



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/2157/>

Monograph:

Wardman, M., Whelan, G.A. and Toner, J.P. (1994) Direct Demand Models of Air Travel: A Novel Approach to the Analysis of Stated Preference Data. Working Paper. Institute of Transport Studies, University of Leeds, Leeds, UK.

Working Paper 421

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



White Rose Research Online

<http://eprints.whiterose.ac.uk/>



[Institute of Transport Studies](#)

University of Leeds

This is an ITS Working Paper produced and published by the University of Leeds. ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

White Rose Repository URL for this paper:

<http://eprints.whiterose.ac.uk/2157/>

Published paper

Wardman, M., Whelan, G.A., Toner, J.P. (1994) *Direct Demand Models of Air Travel: A Novel Approach to the Analysis of Stated Preference Data*. Institute of Transport Studies, University of Leeds. Working Paper 421

UNIVERSITY OF LEEDS
Institute for Transport Studies

ITS Working Paper 421

ISSN 0142-8942

August 1994

**DIRECT DEMAND MODELS OF AIR TRAVEL:
A NOVEL APPROACH TO THE ANALYSIS OF
STATED PREFERENCE DATA**

**M Wardman
G A Whelan
J P Toner**

This work was undertaken on a project sponsored by the Economic and Social Research Council (Grant Ref: R000233791)

ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

CONTENTS

	Page
ABSTRACT	
1. INTRODUCTION AND OBJECTIVES	1
2. BACKGROUND: MODELLING APPROACHES	1
3. MODELLING ISSUES	2
4. ADVANTAGES OF DIRECT DEMAND SP MODELS	4
5. EMPIRICAL RESULTS	11
6. CONCLUSIONS	16
APPENDIX	18

ABSTRACT

Wardman, M., Whelan, G.A. and Toner, J.P. (1994) Direct Demand Models of Air Travel: A Novel Approach to the Analysis of Stated Preference Data. *ITS Working Paper 421*, Institute for Transport Studies, University of Leeds, Leeds.

This paper uses what has been termed the direct demand approach to obtain elasticity estimates from discrete choice Stated Preference data. The Stated Preference data relates to business travellers' choices between air and rail. The direct demand methodology is outlined and some potential advantages over the conventional disaggregate logit model are discussed. However, further research regarding the relative merits of the two approaches is recommended. The direct demand model is developed to explain variations in the demand for air travel as a function of variations in air headway and cost and in train journey time, frequency, interchange and cost. Relatively little has previously been published about the interaction between rail and air and the elasticities and variation in them which have been estimated are generally plausible. In particular, the results show that large improvements in rail journey times can have a very substantial impact on the demand for air travel and that the rail journey time cross-elasticity depends on satisfying a three hour journey time threshold.

KEY-WORDS: Demand modelling; mode choice, elasticities, stated preference

Contact: Mark Wardman, Institute for Transport Studies (0532 335349)

DIRECT DEMAND MODELS OF AIR TRAVEL: A NOVEL APPROACH TO THE ANALYSIS OF STATED PREFERENCE DATA

1. INTRODUCTION AND OBJECTIVES

The research reported here was undertaken as part of an ESRC funded project (R000233791) examining the impact of rail service quality and fare on the demand for air travel and long distance car travel. This paper is concerned with the interaction between air and rail in the business market.

Whilst the research findings reported here are of interest in their own right, since relatively little research has been conducted on the interaction between rail and air in Great Britain, this paper is also concerned with methodological issues; namely the development of what we have termed direct demand models based on Stated Preference (SP) data.

It is our understanding that models developed on SP data have invariably been of the disaggregate logit form, and in particular they have examined individuals' discrete choices amongst alternative modes of transport. We here present models which are calibrated on data collected from SP discrete choice experiments but which aim to explain the variation in collective behaviour rather than the variation in individuals' discrete choices.

Section 2 outlines the broad distinctions we make between what we have termed the aggregate and disaggregate modelling approaches whilst section 3 discusses in more detail the modelling issues relating to the direct demand model. Section 4 advances what we believe to be the main advantages of using a direct demand approach to the analysis of SP data and section 5 presents the empirical findings. The concluding remarks are contained in section 6.

2. BACKGROUND: MODELLING APPROACHES

Although distinctions can be made between alternative demand analysis methodologies on various dimensions, we here distinguish according to the unit of observation that enters the demand model.

What are termed direct demand models are a form of aggregate model in that the dependent variable takes the form of the collective behaviour of travellers. They subsume the various stages of travel decision making into a single stage, represented by the quantum of individuals travelling between two points by a given mode. They are based on recorded ticket sales, as is common in the railway industry (Wardman, 1994), and on counts or surveys of travel behaviour.

The model aims to explain the variation in the demand for a mode of travel as a function of variations in the characteristics of that mode of travel, of competing modes of travel and of other relevant socio-economic factors, although most direct demand models rarely include factors other than those relating to the mode in question. It is therefore the empirical equivalent of the conventional demand function of economic theory. Although some argue that such direct demand models do not have the same behavioural basis as disaggregate models derived from utility

maximising behaviour, we would argue that the axiom that there is a relationship between aggregate demand and modal characteristics cannot be described as anything other than behavioural whilst, quite fundamentally, we will argue that these models are potentially more behaviourally robust since the only assumption required with direct demand models is that there is a well behaved relationship between demand and modal characteristics.

In contrast, disaggregate models take the discrete choices of an individual or some other decision making unit as the unit of observation. The choices are commonly the mode choices individuals are either observed or stated to make and, in contrast to this study, the context is often that of urban travel.

Disaggregate models specify a relationship between the probability that an individual will choose an alternative and the utility of each alternative in the choice set. The utilities are related to relevant variables representing travel circumstances and personal characteristics, with unobserved factors being represented by an error term. It is the purpose of the calibration stage to estimate the weights attached to each variable in the utility expression. The form of model derived depends on the assumptions concerning the error term. Assuming that the errors associated with each alternative are independently and identically distributed and follow a type I extreme value distribution yields the commonly used multinomial logit model. The binomial logit model, in the special case of just two alternatives, does not need to assume independently and identically distributed errors, and is essentially the only model used to examine discrete choices amongst two alternatives. The hierarchical logit model is typically used when choices are expressed amongst more than two alternatives since it avoids the restrictive assumptions of the multinomial logit model in such instances. A more detailed discussion of disaggregate models is provided in Maddala (1983) and Ortuzar and Willumsen (1990).

3. MODELLING ISSUES

Although direct demand SP models are based on the data derived from discrete choice experiments, they differ from disaggregate models in that they operate on the volume (V) of demand rather than discrete choices. The discrete choices for a particular travel alternative are summed for each scenario of the SP experiment. Thus if 100 individuals respond to an SP experiment involving 12 pairwise comparisons of, say, rail and air, we have 12 estimates of the volume of rail demand and of air demand, as opposed to 1200 choice observations that would be available for input to a disaggregate model. Clearly a single SP design, which typically offers between 9 and 16 comparisons, would not provide a sufficient number of observations to develop a robust model. Thus there is a need to offer different designs to increase the number of observations available for modelling. If four designs each involving 12 comparisons were allocated amongst the 100 individuals, we would have 48 observations of the volume of both rail and air demand.

Suppose the relationship between the volume of demand for mode m between i and j at any point in time can be specified as:

$$V_{ijm} = f(G_i, A_j, GC_{ijm}, GC_{ijn}) \quad (1)$$

where GC denotes the generalised cost of mode m and of a competing mode n, G_i represents the generating potential of origin i and A_j represents the attracting potential of destination j.

We are provided with an estimate of V_{ijm} by summing the SP responses as stated above. We cannot estimate the G_i and A_j since our sample forms only a portion of the travellers between i and j and it is highly unlikely that we will know what proportion have been sampled. However, we are often interested in being able to explain changes in demand, rather than demand in absolute, and in estimating elasticities. If we specify the demand function as:

$$V_{ijm} = \alpha_i G_i \beta_j A_j GC_{ijm}^\lambda GC_{ijn}^\tau \quad (2)$$

then taking the ratio of the function for two SP scenarios (1 and 2) yields:

$$\frac{V_{ijm2}}{V_{ijm1}} = \frac{GC_{ijm2}^\lambda GC_{ijn2}^\tau}{GC_{ijm1}^\lambda GC_{ijn1}^\tau} \quad (3)$$

The generating and attracting variables cancel out and the parameters of equation 2 are estimated by multiple regression of:

$$\log \frac{V_{ijm2}}{V_{ijm1}} = \lambda \log \frac{GC_{ijm2}}{GC_{ijm1}} + \tau \log \frac{GC_{ijn2}}{GC_{ijn1}} \quad (4)$$

There is a further reason for estimating the form of model specified in equation 4. Given that we have more than one SP design, in order to provide the direct demand model with a large enough number of observations, the different numbers of individuals replying to each design will distort the relationship between the absolute level of demand and the independent variables. For example, if a relatively large number have been interviewed for a design that relates to longer distance and hence more expensive journeys, direct regression of equation 2 could well obtain estimates which imply that volume increases as fare increases! We therefore need to standardise for the number of individuals completing each SP exercise and this is achieved by the use of equation 4. Thus if we have 12 SP scenarios, we can specify 11 observations of the form of equation 4.

If the survey concentrates exclusively on the users of a particular mode of travel, it is only possible to estimate the demand function for points worse than the current situation. This is because the effect of improvements would, at best, only be represented by the extra trips generated by existing users and we would fail to cover the much more important dimensions of switching from other modes of travel and the attraction of new travellers to the market. This would clearly lead to biased estimates of the demand function's parameters. Indeed, the demand function would be kinked at the base situation. In such instances, the analysis should therefore be restricted to data points reflecting lower demand than the base situation. Subsequent forecasting

applications assume that the behavioural response to, say, a fare increase from x to y can be used to forecast the effect of a fare reduction from y to x and that the demand function is well behaved in the sense that the relationship between demand and explanatory variables obtained from model calibration is continued across points on the demand function beyond the range in which it was calibrated. These assumptions do not seem objectionable and indeed the former assumption is implicit in all demand forecasting models.

Equation 4 is estimated by weighted least squares because the variance of the error term will not be constant. The derivation of the error variance is given in the Appendix.

4. ADVANTAGES OF DIRECT DEMAND SP MODELS

An advantage of disaggregate models over aggregate models which is often claimed is their ability to examine travel behaviour in more detail. It is often not possible to segment aggregate models by important factors such as journey purpose; for example, rail ticket sales data split by purpose is not available and class of travel is only a crude proxy. Moreover, the amount of variation between zones, such as access times to stations, may well be insufficient to allow reliable estimation by aggregate models even though there is considerable intra-zone variation. These problems are not faced by disaggregate models. However, it can easily be seen that a direct demand model based on SP data would not suffer these shortcomings since it can be split by purpose and we can ensure that variables whose effects on demand we wish to estimate exhibit sufficient variation.

Clearly there is little point in developing direct demand SP models if they have no advantages over disaggregate SP choice models. We list below what we see to be the potential advantages of the direct demand approach to modelling SP data. Some of the claims we make are somewhat tentative. However, raising these issues is warranted since we view them as meriting further research.

Design Implications

Since the dependent variable of a direct demand model is a continuous variable, we would suspect that design issues are more straightforward than in the conventional context of discrete choice modelling. Further research is required to substantiate this claim but consider the following point. If a scenario contains more than a negligible change in one or more of the variables relative to some other scenario then it must contain information on the sensitivity of demand to changes in relevant variables. Thus the principle behind the design is simply to vary the travel variables so as to be able to observe what happens to demand. However, the design of SP exercises to be analysed by disaggregate logit models is difficult; it is possible to unwittingly design an SP exercise so that some scenarios have very little information content or indeed provide the same information as some other scenario. Recent research by Holden (1992) shows that SP design is by no means straightforward when the SP responses are to be modelled as discrete data.

Model Assumptions

The direct demand SP model requires that there is a well behaved relationship between demand and relevant travel variables. Regardless of how individuals arrive at their travel decisions, we

expect that, for example, demand falls as fare increases and that any variation in the fare elasticity can be explained by reference to fare levels or variation in other factors.

The assumptions underlying disaggregate models would seem to be more demanding and hence less likely to be satisfied. Such models generally assume that there is no taste variation across individuals and that all individuals have the same decision making process which is typically specified to be based on a compensatory utility function and utility maximising behaviour. Horowitz (1981) shows that the implications of taste variation on logit choice probabilities can be serious. He states that "probabilistic discrete choice models are sensitive to a larger group of specification and data errors than linear econometric models are".

The attractions of the direct demand approach essentially stem from our feeling that it is easier to explain and predict the behaviour of groups of travellers than it is to explain the variation in discrete choices across different individuals' circumstances. If we want to predict changes in aggregate behaviour, it is not surprising that we would have a preference towards models which are calibrated to changes in group behaviour rather than towards cross-sectional comparisons of what different individuals do.

Choice Based Samples

In many circumstances SP exercises are conducted on users of only one of the alternatives which enters the SP; for example, there have been a number of studies of rail users which have used SP exercises involving the choice between rail and some alternative mode. This is done primarily for cost purposes, particularly since direct access to users of other public transport operators is not usually available and car users are expensive to contact during their journey. Other forms of contact, such as household interviews, are expensive, particularly when we are dealing with rare events such as long distance trip making.

Whilst focussing exclusively on users of one mode is quite understandable, it does have implications for the use of disaggregate choice models for forecasting. There are two ways we could approach forecasting with this model. The first approach is to use a sample enumeration approach to forecast what the individuals upon whom the model is calibrated would do in various circumstances. The second approach is to use the model in an aggregate fashion to forecast what the market as a whole would do.

The problem with the first approach is that the probability of using the mode actually used is one but the logit model cannot recover such a probability. Whilst it may forecast individual probabilities close to one, there is the risk that it forecasts probabilities somewhat different to one. Given that the sensitivity of demand to changes in travel variables (elasticities) in logit models depends on the choice probabilities, we would have concerns about using a model which did not closely replicate the base probabilities. Moreover, this approach is limited to the sample of individuals upon which the model was calibrated.

If we use the model in an aggregate fashion to predict overall market shares and changes to them, we can hardly expect a model calibrated on one particular set of users to be appropriate to the whole market. Even if the parameters of the calibrated model were appropriate for the entire market, the mode specific constants would still need to be adjusted because the sample is choice based (Coslett, 1981) yet this adjustment cannot be made when one of the alternatives has a zero market share in the sample of travellers.

Scale Factor Problem

The scale factor problem arises because the coefficients of disaggregate choice models, and hence the forecast choice probabilities and implied elasticities, are estimated in units of residual deviation (Wardman, 1991). The coefficients of a logit model contain a scale factor (Ω) of:

$$\Omega = \frac{\pi}{\sqrt{6}\sigma_e} \quad (5)$$

where σ_e is the standard deviation of the error term. The unobserved (error) component of choice models represents the effect on choice of factors not included in the model. However, SP responses may include error which is unique to the SP exercise and which has no bearing on actual choices yet it will enter σ_e , and hence influence the coefficient estimates and forecasts via the scale factor (Ω), because the model cannot distinguish between such SP specific error and legitimate error.

It is tempting to think that regression based models are free from such problems because the coefficients are not estimated in units of residual deviation. However, this is not the case. We will use a numerical example to illustrate the problem. Let us suppose that the true demand for, say, rail travel amongst 120 rail travellers is expressed as:

$$V = 30P_R^{-0.8}P_C^{0.8} \quad (6)$$

where P_R is the price of rail and P_C is the price of coach and coach is the only other mode. However, in an SP exercise of these 120 rail users, 20% make the wrong choice. Table 1 lists the actual and reported volumes for 10 scenarios.

Table 1: Actual and Reported Volumes

	True Volume	P_R	P_C	Rail Share	Reported Volume	Error
1	108.7	1	5	0.91	89.2	-19.5
2	90.9	1	4	0.76	78.6	-12.4
3	72.2	1	3	0.60	67.3	-4.9
4	52.2	1	2	0.44	55.3	3.1
5	62.4	2	5	0.52	61.5	-1.0
6	52.2	2	4	0.44	55.3	3.1
7	30.0	3	3	0.25	42.0	12.0
8	21.7	3	2	0.18	37.0	15.3
9	17.2	4	2	0.14	34.3	17.1
10	25.1	5	4	0.21	39.1	14.0

The first column in Table 1 is the true volume of rail demand given the rail and coach fare in the second and third columns and the demand function of equation 6. The fourth column gives the rail share, defined as the true rail volume as a proportion of the 120 travellers, whilst the fifth column lists the volumes of rail demand given that 20% of the reported choices are incorrect. The figures in this column are obtained as 80% of those who correctly state that they would use train and 20% of those who would use coach but who wrongly state that they would use rail. Thus in scenario 1 we would have 108.7 actual rail users and 11.3 actual coach users. 80% of the rail users correctly choose rail, giving 87.0, to which is added 2.2 actual coach users who state that they would prefer rail. We therefore have a reported rail volume of 89.2. The final column denotes the error defined as the difference between the reported and true rail volumes.

Where the rail share is greater (less) than 0.5, the reported volume is less (greater) than the actual. The reported rail share will tend to 0.5 relative to the true rail share and hence the former can be taken to indicate the sign of the error in the reported volume and is therefore potentially useful in measures to alleviate the scale factor problem.

The error in the reported volumes is consistent with random error across individuals' SP responses. A respondent who would actually use rail and has an error in the SP response which adds to the relative attractiveness of rail would give a correct answer and hence the error would have no effect. However, the SP choice could be incorrect when the error in the SP response is of the reverse direction. We would expect random error to affect more choices in absolute for the mode which would have the largest actual share. Thus it will operate to understate (overstate) the reported volume of demand for the major (minor) mode.

If we regress the reported volumes of Table 1 on fares, with the suitable logarithmic transformation, the elasticity estimates are around 50% of the true values when 20% of choices are incorrect. When the proportion of incorrect choices falls to 10%, the estimated elasticities are around 72% of the actual values whilst they are around 85% of the actual values when 5% of choices are wrong. The estimates are biased because, as can clearly be seen from Table 1, the errors are correlated with the dependent variable rather than being random and this systematic relationship between dependent variable and error has a systematic effect on the coefficients. This can be seen in Figure 1. The true demand function is represented by line A but the errors in the reported volumes would lead to the estimation of a line such as B which would quite clearly have different parameter estimates.

The regression based direct demand approach therefore also experiences a scale factor problem and the question now is how does it compare with the scale factor problem that arises in the disaggregate logit approach? To compare the two approaches we need to find some way of converting from the error in one model to that in the other model.

Wardman (1991) provides a table showing how logit choice probabilities are affected by increases in the residual variation. We have taken the choice probability distributions from that table which most closely approximate the choice probability distributions obtained from the approach underlying the figures in Table 1 for 5%, 10% and 20% incorrect choices. These are given in Table 2.

Figure 1: Actual and Estimated Demand Functions

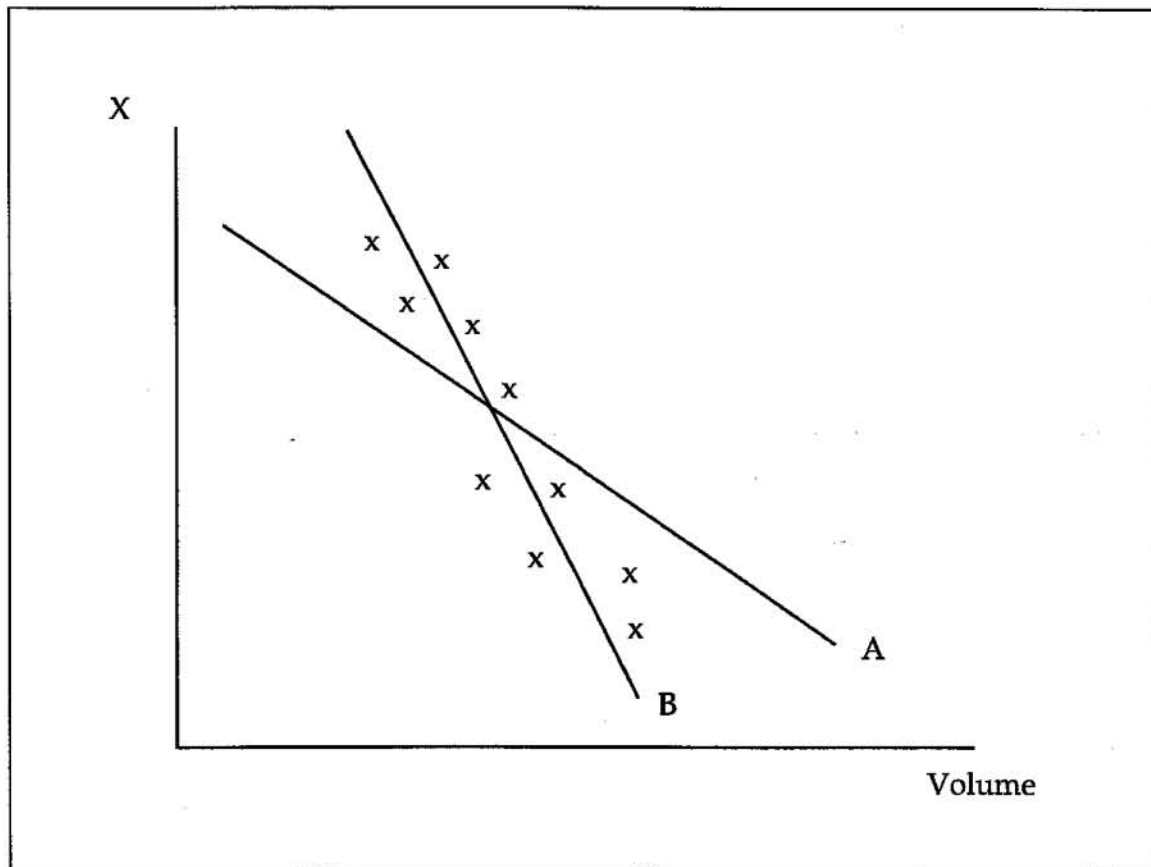


Table 2: Corresponding Market Shares

True Share	5% Error	σ +20%	10% Error	σ +33%	20% Error	σ +75%
0.95	0.905	0.921	0.860	0.901	0.770	0.843
0.90	0.860	0.862	0.820	0.839	0.740	0.778
0.80	0.770	0.760	0.740	0.739	0.680	0.688
0.70	0.680	0.670	0.660	0.654	0.620	0.619
0.60	0.590	0.584	0.580	0.576	0.560	0.558
0.50	0.500	0.500	0.500	0.500	0.500	0.500
0.40	0.410	0.416	0.420	0.424	0.440	0.442
0.30	0.320	0.330	0.340	0.346	0.380	0.381
0.20	0.230	0.240	0.260	0.261	0.320	0.312
0.10	0.140	0.138	0.180	0.161	0.260	0.222
0.05	0.095	0.078	0.140	0.099	0.230	0.157

The first column in Table 2 lists market shares between 0.95 and 0.05 which are taken as the true shares. The second column indicates what the market shares would be if 5% of choices were incorrect. The third column is taken from Wardman (1991) and is the logit choice probability distribution which most closely approximates the probabilities in column 2. This is for a residual deviation (σ) 20% higher than that which would give the true shares in column 1. Thus 5% incorrect choices is equivalent to a residual deviation in a logit model which is 20% too high. The remaining columns in Table 2 show that 10% incorrect choices corresponds with a residual deviation which is 33% too high and the probability distribution for 20% incorrect choices is equivalent to a residual deviation which is 75% too high. We are now in a position to compare the effect of the scale factor problem on the elasticities obtained from each modelling approach.

The figures in Table 3 present the ratio of the estimated elasticity to the actual elasticity for a range of actual market shares and the three levels of additional residual variation identified to best approximate the 5%, 10% and 20% of incorrect choices of the direct demand approach. The logit model elasticities (η_x) are calculated as:

$$\eta_x = \alpha X(1 - P) \quad (7)$$

where α is a coefficient estimate and P is a choice probability or market share. Both of these are affected by the scale factor problem. Thus for a residual variation which is 20% too high, an actual share of 0.95 would become 0.921 and α would be 16.67% too small. For probabilities in excess of 0.5, the effects of the scale factor on the α and $(1 - P)$ terms are offsetting. However, they compound for probabilities less than 0.5.

Table 3: Ratios of Estimated and Actual Logit Point Elasticities

True Share	σ +20%	σ +33%	σ +75%
0.95	1.32	1.49	1.79
0.90	1.15	1.21	1.27
0.80	1.00	0.98	0.89
0.70	0.92	0.87	0.73
0.60	0.87	0.80	0.63
0.50	0.83	0.75	0.57
0.40	0.81	0.72	0.53
0.30	0.80	0.70	0.51
0.20	0.79	0.69	0.49
0.10	0.80	0.70	0.49
0.05	0.81	0.71	0.51

For shares less than 0.5, and a residual variation which is too large, the elasticities of an SP logit model would be too low. For shares greater than 0.5, the forecast elasticities can be too large or too small. Given that we found the error in the direct demand model's elasticities, expressed as ratios of estimated and actual elasticities, to be around 0.85, 0.72 and 0.50 for 5%, 10% and 20% of incorrect choices, comparison with Table 3 shows that the logit model would be preferred since the two approaches have a similar level of accuracy for shares less than 0.5 but that, with the exception of very large probabilities, the logit model would produce more accurate elasticity estimates for shares in excess of 0.5.

However, this is not the whole story; there are two further issues which we can take into account. Firstly, the error in the direct demand model discussed so far is entirely of a systematic nature in the sense that there is a very high correlation between the error term and the level of the dependent variable and this is the source of the biased coefficient estimates. Let us suppose that in addition to this systematic error there is also some random error surrounding the dependent variable. This would not have any effect on the coefficient estimates of the direct demand model but it would affect the coefficient estimates of the disaggregate logit model since additional error will inflate its residual variation.

A second, and more important issue, is that it may be possible to reduce the impact of the scale factor problem on the direct demand model's parameters by suitable transformation of the model prior to estimation. In Table 1 we can see that some of the errors are positive and some are negative. The problem we have encountered above is that the errors are correlated with the dependent variable and hence the coefficients are biased. However, transformation of the model could make the errors of a more random nature. We therefore specified the model to be estimated as:

$$V_1 + V_2 = \alpha P_{R1}^\beta P_{C1}^\lambda + \alpha P_{R2}^\beta P_{C2}^\lambda \quad (8)$$

The subscripts 1 and 2 denote two different scenarios and the functional form is the same as that in equation 7. This additive form is used to allow the errors to offset and the parameters are estimated by non-linear least squares. Specifying a logarithmic form of model in order to allow ordinary least squares estimation did not have the desired effect since the correlation between the dependent variable and the error term was not broken.

We selected from Table 1 the comparisons of scenarios 2 and 7, 1 and 9, 2 and 8, 5 and 6, 2 and 10, and 3 and 5 so that the error term using equation 8 could be both positive and negative. The data was copied to itself to form a data set of 60 observations to allow estimation of equation 8 by non-linear least squares. The estimates of α , β and λ were 33.59, -0.70 and 0.69 respectively. These are certainly much nearer the values of 30, -0.8 and 0.8 which generated the data than was previously the case.

We conclude that the suitable transformation of the direct demand model offers the potential to obtain more accurate elasticity estimates relative to the conventional disaggregate approach in the light of the scale factor problem. However, further research is required to be able to draw firmer conclusions on this issue.

5. EMPIRICAL RESULTS

At the outset we must state that we had not originally envisaged using the direct demand modelling approach and this, coupled with limited access to air travellers at airports (Wardman, Whelan and Toner, 1994), means that we do not have many observations upon which to base the analysis.

SP experiments offering choices between train and plane were conducted on air users on the routes listed in Table 4. The number of observations that enter the direct demand SP model and the number of individuals across which the demand estimates are obtained are also given for each flow. Three routes offered 9 comparisons of train and plane and hence have 8 observations for modelling purposes, whilst two routes offered 12 comparisons and thus have 11 observations available for modelling. All available scenarios were used since the SP experiments concentrated on offering improvements to rail and deteriorations to air and hence all scenarios represent positions on the air demand function which are worse than the current situation.

The SP exercise represented air by cost and frequency, with journey time fixed at the current level, whilst train was characterised by cost, frequency, journey time and, on the Edinburgh to Manchester route, interchange. In particular, we offered rail journey times which were improvements on current performance and which at their best were consistent with very high speed train operation. For example, we offered journey times of 2 hours 40 minutes between Edinburgh and London and 3 hours between Edinburgh and Birmingham compared with current best times of around 4 hours and around 4½ respectively. More details about the design of the SP exercises and the data collection stage are given in Wardman, Whelan and Toner (1994).

Table 4: Sampling of Air Users

Flow	Observations	Individuals
Edinburgh-London	8	72
Edinburgh-Birmingham	8	17
Edinburgh-Manchester	11	14
Newcastle-London	11	18
Manchester-London	8	19

We therefore have 46 observations for modelling purposes. Table 5 reports the air direct demand model which takes the conventional constant elasticity form of equation 4 with the exception of interchange which is specified in difference form since it can only take the value of 1 or 0. The exponential of the interchange coefficient therefore indicates the proportionate change in air demand resulting from the introduction of a train interchange when none was otherwise present.

Table 5: Air Direct Demand Model - Constant Elasticities

CONSTANT	-0.005 (0.13)
HEADWAY-AIR	-0.085 (0.91)
HEADWAY-RAIL	0.146 (1.93)
TIME-RAIL	1.097 (3.80)
INTERCHANGE-RAIL	0.162 (1.11)
COST-AIR	-0.964 (3.57)
COST-RAIL	0.116 (0.73)
ADJ R ²	0.628
OBS	46

The estimated model has a number of satisfactory features, particularly bearing in mind the limited number of observations upon which it is calibrated.

The goodness of fit statistic is not particularly high but in our experience is comparable with that achieved by this type of modelling approach when used to explain changes in the recorded volume of actual rail demand. The highest correlations of estimated coefficients are 0.52 between air headway and air cost and -0.53 between rail headway and rail journey time.

It is encouraging that the constant term is very small since, in an SP experiment, there is no reason for demand to vary except as a result of variation in the explanatory variables in the model. The headway elasticities are of the correct sign and seem plausible. The rail headway elasticity implies that the volume of air demand would increase by 11% if train headway deteriorated from every hour to every two hours. Although the air headway coefficient is not statistically significant, we must bear in mind the small sample size with which we are working.

The rail journey time coefficient relates to variations in train in-vehicle time and, given its importance in the SP design, it is hardly surprising that it is the most precisely estimated coefficient. The coefficient estimate represents the cross-elasticity of air demand with respect to the level of train journey time. At first sight it appears somewhat large. However, it is plausible when we bear in mind the impact of high speed trains such as TGV on air demand and that the SP exercise in many instances offered very substantial rail journey time savings. The evidence from TGV Sud-Est introduction between Paris and Lyon shows that there was a very large proportionate reduction in air demand as a result of a halving of rail journey times which implies very high cross-elasticities as we have here obtained.

Although the interchange coefficient is not statistically significant, this could be because interchange was only a relevant variable on the Edinburgh to Manchester route and only varies in 6 out of the 46 cases. This merely compounds the problems of limited sample size. However, it quite plausibly indicates that air demand would increase by 18% after the introduction of a rail interchange when none otherwise existed.

The rail cost elasticity is statistically insignificant. Although the small sample size may be a contributory factor, we are inclined to feel that it may well be that there is little scope for rail to attract business travellers from air by reducing its fares. On the other hand, the model suggests that air fare increases do have a substantial impact on air demand. We find it plausible that reductions in what are relatively low rail fares do not affect air demand greatly but that increases in relatively high air fares have a somewhat greater effect. It is plausible to advance a form of elimination by aspects decision rule as underlying business travel mode choice. This would take the form of choosing the fastest mode provided that other variables, such as cost, are acceptable. Thus improving rail fares would not influence mode choice since rail is already acceptable in this respect but increasing air fares would have an effect on demand if the higher air fare was deemed unacceptable and therefore rules out air travel. However, we do feel that there is also an element of protest against air fare increases and we regard the air fare elasticity to be too high, although this is not a serious problem since air fare is not a policy relevant variable as far as this study is concerned.

The model in Table 5 is satisfactory in several respects but it should only form a starting point to a more detailed consideration of the appropriate relationship between air demand and the characteristics of air and of competing rail services. It specified the constant elasticity form but a more general functional form is represented by:

$$V = e^{\beta X^{\lambda}} \quad (9)$$

which is estimated as:

$$\log \frac{V_2}{V_1} = \beta (X_2^{\lambda} - X_1^{\lambda}) \quad (10)$$

The implied point elasticity (η_x) is:

$$\eta_x = \beta \lambda X^{\lambda} \quad (11)$$

A value of λ close to zero implies a constant elasticity whereas a value of λ close to 1 implies that the elasticity is proportional to the level of the variable. The attraction of this elasticity function is that it allows a departure from the potentially restrictive constant elasticity position yet it does not require that the elasticity variation is as large as being proportional to the level of the variable.

Whilst β and λ can be simultaneously estimated using non linear least squares, which is available in the SAS statistical package used in this study, we did not consider this level of sophistication to be worthwhile on such a small data set. Instead we estimated the β 's of equation 10 for various values of λ and identified the combination of β 's and λ 's across the variables in the model which provided the best fit.

We would not expect the journey time elasticity to be constant. The main factor in business travellers' mode choices is the relative journey times of air and rail. If rail does not allow a round

trip to be done in a day, it will be a very unattractive option to those business travellers making a day trip. We take this threshold to be around three hours. For journey times below three hours, rail not only becomes a feasible option for those making day trips but it will tend to have quicker door-to-door journey times and hence tend to dominate. It is noticeable that trains are provided every half hour between Newcastle and London where the journey time is a little less than 3 hours yet there are only around 6 air services per day to London Heathrow and, at similar departure times, 5 per day to London Gatwick. For journeys such as Leeds to London, where the rail journey time is between 2 and 2½ hours, rail has a far bigger share of the business market than air.

We have therefore experimented with the function set out in equation 10 and also with the specification of three journey time variables according to whether:

- i) The journey time threshold of 3 hours was not satisfied in both the before and after (1 and 2) scenarios (TIME1).
- ii) The journey time threshold was crossed, that is, the threshold was satisfied in just one of the two scenarios (TIME2).
- iii) The journey time threshold was satisfied in both scenarios (TIME3).

We could not test an additional threshold of 2 hours since this was only satisfied on the London to Manchester flow and there were only four cases where this threshold was crossed and only two cases of journey time changes where this threshold was satisfied in both scenarios.

The coefficient associated with Time1 was insignificant ($t=0.51$). This confirms our hypothesis that rail journey time variations do not affect air demand when the three hour threshold is not satisfied. The removal of the insignificant Time1 variable resulted in coefficients of 1.15 and 1.02 for Time2 and Time3, with respective 95% confidence intervals of $\pm 51\%$ and $\pm 76\%$ and a goodness of fit of 0.649. We therefore proceeded to examine elasticity variation of the form specified by equation 11 for journey time provided that the three hour threshold was satisfied in at least one of the two scenarios. This variable is termed Time4. When it was specified in constant elasticity form, along with all the other variables, the goodness of fit was 0.657.

The preferred direct demand model is listed in Table 6. It contains each variable in the form which provided the best fit overall. The final column of Table 6 indicates whether the constant elasticity is retained for a variable, which is denoted by LOG, or whether the functional form expressed in equation 10 provides a better fit whereupon the value of λ which yields the best fit is given. Since the interchange variable can only take the value of one or zero, it is specified in the same form as for Table 5.

The model reported in Table 6 obtains an improvement in fit from 0.657 to 0.683 in relation to the constant elasticity equivalent. This is quite an appreciable improvement in fit in our experience and stems from some noticeable departures from the constant elasticity position.

Table 6: Preferred Air Direct Demand Model

VARIABLE	ESTIMATE	FUNCTION
CONSTANT	0.0033 (0.09)	
HEADWAY-AIR	-0.00006 (1.86)	1.5
HEADWAY-RAIL	0.2198 (3.27)	LOG
TIME4-RAIL	0.1633 (4.50)	0.5
INTERCHANGE-RAIL	0.2075 (1.54)	DUMMY
COST-AIR	-0.000002 (4.44)	1.3
COST-RAIL	0.1018 (0.71)	LOG
ADJ R ²	0.683	

The air headway elasticity is found to vary strongly with the level of headway whereas the rail headway elasticity is constant. However, the rail headway varies much less than air headway. Rail headway has maximum and minimum values of every two hours and every half hour yet the corresponding values for air are every five hours (3 flights per day) and every hour. Table 7 indicates the implied point elasticities for air headway for a range of frequency levels. It also presents a range of journey time and air cost elasticities.

Table 7: Elasticity Variation

Air Headway (minutes)	Air Headway Elasticity	Rail Time (minutes)	Rail Time Elasticity	Air Cost (pence)	Air Cost Elasticity
60	-0.04	60	0.63	15000	-0.70
120	-0.12	120	0.89	17500	-0.85
180	-0.22	180	1.10	20000	-1.01
240	-0.33	240	1.26	22500	-1.18
300	-0.47			25000	-1.36

It is to be expected that the the benefits of improving air frequency will be greater where there is a poor service than where the service frequency is good. At hourly service intervals, as exist on routes such as between London and Glasgow, Edinburgh and Manchester, the model indicates that there would be little to be gained from improved air service frequency.

The journey time elasticity variation is quite plausible. The point elasticity is only defined where the 3 hour threshold is satisfied but the arc-elasticity would be large for a journey time reduction

which meant that the threshold was crossed. Thus we have also given the point elasticity for 240 minutes. It is to be expected that the rail journey time cross-elasticity falls as rail journey time falls and an increasingly large proportion of air business travellers have switched to rail.

Although we have expressed some reservations about the validity of the absolute air cost elasticity we have obtained, the elasticity variation across a range of full air fares is quite plausible.

The elasticity variation that we have estimated may not be the complete amount. Whilst the air fare elasticity variation is plausible, there may additionally be an effect according to, for example, the competitive position of air and rail. However, we only have a limited sample size and it was not considered worthwhile to examine additional forms elasticity variation. Nonetheless, we did further examine the threshold issue. We can hypothesise that not only does the rail time cross-elasticity depend on satisfying the three hour threshold but also that the effect of rail frequency and rail cost on the demand for air travel will be strongly influenced by whether the threshold is satisfied. However, when we allowed the rail frequency and rail cost elasticities to depend on whether the three hour journey time threshold was satisfied, we could not obtain any convincing evidence that such a relationship exists, although a larger data set would be required to draw more definite conclusions.

6. CONCLUSIONS

The research reported here serves two purposes. Firstly, it has dealt with methodological issues surrounding the analysis of SP data using the direct demand method. Secondly, it provides results regarding the degree of interaction between air and rail travel.

It is our understanding that models developed on SP discrete choice data have been of the disaggregate logit form and that there has been little attempt at developing what we have termed direct demand models whose explicit purpose is to explain changes in collective behaviour as a function of changes in relevant travel variables.

We have cited the potential advantages of the direct demand approach to modelling SP data. Some of the claims require further attention to determine whether the direct demand approach does indeed have significant attractions, at least in some circumstances, in relation to the conventional disaggregate logit model. Whilst we have examined the scale factor problem in some detail, further research is also warranted in this important area.

As far as the empirical results are concerned, relatively little has previously been published regarding the interaction between rail and air. The calibrated models are very satisfactory given that they are only based on 46 observations. We have been able to obtain plausible own and cross-elasticity estimates and the analysis of elasticity variation, although by no means exhaustive, has obtained some interesting results. The calibrated models would allow us to forecast the impact on the demand for air travel of a range of improvements to long distance rail services. In particular, we note that large reductions in rail journey times can have a substantial impact on the volume of demand for air travel and that the rail journey time cross-elasticity is strongly dependent on the satisfaction of a three hour rail journey time threshold.

References

- Coslett, S.R. (1981) 'Efficient Estimation of Discrete Choice Models' in C.F. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data: With Econometric Applications*, MIT Press, Cambridge, Mass.
- Holden, D. (1992) 'Design Procedures for Stated Preference Experiments' PhD Thesis, Institute for Transport Studies, University of Leeds.
- Horowitz, J (1981) 'Sampling, Specification and Data Errors in Probabilistic Discrete Choice Models', Appendix C in *Applied Discrete Choice Modelling*, Hensher, D.A. and Johnson, L.W., Croom Helm, London.
- Maddala, G.S. (1983) 'Limited Dependent and Qualitative Variables in Econometrics' Cambridge University Press.
- Ortuzar, J. de D. and Willumsen, L.G. (1990) 'Modelling Transport' John Wiley and Sons, Chichester.
- Wardman, M. (1991) 'Stated Preference Methods and Travel Demand Forecasting: An Examination of the Scale Factor Problem' *Transportation Research* 25A.
- Wardman, M. (1994) 'Forecasting the Impact of Service Quality Changes on the Demand for Inter-Urban Rail Travel' *Journal of Transport Economics and Policy*
- Wardman, M., Whelan, G.A. and Toner, J.P. (1994) 'Air and Rail Stated Preference Mode Choice Model' Working Paper 422, Institute for Transport Studies, University of Leeds.

APPENDIX

WEIGHTED LEAST SQUARES FOR RATIO MODEL

The true model is:

$$\ln \left(\frac{m_{i2}}{m_{i1}} \right) = \beta' x_i + \varepsilon_i$$

The estimated model is:

$$\ln \left(\frac{\hat{m}_{i2}}{\hat{m}_{i1}} \right) = \beta' x_i + u_i$$

Thus:

$$u_i = \ln \left(\frac{\hat{m}_{i2}}{\hat{m}_{i1}} \right) - \ln \left(\frac{m_{i2}}{m_{i1}} \right)$$

Expanding

$$\ln \left(\frac{\hat{m}_{i2}}{\hat{m}_{i1}} \right) \text{ around } (m_{i1}, m_{i2})$$

in a Taylor series, and ignoring terms of second order and above (and suppressing subscripts i),

$$\ln \left(\frac{\hat{m}_2}{\hat{m}_1} \right) = \ln(\hat{m}_2) - \ln(\hat{m}_1)$$

$$= \ln(m_2 + (\hat{m}_2 - m_2)) - \ln(m_1 + (\hat{m}_1 - m_1))$$

$$= \ln \left(\frac{m_2 + (\hat{m}_2 - m_2)}{m_1 + (\hat{m}_1 - m_1)} \right)$$

$$\approx \ln \left(\frac{m_2}{m_1} \right) + \frac{(\hat{m}_2 - m_2)}{m_2} + \frac{(\hat{m}_1 - m_1)}{-m_1}$$

Now

$$\begin{aligned}
u &= \ln \left(\frac{\hat{m}_2}{\hat{m}_1} \right) - \ln \left(\frac{m_2}{m_1} \right) \\
&\approx \ln \left(\frac{m_2}{m_1} \right) + \frac{(\hat{m}_2 - m_2)}{m_2} - \frac{(\hat{m}_1 - m_1)}{m_1} - \ln \left(\frac{m_2}{m_1} \right) \\
&= \frac{\hat{m}_2}{m_2} - \frac{\hat{m}_1}{m_1}
\end{aligned}$$

and

$$VAR(u) = VAR \left(\frac{\hat{m}_2}{m_2} - \frac{\hat{m}_1}{m_1} \right)$$

Thus:

$$VAR(u_i) = \frac{1}{m_{i2}^2} VAR(\hat{m}_2) + \frac{1}{m_{i1}^2} VAR(\hat{m}_1) - \frac{2COV(\hat{m}_1, \hat{m}_2)}{m_{i1}m_{i2}}$$

Alternatively, using expectation notation:

$$VAR(u_i) = E(u_i^2) - [E(u_i)]^2$$

$$E(u_i) = \frac{E(\hat{m}_2)}{m_{i2}} - \frac{E(\hat{m}_1)}{m_{i1}}$$

$$\begin{aligned}
E(u_i^2) &= E \left[\left(\frac{\hat{m}_2}{m_{i2}} \right)^2 - \left(\frac{2\hat{m}_1\hat{m}_2}{m_{i1}m_{i2}} \right) + \left(\frac{\hat{m}_1}{m_{i1}} \right)^2 \right] \\
&= \frac{E(\hat{m}_2^2)}{m_{i2}^2} + \frac{E(\hat{m}_1^2)}{m_{i1}^2} - \frac{2E(\hat{m}_1\hat{m}_2)}{m_{i1}m_{i2}}
\end{aligned}$$

Thus

$$VAR(u_i) = \frac{E(\hat{m}_2^2)}{m_{i2}^2} + \frac{E(\hat{m}_1^2)}{m_{i1}^2} - \frac{2E(\hat{m}_1\hat{m}_2)}{m_{i1}m_{i2}} - \left(\frac{[E(\hat{m}_2)]^2}{m_{i2}^2} + \frac{[E(\hat{m}_1)]^2}{m_{i1}^2} - \frac{2E(\hat{m}_1\hat{m}_2)}{m_{i1}m_{i2}} \right)$$

$$\left(= \frac{1}{m_{i2}^2} VAR(\hat{m}_2) + \frac{1}{m_{i1}^2} VAR(\hat{m}_1) - \frac{2COV(\hat{m}_1, \hat{m}_2)}{m_{i1}m_{i2}} \right)$$

$\hat{m}_1$ and $\hat{m}_2$ are used as estimates of the unknown m_1 and m_2 .