Comparing Forecasting Effectiveness Through Air Travel Data

Suzilah Ismail[a] and Tuan Zalizam Tuan Muda[b]

[a]*Faculty of Quantitative Sciences*
Universiti Utara Malaysia, 06010 Sintok, Kedah.
Email: halizus@uum.edu.my

[b]*Faculty of Information Technology*
Universiti Utara Malaysia, 06010 Sintok, Kedah.
Email: zalizam@uum.edu.my

ABSTRACT

Airline traffic forecasting in the medium term is important to airlines and regulatory authorities that attempt to plan and schedule capacity. This study examines a number of alternative approaches to forecasting short to medium term (1 to 3 years) air traffic flows. The data examined are flows between the UK and six other countries over the period of 1961-2002, which has seen substantial changes in both transport technology and economic development. The economic drivers, under consideration, are price, income and bilateral trade. The forecasting models employed include autoregressive models, autoregressive distributed lag models specified using various statistical and economic criteria and a newly developed automatic method for model specification (PcGets), as well as time varying parameter models. Various approaches to including interactions between the contemporaneous air traffic flows are examined including pooled autoregressive distributed lag models and the inclusion of a 'world' variable that measures overall trade growth in the world economy. Based on the analysis of forecasting error measures, it is concluded that time varying parameter models that include the 'world' variable with an average error of around 2.5% outperform alternative forecasting models. This is perhaps explained by the dramatic structural changes seen in the air traffic market.

Keywords:

Airline traffic; Comparative forecasting accuracy; Econometric model building; Time varying parameter; Pooled cross-section time series

1.0 INTRODUCTION

This study is motivated by the limited evidence on forecasting accuracy of demand models of air travel and the comparative performance of alternative model specifications. In particular, earlier evaluations(e.g. by Young (1972), Anderson and Kraus (1980), Ippolito (1980), Friedstrom and Thune-Larson (1989), Kaemmerle (1991), Jorge-Calderon (1997) and Grubb and Mason (2001)) have been limited to only short forecast horizons, which are too short to be valuable for planning purpose. This problem is exacerbated by comparative performance that may be dependent on forecast time horizon. In addition, recent years have seen the development of a number of alternative econometric modelling approaches that offer the prospect of enhanced forecasting accuracy but have rarely been compared (Allen and Fildes, 2001). These approaches include:

- Hendry's general to specific approach to model building with full lag structure (Hendry,1986);
- A model that includes full lag structure and pools cross-sectional evidence. Garcia-Ferrer *et al.* (1987) show that using pooled data helps improving modelling performance;
- A PcGets automatic selection model based on Hendry's methodology (Hendry & Krolzig, 2001);
- A time varying parameter model (Zellner, Hong & Min, 1991);
- An enhanced version of the above approaches that includes a 'world' variable to capture the overall growth in demand across all countries.

In the light of the importance of airline forecasting on one hand and the paucity of evidence on the relative benefits of these alternative approaches on the other, this research compares the accuracy of the above five approaches in contrast to 'naïve' autoregressive alternatives.

It is hoped this research will benefit forecasting practitioners and econometricians. The 'best' model obtain will lead to better forecast, there fore the airlines can increase their profit through transport planning. On the other hand, the comparison of recent alternatives approaches can contribute to the limited research in this area.

The plan of the paper is as follows. Section 2 describes variables and data. Section 3 presents models, methodology and error measures. The results based on ex-post forecasting up to three years ahead are discussed in section 4. Section 5 offers some conclusions on the relative accuracy of the particular methods, noting how structural change in the air traffic market needs to be taken into account to achieve the most accurate forecasts.

2. 0 VARIABLES AND DATA

The dependent variable is the demand for air travel. It is measured by the total of annual international passenger traffic (000's) between UK reporting airports and Germany, Sweden, Italy, Japan, USA and Canada. These countries are chosen to avoid routes of high tourist intensity, e.g. Spain. Data are obtained from the Annual Statement of Movements (Passenger and Cargo) published by UK's Civil Aviation Authority (CAA) for 1961-2002. They include all passengers carried on scheduled and chartered services by all airlines flying between UK and each of the six countries, excluding those carried on chartered services by government departments. The figures do not make distinction between economy and business classes.

The process of selecting the explanatory variables could start with several potential variables such as GDP, population growth, employment/unemployment rate, airfares, and volume of trade. To maintain simplicity of model formulation, the number of explanatory variables is narrowed down to include only income, trade and price, which have been proved important in earlier studies of the demand for air travel (see Quandt & Baumol, 1966; Kaemmerle, 1991). When income is high, more people may want to fly to destinations rather than use other modes of transportation. As a result, demand for air travel may be high. When airfares go up, demand for air travel may be low. Trade is included to capture the effects of passengers who travel for business purpose. Swan (2002) notes that travel grows with trade. The explanatory variables are aggregated to provide a single route based measure using a weighted average in line with passenger flows. Details about data sources, data transformation and weighting procedure can be obtained from the author.

Prior to estimation, the log-transformed data are tested for stationarity using the Augmented Dickey-Fuller (1981) (ADF) test. The results of the test show that all variables under investigation are I(1) and thus need to be differenced to achieve stationarity.

3.0 METHODOLOGY

The first 31 observations (1961-1991) are used to fit models and the remaining 11 observations (1992-2002) are used to evaluate the models. Various econometric methodologies have been proposed as suitable for developing effective forecasting models in the literature. Here we consider those that have proved empirically most effective, including the general-to-specific approach associated with the London School of Economics (Pagan, 1987; Clements & Hendry, 1998), Zellner's Structural Equation Modelling, Time Series Analysis Approach (SEMTSA) (Zellner & Palm, in press) and time varying parameters models. In addition, a benchmark autoregressive model is included along with two naïve models in order to measure the improvements arising from the use of more complex models and additional explanatory variables.

The general-to-specific methodology supposes 'a general class of possible models' can be specified and then proceeds to select the 'parsimonious model from that class' by the process of elimination, at each stage sequentially testing any parameter restrictions as suggested by the data or alternatively economic theories. Initially, a 'full' lag structure is included. Insignificant variables are then dropped from the model. Our starting point is therefore a general autoregressive distributed lag model defined as follows:

- **Autoregressive Distributed Lag (ADL) Model**

The general unrestricted Autoregressive Distributed Lag (ADL) model can be written in the form of,

$$\Delta y_{it} = \alpha_{i0} + \sum_{j=1}^{J} \alpha_{ij} \Delta y_{i(t-j)} + \sum_{k=1}^{K} \sum_{j=0}^{J_k} \phi_{ikj} \Delta x_{ik(t-j)} + \varepsilon_{it} \tag{1}$$

where j is the lag length, i = 1,2,...,N (countries), t = 1,2,...,T (time periods) and Δy_{it} is the growth rate of the number of passenger in year t for route i. Δx_{ikt} is the growth rate of the kth explanatory variable in year t for route i, ε_{it} are identically independently distributed random errors with mean zero and variance $\sigma^2_{\varepsilon_{it}}$, α and ϕ are unknown parameter vectors to be estimated.

- **Pooled ADL Model**

An extended version of ADL model is derived from pooling the data and then estimating all models simultaneously using the seemingly unrelated regression (SUR) approach. This is based on the assumption that there exists common (but unmeasured) influence on air traffic flows and this induces contemporaneous correlation among the errors of the individual routes. The SUR approach allows intercepts as well as slope coefficients to vary over all routes. SUR model can be compactly represented as,

$$\Delta Y = \Delta X\beta + \varepsilon \tag{2}$$

where ΔX is a block diagonal matrix of N cross sections and each one of ΔX_t has a dimension $T \times K'$. The K' explanatory variables include the constant term (and may include the lagged dependent variables). Generalized least squares (GLS) provides an efficient estimator in this approach.

- **PcGets Automatic Econometric Model Selection**

PcGets is a software programme created to overcome some of the subjectivities inherent in model building, which typically relies overly on the tacit knowledge of the model builder (Pagan, 1999; Allen & Fildes, 2004). Starting with a ADL model with 'full' lag structure (known as the general unrestricted model - GUM), the selection process uses the recommended non-stringent significance level (e.g. 0.9) to simplify the GUM whilst ensuring congruence (where the model passes all the diagnostic tests) is maintained throughout the process until the final stage where parameter restrictions are rejected and the final model is identified (Hendry & Krolzig, 2001).

The difference between the ADL model and PcGets lies in the selection process. The former is carried out by the researcher based on such factors as the correct signs of the coefficients according to economic theory, the satisfaction of individual t-tests and all the diagnostics tests, significant overall F-values for groups of coefficients etc. In contrast, with PcGets the model is chosen automatically based on several stages, which are pre-search reductions (involving both a top-down and bottom-up specification process) with multiple search paths, encompassing tests and sub sample reliability tests.

- **Time Varying Parameter (TVP) Modelling**

In recent years there has been increased questioning of the assumption that the regression coefficients are constant in models such as those described above. The shift from the assumption of constancy in the effects of the economic drivers to time-varying parameters has occurred because of the recognition on, among other factors, changes in aggregation effects and policies, adaptive optimisation on the part of economic agents (Garcia-Ferrer *et al.*, 1987; Harvey, 1989; Riddington, 1993; Zellner & Min, 1993), and changes in tastes. It is particularly relevant in an industry such as the airline industry where preferences for travel destinations and its emerging economic role, among other factors, change over time. The TVP model is represented as follows,

$$\boldsymbol{\Delta y}_{it} = \boldsymbol{\Delta x}'_{it}\boldsymbol{\beta}_{it} + \boldsymbol{\varepsilon}_{it} \tag{3}$$

$$\boldsymbol{\beta}_{it} = \boldsymbol{\beta}_{it-1} + \boldsymbol{v}_{it} \tag{4}$$

where Δx_{it} is a vector of explanatory variables, β_{it} is a vector of parameters subject to time-dependent variation, v_{it} are identically independently distributed random errors with zero mean and constant variance. The βs are assumed to be adaptive in nature, subject to permanent and transitory changes and are modelled in equation (4) as a multivariate random walk without drift.

- **Naïve Models**

Using a similar modelling technique to that used by Garcia-Ferrer *et al.* (1987), two benchmarks models are used as basis of comparison against all other estimated models. The two benchmark models are naïve model one (NM1) which assumes zero growth and naïve model two (NM2) which assumes that the future growth rate equals the rate of last period. They are defined respectively as,

$$NM1: \Delta \hat{y}_t = 0 \tag{5}$$

$$NM2: \Delta \hat{y}_t = \Delta y_{t-1} \tag{6}$$

- **Autoregressive of Order 3 [AR(3)] Model**

Garcia-Ferrer et al (1987) argue that the poor performance of NM1 and NM2 models in their study could be improved by fitting an autoregressive of order 3 [AR(3)]. This model is also included here and is defined as,

$$\Delta y_{it} = \beta_{0i} + \beta_{1i}\Delta y_{i(t-1)} + \beta_{2i}\Delta y_{i(t-2)} + \beta_{3i}\Delta y_{i(t-3)} + \varepsilon \tag{7}$$

Various computer programmes have been used to estimate the models, Microfit for the ADL models, EViews for the time varying parameter models and PcGets for the automatic model specification. After estimating the regressions, the ex-post forecasts for one, two and three years ahead are generated recursively with the models updated. In order to evaluate the models, two error measures are used, i.e. Root Mean Square Error (RMSE) and Geometric Root Mean Square Error (GRMSE).

RMSE for *l*-step-ahead forecasts is generated over n out of sample data points ($T_0+1, T_0+2, \ldots, T_0+n$) where T_0 is taken as 1991 and is written as,

$$RMSE = \sqrt{\frac{\sum_{t=T_0}^{T_0+n-l} e_t^2(l)}{n+1-l}} \tag{8}$$

where $e_t(l) = y_{t+l} - \hat{y}_t(l)$, y_{t+l} is the actual observation at t+*l*, and $\hat{y}_t(l)$ is the *l*-step-ahead forecast of y_{t+l} generated from the models using T_0 to T_0+n-l observations. RMSE can be greatly affected by outliers (Fildes, 1992). As a means of overcoming such a problem, when confronting a significantly large error term due to a particularly bad forecast, Fildes suggests using GRMSE, which is written as,

$$GRMSE = [\prod_{t=T_0}^{T_0+n-l} e_t^2(l)]^{\frac{1}{2(n+1-l)}} \qquad (9)$$

Between the two error measures, RMSE is commonly favoured among both practitioners and academics (Armstrong & Collopy, 1992) despite its lack of robustness. In an empirical study of forecast error characteristics Fildes shows GRMSE does not suffer such an acute problem.

4.0 ESTIMATED MODELS AND FORECASTING PERFORMANCE

After appropriate simplification the estimated ADL models for each route (over the period of 1961 to 1991) are shown in Table 1. All the variables show the expected signs (i.e. positive sign for income and trade and negative sign for ticket price). The F tests indicate an overall significance of the coefficients in their respective models. The values of $\overline{R}^2$ range from 0.11 for UK-Japan route to 0.38 for UK-Canada route which is small. Perhaps due to the log transformation and/or the models are missing important variables. All models pass the tests of normality, serial correlation, functional form and heteroscedasticity. Only UK-Germany and UK-Canada models include one period lagged dependent variable, which is significant at the 5% level. The income variable is included for UK-Italy only. This indicates that the higher the income, the more the tendency for people travelling abroad by air, but the demand for air travel does not rise proportionally with income given the coefficient is less than one. In all other routes, income does not directly appear to influence the demand for air travel significantly. This suggests that, with the rapid development of transportation technologies, air travel is becoming a necessity in people's lives and business activities. Three out of six routes show that trade is a more reliable predictor compared to income. More bilateral trade implies a higher level of integration between countries; as a result, an increasing need for air travels between countries. This is consistent with Swan's (2002) study on the US. Price is only included in UK-Italy and UK-Canada routes. And it appears that air travel is price inelastic. This is perhaps because countries included in this study are not high tourist. Another possible explanation is that there are inevitable errors in the measurement of the price variable as suggested by Swan (2002) who used yield data instead as a measure for the US, an option not available here due to data limitations.

One path towards improving the performance of the ADL models is to take into consideration the possible interdependencies among the routes. This could arise from common economic and/or social factors across the world such as political upheaval (e.g. Iraq's invasion of Kuwait). A solution to this problem is to consider a common variable believed to affect all the routes. One possible variable is the total trade of all industrial countries, which here is called 'world' variable (Garcia-Ferrer *et al.*, 1987 used 'world' variable to improve their forecasts for GDP growth rates for 9 countries). Another reason to include 'world data' is to capture the effects of through traffic passengers. 'Through traffic' is defined as traffic originating from another country or airport, passing through the country in question, on the way to another (third) country.

As shown in Table 2, the inclusion of 'world' variable appears to have improved the ADL models in terms of $\overline{R}^2$s. Individually, the 'world' variable appears significant in all routes except UK-Sweden and UK-Japan. The same variables that have been included in the ADL and extended ADL models are then used to build up models for the pooled ADL and TVP models (the results of which are not included in the table).

Table 3 provides summary information on one, two and three-years-ahead forecasts, respectively. The one-year ahead forecasts show that pooled ADL models with the 'world' variable perform better compared to the unpooled world ADL in terms of both errors measures but the results switch for two and three years ahead forecast. It is noteworthy that both error measures for one to three-years-ahead forecasts identify the TVP model, which included the 'world' variable, as the best performing model. The NM1 remains the worst.

The models permit us to answer the question of the effects of September 11th. Using the best model (i.e. TVP with the 'world' variable included) estimated using data to 2000 gives residuals for 2001 and 2002 with median value for the residuals for 2001 and 2002 significantly negative at the 5% level with growth down an unexpected 7.9% and 5.0% respectively. By 2002 the indirect effects of the catastrophe on trade and income were coming through so the overestimate was less (based on the realised values of the explanatory variables).

5.0 CONCLUSIONS

In this research, several models have been estimated for air passenger traffic demand between the UK and six selected countries. The demand models developed were then evaluated for their (conditional) out-of-sample forecasting accuracy. The one, two and three-year-ahead ex-post forecasts were generated and compared against the benchmark models namely, naïve model 1 (NM1) and naïve model 2 (NM2). Two different criteria – RMSE and GRMSE were used in evaluating their performances. Different error measures tended to produce different results of model evaluations, and therefore any discussion of comparative forecasting performance must be supported with evidence obtained from several error measures.

The ADL models with fixed parameter outperformed both naïve models but lost out to the autoregressive AR(3) models in terms of GRMSE despite the apparently well specified economic basis of the model. This offers a further example of why a pure time series model may be more accurate than a causal model and suggests that the proposed causal model is miss-specified with important drivers missing, in particular measures of structural change. The ADL models (with or without the 'world' variable) proved better performers compared to PcGets, contradicting our first hypothesis as to the benefits of removing subjectivity in model building.

Pooling the data and simultaneously estimating the models, have improved the forecasting performance. This confirms the conclusions of Garcia-Ferrer et al (1987) and Zellner et al (1991) where forecasting models developed using the seemingly unrelated regression (SUR) approach and recognizing contemporaneous co-variation showed improvements over those models estimated individually.

When the 'world' variable was included in the models, the time varying parameters (TVP) shows significant improvements over fixed parameters models as well as over the autoregressive benchmark models. This supports the conclusions of Riddington's (1993) survey as well as Zellner et al (1991) and underlines the importance of recognising slowly changing economic and market structures.

September 11th, 2001 has a major negative impact on the airline industry but 2002 data show that the recovery process was already underway although not back to its long term growth path.

In summary, the results of this research show that employing 'advanced' econometric methods leads to improve forecasting performance for air traffic with error levels up to three years that are probably acceptable for some (if not all) planning purposes. Structural changes with continuous changes in taste and a discontinuous break arising from the terrorism attack in 2001 demonstrate the importance of selecting adaptive models with time varying parameters as well as including the 'world trade variable' that captures those concurrent economic developments that affect air traffic routes.

REFERENCES

ABC World Airways Guide, Dunstable, UK, various issues.

Allen, P.G., & Fildes, R. (2003). Level, differences and ecms – principles for improved econometric forecasting. *Lancaster University Management School working paper*.

Allen, P.G., & Fildes, R. (2001). Econometric forecasting: strategies and techniques. In: *Principles of Forecasting: A Handbook for Researchers and Practitioners*, J.Scott Armstrong (ed.): Norwell, MA:Kluwer Academic Publishers.

Anderson, J.E., & Kraus, M. (1980). Quality of service and the demand for air travel. *The Review of Economics and Statistics*, 63, 533-540.

Armstrong, J.S., & Collopy, F. (1992). Error measures for generalising about forecasting methods: empirical comparisons. *International Journal of Forecasting*, 8, 69-90.

Business Monitor. (1992). Central Statistical Office (Quarter Two).

Charemza, W.W., & Deadman, D.F. (1992). New *direction in econometric practice*. Edward Elgar Publishing, Aldershot, UK.

Clements, M.P. & Hendry D.F. (1998). *Forecasting Economic Time Series*. Cambridge University Press, Cambridge.

Cooley, T.F., & Prescott, E.C. (1973). Systematic (non-random) variation models varying parameter regression: a theory and some applications. *Annals of Economics and Social Measurement*, 2, 463-473.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.

Dickey, D.A., and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with unit root. *Econometrica*, 49, 1057-1072.

Du Preez J. & Witt, S.F. (2003). Univariate versus multivariate time series forecasting: an application to international tourism demand. *International Journal of Forecasting*, 19, 435-451.

Engle, R.F., & Granger, C.W.J. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*, 55, 251-276.

Fildes, R., (1992). The evaluation of extrapolative forecasting methods. *International Journal of Forecasting*, 8, 81-98.

Friedstorm, L., & Thune-Larsen, H. (1989). An econometric air travel demand model for the entire conventional domestic network: the case of norway. *Transportation Research 23B*, 213-223.

Garcia-Ferrer, A. (1987). Macroeconomic forecasting using pooled international data. *Journal of Business & Economic Statistics*, 5, 53-67.

Grubb, H., & Mason, A. (2001). Long lead-time forecasting of uk air passengers by holt-winters methods with damped trend. *International Journal of Forecasting*, 17, 71-82.

Harvey, A.C. (1989). *Forecasting structural time series models and the kalman filter*. Cambridge University Press, Cambridge.

Hendry, D.F. (1986). Econometric modelling with cointegrated variables: an overview. *Oxford Bulletin of Economics and Statistics*, 48, 201-211.

Hendry, D.F., & Krolzig, H-M. (2001). *Automatic econometric model selection using PcGets 1.0*. Timberlake Consultants Ltd.

Ippolito, R.A. (1981). Estimating airline demand with quality of service variables. *Journal of Transport Economic and Policy*, 15:1, 7-15.

Jorge-Caleron, J.D., (1997). A demand model for scheduled airline services on international european routes. *Journal of Air Transport Management*, 3, 23-35.

Judge, G.G., Griffiths, W.E., Hill, R.C., & Lee, T.C., 1985. *The theory and practice of econometric*, 2nd Ed. Wiley, New York, Chap. 7.

Kaemmerle, K.C. (1991). Estimating the demand for small community air service. *Transportation Research 25A*, 101-112.

Nelson, C.R., & Plosser, C.I. (1982). Trends and random walks in macroeconomic time series, some evidence and implications. *Journal of Monetary Economics*, 10, 139-162.

O'Conner, W.E. (1989). *An introduction to airline economics*. 4th. Ed., Praeger, London/New York.

Pagan, A.R. (1987). Three econometric methodologies: a critical appraisal. *Journal of Economic Surveys*, 1, 3-24.

Pagan, A.R. (1999). The Tilburg experiments: impressions of a drop-out. In J.R. Magnus & M.S. Morgan (Eds.) *Methodology and Tacit Knowledge*, pp. 369-374. Wiley, Chichester, UK.

Quandt, R.E., & Baumol, W.J. (1966). The demand for abstract transport modes: theory and measurement. *Journal of Regional Science*, 6:2, 13-26.

Riddington, G.L. (1993). Time varying coefficient models and their forecasting performance. *Omega*, 21, 573-583.

Rodriquez, L.L. (1981). Traffic forecast for planning purposes. Presented at *an ICAO Regional Workshop on Aviation Forecasting and Economic Planning*, Mexico City.

Shaw, S. (1995). *Airline marketing and management*. 3rd. Ed., Pitman Publishing, London.

Swan, W., (2002). Three mistakes we made in forecasting. Presented at *International Symposium on Forecasting*, Dublin.

Teodorovic, D. (1988). *Airline operations research, transportation studies*. 10, Gordon and Breach Science Publishers.

Young, K.H. (1972). A synthesis of time series and cross-section analyses: demand for air transportation service. *Journal of the American Statistical Association*, 67, 560-566.

Zellner, A. (1987). An efficient method of estimating seemingly unrelated regressions and test for aggregation bias. *Journal of the American Statistical Association*, 57, 348-368.

Zellner, A., Hong, C., & Min, C., 1991. Forecasting turning points in international output growth rates using bayesian exponentially weighted autoregression, time varying parameter and pooling techniques. *Journal of Econometrics*, 49, 275-304.

Zellner, A., & Min, C., 1993. Bayesian and non-bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *Journal of Econometrics*, 56, 89-118.

Zellner, A., & Palm F.C. (eds.), forthcoming. *Structural Econometric Modelling, Time Series Analysis (SEMSTA) Approach*. Cambridge University Press, Cambridge.

Table 1: Estimated ADL Models, 1961-1991

Variables	Routes					
	UK-Germany	UK-Sweden	UK-Italy	UK-Japan	UK-USA	UK-Canada
Intercept	0.05* (0.05)	0.09** (0.00)	0.06** (0.01)	0.16** (0.00)	0.08** (0.00)	0.02 (0.40)
Δy_{t-1}	0.41** (0.04)	-	-	-	-	0.60** (0.00)
Δx_{1t} (Income)	-	-	0.49** (0.00)	-	-	-
Δx_{2t} (Trade)	-	0.34** (0.04)	-	0.42** (0.04)	0.45** (0.02)	-
Δx_{3t} (Ticket price)	-	-	-0.13* (0.07)	-	-	-0.17** (0.02)
$\overline{R}^2$	0.11	0.11	0.30	0.11	0.15	0.38
Diagnostics Test Results						
Normality	1.34 (0.51)	0.94 (0.63)	0.65 (0.72)	0.15 (0.93)	0.29 (0.87)	0.82 (0.66)
Autocorrelation	0.23 (0.63)	0.62 (0.43)	2.70 (0.10)	0.05 (0.83)	0.01 (0.95)	1.66 (0.20)
Functional Form	0.72 (0.40)	1.44 (0.23)	0.66 (0.42)	0.06 (0.80)	1.07 (0.30)	0.04 (0.84)
Heteroscedasticity	0.99 (0.32)	1.03 (0.31)	0.78 (0.38)	0.01 (0.93)	0.28 (0.60)	0.84 (0.36)
F	4.53** (0.04)	4.60** (0.04)	6.66** (0.00)	2.69* (0.09)	7.24** (0.00)	9.64** (0.00)

( ) – p-value; ** Significant at 5% level ; * Significant at 10% level

Table 2: Estimated ADL Models with World Variable, 1961-1991

Variables	Routes					
	UK-Germany	UK-Sweden	UK-Italy	UK-Japan	UK-USA	**UK-Canada**
Intercept	0.01 (0.65)	0.05* (0.09)	0.02 (0.38)	0.13** (0.02)	0.03 (0.29)	-0.01 (0.69)
Δy_{t-1}	0.11 (0.59)	-	-	-	-	0.39** (0.04)
Δx_{1t} (Income)	-	-	0.44** (0.00)	-	-	-
Δx_{2t} (Trade)	-	0.24 (0.16)	-	0.40* (0.06)	0.29 (0.10)	-
Δx_{3t} (Ticket price)	-	-	-0.11 (0.13)	-	-	-0.17** (0.01)
Δx_{4t}(World trade)	1.02** (0.01)	0.64 (0.14)	0.64* (0.05)	0.80 (0.43)	0.98** (0.01)	0.68** (0.03)
$\overline{R}^2$	0.32	0.15	0.36	0.12	0.32	0.46
Diagnostics Test Results						
Normality	1.47 (0.48)	2.30 (0.32)	0.54 (0.76)	0.19 (0.91)	3.32 (0.19)	1.62 (0.44)
Autocorrelation	1.80 (0.18)	0.13 (0.72)	1.29 (0.26)	0.50 (0.48)	0.16 (0.90)	1.88 (0.17)
Functional Form	0.38 (0.54)	0.22 (0.63)	0.06 (0.80)	0.07 (0.80)	0.83 (0.36)	0.07 (0.79)
Heteroscedasticity	0.27 (0.60)	0.04 (0.84)	1.37 (0.24)	0.01 (0.93)	1.29 (0.26)	0.43 (0.51)
F	7.59** (0.00)	3.55** (0.04)	6.40** (0.00)	2.58* (0.09)	7.92** (0.00)	9.10** (0.00)

( ) – p-value; ** Significant at 5% level ; * Significant at 10% level

Table 3: Comparative forecasting performance of one, two and three years ahead of air traffic models, measured by RMSE and GRMSE in percentage points (ranking).

Model	Median (ranking) 1 year ahead forecasts	Median (ranking) 2 years ahead forecasts	Median (ranking) 3 years ahead forecasts	Overall rank
		RMSE		
1. NM1	8.87(11)	8.04(11)	8.94(11)	11
2. NM2	8.85(10)	7.75(8)	8.12(10)	10
3. AR(3)	7.58(8)	7.49(6)	8.07(9)	8
4. ADL	7.36(7)	7.96(10)	6.95(7)	7
5. Pooled ADL	7.21(5)	7.05(5)	6.79(6)	6
6. PcGets	8.18(9)	7.91(9)	8.07(8)	9
7.TVP	6.87(3)	6.41(4)	6.74(5)	5
World data included:				
8. ADL	6.93(4)	5.47(2)	5.46(2)	2
9. Pooled ADL	6.48(2)	7.50(7)	6.15(3)	3
10. PcGets	7.35(6)	6.15(3)	6.60(4)	4
11. TVP	***5.86(1)***	***5.15(1)***	***5.14(1)***	**1**
		GRMSE		
1. NM1	5.86(11)	5.30(11)	5.50(11)	11
2. NM2	3.79(4)	4.67(9)	4.88(9)	9
3. AR(3)	3.51(2)	3.32(2)	3.95(3)	2
4. ADL	4.80(9)	3.87(4)	4.60(8)	8
5. Pooled ADL	4.28(8)	4.69(10)	4.04(5)	6
6. PcGets	4.93(10)	4.46(7)	5.18(10)	10
7.TVP	4.10(6)	4.20(6)	4.17(6)	5
World data included:				
8. ADL	3.89(5)	3.71(3)	3.51(2)	3
9. Pooled ADL	3.84(3)	4.54(8)	4.33(7)	7
10. PcGets	4.19(7)	4.01(5)	4.00(4)	4
11. TVP	***2.43(1)***	***2.48(1)***	***2.31(1)***	**1**

Note: The values in the parentheses indicate the ascending ranks of the RMSE and GRMSE, respectively