, the variable was highly correlated with the dummy variable representing the change in road service levels in both cases (the sign of the dummy was also negative, and it was significant at the 5% level).

**Seasonal factors**

The hypothesis that the demand for inter-city rail travel is subject to systematic seasonal influences is supported by the results of the study. Both the amplitude and the timing of these fluctuations vary quite widely between flows of different kinds. Broadly speaking, the results indicate that the seasonal peaks on short-medium distance flows occur in late summer and autumn. The very long-distance flows between London and Scotland exhibit seasonal peaks in the summer and around Christmas and the New Year. The amplitude of the fluctuations on the Scottish flows is substantially higher than on the shorter distance flows.

4. CONCLUSIONS

The empirical results obtained in this study are, by and large, consistent with *a priori* hypotheses about the main determinants of the demand for inter-city rail travel between London and the set of provincial centres represented in the sample.

Thus, we have found that demand is significantly affected by the level of average fare paid, deflated by the retail price index, by rail journey time, by the level of service offered on competing modes, by the level of cyclical economic activity, and by seasonal factors. We have found some very limited evidence of a positive relationship between the demand for travel and population at the provincial end of the journey.

We have found no evidence that the demand is significantly responsive to changes
in train departure frequency within the range of the experiments observed in our sample, or to the level of petrol prices deflated by the RPI. We have found no consistent evidence that the demand for rail travel is either positively or negatively related to variations in real GDP.

Any discussion about the possible implication of these results for the management of BR’s inter-city services must be subject to the proviso that the set of flows covered is not a representative sample of the traffic. Nevertheless the ticket sales covered in the study might generate perhaps 15–20% of BR’s inter-city revenues, and our results should offer at least some clues on options which might be open to railway management in seeking to improve the financial performance of inter-city services.

The results suggest that, during the period of estimation, BR’s finances could have been improved by some increases in real fares above the levels then prevailing, and by some reduction in frequency of train departures. The financial benefit from higher train speeds is uncertain in the absence of any estimates of their effect on capital and operating costs.

Whether the same prescription would be appropriate in BR’s present circumstances is uncertain. In particular, the combined effects of developments in BR’s marketing strategy (the “railcard revolution”) and in inter-city bus services in the wake of deregulation may have been to make the inter-city rail market more price-elastic in total. If so, efforts to improve financial performance may need to be focused more specifically on achieving a better match between traffic and capacity on offer. In the light of the findings reported in the recent report of the Monopolies and Mergers Commission (1980) on the operation of BR’s London and South-Eastern services (see in particular chapter 9), it would be surprising if there were no scope for significant savings in this direction.

APPENDIX A

Elasticity of Demand with respect to Frequency

An indication of the range of possible outcomes involved may be gained by using data in Table 1 of Price Commission (1978). Unfortunately, the data in that Table do not relate the type of expense (fuel, materials, staff, etc.) to the direct-indirect classification, so some assumptions must be made about the matching. For present purposes we make the following assumptions: (1) that fuel and power and depreciation are exclusively direct costs; (2) that about 50% of “materials” are direct costs; (3) that the cost structure with respect to the type of expense is the same for inter-city services as for passenger services as a whole. These assumptions yield the following breakdown of costs for inter-city services: staff 72%, fuel 17%, materials 7%, depreciation 4%.

A lower limit estimate of cost elasticity may be obtained by assuming that the change in departure frequency is off-peak and that no additional crew time inputs are required. The only variable items then will be fuel, the costs of mileage-related vehicle maintenance, and depreciation. A 10% increase in off-peak mileage might then result
in a 10% increase in fuel costs, a 5% increase in materials costs, a 5% increase in depreciation costs and a 3% increase in staff costs on account of increased maintenance. In total, direct costs might increase by about 4½%. Given the relationship between direct costs and revenues, this would indicate a "required" elasticity of demand of about −0.2 to −0.25. We would expect that a 10% increase in "peak" frequency would involve roughly a 10% change in costs.

APPENDIX B

Criteria for Inclusion of Variables

With the exception of the work on lag structures referred to earlier, the models have been estimated using the SPSS stepwise regression package. This requires the prior specification of a minimum significance level for inclusion of variables in the regression and of a minimum tolerance level. Included variables must satisfy both criteria. In the present work the minimum value of the $F$ ratio for the variable $X$, defined as:

$$F = \frac{\text{(incremental SS due to inclusion of } X)}{\{SS \text{ (residual)}\}/(N-K-1)}$$

$K = 1 \cdots (i - 1)$

has been set at 0.5. The tolerance level for an independent variable being tested for inclusion is defined as the proportion of the variance of the additional variable not explained by variables already in the regression. It therefore acts as a kind of screening device to avoid the difficulties introduced by severe multicollinearity. In the present work this minimum tolerance level has been set at 0.2. It follows that the term "significant" in the discussion of the empirical results is used in two rather different ways. The term "not significant" is used when a variable has not been included in the set of explanatory variables because it has failed the tests specified above for inclusion. In the table of results, however, we indicate which of the two criteria the variable has failed to satisfy. The term "not significantly different from zero" is used when the variable in question has been included in the set of explanatory variables but is not significantly different from zero at the specified level of significance (1%, 5%).

REFERENCES

DEMAND FOR INTER-CITY RAIL TRAVEL IN UK

Ian S. Jones et al