WORKING PAPERS

Investigação - Trabalhos em curso - nº 109, Julho de 2001

THE FORECASTING ABILITY OF A COINTEGRATED VAR DEMAND SYSTEM WITH ENDOGENOUS VS. EXOGENOUS EXPENDITURE VARIABLE

Margarida de Mello

Kevin S. Nell

FACULDADE DE ECONOMIA

UNIVERSIDADE DO PORTO

www.fep.up.pt

Faculdade de Economia do Porto - R. Dr. Roberto Frias - 4200-464 - Porto - Portugal
Tel . +351 225 571 100 - Fax. +351 225 505 050 - http://www.fep.up.pt

THE FORECASTING ABILITY OF A COINTEGRATED VAR DEMAND SYSTEM WITH ENDOGENOUS VS. EXOGENOUS EXPENDITURE VARIABLE: AN APPLICATION TO THE UK IMPORTS OF TOURISM FROM NEIGHBOURING COUNTRIES

MARGARIDA DE MELLO

Faculdade de Economia do Porto
Rua Dr. Roberto Frias
4200-464 Porto, Portugal

email: mmdm@clix.pt

KEVIN S. NELL

National Institute of Economic Policy
P. O. Box 32848
Braamfontein 2017, Johannesbourg, South Africa

ABSTRACT

This paper uses Sims's VAR methodology, as an alternative to Deaton and Muellbauer's AIDS approach, to establish the long-run relationships between I(1) variables: tourism shares, tourism prices and UK tourism budget. The VAR deterministic components and sets of exogenous and endogenous variables are established, and the Johansen's rank test is used to determine the cointegrated vectors in the system. The structural form of the cointegrated VAR is identified and the long-run parameters are estimated under several theoretical restrictions. The restricted cointegrated VAR reveals itself a theoretically consistent and statistically robust means to analyse the long-run demand behaviour of UK tourists and an accurate forecaster of the destinations' shares.

1. INTRODUCTION

The Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980a, 1980b) has been used, in tourism demand contexts, for estimating destinations' tourism shares and for verifying consumers theory restrictions of homogeneity and symmetry (e.g. Papatheodorou, 1999; De Mello *et al*., 2001). The model's specification includes an assumed endogenous-exogenous division of variables that may be questionable and the use of nonstationary time series. When dealing with nonstationary data, the concept of cointegration is synonymous with the concept of long-run equilibrium. Failure to establish cointegration often means the non-existence of a steady state relationship among the variables and "*it is important to recognise the effect of unit roots on the distribution of estimators*" (Harvey, 1990, p.83). Hence, the estimation results obtained with an AIDS model can be deemed spurious and the statistical inference invalid, if the assumption of exogenous regressors does not hold and/or no cointegrated relationship(s) exist. Consequently, there seems to be a risk involved in the estimation of systems with nonstationary data which regress endogenous variables on several assumed exogenous variables, without sanctioning their statistical validity with appropriate testing and cointegration analysis. Given that the number of cointegrated vectors is unknown, and given the possibility of simultaneously determined variables, empirical analysis must go one step further and specify econometric models which can be efficiently estimated and validly tested within a system of equations approach.

The main goal of this paper is to contribute an empirical basis for the validation or otherwise, of the estimation and inference procedures implemented with an AIDS model for the UK tourism demand. We do so by using Sims' (1980) vector auto-regressive (VAR) approach to specify the relationship between destinations' tourism shares and their determinants, Johansen's (1988) reduce rank test to establish the number of cointegrated vectors in the system, the findings of Pesaran, Shin and Smith (2000) regarding the structural analysis of cointegrated VARs with exogenous I(1) variables, and the procedures of Pesaran and Shin (1998) to exactly-identify the long-run coefficients of the cointegrated vectors in accordance to the theoretical principals underlying the AIDS approach. We provide empirical evidence for sanctioning the cointegrated VAR and AIDS models as statistically robust and theoretically consistent means to produce valid and reliable estimates of the structural parameters underlying the relationships between destinations' tourism shares, tourism prices and an origin's real per capita tourism budget. In addition, we confirm the competence of the VAR model, both in its general unrestricted reduce-form and under the full set of theoretical restrictions, for providing accurate forecasts of the destinations' shares.

The paper proceeds as follows. Section 2, addresses the main features of the VAR methodology, establishes the order of integration and the appropriate lag-length of the variables, specifies the (unrestricted) VAR model for the UK tourism demand and presents the forecasting results obtained with this specification. Section 3, applies the Johansen rank test to determine the number of cointegrated vectors and presents the cointegrated structural VAR estimates under exactly- and over-identifying restrictions. Section 4, provides the forecasts obtained with the cointegrated VAR and compares them with those obtained with the unrestricted VAR and AIDS model of De Mello *et al* (2001). Section 5 concludes.

2. VAR MODELLING OF THE UK TOURISM DEMAND

The problem of simultaneous bias is often present in a structural system because this specification expresses each endogenous variable as a separate function of other endogenous variables, alongside predetermined variables. Since the endogenous variables are correlated with the error terms, the structural coefficients cannot be consistently estimated by OLS. This problem can be removed if the structural equations are solved for the existing endogenous variables, making them dependent solely on predetermined variables and stochastic disturbances. Since the former are assumed to be uncorrelated with the latter, OLS applied to the reduced-form equations generates consistent and asymptotically efficient estimates. Yet, the estimates obtained with this procedure are those of the reduced-form and not of the structural-form coefficients which are ultimately of interest. Since the latter are combinations of the former, the possibility exists that the structural coefficients can be retrieved from the reduced-form coefficients. Whether this is the case, brings about the problem of identification.

Frequently, the "secret" for identification, is related to the use of 'zero restrictions'. Often, however, models seem to be formulated with variables added to equations and deleted from others merely to achieve identification, and without much economic justification. Critics to multi-equation structural modelling have been centred on the role of zero restrictions and on the assumed exogenous/endogenous division of variables. Sims (1980), regards zero restrictions as "incredible" and devises a new approach for the specification and estimation of multi-equation systems, known as vector autoregressive (VAR) methodology, where all variables are endogenous and no zero restrictions are imposed. However, Sims' VAR is often labelled an a-theoretical approach to long-run equilibrium analysis since much of the long-run analysis within this 'purely-statistical' approach is conducted "*without providing an explicit account of the type of equilibrium theory that may underlie it*", and "*empirical applications of this methodology have focused on the statistical properties of the underlying economic time series, often at the expense of theoretical insights and economic reasoning*" (Pesaran, 1997, p. 178). Hence, the features that make the VAR a flexible and simple tool, also mark an area of weaknesses: a reduced-form VAR is both a truly simultaneous system and a simple specification since all it's variables are regarded as endogenous and it requires little more than the choice of appropriate variables and a suitable lag-length. Yet, unlike the traditional structural systems, an unrestricted VAR does not use any *a priori* information and, unless the underlying structural model can be identified from the reduced-form, the interpretation of its estimates is difficult.

If autocorrelation exists in the error terms of a VAR, the predetermined right-hand side variables can be correlated with the error terms leading to inconsistent estimators. So, proper lag-length selection is crucial in VAR modelling. However, the longer the lag-length, the faster degrees of freedom are eroded and, given the limited number of observations generally available in most empirical analysis, the introduction of several lags for each variable can be a problem. Still, if an appropriate lag-length can be established, the error terms of each equation are serially uncorrelated and, as a VAR expresses the current values of each endogenous variable as a function solely of predetermined variables and these are not correlated with the error terms, each equation can be estimated by OLS providing consistent and asymptotically efficient estimates.[1]

A VAR can be viewed as a reduced-form system with no exogenous variables specially adapted for forecasting. Even so, an unrestricted VAR model may be over-parameterised in the sense that some lagged variables in the model could be properly deleted on the basis of statistical insignificance. Yet, the advocates of this methodology advise against this procedure arguing that the imposition of zero restrictions may suppress important information, and that the regressors in a VAR are likely to be highly

[1] If the appropriate lag-length is not the same for all variables, or if the equations include different regressors, Zellner's (1962) SURE method is more efficient than OLS for estimating a VAR.

collinear so that the t-tests on individual coefficients are not reliable guides for down-sizing the model. In practice, however, it is impossible to avoid the consideration of prior restrictions. Sample size constraints mean that there is always a limit to the number of variables and lags included. Even if the sample size is not a problem, the possibility of giving some structure to a VAR and using it for economic analysis alongside forecasting purposes, requires the imposition of restrictions. Well-founded theoretical constraints may help to transform an unrestricted VAR into a restricted model "*consistent with even highly detailed economic theories*" (Charemza and Deadman, 1997, p.157). The consideration of such restrictions allows for identification and economic interpretation of the structural parameters in a way not possible with the reduced-form. Furthermore, *a priori* information concerning the parameters allows for testing restrictions which can improve the precision of estimates and reduce the forecast error variance. Hence, even if forecasting is the main objective, the down-sizing of an over-parameterised VAR can help to improve results.

When empirical analysis on the existence of long-run relationships among more than two non-stationary series is conducted and there are doubts about the exogenous nature of regressors, an appropriate modelling strategy consists of treating all variables as endogenous within a reduced-form VAR framework. Then, using one or more of the methods available, tests for the exogeneity of the set of variables in doubt can be carried out. Once the endogenous-exogenous division is established, the Johansen approach can be used to test for the existence of cointegrated relationship(s). The number of cointegrated vectors found establishes the number of meaningful long-run relations in the system. The estimates of the long-run coefficients can be assessed by imposing restrictions to exactly-identify the underlying structural VAR. Once the structural form is identified, additional restrictions making the VAR compatible with specific economic theories, can also be tested.

We specify a general unrestricted VAR model of the UK tourism demand for France, Spain and Portugal which includes the variables in vector $z_t = [WF_t, WS_t, WP_t, PF_t, PS_t, PP_t, E_t]$. Where WF, WS and WP represent, respectively, the UK tourism expenditure shares of France, Spain and Portugal; PF, PS and PP represent tourism prices in France, Spain and Portugal and E represents the UK real per capita tourism budget. All variables in z_t, and data sources are described in Appendix A, according to the definitions and data sources used in De Mello *et. al*. (2001).

The specification of a VAR model starts with establishing the order of integration and the appropriate lag-length of its variables. Hence, we start by determining whether the time series in z_t are stationary and the appropriate lag-length of the VAR. All estimations and statistical tests implemented below were computed using Pesaran and Pesaran (1997) Microfit 4.0.

3.1. Order of integration of the variables included in the VAR

Table 1 shows the statistics and respective critical values at the 5% significance level of the Dickey-Fuller (1979, 1981) (DF) and Augmented Dickey-Fuller (ADF) unit root tests for the levels and first differences of variables WF, WS, WP, PP, PS, PF and E. It also shows the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) for the lag-length selection of the DF test equations.

****
Insert Table 1 here
****

The tests indicate that all variables in levels are non-stationary and, except for ΔPP, all variables in first differences are stationary. This means that according to the DF and ADF tests, all variables, except PP, can be considered I(1) variables. Doubts about the order of integration of variables PP and ΔPP can be cleared with the Phillips-Peron (1988) test. The Phillips-Peron test is based on a simple DF regression, such that

$$\Delta PP_t = \beta_0 + \beta_1 PP_{t-1} \tag{1}$$

The OLS estimation results for equation (1) are (t-ratios in brackets),

$$\Delta\hat{P}P_t = \underset{(0.536)}{0.0079} - \underset{(-1.4177)}{0.14139}\, PP_{t-1}$$

To carry out the non-parametric correction of the t statistic proposed by Phillips and Peron, we use the White and Newey-West covariance matrix and compute the adjusted variances. The estimates of regression (1) using the adjusted covariance matrix are

$$\Delta\hat{P}P_t = \underset{(0.5112)}{0.0079} - \underset{(-2.9039)}{0.14139}\, PP_{t-1}$$

The adjusted t ratio for β_1 coefficient is now the valid statistic to compare with the critical value. As $|-2.9039| < |-2.971|$ we cannot reject the hypothesis of PP being non-stationary. This confirms the result obtained with the ADF test in Table 1 for the same variable. However, the problem resides in that the ADF test indicates ΔPP to be non-stationary as well. So, we perform the Phillips-Peron test for ΔPP, with regression:

$$\Delta\Delta PP_t = \beta'_0 + \beta'_1 \Delta PP_{t-1} \tag{2}$$

The OLS estimation results for regression (2) are

$$\Delta\Delta\hat{P}P_t = \underset{(-0.1211)}{-0.001692} - \underset{(-2.4803)}{0.51046}\,\Delta PP_{t-1}$$

The estimation results for (2) using the adjusted covariance matrix are

$$\Delta\Delta\hat{P}P_t = \underset{(-0.1576)}{-0.001692} - \underset{(-3.9587)}{0.51046}\,\Delta PP_{t-1}$$

As $|-3.9587| > |-2.975|$ we cannot reject the hypothesis of ΔPP being stationary. Hence, the variable in levels PP is I(1) and we can consider all variables in levels as I(1).

3.2 Determination of the order of the VAR

Given the number of variables and sample size, the lag-length (p) for this VAR model cannot exceed two. Given these limitations, we used the AIC and SBC criteria and the adjusted (for small samples) Likelihood Ratio (LR) test for selecting the order of the VAR with the maximum lag-length permitted. The LR test rejects order zero but cannot reject a first order VAR. The SBC criterion clearly indicates the order of the VAR to be one. The AIC criterion indicates p to be two, but by a very small margin. So, we select a order one VAR. Yet, it is always prudent to examine the residuals of the equations in order to check for serial correlation. Table 2 presents the LR statistic and the two selection criteria for choosing the lag-length (p).

****
Insert Table 2 here
****

3.3. The unrestricted VAR specification of the UK demand for tourism

The reduced form of a first order unrestricted VAR of the UK tourism demand for France, Spain and Portugal (denoted VAR I) can be written as: [2]

$$z^*_t = A_0 + A_1 z^*_{t-1} + \in_t \tag{3}$$

where $z^*_t = [WF_t , WS_t , PP_t , PS_t , PF_t , E_t]$

Since the share variables observations (WF_t, WS_t and WP_t) sum up to unity, one of the share equations is omitted. We omit the equation for the share of Portugal (WP_t). The estimation results are invariant whichever equation is excluded and, by the adding-up property, all coefficient estimates of the omitted equation can be retrieved from the coefficient estimates of the other two.

The statistical quality of VAR I model is accessed by estimating (3) and computing relevant diagnostic statistics. Table 3 shows the estimation results (t ratios in brackets), the AIC and SBC criteria and a set of statistics - adjusted R^2, F statistic and χ^2 statistic for diagnostic tests of serial correlation, functional form, error normality and heteroscedasticity (p values in brackets) - for all equations included in the basic unrestricted VAR I structure.

****
Insert Table 3 here
****

There is no statistical evidence of serial correlation in VAR I equations. Hence, the lag-length selected seems to be adequate. However, the tests indicate problems in the functional form and error normality for some equations. We are interested in the expenditure share equations and functional form problems for these, appear severe.

In an AIDS system, the tourism shares of France (WF), Spain (WS) and Portugal (WP) would be the only endogenous variables, and changes in these variables would be explained by a set of exogenous regressors which includes tourism prices (PF, PS and PP) and the UK real per capita tourism expenditure (E). Within a VAR specification, we question the assumed exogeneity of the price variables. However, there seems to be no obvious theoretical or empirical basis for challenging the multi-stage budgeting process underlying the rationality of an AIDS expenditure share system, which sets the variable 'UK real per capita expenditure' as an exogenous determinant of the demand shares in (3). The reasons are as follows.

A VAR model sets all its variables as endogenous. Hence, a bi-directional cause-effect relationship (feedback) between them should exist. In a tourism demand context, however, even if it is reasonable to consider that changes in the UK real per capita expenditure affect the tourism shares of important UK holiday destinations such as

[2] This basic specification of the unrestricted VAR I is later subjected to modifications due to the introduction of dummy variables and the exogeneity assumption concerning some regressors.

France, Spain or Portugal, it does not seem realistic to expect that changes in these shares affect the way in which UK consumers allocate their budgets. Empirical evidence seems to support this line of reasoning. As it can be inferred from the estimation results of the equation for E_t in Table 3, the 99% of this variable's variations explained by the model lie exclusively on its own lagged value. No other variable in that equation is individually or jointly statistically significant. In fact, the F test for the joint significance of all explanatory variables, excluding E_{t-1}, presents a value of 0.86 implying that the hypothesis of these variables' coefficients being zero cannot be rejected. Furthermore, the estimation results of the VAR equations for WF_t and WS_t indicate that the lagged value of E_t does not affect significantly the current values of these tourism shares.

To investigate further the features of the link between the expenditure shares (Wi_t) and the UK real per capita tourism budget (E_t), we analyse the relationships between the error term of the conditional model for Wi_t and the stochastic disturbance of the assumed data generating process (d.g.p.) of E_t.

Consider that the i^{th} share equation is

$$Wi_t = \alpha_0 + \alpha_1 E_t + \alpha_2 Wi_{t-1} + ui_t \qquad \text{where i = F, S, P} \qquad (4)$$

and assume that E_t is a stochastic variable which underlying d.g.p. is

$$E_t = \beta_1 E_{t-1} + \varepsilon_t \ ; \ \beta_1 < 1 \text{ and } \varepsilon_t \rightarrow N(0, \sigma^2) \qquad (5)$$

If ui_t and ε_t are not correlated we can say that $EV(ui_t, \varepsilon_s)=0$ for all t, s (where *EV* stands for expected value, not to be confused with variable E). Then, it is possible to treat E_t as if it was fixed, that is, E_t is independent of ui_t such that $EV(E_t, ui_t)=0$. Hence, we can treat E_t as exogenous in terms of (4), with the *current* value of the UK real per capita expenditure (E_t) being said to Granger-cause Wi_t. Equation (4) is a conditional model since Wi_t is conditional on E_t, with E_t being determined by the marginal model (5).[3]

"A *variable cannot be exogenous per se*" (Hendry, 1995, p.164). A variable can only be exogenous with respect to a set of parameters of interest. Hence, if E_t is deemed to be exogenous with respect to parameters α_j (j = 0, 1, 2) in (4), the marginal model (5) can be neglected and the conditional model (4) is complete and sufficient to sustain valid

[3] As noted in Harris (1995), if (5) is reformulated as $E_t = \beta_1 E_{t-1} + \beta_2 Wi_{t-1} + \varepsilon_t$, $EV(E_t, ui_t)=0$ still holds. However, as past values of Wi_t now determine E_t, E_t can only be considered weakly exogenous in (4). The current value E_t still causes Wi_t but not in the Granger sense, since lags of Wi_t now determine E_t.

inference. So, knowledge of the marginal model will not significantly improve the statistical or forecasting performance of the conditional model. Following this line of reason, we run regression (4) for the expenditure shares of France, Spain and Portugal and regression (5) as a representation of the d.g.p. of E_t. We retrieve the residual series of these four regressions, namely u_{Ft} , u_{St} , u_{Pt} standing for the residual series of (4) for, respectively, France, Spain and Portugal, and ε_t standing for the residuals of (5). We then run regressions of the current and lagged values (up to the fifth lag) of the residuals u_{it} (i = F, S, P) on the current and lagged (up to the fifth lag) values of ε_t. The estimation results of all 90 regressions indicate no significant relationship linking the current or lagged residuals of the conditional models to the current and lagged residuals of the marginal model. These results can be viewed as an indication that knowledge of the marginal model does not improve the statistical or forecasting performance of the conditional (on E_{t-1}) equations for WF_t, WS_t, PP_t, PS_t and PF_t in the VAR and so E_t may be treated as exogenous in the VAR given by (3).

The claim that feedback effects might be absent in the relationships between variables 'UK real per capita expenditure' (E_t) and expenditure shares (Wi_t, i = F, S), can be further investigated using the causality concept proposed by Granger (1969). A test for causality is related to whether the lags of a variable are statistically significant in the equation of another variable. Using Griffiths *et al.*'s (1992, p. 695), definition "*a variable $y1_t$ is said to be Granger-caused by $y2_t$, if current and past information on $y2_t$ help improve forecasts in $y1_t$*".[4] Hence, if the lagged values of $y2_t$ do not improve the forecasts of $y1_t$, i.e., if the lagged values of $y2_t$ are statistically insignificant in the reduced form equation for $y1_t$, then $y2_t$ does not Granger-cause $y1_t$.

A multivariate generalisation of the Granger-causality test can be used to establish if one or more variables in a VAR should or should not integrate the group of endogenous variables. The use of the LR statistic for testing the null that the coefficients of a subset of variables in a VAR are zero, is called 'block Granger *non*-causality test'. This test provides a statistical measure of the extent to which lagged values of a set of variables (say E_t), are important in predicting another set of variables (say Wi_t), once lagged values of the latter (Wi_{t-1}) are included in the model. The LR statistic value for block Granger non-causality of variable E_t, testing the null that the coefficients of E_{t-1} are zero

[4] Granger's causality concept does not imply a cause-effect relation. Indeed, "c*ausality has a meaning more on the lines of 'to predict' rather than 'to produce*" (Charemza and Deadman, 1997, p. 165).

in the block equations for WF_t, WS_t, PP_t, PS_t and PF_t, is $\chi^2(5)=10.578$. As the 5% critical value is 11.071, the null cannot be rejected.

Given the above, we exclude E_t from the set of endogenous variables and re-formulate the VAR under this assumption. This new version of the model is denoted VAR II. Table 4 shows the model's estimates and the selection criteria and diagnostic statistics previously used to access its quality. The selection criteria and diagnostic tests indicate VAR II to be statistically more robust than VAR I. However, problems with the functional form of the share equations still remain.

****

Insert Table 4 here

****

There is cause to believe that events in the 1970s (change of political regimes in Portugal and Spain and the oil crises) may have affected the time path of variables included in the VAR. Additionally, the integration process of Spain and Portugal in the European Union (EU) in 1986 is likely to have affected the UK tourism demand for its southern neighbours. To account for the 1970s events we add a dummy variable D1 taking the value of unity in the period 1975-1981 and zero otherwise. To account for the integration of the two Iberian countries in the EU, we use dummy variables D2 (taking the unity value in the period 1982-1997 and zero otherwise) and D3 (taking the unity value in the period 1989-1997 and zero otherwise) to split the integration process into two sub-periods: the integration period (1982-1988) and the pos-integration period (1989-1997). These dummies are assumed to be exogenous. This third version of the VAR is denoted VAR III. Table 5 shows the AIC and SBC criteria and the same set of diagnostic statistics used before to access the quality of the VAR III.

****

Insert Table 5 here

****

The results in Table 5 indicate that the statistical quality of VAR III over-performs those of VAR I and VAR II, particularly with respect to the expenditure share equations. Additionally, the LR test for block Granger non-causality of variable E_t performed on VAR III specification, confirms the results of the similar test performed on VAR I. The LR test presents now the statistic value of $\chi^2(5)=8.376$, which is well below the 5% critical value (11.071), further supporting the null that the coefficients of E_{t-1} are statistically insignificant in the block equations of the VAR. The tests results concerning the significance of the intercept, the joint significance of the dummy variables, and the Granger block non-causality of variable E_{t-1} are shown in Table 6. The first column presents the null hypothesis for each test and shows the variables entering the VAR under the "unrestricted" hypothesis (U), and the "restricted" null hypothesis (R) (for simplicity, the time subscripts are omitted). In each case, the set of endogenous variables is separated from the set of deterministic components and exogenous variables by the symbol '&'. The second column presents the maximum value of the likelihood function (ML) for the unrestricted and restricted alternatives. In the third column the LR statistic is computed. Under the null, LR is asymptotically distributed as χ^2 with degrees of freedom (i) equal to the number of restrictions. The null is rejected if LR statistic is larger than the relevant critical value.

****
Insert Table 6 here
****

Although VAR III is statistically more robust than VAR I, the latter is a more general model, not being restricted in any way, while the former is a specific model, being a partial system conditioned on exogenous variables. Hence, it is interesting to use both models for comparison purposes in accessing the features of the UK tourism demand for France, Spain and Portugal, within a reduced-form VAR specification. However, a reduced-form VAR may not be the ideal means to conduct economic analysis. First, because an *a-theoretical* model is unlike to produce estimation results interpretable within the limits of economic theory. Second, because the economic interpretation of the structural parameters is only possible if the underlying structural model is identified from the reduced-form. Finally, because the lag-structure of the reduced-form is likely to be over-parameterised causing imprecision in the coefficients' estimates and obscuring the economic meaning of the long-run parameters. However, for forecasting purposes, a simple reduced-form VAR has its advantages. Hence, we use VAR I and VAR III reduced-forms for forecasting rather than for economic analysis.

3.4. Forecasting ability of the reduced-form VAR I and VAR III models

To obtain forecasts for the UK tourism shares of France, Spain and Portugal from VAR I and VAR III, we estimate these models for the period 1969-1993, using the last four observations for forecasting. Tables 7, 8 and 9 report the actual and forecasted values and the forecast errors for, respectively, the shares of France, Spain and Portugal, and a number of summary statistics (mean absolute prediction errors – MAE; mean squared prediction errors – MSE; root mean squared prediction errors – RMSE) to evaluate the models' forecasting accuracy. The forecast errors and all quality criteria indicate VAR I to be a better forecaster than VAR III.

****
Insert Tables 7, 8, and 9 here
****

The fact that the variables included in the VAR models are I(1) implies that estimation, statistical tests and forecasting procedures are strictly valid if cointegrating relationship(s) exist linking the variables. Therefore, the next step for proceeding the empirical analysis is to establish whether cointegrated vectors exist within the VARs. To do so we use the Johansen's cointegration rank test.

Besides establishing the procedure for cointegration rank test, the Johansen approach provides a general framework for identification, estimation and hypothesis testing in cointegrated systems (see Johansen, 1988, 1991, 1995 and 1996 and Johansen and Juselius, 1990 and 1992). Hence, given a cointegrated system, this approach provides a method for identifying the structural relationships underlying the unrestricted reduced-form. However, the Johansen's 'empirical process' to exactly-identify the long-run coefficients may not always be adequate, particularly in economic contexts where theory provides strong, sensible and testable restrictions (Pesaran and Shin, 1998). In these cases, the cointegrated vectors must be subject to exact- and over-identifying restrictions suggested by theory and other relevant *a priori* information, rather than being subject to some normalisation process that does not consider the theoretical and empirical framework within which the phenomenon evolves. As Harris (1995, p.117) points out "*what is becoming increasingly obvious is the need to ensure that prior information motivated by economic arguments forms the basis for imposing restrictions, and not the other way around*". This is particularly important in a tourism context involving an AIDS system of equations regressing (m-1) tourism shares on m tourism prices and an origin's per capita tourism budget. In this case, the number of long-run relationships predicted by theory is (m−1), i.e., the number of equations in the system.

Hence, if theory is right, the cointegration tests involving a VAR of the UK tourism demand with variables WF_t, WS_t, PP_t, PS_t, PF_t and E_t should indicate that the model's relevant long-run relationships are those established by the share equations.

Furthermore, the identification process of the structural equations should confirm that their steady-state form is that of an AIDS model's equations. Consequently, in either VAR I and VAR III, we expect to find exactly two cointegrated vectors and to identify the structural parameters with restrictions that match those of the normalisation process used for identifying the share equations of an AIDS model. If this is the case, then we can subject the long-run relationships to further hypothesis testing, such as homogeneity and symmetry, and contribute an empirical basis for the validation of the principals of consumer theory within an AIDS system framework using a VAR specification.

4. JOHANSEN'S REDUCED RANK TEST

Let z_t be a vector of n potentially endogenous variables. Then, it is possible to model z_t as an unrestricted VAR involving up to p-lags such that

$$z_t = A_1 z_{t-1} + \ldots + A_p z_{t-p} + u_t \qquad u_t \to IN(0, \Sigma) \tag{6}$$

where z_t is a (n x 1) vector and each of the A_i is a (n x n) matrix of parameters. The VAR model in (6) can be reformulated as a vector error-correction (VEC) such that:

$$\Delta z_t = \Gamma_1 \Delta z_{t-1} + \ldots + \Gamma_{p-1} \Delta z_{t-p+1} + \Pi z_{t-p} + u_t \tag{7}$$

where $\Gamma_i = -(I - A_1 - \ldots - A_i),\ i = 1, \ldots, p-1$ and $\Pi = -(I - A_1 - \ldots - A_p) = \alpha\beta'$

α represents the speed of adjustment matrix and β is the matrix of long-run coefficients such that, $\beta' z_{t-p}$ implicit in (7) represents up to (n−1) cointegration vectors in the multivariate VAR. These cointegrating vectors define the relationships which ensure that the variables in z_t converge to their long-run solutions. If the variables in z_t are I(1) then, $\prod z_{t-p}$ in (7) must be stationary for $u_t \sim I(0)$ to be a white noise. This happens when $\prod = \alpha\beta'$ has reduced rank, that is, when there are $r \leq (n-1)$ cointegrating vectors in β alongside (n−r) nonstationary vectors. Hence, testing for cointegration amounts to determining the number of r linearly independent columns in $\prod$.

The hypothesis that there are at most r cointegrating vectors (n−r nonstationary relationships) can be tested with the eigenvalue trace statistic (λ_{trace}) which null is r=q (q=0, 1,…, n−1) against the alternative r≥q+1, and/or the maximum eigenvalue statistic (λ_{max}) which null is that there are r=q cointegrating vectors against the alternative r=q+1 exist. The results of these tests are given in Table 10. The first column shows the eigenvalues associated with each of the I(1) variables, ordered from highest to lowest, necessary to compute λ_{max} and λ_{trace}. The second column shows the various hypothesis to be tested, starting with no cointegration (r=0 or n−r=6 in the case of VAR I, and r=0 or n−r=5 in the case of VAR III) and followed by increasing numbers of cointegrated vectors. The following columns show the estimated λ_{max} and λ_{trace} and respective 5% and 10% critical values. The last column presents the SBC model selection criterion.

****
Insert Table 10 here
****

For VAR I, both the λ_{max} and λ_{trace} statistics give evidence, at the 5% level, of two cointegrated vectors corresponding to the higher eigenvalues attached to the share equations. Both statistics associated with the null of r=0 and r=1 reject these hypothesis (statistic value>critical value) but cannot reject r=2 (statistic value<critical value). The SBC criterion further supports these findings by selecting the model with two cointegrated vectors. For VAR III, at the 5% level, λ_{max} statistic suggests the existence of two cointegrating vectors while λ_{trace} statistic does not reject the hypothesis of only one vector. This disagreement is not uncommon, particularly in cases of small samples and added dummy variables. However, we have enough evidence supporting the choice of r=2: we have λ_{max} statistic clearly rejecting the existence of only one in favour of two cointegrated vectors; we have SBC criterion selecting VAR III model with two cointegrating vectors; we have theory suggesting the existence of two and not one long-run relationships which is unmistakably supported by both test statistics and selection criterion in the more general VAR I model. Hence, given that evidence against theory prediction seems weak, we proceed by setting r=2 for both VAR I and VAR III.

To identify the structural form of the two cointegrated vectors we follow Pesaran and Shin (1998) and use the theoretical exact-identifying restrictions implicit in the share equations' of an AIDS model. Consider the following notation for matrix β of the two cointegrating vectors in VAR I with variables WF_t WS_t PP_t PS_t PF_t E_t and intercept,

$$\beta'_I = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} & \beta_{61} & \beta_{71} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} & \beta_{52} & \beta_{62} & \beta_{72} \end{bmatrix}$$

and for matrix β of the two cointegrating vectors in VAR III with variables WF_t, WS_t, PP_t, PS_t, PF_t, E_t, D1, D2, D3 and intercept

$$\beta'_{III} = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} & \beta_{51} & \beta_{61} & \beta_{71} & \beta_{81} & \beta_{91} & \beta_{101} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} & \beta_{52} & \beta_{62} & \beta_{72} & \beta_{82} & \beta_{92} & \beta_{102} \end{bmatrix}$$

The theoretical restrictions which exactly-identify the cointegrated vectors as the share equations of an AIDS specification in both VAR I and VAR III models are given by:

$$H_{AIDS}: \begin{Bmatrix} \beta_{11} = -1 & \beta_{12} = 0 \\ \beta_{21} = 0 & \beta_{22} = -1 \end{Bmatrix}$$

The coefficients estimates of the two exactly-identified cointegrated vectors in VAR I and VAR III models are presented in Table 11 (asymptotic t ratios in brackets). The 'third vector' corresponds to the share equation for Portugal, and was retrieved from the estimates of the other two applying the adding-up property.

****
Insert Table 11 here
****

There is a sharp difference between the estimates of VAR I and VAR III models, both in magnitude, statistical relevance and in their expected signs. For instance, VAR I estimates indicate all coefficients as statistically irrelevant at the 5% level in the share equation for Portugal, and in the share equations for France and Spain only the price of Spain and the intercept as significant at the 5% level. In the equation for Portugal, the own-price and intercept estimates have 'wrong' signs and implausible magnitudes and in the equation for France, implausible magnitude is also the case for the intercept estimate. In contrast, the estimates of the cointegrated VAR III model are overall statistically relevant, present the expected signs and magnitudes and give plausible information about how the events represented by the dummy variables affected the UK tourism demand for the destinations. Indeed, the coefficients of D1 indicate that the oil crises and political changes in Portugal and Spain, affected significantly and negatively UK tourists preferences for these destinations, favouring France instead in 1975-1981. The coefficients of D2 indicate that Spain and Portugal's integration process in the EU caused UK tourism flows to divert from France to the Iberian peninsula, although favouring more Spain than Portugal. The D3 coefficients, representing the post-integration period, indicate a recovery of the share for France at the expense of Spain's which steadily declines (the opening of the channel tunnel in 1994 may also have contribute for this result). In the same period, the share of Portugal shows an increase although not statistically significant.

We showed that the intercept is a relevant deterministic component of the VAR specifications. We considered that time trends should not be included as there is no statistical support for their presence. We found statistical support for not rejecting the possibility of treating E_t as exogenous, and for including variables D1, D2 and D3 as statistically relevant regressors. VAR III model integrates all these features in its specification while VAR I considers only the former two. Hence, the consistency and plausibility of results obtained with VAR III in contrast with those of VAR I, should be expected. Hence, we consider VAR I not to be an appropriate means to supply reliable information on the long-run demand behaviour of UK tourists and proceed the analysis with the cointegrated VAR III model.

In many cases, and particularly in demand systems, "*the focus of interest is on a set of hypothesis relating to the long-run structure of the model, which is quite independent of any short-run dynamics fitted to the empirical model*" (Chambers 1993, p.727). Likewise, in the case of a VAR system for the UK tourism demand, our interest is focused on the structural equations upon which parameter restrictions are to be imposed to test theoretical hypothesis on the long-run equilibrium relationships. Theory suggests that homogeneity and symmetry, reflecting the rationality of consumers behaviour, should hold in a system of demand equations. In addition, and according to De Mello *et al.*'s (2001) assumption about the competitive behaviour of neighbouring destinations, price changes in France (Portugal) should not affect UK tourism demand for Portugal (France), while price changes in Spain should affect significantly UK tourism demand for both France and Portugal. To test these hypothesis we use the LR statistic on the cointegrated VAR III model. Table 12 shows the tests results for the hypothesis of null cross-price effects between France and Portugal, homogeneity and symmetry and all these hypothesis simultaneously. The tests indicate that the set of hypothesis cannot be rejected. Hence, the cointegrated VAR III complies with the theoretical restrictions of homogeneity and symmetry and with the assertion of null cross-price effects between the share equations for France and Portugal.

****
Insert Table 12 here
****

The coefficients estimates for the two cointegrated vectors of VAR III under the full set of restrictions are given in Table 13 (asymptotic t ratios in brackets). The results show the cointegrated VAR III as a statistically robust, theoretically consistent and empirically plausible model, indicating it as an adequate basis for analysing the long-run behaviour of the UK tourism demand for France Spain and Portugal.

****
Insert Table 13 here
****

Even so, a more detailed investigation requires the analysis of the relevant elasticity values. We compute the expenditure, and uncompensated own- and cross-price elasticities, using the long-run estimates of the cointegrated VAR III under the full set of restrictions and the formulae given in Appendix A. Table 14 presents these and the corresponding estimates obtained with the AIDS model of De Mello *et al*. (2001).[5]

****
Insert Table 14 here
****

The elasticities estimates of the VAR and AIDS models are similar. The expenditure elasticities are close to unity for all destinations in both models and except in the VAR equation for Spain, the own-price elasticity estimates are close to –2. In the VAR equation for Spain this estimate is roughly half of that of its neighbours. The VAR cross-price elasticities estimates give the same indications as those of the AIDS model. For instance, insignificant cross-price effects between the equations for Portugal and France, indicating that the UK demand for Portugal (France) is not sensitive to price changes in France (Portugal), and significant cross-price effects between Spain and France and Spain and Portugal, indicating bilateral competitive behaviour. These cross-price estimates also imply that the UK demand for Portugal or France is more sensitive to price-changes in Spain than that for Spain is to price changes in its neighbours.

The results substantiate the importance of the VAR approach and the Johansen's rank test in finding statistical support for the existence of the two cointegrated vectors which endorse the share equations of the AIDS system as the only meaningful long-run relationships. In addition, as showed by Pesaran and Shin (1998), also in the case of tourism share equations the identification of the long-run structural model requires *a priori* information provided by economic theory and knowledge of pragmatic aspects concerning the relationships between the countries involved. Exogeneity of regressors, market conditions, policy regulations, institutional aspects, geographical attributes,

[5] We use De Mello *et al*.'s (2001) estimates for the 'second period' (1980-1997), as these correspond to more recent behaviour of the UK tourism demand.

political upheavals and economic instability, need to be appropriately modelled and tested within a restricted VAR approach to obtain theoretically consistent, empirically plausible and statistically reliable estimates for the structural parameters. As the cointegrated VAR III model fully complies with the theory predictions underlying the AIDS approach, the similarity of their estimates is not surprising. This similarity gives further support to the AIDS approach as a theoretical and empirical robust means for economic analysis of long-run tourism demand. However, the analysis is not complete without accessing the forecasting ability of the cointegrated VAR III.

5. Forecasting with a cointegrated VAR specification

For assessing the forecasting ability of cointegrated VAR III under the full set of restrictions (denoted hereafter by 'over-VAR') we estimate it for the period 1969-1993, leaving the last four observations for forecasting purposes. To compare the forecasting accuracy of this model with that of the unrestricted VAR I of section 2 (denoted hereafter 'pure-VAR'), and that of the De Mello *et al.*'s (2001) AIDS model, tables 15, 16 and 17 show the actual values, forecasts and forecast errors of, respectively, the tourism shares of France, Spain and Portugal.[6] The corresponding summary statistics MAE, MSE and RMSE are reported in table 18.

****
Insert Tables 15, 16, 17, and 18 here
****

For an overview of the three models' forecasting performance we display Figures 1, 2 and 3 showing a plot of the actual and forecasted values of the destinations' tourism shares obtained with the three models.

****
Insert Figures 1, 2, and 3 here
****

For the shares of France and Spain, the over-VAR is the best forecaster, followed by the AIDS model. For the share of Portugal, the pure-VAR is the best forecaster followed by the over-VAR. For this share, however, the differences between the two VAR models' forecasts are so small that they can be considered equivalently as excellent forecasters.

[6] The forecasts reported in De Mello *et al*. (2001) were computed for the last three observations (1995-1997). The VAR forecasts in this study are computed for the last four observations (1994-1997). Hence,

Hence, we find VAR III specification under the full set of theoretical restrictions as the best forecasting device. The forecasting performance of the AIDS model is also remarkable, particularly in the cases of the share equations for France and Spain.

Although the pure-VAR predictions are not as accurate as the other models' they are, even so, fairly precise. This brings out the unrestricted VAR forecasting ability alongside its remarkable qualities of form simplicity, estimation ease and assumption minimalism. Indeed, if the main purpose of a research is to provide fairly accurate forecasts of expenditure shares, the estimation of a unrestricted reduce-form VAR with the appropriate variables and lag-length, might be considered the preferable approach, since the quality differences do not seem to justify undergoing the complexities of cointegration testing and identification procedures of the structural model.

6. Conclusion

Economic models portraying long-run relationships of time series variables are a central aspect of theoretical and empirical research. Part of the role of applied work is to establish appropriate formal specifications which can be considered valid means to estimate the equilibrium path of relevant variables and to test competing theoretical hypothesis underlying those specifications. However, estimation of equation systems involving long-run relationships of nonstationary data, dubious assumptions on the exogeneity of regressors and the inclusion of zero restrictions with poor economic basis, may lead to invalid inference and forecasting procedures. Yet, inference based on such systems can be validated if one or more cointegrated relationships exist and the exogeneity assumptions and zero restrictions have theoretical and statistical support. Hence, tests must be performed to determine whether a system is cointegrated and its underlying assumptions are statistically valid and, if so, proceed with suitable estimation techniques which can provide valid estimates for its structural parameters.

Besides establishing a test procedure for determining the number of cointegrated vectors, the Johansen approach provides a general framework for identification, estimation and hypothesis testing of the structural model within a VAR specification. However, the Johansen's identification process may not be appropriate in economic contexts where theory provides specific 'rules' to identify the structural parameters of

to compare the forecasting ability of the VAR and AIDS models on an equal basis, we re-estimate the AIDS model of De Mello *et al.* for the period 1969-1993, and obtain forecasts for the period 1994-1997.

the cointegrated vectors. The AIDS model of the UK tourism demand for France, Spain and Portugal, is a system of equations which includes nonstationary I(1) series and assumes exogeneity for all its right-hand side variables. The implications of these features can risk the validity of its estimation results, statistical inference and forecasts, if no cointegrating relationships exist among the variables and/or the regressors' exogeneity assumption does not hold. As an alternative to such an AIDS model, we specified a reduced-form unrestricted VAR system and used statistical tests to establish the lag-length, deterministic components and the endogenous/exogenous division of its variables. Once the appropriate form of the VAR was in place, we used the Johansen rank test to determine the number of cointegrated vectors. Theory underlying a system of equations regressing two destinations' tourism shares on tourism prices and an origin's tourism budget, predicts the existence of exactly two long-run (cointegrated) relationships. Therefore, in a VAR system with the same variables, two cointegrated vectors should be accounted for. The cointegrated rank test provided statistical evidence to support the theory predictions. The theoretical framework of an AIDS system of the UK tourism demand also establishes the structural form of the long-run relationships it predicts. Hence, the structural parameters of the cointegrated vectors in the VAR should be exactly-identified with restrictions matching those of the normalisation process used to identify the tourism share equations of an AIDS system. This was done and the resulting structural form of the cointegrated VAR was then subject to additional restrictions such as homogeneity, symmetry and null cross-price effects between the share equations for Portugal and France. These restrictions could not be rejected. Consequently, evidence was obtained on the capability of the cointegrated VAR to comply with theoretical predictions underlying the rationality of the UK tourism demand behaviour and the destinations' competitive conduct. The estimates of the structural coefficients of the cointegrated VAR under the full set of restrictions were then used to compute the expenditure, own- and cross-price elasticities estimates of the UK tourism demand. These estimates, similar to the corresponding ones obtained with the AIDS model of De Mello *et al.* (2001), proved to be statistically relevant and empirically plausible.

Given the theoretical consistency and statistical robustness of the cointegrated VAR model under the full set of restrictions, its excellent predictive ability was expected. Indeed, the quality criteria measuring the forecasts' accuracy of this model's indicate it as the best forecasting model, when compared with the reduce-form "pure" VAR and the AIDS models. Nevertheless, these same quality criteria suggest the reduced-form (unrestricted) VAR to be a fairly accurate forecasting device. This finding gives support to the claimed competence of the VAR approach for forecasting purposes, and emphasizes the valuable qualities of modelling simplicity and estimation ease of a "pure" unrestricted VAR for forecasting purposes.

REFERENCES

BUSINESS MONITOR MA6 (1970-1993) *Overseas Travel and Tourism*, Government Statistical Service, Central Statistical Office, HMSO, London.

CHAMBERS, M. J. (1993), Consumers' demand in the long-run: some evidence from UK data, *Applied Economics*, **25**, 727-733

CHAREMZA, W. W. and DEADMAN D. F. (1997), *New directions in econometric practice*, 2nd edition, Edward Elgar Publishing Limited, Cheltenham.

DE MELLO, M., SINCLAIR, M. T. and PACK, A. (2001) A system of equations model of UK tourism demand in neighbouring countries, *Applied Economics*, forthcoming.

DEATON, A. and MUELLBAUER, J. (1980a) *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.

DEATON, A. and MUELLBAUER, J. (1980b) An Almost Ideal Demand System*, The American Economic Review*, **70**, 312-26.

DICKEY, D. A. and W. A. FULLER (1979) Distributions of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427-31.

DICKEY, D. A. and W. A. FULLER (1981) Likelihood ratio statistics for autoregressive time series with a unit root *,* *Econometrica*, **49**, 1057-72

GRANGER, C. W. J. (1969) Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, **37**, 24-36.

GRIFFITHS, W. E., HILL R. C. and JUDGE G. G. (1993) *Learning and practicing econometrics*, John Wiley & Sons, Inc., New York.

HARRIS, R. I. D. (1995) *Using cointegration analysis in econometric modelling*, Prentice Hall, Hemel Hempstead.

HARVEY, A. C. (1990) *The econometric analysis of time series*, 2nd edition, Philip Allan, Hemel Hempstead.

HENDRY, D. F. (1995) *Dynamic Econometrics*, Oxford University Press, Oxford.

INTERNATIONAL MONETARY FUND (1984-1998) *International Financial Statistics Yearbook*, International Monetary Fund, Washington, D.C.

JOHANSEN, S. (1988) A statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, **12**, 231-54.

JOHANSEN, S. (1991) Estimation and hypotheses testing of cointegrating vectors in Gaussian vector autoregressive models, *Econometrica*, **59**, 1551-80

JOHANSEN, S (1995) Identifying restrictions of linear equations with applications to simultaneous equations and cointegration, *Journal of Econometrics*, **69**, 111-33.

JOHANSEN, S (1996) Likelihood-based inference in cointegration vector autoregressive models, 2nd edition, Oxford University Press, Oxford.

JOHANSEN, S. and JUSELIUS, K. (1990) Maximum likelihood estimation and inference on cointegration with application to the demand for money. *Oxford Bulletin of Economics and Statistics*, **52**, 169-210.

JOHANSEN, S. and JUSELIUS, K. (1992) Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for the UK, *Journal of Econometrics*, **53**, 211-44.

PAPATHEODOROU, A. (1999) The demand for international tourism in the Mediterranean region, *Applied Economics*, **31**, 619-30.

PESARAN, M. H. (1997) The role of economic theory in modelling the long-run. *Economic Journal*, **107**, 178-191.

PESARAN, M. H. and PESARAN, B. (1997) *Working with microfit 4.0: interactive econometric analysis*, Oxford University Press, New York.

PESARAN, M. H. and SHIN, Y. (1998) Long-run structural modelling, *DAE Working Papers Amalgamated Series* no.9812, Department of Applied Economics, University of Cambridge. Revised last in April 2001.

PESARAN, M. H., SHIN, Y. and SMITH, R. J. (2000) Structural analysis of vector error correction models with exogenous I(1) variables, *Journal of Econometrics*, 97, 293-343.

PHILLIPS, P. C. B. and PERRON, P. (1988) Testing for a unit root in time series regression, *Biometrika*, **75**, 335-46.

SIMS, C. (1980) Macroeconomics and reality. *Econometrica*, **48**, 1-48

Travel Trends: A report on the 1994 (1995/6/7/8) international passenger survey, *Government Statistical Service*, Central Statistical Office, HMSO, London.

ZELLNER, A. (1962) An efficient method of estimating seemingly unrelated regressions and test for aggregation bias, *Journal of the American Statistical Association*, 57, 348-68.

Table 1: Unit root DF and ADF tests for the variables WF, WS, WP, PF, PS, PP and E

Variable	Test	Statistic	AIC criterion	SBC criterion	Critical value
WF	ADF(0)	**-2.100**	54.21*	52.87*	-2.971
	ADF(1)	-2.008	50.81	48.87	-2.975
ΔWF	ADF(0)	**-5.163**	49.71*	48.42*	-2.975
	ADF(1)	-3.284	46.36	44.47	-2.980
WS	ADF(0)	**-1.965**	52.09*	50.76*	-2.971
	ADF(1)	-1.756	48.71	46.77	-2.975
ΔWS	ADF(0)	**-5.187**	48.08*	46.72*	-2.975
	ADF(1)	-3.510	44.78	42.89	-2.980
WP	ADF(0)	**-1.738**	86.03*	84.70*	-2.971
	ADF(1)	-1.458	81.64	79.69	-2.975
ΔWP	ADF(0)	**-5.429**	81.49*	80.20*	-2.975
	ADF(1)	-4.827	78.25	76.36	-2.980
PF	ADF(0)	**-1.869**	32.98*	31.64*	-2.971
	ADF(1)	-2.311	32.24	30.30	-2.975
ΔPF	ADF(0)	**-3.435**	30.59*	29.23*	-2.975
	ADF(1)	-3.417	28.69	26.80	-2.980
PS	ADF(0)	**-2.083**	33.50*	32.17*	-2.971
	ADF(1)	-2.428	31.74	29.80	-2.975
ΔPS	ADF(0)	**-3.788**	29.78*	28.48*	-2.975
	ADF(1)	-2.968	27.24	25.35	-2.980
PP	ADF(0)	-1.418	31.74	30.41	-2.971
	ADF(1)	**-2.605**	34.17*	32.23*	-2.975
ΔPP	ADF(0)	**-2.480**	31.81*	30.51*	-2.975
	ADF(1)	-2.245	29.16	27.27	-2.980
E	ADF(0)	**-2.362**	20.51*	19.18*	-2.971
	ADF(1)	-1.624	18.48	16.53	-2.975
ΔE	ADF(0)	**-3.696**	18.07*	16.77*	-2.975
	ADF(1)	-2.675	16.94	15.06	-2.980

Table 2: AIC and SBC criteria and adjusted LR test for selecting the order of the VAR

Order (p)	AIC	SBC	Adjusted LR test
2	296.91	246.37	-------
1	296.59	269.37	$\chi^2(36) = 37.67$ (0.393)
0	185.90	182.01	$\chi^2(72) = 366.02$ (0.000)

Table 3. Estimation results and statistical performance of the unrestricted VAR I model

REGRESSORS	EQUATIONS					
	WF_t	WS_t	PP_t	PS_t	PF_t	E_t
WF_{t-1}	0.9518 (2.10)	-0.5617 (-1.14)	-2.0208 (-1.57)	-1.0503 (-0.75)	-1.3501 (-1.32)	-0.3106 (-0.15)
WS_{t-1}	0.6820 (1.44)	-0.2932 (-0.57)	-1.9954 (-1.48)	-0.7034 (-0.48)	-0.9772 (-0.78)	-1.3497 (-0.61)
PP_{t-1}	-0.0712 (-1.14)	0.1062 (1.56)	1.2525 (7.05)	0.2108 (1.10)	0.5033 (3.06)	-0.4728 (-1.62)
PS_{t-1}	0.4620 (4.85)	-0.4622 (-4.45)	-0.1341 (-0.49)	0.7453 (2.54)	-0.0014 (-0.01)	0.0602 (0.14)
PF_{t-1}	-0.3195 (-2.38)	0.3157 (2.16)	-0.4922 (-1.29)	-0.2276 (-0.55)	0.0598 (0.17)	0.6837 (1.09)
E_{t-1}	-0.0071 (-0.79)	0.0000 (0.00)	-0.0312 (-1.23)	-0.0027 (-0.10)	-0.0318 (-1.35)	0.9315 (22.40)
Intercept	-0.2630 (-0.60)	0.8449 (1.77)	1.9533 (1.57)	0.7627 (0.57)	1.1671 (1.01)	1.1990 (0.59)
SLECTION CRITERIA AND DIAGNOSTIC STATISTICS						
AIC	61.24	58.82	31.99	29.77	34.10	18.12
SBC	56.58	54.15	27.33	25.11	29.44	13.45
Adjusted R^2	0.789	0.824	0.771	0.562	0.612	0.991
F statistic	17.88	22.12	16.11	6.79	8.09	511.34
Serial Correlation	0.78(0.38)	2.48(0.12)	0.73(0.39)	0.19(0.67)	0.37(0.54)	0.58(0.45)
Functional Form	3.81(0.05)	8.66(0.00)	1.85(0.17)	7.39(0.01)	0.00(0.98)	0.18(0.67)
Normality	1.02(0.60)	1.56(0.46)	0.56(0.75)	4.82(0.09)	11.65(0.00)	1.66(0.44)
Heteroscedasticity	0.39(0.53)	1.19(0.28)	1.33(0.25)	0.28(0.60)	0.03(0.87)	0.21(0.65)

Table 4. AIC and SBC selection criteria and diagnostic tests for the VAR II model

SELECTION CRITERIA AND DIAGNOSTIC TESTS					
	EQUATIONS				
	WF_t	WS_t	PP_t	PS_t	PF_t
AIC	61.39	58.83	33.58	30.02	34.62
SBC	56.73	54.17	28.92	25.36	29.96
Adjusted R^2	0.792	0.825	0.795	0.570	0.656
F statistic	18.11	22.15	18.47	6.97	8.53
Serial Correlation	0.83(0.38)	2.16(0.14)	0.27(0.50)	0.46(0.67)	0.47(0.49)
Functional Form	3.28(0.07)	8.49(0.00)	0.70(0.40)	5.56(0.02)	0.02(0.90)
Normality	0.86(0.65)	1.55(0.46)	0.05(0.97)	3.04(0.22)	6.49(0.04)
Heteroscedasticity	0.42(0.52)	1.23(0.27)	0.84(0.36)	0.21(0.65)	0.00(0.95)

Table 5. AIC and SBC selection criteria and diagnostic tests for the VAR III model

SELECTION CRITERIA AND DIAGNOSTIC TESTS					
EQUATIONS					
	WF_t	WS_t	PP_t	PS_t	PF_t
AIC	69.79	65.69	32.27	27.28	33.69
SBC	63.13	59.03	25.61	20.62	27.02
Adjusted R^2	0.892	0.899	0.788	0.508	0.623
F statistic	25.87	27.64	12.16	4.10	5.96
Serial Correlation	0.35(0.55)	0.02(0.90)	0.04(0.85)	0.21(0.64)	0.10(0.66)
Functional Form	0.04(0.84)	0.21(0.64)	0.57(0.45)	7.24(0.01)	0.14(0.71)
Normality	1.28(0.53)	0.63(0.73)	0.08(0.96)	3.50(0.17)	11.64(0.03)
Heteroscedasticity	0.10(0.75)	0.01(0.92)	0.68(0.41)	0.26(0.61)	0.03(0.86)

Table 6. LR tests for restrictive hypothesis on the intercept and variables of the VAR

MODEL	ML	LR → $\chi^2(i)$	Critical value (5%)	Result
H_0: Non-significance of intercept (INT)				
U: WF WS PP PS PF E **&** INT	351.68	$\chi^2(6)=18.05$	12.59	Rejected
R: WF WS PP PS PF E **&**	342.05			
H_0: Non-significance of dummy variables D1 D2 D3				
U: WF WS PP PS PF E **&** D1 D2 D3 INT	385.40	$\chi^2(18)=67.43$	28.87	Rejected
R: WF WS PP PS PF E **&** INT	351.68			
H_0: Block non-causality of E without dummy variables				
U: WF WS PP PS PF E **&** INT	351.68	$\chi^2(5)=10.58$	11.07	Not rejected
R: WF WS PP PS PF **&** E INT	346.39			
H_0: Block non-causality of E with dummy variables				
U: WF WS PP PS PF E **&** D1 D2 D3 INT	385.40	$\chi^2(5)=8.37$	11.07	Not rejected
R: WF WS PP PS PF **&** E D1 D2 D3 INT	381.21			

Table 7: Forecasting results for the UK expenditure share of France

FRANCE		1994	1995	1996	1997
Actual values		**0.39700**	**0.38540**	**0.38967**	**0.40481**
VAR I	Forecast	0.34422	0.36840	0.39676	0.42782
	Forecast error	0.052780	0.017006	-0.007086	-0.023007
VAR III	Forecast	0.34274	0.44818	0.41524	0.40121
	Forecast error	0.054261	-0.062778	-0.025567	0.003595

SUMMARY STATISTICS FOR RESIDUAL AND FORECAST ERRORS			
		Estimation period: (1970 -1993)	Forecast period: (1994 –1997)
VAR I	MAE	0.015321	0.024970
	MSE	0.000315	0.000914
	RMSE	0.017734	0.030226
VAR III	MAE	0.008510	0.036550
	MSE	0.000119	0.001888
	RMSE	0.010917	0.043451

Table 8: Forecasting results for the UK expenditure share of Spain

SPAIN		1994	1995	1996	1997
Actual values		**0.51857**	**0.52625**	**0.52292**	**0.50691**
VAR I	Forecast	0.57000	0.54972	0.52281	0.49058
	Forecast error	-0.051431	-0.023469	0.000108	0.016328
VAR III	Forecast	0.56593	0.45798	0.48770	0.50172
	Forecast error	-0.047361	0.068264	0.035218	0.005191

SUMMARY STATISTICS FOR RESIDUAL AND FORECAST ERRORS			
		Estimation period: (1970 -1993)	Forecast period: (1994 –1997)
VAR I	MAE	0.016475	0.022834
	MSE	0.000301	0.000866
	RMSE	0.019780	0.029422
VAR III	MAE	0.010340	0.039009
	MSE	0.000192	0.002043
	RMSE	0.013847	0.045195

Table 9: Forecasting results for the UK expenditure share of Portugal

PORTUGAL		1994	1995	1996	1997
Actual values		**0.08443**	**0.08835**	**0.08741**	**0.08828**
VAR I	Forecast	0.08578	0.08189	0.08043	0.08160
	Forecast error	-0.001349	0.006462	0.006977	0.006679
VAR III	Forecast	0.09133	0.09384	0.09706	0.09797
	Forecast error	-0.006900	-0.005488	-0.009651	-0.008786

SUMMARY STATISTICS FOR RESIDUAL AND FORECAST ERRORS			
		Estimation period: (1970 -1993)	Forecast period: (1994 –1997)
VAR I	MAE	0.008054	0.005367
	MSE	0.000088	0.000034
	RMSE	0.009365	0.005850
VAR III	MAE	0.006276	0.007706
	MSE	0.000060	0.000062
	RMSE	0.007770	0.007875

Table 10: Tests for the cointegration rank of VAR I and VAR III models

Eigen values	H_0		$\hat{\lambda}_{max}$	λ_{max} critical		$\hat{\lambda}_{trace}$	λ_{trace} critical		SBC
	r	n – r		5%	10%		5%	10%	
VAR I									
λ_1=0.9201	r = 0	n–r =6	**70.76**	40.53	37.65	**153.96**	102.56	97.87	274.71
λ_2=0.7489	r = 1	n–r =5	**38.69**	34.40	31.73	**83.20**	75.98	71.81	290.09
λ_3=0.5889	r = 2	n–r =4	24.89	28.27	25.80	44.50	53.48	49.95	**292.78**
λ_4=0.3273	r = 3	n–r =3	11.10	22.04	19.86	19.61	34.87	31.93	291.89
λ_5=0.1838	r = 4	n–r =2	5.69	15.87	13.81	8.51	20.18	17.88	287.45
λ_6=0.0958	r = 5	n–r =1	2.82	9.16	7.53	2.82	9.16	7.53	283.63
VAR III									
λ_1=0.8798	r = 0	n–r =5	**59.31**	46.77	43.80	**142.48**	119.77	114.38	265.82
λ_2=0.7734	r = 1	n–r =4	**41.57**	40.91	38.03	83.16	90.60	85.34	272.15
λ_3=0.5833	r = 2	n–r =3	24.51	34.51	31.73	41.59	63.10	59.23	**272.94**
λ_4=0.2767	r = 3	n–r =2	9.07	27.82	25.27	17.08	39.94	36.84	268.54
λ_5=0.2489	r = 4	n–r =1	8.01	20.63	18.24	8.01	20.63	18.24	259.74

Table 11: Long-run coefficients estimates of the exactly-identified share equations

Variables	Cointegrated VAR I			Cointegrated VAR III		
	Vector 1 (WF)	Vector 2 (WS)	'Vector' 3 (WP)	Vector 1 (WF)	Vector 2 (WS)	'Vector' 3 (WP)
WF	-1	0		-1	0	
WS	0	-1		0	-1	
PP	-0.4965 (-0.86)	0.1781 (0.60)	0.3184 (0.95)	-0.0027 (-0.05)	0.1090 (1.69)	-0.1062 (-3.28)
PS	0.8214 (2.09)	-0.5937 (-2.97)	-0.2277 (-0.97)	0.2256 (4.68)	-0.3075 (-5.22)	0.0820 (2.90)
PF	-0.2244 (-0.48)	0.3119 (1.20)	-0.0875 (-0.33)	-0.3394 (-3.59)	0.3044 (2.57)	0.0350 (0.56)
E	-0.0684 (-1.12)	0.0159 (0.51)	0.0525 (1.45)	0.0153 (2.33)	-0.0183 (2.29)	0.0030 (0.78)
D1				0.0380 (5.27)	-0.0154 (-1.74)	-0.0226 (-5.18)
D2				-0.0565 (-3.65)	0.0528 (2.78)	0.0037 (0.41)
D3				0.0574 (5.17)	-0.0590 (-4.34)	0.0016 (0.24)
INT	0.8309 (2.20)	0.3818 (1.97)	-0.2127 (-0.95)	0.3687 (16.26)	0.5443 (19.50)	0.0870 (6.33)

Table 12: Tests of over-identifying restrictions for cointegrated VAR III model

Hypothesis	$LR \rightarrow \chi^2(i)$	5% Critical value
Null cross-price effects	$\chi^2(2)=0.36$	5.99
Homogeneity and symmetry	$\chi^2(3)=7.28$	7.81
Homogeneity, symmetry and null cross-price effects	$\chi^2(4)=7.53$	9.49

Table 13 Long-run coefficients estimates of the cointegrated restricted VAR III

Variables	VAR III		
	Vector 1 (WF)	Vector 2 (WS)	'Vector 3' (WP)
WF	-1	0	
WS	0	-1	
PP	0	0.0895 (5.80)	-0.0895 (-5.80)
PS	0.2891 (4.79)	-0.3785 (-5.97)	0.0895 (5.80)
PF	-0.2891 (-4.79)	0.2891 (4.79)	0
E	0.0091 (1.16)	-0.0121 (-1.38)	0.0030 (0.95)
D1	0.0354 (3.83)	-0.0115 (-1.13)	-0.0239 (-7.05)
D2	-0.0373 (-2.36)	0.0360 (2.02)	0.0013 (0.20)
D3	0.0429 (4.56)	-0.4184 (-3.95)	-0.0011 (-0.27)
INT	0.3849 (13.17)	0.5230 (16.96)	0.092 (11.69)

Table 14. Expenditure and uncompensated own- and cross-price elasticities estimates

	Models	Expenditure elasticities	Own-price elasticities	Cross-price elasticities		
				PP	PS	PF
WP	Cointegrated VAR III	1.039	-2.158	X	1.137	-0.017
	AIDS (2nd period)	0.947	-1.797	X	0.830	0.019
WS	Cointegrated VAR III	0.979	-1.057	0.161	X	0.523
	AIDS (2nd period)	1.150	-1.933	0.124	X	0.658
WF	Cointegrated VAR III	1.026	-1.817	-0.002	0.793	X
	AIDS (2nd period)	0.808	-1.901	0.017	1.077	X

Table 15: Forecasting results for the UK expenditure share of France

Actual values		1994	1995	1996	1997
		0.39700	**0.38540**	**0.38967**	**0.40481**
Pure-VAR	Forecast	0.34422	0.36840	0.39676	0.42782
	Forecast error	0.052780	0.017006	-0.007086	-0.023007
Over-VAR	Forecast	0.39505	0.39260	0.39485	0.39475
	Forecast error	0.001942	-0.007194	-0.005178	0.010054
AIDS	Forecast	0.37775	0.38044	0.40188	0.38852
	Forecast error	0.019251	0.004959	-0.012202	0.016484

Table 16: Forecasting results for the UK expenditure share of Spain

Actual values		1994	1995	1996	1997
		0.51857	**0.52625**	**0.52292**	**0.50691**
Pure-VAR	Forecast	0.57000	0.54972	0.52281	0.49058
	Forecast error	-0.051431	-0.023469	0.000108	0.016328
Over-VAR	Forecast	0.51884	0.52718	0.52568	0.52543
	Forecast error	-0.000264	-0.000932	-0.002761	-0.018518
AIDS	Forecast	0.52613	0.52235	0.49786	0.51260
	Forecast error	-0.007558	0.003893	0.025054	-0.005696

Table 17: Forecasting results for the UK expenditure share of Portugal

Actual values		1994	1995	1996	1997
		0.08443	**0.08835**	**0.08741**	**0.08828**
Pure-VAR	Forecast	0.08578	0.08189	0.08043	0.08160
	Forecast error	-0.001349	0.006462	0.006977	0.006679
Over-VAR	Forecast	0.08611	0.08022	0.07947	0.07982
	Forecast error	-0.001678	0.008124	0.007939	0.008463
AIDS	Forecast	0.096126	0.097201	0.100263	0.099071
	Forecast error	-0.011693	-0.008853	-0.012852	-0.010788

Table 18: Summary statistics for forecast errors

	FRANCE			SPAIN			PORTUGAL		
	Pure-VAR	Over-VAR	AIDS	Pure-VAR	Over-VAR	AIDS	Pure-VAR	Over-VAR	AIDS
MAE	0.0250	0.0061	0.0132	0.0228	0.0056	0.0106	0.0054	0.0066	0.0110
MSE	0.0009	0.0000	0.0002	0.0009	0.0001	0.0002	0.0000	0.0001	0.0001
RMSE	0.0302	0.0068	0.0143	0.0294	0.0094	0.0135	0.0059	0.0071	0.0112

Figure 1: Actual values and forecasts for the tourism share of France

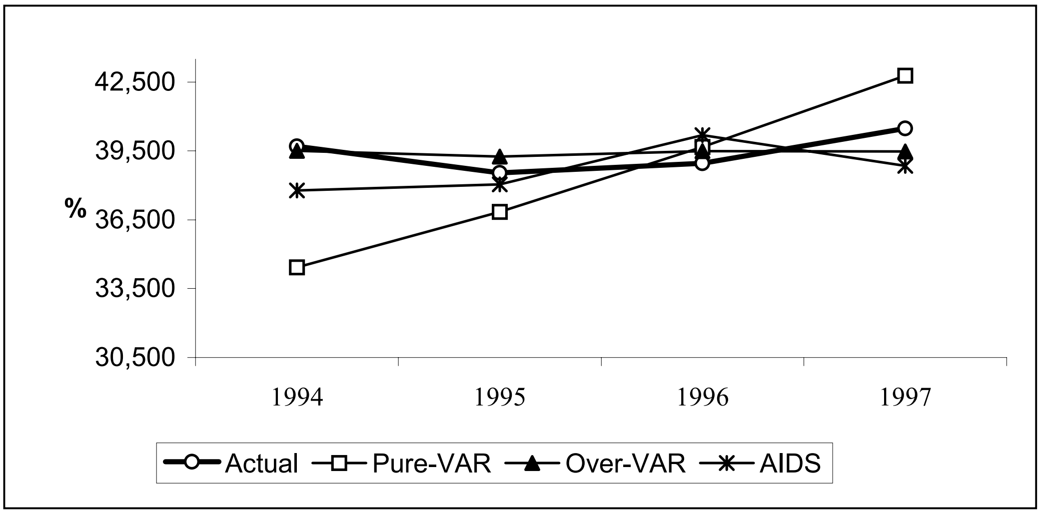


Figure 2: Actual values and forecasts for the tourism share of Spain

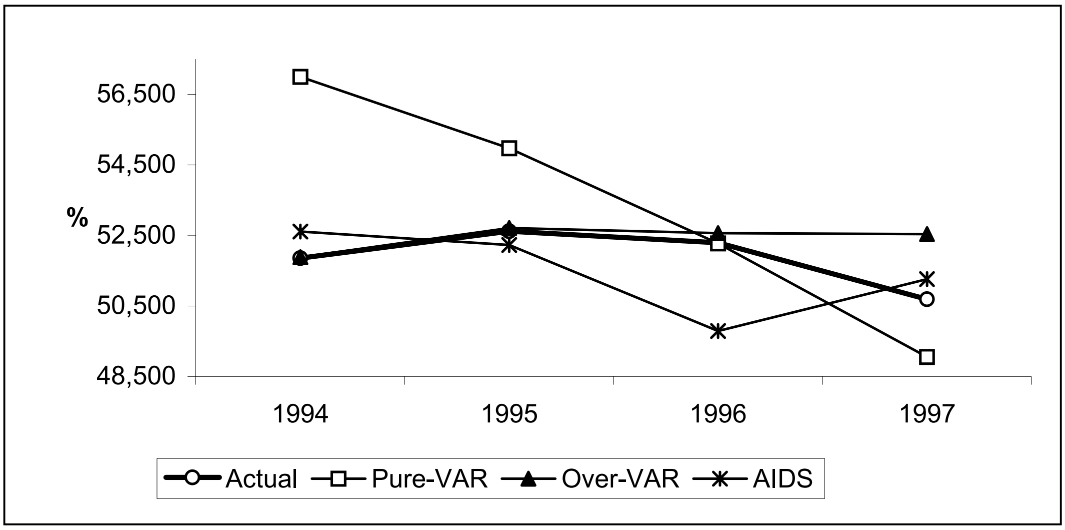
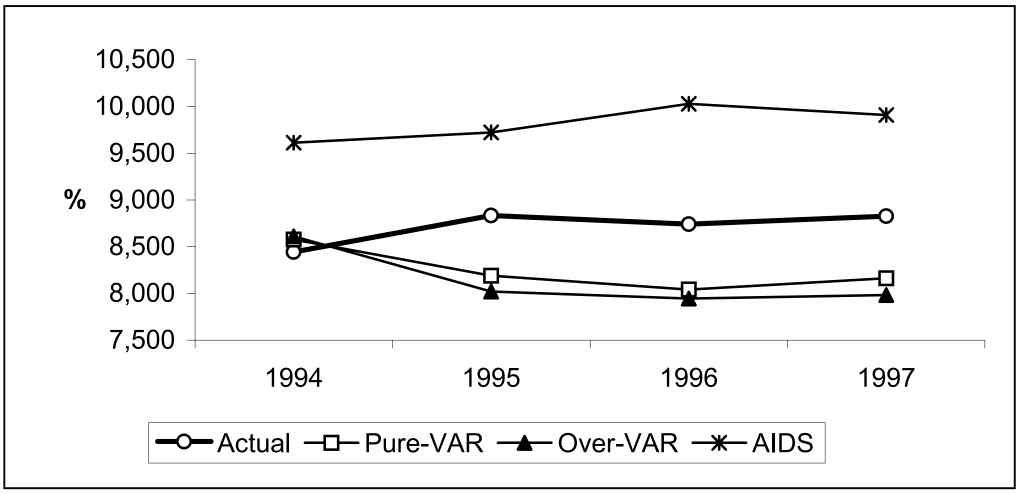


Figure 3: Actual values and forecasts for the tourism share of Portugal



APPENDIX A

A1. Derivation of the AIDS model

Let x be the exogenous budget or total expenditure which is to be spent within a given period on some or all of n products. These products can be bought in nonnegative quantities q_i at given prices p_i, i=1, …, n. Let $\mathbf{q} = (q_1, q_2, ..., q_n)$ be the quantities vector of the n products purchased, and $\mathbf{p} = (p_1, p_2, ..., p_n)$ the price vector. The budget constraint of the representative consumer is $\sum_{i=1}^{n} p_i q_i = x$ .Defining the utility function as $u(\mathbf{q})$, the consumer's aim is to maximise the utility subject to the budget constrain:

$$\max u(\mathbf{q}) \text{ subject to } \sum_{i=1}^{n} p_i q_i = x \tag{A1}$$

The solution for this maximisation problem leads to the Marshallian (uncompensated) demand functions $q_i = g_i(\mathbf{p}, x)$. Alternatively, the consumer's problem can be defined as the minimum total expenditure necessary to attain a specific level of utility u*, at given prices:

$$\min \sum_{i=1}^{n} p_i q_i = x \text{ subject to } u(\mathbf{q}) = u^* \tag{A2}$$

The solution for this minimisation problem leads to the Hicksian (compensated) demand functions $q_i = h_i(\mathbf{p}, u)$. Therefore, a cost function can be defined as

$$C(\mathbf{p}, u) = \sum_{i=1}^{n} p_i h_i(\mathbf{p}, u) = x \tag{A3}$$

Given the total expenditure x and prices **p**, the utility level u* is derived from the solution of the problem stated in equation (1). Solving (3) for u, an indirect utility function is obtained such that $u = v(\mathbf{p}, x)$.

The AIDS model specifies a cost function which is used to derive the demand functions for the commodities under analysis. The process of derivation can be summarised in the following three steps:

1st $\dfrac{\partial C(\mathbf{p}, u)}{\partial p_i} = h_i(\mathbf{p}, u)$ is derived establishing the Hicksian demand functions.

2nd solving (3) for u, the indirect utility function is obtained, such that $u = v(\mathbf{p}, x)$.

3rd $h_i[\mathbf{p}, v(\mathbf{p}, x)] = g_i(\mathbf{p}, x)$ is retrieved stating the Hicksian and the Marshallian demand functions as equivalent.

The Hicksian and Marshallian demand functions have the following properties:

1. Adding-up: $\sum_i p_i h_i(\mathbf{p}, u) = \sum_i p_i g_i(\mathbf{p}, x) = x$; all budget shares sum to unity;

2. Homogeneity: $h_i(\mathbf{p}, u) = h_i(\theta\mathbf{p}, u) = g_i(\mathbf{p}, x) = g_i(\theta\mathbf{p}, \theta x) \quad \forall\theta > 0$; a proportional change in all prices and expenditure has no effect on the quantities purchased;

3. Symmetry: $\frac{\partial h_i(\mathbf{p}, u)}{\partial p_j} = \frac{\partial h_j(\mathbf{p}, u)}{\partial p_i}, \quad \forall i \neq j$; consumer's choices are consistent;

4. Negativity: The (nxn) matrix of elements $\frac{\partial h_i(\mathbf{p}, u)}{\partial p_j}$ is negative semidefinite, that is, for any n vector ξ, the quadratic form $\sum_i \sum_j \xi_i \xi_j \frac{\partial h_i(\mathbf{p}, u)}{\partial p_j} \leq 0$. This means that a rise in prices results in a fall in demand as required when the commodities under analysis are normal goods.

The AIDS model specify the following cost function:

$$\ln C(\mathbf{p}, u) = a(\mathbf{p}) + u b(\mathbf{p}) \tag{A4}$$

where $a(\mathbf{p}) = \alpha_0 + \sum_i \alpha_i \ln p_i + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j$ and $b(\mathbf{p}) = \beta_0 \prod_i p_i^{\beta_i}$

The derivative of (4) with respect to ln p_i is:

$$\frac{\partial \ln C(\mathbf{p}, u)}{\partial \ln p_i} = \alpha_i + \sum_j \gamma_{ij} \ln p_j + u\beta_i \beta_0 \prod_i p_i^{\beta_i} \tag{A5}$$

As $C(\mathbf{p}, u) = x \Leftrightarrow \ln C(\mathbf{p}, u) = \ln x$, then

$$\ln x = a(\mathbf{p}) + u\, b(\mathbf{p}) \tag{A6}$$

Solving (6) for u we obtain

$$u = \frac{\ln x - a(\mathbf{p})}{b(\mathbf{p})} \tag{A7}$$

Substituting (7) in (5) we have

$$\frac{\partial \ln C(\bullet)}{\partial \ln p_i} = \frac{\partial C(\bullet)}{\partial p_i}\frac{p_i}{C(\bullet)} = h_i(\bullet)\frac{p_i}{C(\bullet)} = \frac{p_i q_i}{x} = w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i [\ln x - a(\mathbf{p})]$$

If we define a price index P such that ln P = a(**p**), then

$$\frac{\partial \ln C(\mathbf{p},u)}{\partial \ln p_i} = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i [\ln x - \ln P]$$

or $$w_i = \alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{x}{P}\right) \qquad \text{(A8)}$$

where $$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2}\sum_k \sum_\ell \gamma^*_{k\ell} \ln p_k \ln p_\ell \qquad \text{(A9)}$$

equations (8) and (9) are the basic equations of the AIDS model.

In a tourism analysis context, i is a destination country among a group of n alternative destinations demanded by tourists of a given origin. The dependent variable w_i, represents destination i share of the origin's tourism budget allocated to the set of n destinations. This share's variability is explained by tourism prices (p) in i and alternative destinations j and by the per capita expenditure (x) allocated to the set of destinations, deflated by price index P. The model has the following assumptions:

1. the adding-up restriction requiring that all budget shares sum up to unity:

   $\sum_i \alpha_i = 1, \quad \sum_i \beta_i = 0, \quad \sum_i \gamma_{ij} = 0$, for all j;

2. the homogeneity restriction requiring that a proportional change in all prices and expenditure has no effect on the quantities purchased: $\sum_j \gamma_{ij} = 0$, for all i;

3. the symmetry restriction requiring consistent consumers' choices: $\gamma_{ij} = \gamma_{ji}$, $\forall$i, j;

4. the negativity restriction requiring that a rise in prices result in a fall in demand, i.e., the condition of negative own-price elasticities for all destinations.

The restrictions imposed on α and γ comply with these assumptions and ensure that equation (9) defines P as a linear homogeneous function of individual prices. If prices are relatively collinear, then "P *will be approximately proportional to any appropriately defined price index, for example, the one used by Stone, the logarithm of which is*

$\sum w_k \ln p_k = \ln P^*$" (Deaton and Muellbauer, 1980a, p.76). Hence, the deflator P in equation (9) can be substituted by the Stone price index ln P* such that,

$$\ln P^* = \sum_i w_i^B \ln p_i \tag{A10}$$

where w_i^B is the budget share of destination i in the base year. With this simplification for P, equation (8) can be rewritten and estimated in the following form:

$$w_i = \alpha_i^* + \sum_j \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{x}{P^*}\right) \tag{A11}$$

A.2 Expenditure, own- and cross-price elasticities formulae

Equation (11) specifies a model in the linear-log form which prevents the direct interpretation of its coefficients as elasticities. The elasticities values are computed using the following formulae:

Expenditure elasticity: $\varepsilon_i = \frac{1}{\overline{w}_i}\frac{dw_i}{d\ln x} + 1 = \frac{\beta_i}{\overline{w}_i} + 1$

Uncompensated own-price elasticity: $\varepsilon_{ii} = \frac{1}{\overline{w}_i}\frac{dw_i}{d\ln p_i} - 1 = \frac{\gamma_{ii}}{\overline{w}_i} - \beta_i \frac{w_i^B}{\overline{w}_i} - 1$

Uncompensated cross-price elasticity: $\varepsilon_{ij} = \frac{1}{\overline{w}_i}\frac{dw_i}{d\ln p_j} = \frac{\gamma_{ij}}{\overline{w}_i} - \beta_i \frac{w_j^B}{\overline{w}_i}$

Compensated own-price elasticity: $\varepsilon^{\bullet}_{ii} = \varepsilon_{ii} + w_i^B \varepsilon_i = \frac{\gamma_{ii}}{\overline{w}_i} + w_i^B - 1$

Compensated cross-price elasticity: $\varepsilon^{\bullet}_{ij} = \varepsilon_{ij} + w_i^B \varepsilon_i = \frac{\gamma_{ii}}{\overline{w}_i} + w_j^B$

where $\overline{w}_i$ is the sample's average share of destination i (i=1,…,n) and w_j^B is the share of destination j (j=1,…,n) in the base year.

The model assumes that consumers allocate their budget to commodities in a multi-stage budgeting process implying independent preferences. Thus, it is assumed that the expenditure allocated by UK tourists to France, Spain and Portugal is separable from expenditure allocated to other destinations and that the decision to spend money in those countries is made in several stages. First, UK tourists allocate their budget to tourism and other goods; then to tourism in their southern neighbouring countries and other parts of the world; finally they decide between France, Spain and Portugal. The AIDS model is applied to this last stage of expenditure allocation.

A.3 Variables' definition

The variables integrating the VAR model are the destinations' shares of the UK tourism budget: WP, WS and WF; each destination tourism price: PP, PS, PF; and the UK real per capita tourism budget E.

Each destination share of the UK tourism budget allocated to the three countries is Wi, where i = F (France), S (Spain) and P (Portugal), and is defined as

$$Wi = \frac{EXP_i}{EXP_F + EXP_S + EXP_P}$$

where EXP_i is the nominal tourism expenditure allocated by UK tourists to destination i.

The effective prices of tourism in country i is defined as

$$Pi = \ln\left(\frac{CPI_i / CPI_{UK}}{R_i}\right)$$

where CPI_i is the consumer price index of destination i, CPI_{UK} is the consumer price index of the UK, R_i is the exchange rate between country i and the UK.

The UK real tourism expenditure allocated to all destinations per capita of the UK population is given by

$$E = \ln\left(\frac{\sum_i EXP_i / UKP}{P*}\right)$$

where UKP is the UK population and lnP* is the Stone index defined in equation (A10).

A.4 Data sources

The data for UK tourism expenditure, disaggregated by destinations and measured in £ million sterling, were obtained from one common source, the *Business Monitor MA6* (1970-1993), continued as *Travel Trends* (1994-1998). Data on the UK population, price indexes and exchange rates were obtained from the *International Financial Statistics* (IMF) *Yearbooks* (1984, 1990 and 1998).