THE ROLE OF COINTEGRATION IN THE FORECAST ACCURACY OF VAR MODELS

Maria M. De Mello
mmello@fep.up.pt

CEF.UP∗, Faculdade de Economia, Universidade do Porto
Rua Dr. Roberto Frias s/n, 4200 – 464, Porto, Portugal

August 2009

ABSTRACT

This paper assesses the forecast performance of a set of VAR models under a growing number of restrictions. With a maximum forecast horizon of 12 years, we show that the farther the horizon is, the more structured and restricted VAR models have to be to produce accurate forecasts. Indeed, unrestricted VAR models, not subjected to integration or cointegration, are poor forecasters for both short and long run horizons. Differenced VAR models, subject to integration, are reliable predictors for one-step horizons but ineffectual for multi-step horizons. Cointegrated VAR models including appropriate structural breaks and exogenous variables, as well as being subjected to over-identifying theory consistent restrictions, are excellent forecasters for both short and long run horizons. Hence, to obtain precise forecasts from VAR models, proper specification and cointegration are crucial for whatever horizons are at stake, while integration is relevant only for short run horizons.

Keywords: VAR demand systems; structural breaks, exogenous regressors, integration; cointegration; forecast accuracy.

JEL classification: C320 C530

1. INTRODUCING THE SUBJECT MATTER

Accurate modelling and forecasting of tourism demand are invaluable for the decision-making process of private and public entities regarding investment and planning in the tourism industry. Reliable forecasts also constitute a solid basis for the implementation of policies and business strategies in this area. Since forecast accuracy comparisons of alternative models allow for deciding which are best to supply that crucial information, assessing models' forecasting performance is an important prior task in advising course of action to business agents. Thus, the rapid growth of devoted researchers in this area, and the proliferation of tourism forecasting surveys are not surprising. Indeed, from Archer (1976) to Song and Li (2008), an average of one every couple of years has been published.

The history of economic forecasting with econometric models is regularly plagued by reports of their failure and under-performance when compared with univariate, no-change or other naïve prediction devices. Given that econometric models can incorporate dynamic causal information allegedly able to track the underlying data generating process, it would seem reasonable to expect smaller prediction errors from these specifications than from purely extrapolative devices. Yet, this has not been the case in a large amount of cases, many of which are reported in Mills (1999).[1]

Early explanations for the poor forecasting performance of econometric models are mostly anchored in formal misspecifications such as spurious relationships interpreted as meaningful, dynamic processes modelled as static, use of inappropriate functional forms, and unsuitable lag lengths. The poor predictive ability of such models is summarised by Clements and Hendry (1998, 1999) as the forecasting failure of misspecified models, where non-stationary processes and structural breaks are not accounted for.

In time, Granger (1981, 1986), Granger and Weiss (1983), Engle and Granger (1987), Johansen (1988, 1995 and 1996), Banerjee *et al.* (1993), Harris (1995) and others, established the basis for cointegration analysis, which led to the ascendancy of 'equilibrium-correction' models in econometric modelling. These systems, include "*almost all regression equations and simultaneous systems, … vector autoregressions (VARs), dynamic stochastic general-equilibrium models and many variance models such as GARCH*" (Hendry 2005, p. 400).

---

1 For related studies in tourism economics see, for example, Witt and Witt (1992), Kulendran and King (1997), Kulendran and Witt (2001) and du Preez and Witt (2003), stating that univariate models outperform econometric models; Kim and Song (1998), Song *et al.* (2000) and Song *et al.* (2003), stating that econometric models

Accordingly, more recent research on tourism demand using this broad class of vector 'equilibrium-correction' (VEC) models has been producing important advances in overcoming the above mentioned modelling faults and resulting forecast failure. For instance, Kulendran (1996), Song *et al.* (2000), and De Mello (2001) use autoregressive distributed lag equations, error-correction models and cointegration analysis; De Mello *et al.* (2002), De Mello and Fortuna (2005) and Li *et al.* (2006), use static and dynamic almost ideal demand systems; De Mello (2001), De Mello and Nell (2005), Song and Witt (2006) and Zhou *et al.* (2007) use reduced form and cointegrated VAR systems.

Although some studies (e.g. Clements and Hendry, 1995; Pesaran *et. al.*, 2000; De Mello and Nell, 2005) suggest that forecasts from cointegrated VEC models, should outperform simpler benchmark models, others (e.g. Clements and Hendry 1998, 1999; Makridakis and Hibon, 2000; Fildes and Ord, 2002) report less favourable results for the cointegrated forms.

Indeed, Hendry and Doornik (1997), Clements and Hendry (2003, 2005), Hendry and Mizon (2005) and Hendry (2000, 2004, 2005) find that parameter instability and structural breaks are among the key factors of forecast failure, and that the poor performance of VEC models is mainly due to their lack of robustness to 'location shifts'. In particular, Hendry (2005) states that VEC models are useless if structural breaks occur, regardless of their excellence with stationary generating processes after differencing and cointegration. Moreover, an explanation for why naive models (even being such poor translators of the in-sample generating process), may outperform VEC models is offered: naïve devices forecast better than VEC systems when the former are adaptive to precisely those structural breaks that undermine the latter.

To overcome these problems, one line of research consists in developing non-structural forecasting methods (e.g. over-differentiation; intercept shifts; forecast pooling) allegedly robust to parameter instability. Given that non-structural economic forecasting is mainly a-theoretical, this line of research is not restrained by any theoretical assumptions. As a result, it is more likely to progress faster, supported by the increasingly sophistication of computers and simulation techniques. In contrast, structural econometric forecasting is based on explicit theory and, therefore, "*it rises and falls with new theories, typically with a lag*" (Diebold and Rudebusch, 1999, p. 268). Its progress is generally slower and far more laborious.

---

outperform univariate models. For other studies on this subject see also Witt and Witt (1995), Song and Witt (2000), Witt *et all.* 2003) and Li *et al.* (2005).

Even so, we believe that theory-based econometric models should be favoured by forecasters because they present two inseparable properties which are of particular value for practical business purposes. When well specified, these models can provide not only a precise description of the data generating process but also, and subsequently, accurate predictions for the values of the variable of interest. Thus, the building of econometric models, even for the sole purpose of forecasting, must not disregard the fundamental dimension of an appropriate formal specification. It is this dimension that allows econometric models to provide reliable information for both explaining the past and predicting the future. If the trade off for attaining more precise forecasts with econometric models is to strip them of their theory-based structure, rather use mere extrapolative devices, which are much simpler specifications besides not making any theory/economic sense as well.

In this paper, we provide empirical evidence showing that it is possible to reconcile the ability to forecast well, with the mandatory theoretical structure of econometric specifications. For that purpose, using data from 1969 to 1993, we estimate alternative VAR specifications and obtain out-of-sample (1994-2005) forecasts of the UK tourism expenditure shares for France, Spain and Portugal. These specifications are *'congruently'* modelled (Hendry, 2004), to include increasing number of tested theoretical restrictions and, therefore, growing levels of structural complexity. In time series contexts, meticulous modelling requires a detailed knowledge of the particular conditions surrounding the headway of the variable of interest. Because forecasting failure can also be attributed to ignored structural breaks, it cannot be stressed enough how important it is to correctly incorporate them in the models specification. Indeed, there is no statistical substitute for in-depth, detailed and scrupulous gathering of information about which, how, and for how long, events disrupt coefficients' structural stability. This is time consuming and hard work, but once done, it usually compensates.

This approach of '*congruent modelling for forecasting*' using rigorous in-sample modelling (which can avoid distorted forecasts), generates econometric specifications capable of outperforming other forecasting devices less rigorously modelled. Indeed, after comparing the forecast performance of the alternative VAR specifications, we verify that the models subject to cointegration and including exogenous restrictions, structural breaks and over-identifying theoretical assumptions, outperform unrestricted reduced form VAR models both for short and long run horizons. Additionally, we also confirm that the former outperform the

unrestricted differenced VAR model (benchmark) for multi-step horizons, while being equally precise for one-step horizons.

The account of these results unfolds in the following way: section 2 describes the specification of VAR models used in this forecast comparison exercise and present the results of the cointegration tests performed for some of them. Section 3, explains which models are assessed and why. Section 4 presents the forecasting results obtained. Section 5 concludes.

2. DESCRIBING THE VAR MODELS[2]

2.1. Defining the variables in the VAR systems

As stated previously, the aim of this paper is to compare the predictive ability of VAR models, some of which are defined in De Mello and Nell (2005) while others are constructed here. The VAR models for the UK tourism demand are estimated for the period 1969-1993 with the variables in vector $\mathbf{V}_t = \begin{bmatrix} WF & WS & PP & PS & PF & E \end{bmatrix}$, where WF, WS and WP represent the UK tourists expenditure shares for France, Spain and Portugal, respectively; PF, PS and PP stand for tourism effective prices for the same destinations and E is the UK real per capita tourism budget.[3]

With data until 1997, De Mello and Nell (2005) estimate their models for 1969-1993 leaving the last 4 observations for forecasting purposes. Although the authors argue that some models perform better than others, they also recognise that the small number of out-of-sample forecasts prevents them from drawing more definite conclusions.

In this paper, we gathered additional data up until 2005 and, with models estimated for 1969-1993, obtain 12 out-of-sample forecasts for the years ahead. We believe that this extended sample of forecasts is sufficient to test the prediction accuracy of all the competing models and, therefore, draw unambiguous statements about both the role of cointegration and that of proper model specification in the precision of VAR forecasts.

---

[2] This section draws from sections 2 and 3 in De Mello and Nell (2005).
[3] The definitions and data sources of the variables are described in Appendix 1, following De Mello and Nell (2005) and De Mello *et. al* (2002). The extended data used here were obtained from the same sources.

To demonstrate that integration and cointegration play different and relevant roles in improving forecast accuracy and to show that well specified econometric models, under the right conditions, are hard to beat in a forecast performance contest, we obtain forecasts from simpler models (first differenced and reduced forms) and compare them with those including all proper structural features. To gather the forecasts, we start by estimating the most basic of VAR forms, the reduced form unrestricted VAR. Then, we add three exogenous dummy variables (D1, D2 and D3) representing observed structural breaks in the data. We carry on imposing exogeneity on the variable E, and finally we include both the appropriate exogenous restrictions and the dummy variables simultaneously.[4]

The dummy variables are defined as follows: D1=1 for 1974-1981 and zero otherwise, stands for the 70s oil crises and the political changes that occurred in Portugal and Spain in the same decade. D2 and D3 account for the integration process of the Iberian countries in the EU. This process is split into two sub-periods: the pre-integration period (1982-1988), where D2=1 and zero otherwise, and the integration period (1989-1999), where D3=1 and zero otherwise.[5]

Table 1: VAR models description, codification and included variables

Models' description and codes	Variables in vector V_t	Integration	Cointegration
00 - Reduced Form VAR	WF WS WP PP PS PF E	**NO**	**NO**
01 - Reduced Form VAR	WF WS WP PP PS PF **&** E D1 D2 D3	**NO**	**NO**
10 - Differenced VAR	∆WF ∆WS ∆WP ∆PP ∆PS ∆PF ∆E	**YES**	
20 - Cointegrate VAR	WF WS WP PP PS PF E		**YES**
21 - Cointegrate VAR	WF WS WP PP PS PF **&** E D1 D2 D3		**YES**
21shp - Model **21**, under symmetry homogeneity, and null cross-price effects	WF WS WP PP PS PF **&** E D1 D2 D3		**YES**

Table 1 reports the VAR models under scrutiny. The first column provides the codes and summary description of the models. The second, shows the variables included in the models with symbol "**&**" separating endogenous from exogenous regressors. The third and fourth indicate which models are subject to integration (differentiation), cointegration or none. Given the codes in Table 1, we are interested in comparing the forecast performance of the

---

[4] The exogenous dummy variables D1, D2 and D3 and the exogenous restriction on E were established as relevant in De Mello and Nell (2005).
[5] Although Portugal and Spain joined the European Union (EU) in 1986, the full implications of communitarian legislation for some sectors were only attained some years later. Furthermore, we believe that the integration period should be extended up until the adoption of the Euro as the common currency in 1999.

models within three different set-ups: first, we compare models from which dummy variables and exogeneity restrictions are missing (**00** and **20)** with models where the dummy variables and the exogeneity restriction on E are included (**01** and **21**).

Second, we compare models subject to simple integration (**10**) with models subject to cointegration (**20** and **21**), and models not subject to either integration or cointegration (**00** and **01**). This set-up allows for a comparison of reduced form unrestricted VAR models (**00** and **01**), repeatedly acclaimed as excellent forecasters, with differenced (integrated) VAR models (**10**), regarded as benchmarks for forecast accuracy ruling (Hendry and Clements, 1995), and with cointegrated structural VAR models, distrusted as even acceptable forecasters by some (Christoffersen and Diebold, 1998), but praised as fair predictors by others (Engle and Yoo, 1987; Hendry and Mizon, 1993; Clements and Hendry, 1998).

A third level of comparison confronts the cointegrated VAR models **21** and **21shp**. At this level, we want to know if the restrictions of symmetry, homogeneity and null cross-price effects between the share equations of Portugal and France[6], which are imposed on the cointegrated VAR **21shp**, make it a better forecaster than the cointegrated VAR **21**, which ignores these over-identifying restrictions.

Within these main set-ups, we also consider different forecasting horizons, given that forecast accuracy may vary with the horizon's dimension. Hence, given the 12 out-of-sample point forecasts available, we consider the shorter (1-step) and the longer possible (12-step) horizon ranges for this comparison exercise.

2.2. Specifying the VAR systems

The shares of France (WF), Spain (WS) and Portugal (WP), of the UK tourists spending are functions of prices (PF, PS and PP), the UK real per capita tourism budget (E) and a set of dummy variables (D1, D2 and D3), such that: $Wi = f(PP, PS, PF, E, D1, D2, D3); i = F, S, P$. Using data from 1969 to 2005 for the variables in vector $\mathbf{V}_t = \begin{bmatrix} WF & WS & PP & PS & PF & E \end{bmatrix}$, the reduced form of a first order unrestricted VAR system of the UK tourism demand for France, Spain and Portugal (model **00**) is estimated with:

---

[6] This restriction was tested and not rejected in De Mello and Nell (2005).

$$\mathbf{z}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t \qquad (1)$$

Where $\mathbf{z'}_t = \begin{bmatrix} WF_t & WS_t & PP_t & PS_t & PF_t & E_t \end{bmatrix}$, $\mathbf{A}_0$ is a (6x1) vector of intercepts, $\mathbf{A}_1$ is a (6x6) matrix of parameters, and $\boldsymbol{\varepsilon}_t$ is a vector of well behaved disturbances.[7]

The corresponding cointegrated vector error correction model (VECM) with endogenous and exogenous I(1) variables, intercept and no trend, is:

$$\Delta \mathbf{y}_t = \mathbf{a}_{0y} - \boldsymbol{\Pi}_y \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{iy} \Delta \mathbf{z}_{t-i} + \mathbf{e}_t \qquad (2)$$

In (2), $\mathbf{z}_t = (\mathbf{y'}_t ; \mathbf{x'}_t)'$ is a vector of m variables where vector $\mathbf{x'}_t \, (m_x \times 1) = \begin{bmatrix} E\,D1\,D2\,D3 \end{bmatrix}$ includes the exogenous variables and vector $\mathbf{y'}_t \, (m_y \times 1) = \begin{bmatrix} WF\,WS\,PF\,PS\,PP \end{bmatrix}$ includes the endogenous variables. The implicit VECM in $\mathbf{x'}_t$ is given by:

$$\Delta \mathbf{x}_t = \mathbf{a}_{0x} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_{ix} \Delta \mathbf{z}_{t-i} + \mathbf{v}_t \qquad (3)$$

Combining (2) and (3) gives: $\Delta \mathbf{z}_t = \mathbf{a}_0 - \boldsymbol{\Pi} \mathbf{z}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\Gamma}_i \Delta \mathbf{z}_{t-i} + \mathbf{u}_t \qquad (4)$

In the combined model (4), $\mathbf{a}_0 = \begin{pmatrix} \mathbf{a}_{0y} \\ \mathbf{a}_{0x} \end{pmatrix}$; $\boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{\Pi}_y \\ \mathbf{0} \end{pmatrix}$; $\boldsymbol{\Gamma}_i = \begin{pmatrix} \boldsymbol{\Gamma}_{iy} \\ \boldsymbol{\Gamma}_{ix} \end{pmatrix}$; $\mathbf{u}_t = \begin{pmatrix} \mathbf{e}_t \\ \mathbf{v}_t \end{pmatrix} \to IN(\mathbf{0}, \Sigma)$.

The disturbances are independent and normally distributed; Σ is a positive-definite symmetric matrix; $\mathbf{a}_{0y}(m_y \times 1)$ and $\mathbf{a}_{0x}(m_x \times 1)$ are vectors of intercepts; $\boldsymbol{\Pi}_y(m_y \times m)$, is the long-run multiplier matrix of order $(m_y \times m)$, where $m = m_y + m_x$; $\boldsymbol{\Gamma}_{1y}, \boldsymbol{\Gamma}_{2y}, \dots, \boldsymbol{\Gamma}_{p-1,y}$, are coefficient matrices of order $(m_y \times m)$, capturing the short-run dynamic effects.

---

[7] Likewise, the form in differences can be expressed as: $\Delta \mathbf{z}_t = \mathbf{A}_0^* + \mathbf{A}_1^* \Delta \mathbf{z}_{t-1} + \boldsymbol{\upsilon}_t$, where $\Delta \mathbf{z}_t$ is written as $\Delta \mathbf{z'}_t = \begin{bmatrix} \Delta WF_t & \Delta WS_t & \Delta PP_t & \Delta PS_t & \Delta PF_t & \Delta E_t \end{bmatrix}$ and $\mathbf{A}_0^*$, $\mathbf{A}_1^*$ and $\boldsymbol{\upsilon}_t$, are defined as previously; for forms in levels with dummy variables, $\mathbf{z}_t$ is $\mathbf{z'}_t = \begin{bmatrix} WF_t & WS_t & PP_t & PS_t & PF_t & E_t & D1_t & D2_t & D3_t \end{bmatrix}$.

The system of cointegrated long run equilibrium demand equations for the three destinations can be written in the following form:

$$\begin{cases} WP_t = \alpha_P + \gamma_{PP}PP_t + \gamma_{PS}PS_t + \gamma_{PF}PF_t + \beta_P E_t + \delta_{P1}D1 + \delta_{P2}D2 + \delta_{P3}D3 + up_t \\ WS_t = \alpha_S + \gamma_{SP}PP_t + \gamma_{SS}PS_t + \gamma_{SF}PF_t + \beta_S E_t + \delta_{S1}D1 + \delta_{S2}D2 + \delta_{S3}D3 + us_t \\ WF_t = \alpha_F + \gamma_{FP}PP_t + \gamma_{FS}PS_t + \gamma_{FF}PF_t + \beta_F E_t + \delta_{F1}D1 + \delta_{F2}D2 + \delta_{F3}D3 + uf_t \end{cases}$$

2.2. Testing the VAR systems for cointegration with Johansen's reduced rank test.

All variables in the VAR models were previously tested for unit roots and all were found to be $I(1)$. This implies that estimation, inference and forecasting procedures are strictly valid if cointegrated relationship(s) exist. De Mello and Nell (2005) report Johansen's reduced rank tests for models **00** and **01**[8] estimated for the period 1969-1997, but do not report these tests for the models estimated until 1993.[9] Since the models estimated until 1993 are the base for generating all the forecasts used in this study, we think that it is important to report the reduced rank tests of models **00** and **01** estimated for the in-sample period 1969-1993.

In a tourism demand context involving a system of two destination share equations, which depend on prices and per capita tourism budget of the origin country, the number of existing long-run relationships must be equal to the number of share equations in the system. Consequently, for both models **00** and **01**, we expect to find exactly two cointegrated vectors, confirming the existence of two long-run equilibrium relations.

Johansen's hypothesis (Johansen, 1991; Johansen and Juselius, 1992), which assumes that there are at most r cointegrating vectors in the system, can be tested with either the *eigenvalue trace statistic* (λ_{trace}), which null is $r = q$ ($q = 0, 1, \ldots, n-1$) against the alternative $r \geq q+1$, or the *maximum eigenvalue statistic* (λ_{max}) which null is $r = q$ against the alternative of $r = q+1$. Table 2 shows the cointegration test results. The first column of Table 2 displays the eigenvalues associated with each I(1) endogenous variable, ordered from highest to lowest, required to compute λ_{max} and λ_{trace}. The second column shows the various hypotheses to be tested. The remaining columns give λ_{max} and λ_{trace} estimates and their respective 5% and

---

[8] Denominated "Purevar" and "Wholevar" respectively, in De Mello and Nell (2005).
[9] Although the authors state, in footnote 7, that such tests were performed with satisfactory outcomes.

10% critical values. For both models, at the 5% level, both λ_{max} and λ_{trace} reject the null of $r = 0$ and $r = 1$ (statistic value>critical value), but do not reject $r = 2$ (statistic value<critical value). Hence, both statistics unequivocally support two cointegrated vectors. Thus, the restriction of $r = 2$ is included in the system imposing the existence of two long-run cointegrated relationships on both VAR models. These models, once subject to cointegrating restrictions, acquire different features and, hence, different codes which are set in Table 1. The first column of Table 2 displays the eigenvalues associated with each I(1) endogenous variable, ordered from highest to lowest, required to compute λ_{max} and λ_{trace}. The second column shows the various hypotheses to be tested. The remaining columns give λ_{max} and λ_{trace} estimates and their respective 5% and 10% critical values. For both models, at the 5% level, both λ_{max} and λ_{trace} reject the null of $r = 0$ and $r = 1$ (statistic value>critical value), but do not reject $r = 2$ (statistic value<critical value).

Table 2: Cointegration rank tests for models **00** and **01.**

Eigen values	H_0		$\hat{\lambda}_{max}$	λ_{max} critical		$\hat{\lambda}_{trace}$	λ_{Trace} critical	
	r	m−r		5%	10%		5%	10%
MODEL **00**								
λ_1=0.9526	= 0	= 6	**73.19**	40.53	37.65	**166.82**	102.56	97.87
λ_2=0.8270	= 1	= 5	**42.11**	34.40	31.73	**93.62**	75.98	71.81
λ_3=0.6889	= 2	= 4	28.02	28.27	25.80	51.51	53.48	49.95
λ_4=0.4204	= 3	= 3	13.09	22.04	19.86	23.49	34.87	31.93
λ_5=0.2956	= 4	= 2	8.41	15.87	13.81	10.40	20.18	17.88
λ_6=0.0796	= 5	= 1	1.99	9.16	7.53	1.99	9.16	7.53
MODEL **01**								
λ_1=0.8798	= 0	= 5	**53.00**	46.77	43.80	**147.61**	119.77	114.38
λ_2=0.7734	= 1	= 4	**42.55**	40.91	38.03	**94.61**	90.60	85.34
λ_3=0.5833	= 2	= 3	28.55	34.51	31.73	52.06	63.10	59.23
λ_4=0.2767	= 3	= 2	15.82	27.82	25.27	23.51	39.94	36.84
λ_5=0.2489	= 4	= 1	7.69	20.63	18.24	7.69	20.63	18.24

Hence, both statistics unequivocally support two cointegrated vectors. Thus, the restriction of $r = 2$ is included in the system imposing the existence of two long-run cointegrated relationships on both VAR models. These models, once subject to cointegrating restrictions, acquire different features and, hence, different codes which are set in Table 1. Accordingly, the reduced form VAR **00** subject to cointegration is named VAR **20**; and the reduced form VAR **01** subject to cointegration is named VAR **21**. In the same way, when the reduced form VAR **00** is subject to differentiation the resulting first differenced (integrated) VAR is labelled model **10.**

3. EXPLAINING WHICH MODELS ARE GOING TO BE ASSESSED AND WHY

De Mello and Nell (2005) consider model **20**[10] unsuitable to provide reliable information on the long-run demand behaviour of UK tourists and dismiss the model altogether, not even considering it for forecasting purposes. However, we believe that this model should be used for forecasting purposes, not only because it allows us to contrast the predictive ability of models which do not include the proper structural breaks and exogenous restrictions, against that of models that include them (**00** against **01** and **20** against **21**), but also because we can compare the forecasting ability of models not subject to either cointegration or integration (**00** and **01**) with that of models subject to integration (**10**), and to cointegration (**20** and **21**). So, model **20** becomes indispensable and we use it for this purpose.

Model **21** is statistically robust (passes all diagnostic tests) and theoretically consistent (provides estimates consistent with theory predictions). Hence, it is this model that is further subject to over-identifying restrictions on the equilibrium relationships. The theoretical restrictions of homogeneity and symmetry, reflecting the logic of consumers' behaviour, and the hypothesis suggesting that price changes in France (Portugal) do not affect UK tourism demand for Portugal (France) are not rejected by the data and, therefore, are introduced into the cointegrated VAR **21**. This model, subject to the additional restrictions of symmetry, homogeneity and null cross-price effects between France and Portugal, is labelled **21shp**.

To summarize, we are interested in comparing the forecast performance on three different levels. On a first level, we want to compare simpler models, from which relevant variables and exogeneity restrictions are missing, with more complex models where all relevant variables and proper exogeneity restrictions are included, that is: model **00** with model **01**; and model **20** with models **21**. On a second level, we want to compare models not subject to either cointegration or integration, to models subject to integration, and models subject to cointegration. That is, we compare the benchmark model **10** (subject to integration) with models **00** and **01** (not subject to either integration or cointegration) on the one hand, and models **20** and **21** (subject to cointegration) on the other hand. The third and final level of comparison considers the possibility of over-identifying restrictions, such as homogeneity,

---

[10] Denominated "*Cointegrated Purevar*" in De Mello and Nell (2005).

symmetry and null cross-price effects, making a difference in the forecast accuracy of cointegrated VAR models. Accordingly, we compare the performance of the cointegrated just-identified VAR **21,** with that of the cointegrated over-identified VAR **21shp**.

All models are estimated for the in-sample period 1969-1993 leaving 12 out-of-sample observations from 1994 to 2005, for forecast purposes only. Since the forecast comparisons must be established within the same horizon we obtain, for each of the six models, twelve one-step and twelve multi-step forecasts for short- and long-run comparisons, respectively.

4. ASSESSING THE MODELS FORECASTING PERFORMANCE

To compare the forecasting performance of the models, Tables 4, 5, 6 and 7 show the root mean squared prediction error (RMSE), and mean absolute percentage error (MAPE), of 1-step and multi-step forecasts for the tourism shares of France (WF) Spain (WS) and Portugal (WP). The last lines of all tables show, for each model, the weighted average of the three share equations MAPEs (3EQAV). The measure 3EQAV evaluates each system as a whole by weighting the MAPE of each share equation by its market worth. Consequently, this measure is used as a yardstick to rank all models. The weights assessing the relative importance of the three destinations are their average market shares in the in-sample period 1969-1993. The average shares of Portugal, France and Spain are, respectively, 7.75%; of is 35.9%, is 56.35%. Hence, $3EQAV = 0,5635 \times MAPE_{WF} + 0,3590 \times MAPE_{WS} + 0,0775 \times MAPE_{WP}$.

We also display Figures 1 and 2 showing a plot of the actual values, one-step and multi-step forecasts for the shares of France (WF), Spain (WS) and Portugal (WP). Figure 1 depicts one-step forecasts obtained from the integrated model **10** (best predictor), the cointegrated over-identified model **21hsp** (second best) and the reduced form not subject to either integration or cointegration, model **01** (worst predictor). Figure 2, shows the multi-step forecasts obtained from cointegrated over-identified model **21hsp** (best predictor), cointegrated just-identified model **21** (second best) and integrated model **10** (worst predictor).

Evoking the stages for which we want to carry out the forecast accuracy assessment, first we compare models from which the structural breaks and exogeneity restrictions are missing with models where these features are included; than we compare the benchmark VAR model subject to integration, with models subject to cointegration and not subject to either

integration or cointegration; finally, we compare cointegrated just- and over-identified models. All these comparisons are carried out for one- and multi-step horizons. The details of such analysis unfold in the next sub-sections.

Table 4: Performance of models **00, 10** and **20** for **one-step** forecasts (1994-2005)

MODELS with variables WF, WS, WP, PF, PS, PP, E					
(Reduced form) MODEL 00		(Integrated) MODEL 10		(Cointegrated) MODEL 20	
WF		**WF**		**WF**	
RMSE	0,0659710	RMSE	0,0200069	RMSE	0,0634559
MAPE	0,1476363	MAPE	0,0380789	MAPE	0,1456248
WS		**WS**		**WS**	
RMSE	0,0602794	RMSE	0,0220751	RMSE	0,0661951
MAPE	0,0973918	MAPE	0,0286661	MAPE	0,1166185
WP		**WP**		**WP**	
RMSE	0,0117008	RMSE	0,0099011	RMSE	0,0068995
MAPE	0,1261156	MAPE	0,0875719	MAPE	0,0737406
3 EQ AV 11,77%	0,1176557	3 EQ AV 3,66%	0,0366105	3 EQ AV 12,73%	0,1237087

Table 5: Performance of models **01**, **21** and **21shp** for **one-step** forecasts (1994-2005)

MODELS with variables: WF, WS, WP, PF, PS, PP & E, D1, D2, D3					
(Reduced form) MODEL 01		(Cointegrated) MODEL 21		(Cointegrated) MODEL 21shp	
WF		**WF**		**WF**	
RMSE	0,133521	RMSE	0,044822	RMSE	0,035051
MAPE	0,326401	MAPE	0,100029	MAPE	0,073017
WS		**WS**		**WS**	
RMSE	0,126751	RMSE	0,033065	RMSE	0,029885
MAPE	0,222832	MAPE	0,055678	MAPE	0,044653
WP		**WP**		**WP**	
RMSE	0,008012	RMSE	0,025773	RMSE	0,008047
MAPE	0,080179	MAPE	0,294285	MAPE	0,081551
3 EQ AV 24,90%	0,248958	3 EQ AV 9,00%	0,090092	3 EQ AV 5,77%	0,057695

Table 6: Performance of models **00**, **10** and **20** for **multi-step** forecasts (1994-2005)

MODELS with variables WF, WS, WP, PF, PS, PP, E					
(Reduced form) MODEL 00		(Integrated) MODEL 10		(Cointegrated) MODEL 20	
WF		**WF**		**WF**	
RMSE	0,051087	RMSE	0,072245	RMSE	0,061082
MAPE	0,122937	MAPE	0,163151	MAPE	0,145374
WS		**WS**		**WS**	
RMSE	0,060498	RMSE	0,102375	RMSE	0,058979
MAPE	0,100146	MAPE	0,164761	MAPE	0,105616
WP		**WP**		**WP**	
RMSE	0,011194	RMSE	0,032542	RMSE	0,004708
MAPE	0,126426	MAPE	0,378562	MAPE	0,049234
3 EQ AV 11,04%	0,110365	3 EQ AV 18,08%	0,180753	3 EQ AV 11,55%	0,115520

Table 7: Performance of models **01**, **21** and **21shp** for **multi-step** forecasts (1994-2005)

MODELS with variables: WF, WS, WP, PF, PS, PP & E, D1, D2, D3					
(Reduced form) MODEL 01		(Cointegrated) MODEL 21		(Cointegrated) MODEL 21shp	
WF		**WF**		**WF**	
RMSE	0,05898	RMSE	0,02284	RMSE	0,02578
MAPE	0,13211	MAPE	0,05394	MAPE	0,04944
WS		**WS**		**WS**	
RMSE	0,09347	RMSE	0,03382	RMSE	0,02364
MAPE	0,13162	MAPE	0,05623	MAPE	0,03093
WP		**WP**		**WP**	
RMSE	0,04940	RMSE	0,01954	RMSE	0,00537
MAPE	0,48191	MAPE	0,22601	MAPE	0,05439
3 EQ AV 15,89%	0,15896	3 EQ AV 6,86%	0,06853	3 EQ AV 3,94%	0,03938

Figure 1: Actual values and **one-step** forecasts for France, Spain and Portugal shares
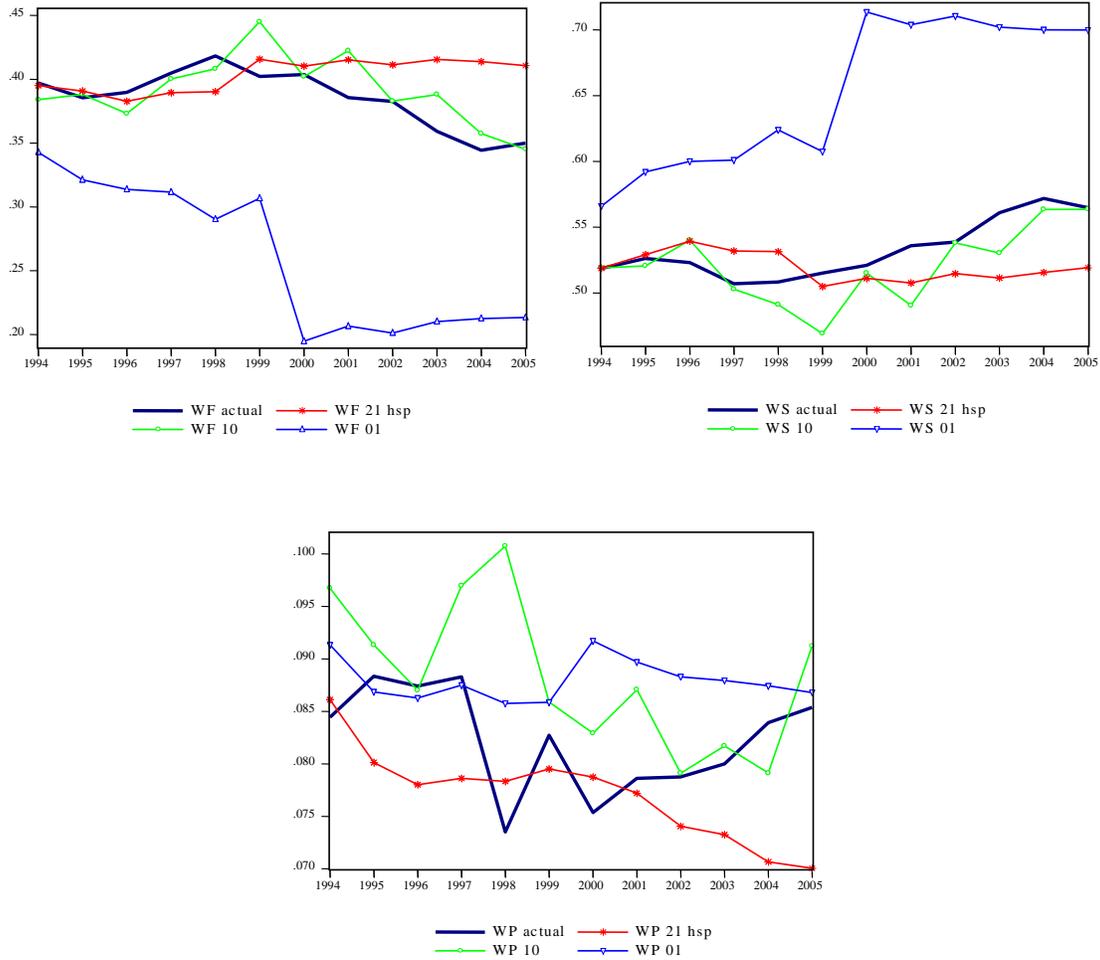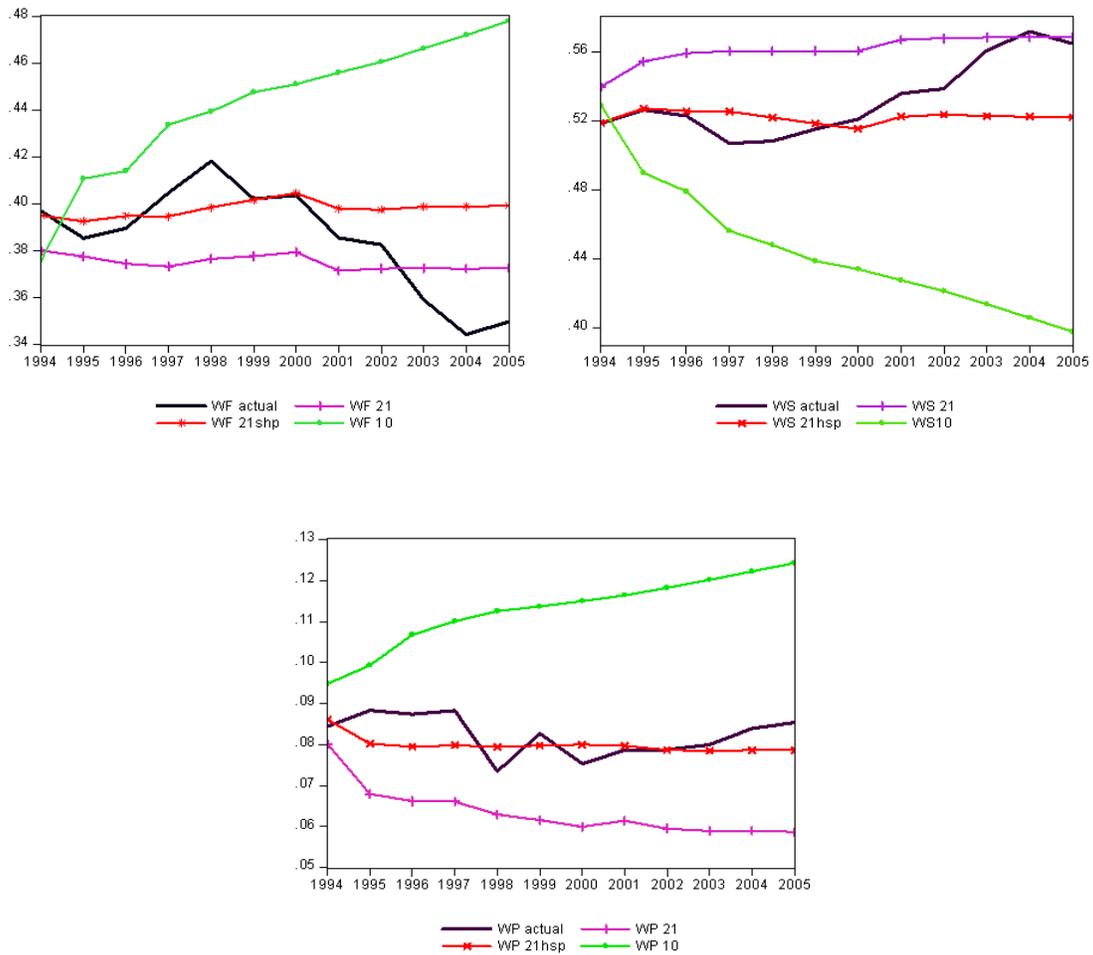
Figure 2: Actual values and **multi-step** forecasts for France, Spain and Portugal Shares

4.1. First stage: VARs with and without exogenous restrictions and dummy variables.

 One-step a-head horizons: Tables 4 and 5 results show that the added dummy variables and exogeneity restrictions do not improve the forecasting performance of the reduced form a VAR models. Quite the opposite, their inclusion doubles the forecast imprecision of these models. Indeed, the imprecision of the reduced form VAR, measured by the weighed average 3EQAV, increases from 11.8% in model **00** (Table 4), to 24.9% in model **01** (Table 5).

The opposite occurs with the cointegrated VAR models. The imprecision of model **20** (Table 4) without the dummy variables and exogeneity restrictions is reduced from 12.4% to 9.0% when the dummy variables and exogeneity restrictions are added in model **21**(Table 5). The forecast imprecision is further reduced to 5.8% in model **21shp** (Table 5), when the additional restrictions of homogeneity, symmetry and null cross-price effects are incorporated.

15

From the results of Tables 4 and 5 it is apparent that the best forecast performance for one-step horizons belongs to the differenced (benchmark) VAR model **10** (Table 4), with a 3EQAV of 3.7%. The second best position belongs to the cointegrated VAR model **21shp** (Table 5), with a 3EQAV of 5.8%. Yet, the accuracy measures for each of the destination shares in both these models do not differ much. Indeed, the forecasts of model **21shp** for the share of France show a MAPE of 7.3%, against that of 3.8% obtained with model **10;** the forecasts for the share of Spain in model **21shp** show a MAPE of 4.5%, against that of 2.9% in model **10**; and the forecasts for the share of Portugal in model **21shp** show a MAPE of 8.2%, against that of 8,8% in model **10**. Thus, it is possible that the forecast accuracy of model **10** is equivalent to that of model **21hsp**. To prove the statistical validity of this equivalence we subject the forecasts of both models to the scrutiny of equal accuracy tests of Diebold and Mariano (1995) and Harvey *et al.* (1997, 1998).

Equal accuracy of two competing forecast series, (i) and (j), can be judged by testing the significance of the difference d_{ij} between economic losses associated with forecast error series e_i and e_j. Assuming that the loss related with prediction failure is a symmetric function of the forecast error, we allow time t loss associated with a series of n forecasts to be a direct function of the forecast error $g(e)$ such that $g(e) = e^2$. The null of equal accuracy of two competing h-step forecast series (i) and (j) is: $E(d_{ij_t}) = 0$, where $d_{ij_t} = g(e_{i_t}) - g(e_{j_t})$; t=1,…, n. For testing the null, we use Diebold and Mariano's (1995) **S1** and Harvey *et al.*'s (1997) **S1*** test statistics. The latter is a modified version of the Diebold and Mariano's (DM) **S1** statistic. The DM statistic is defined as: $\mathbf{S1} = \overline{d}\left[\hat{var}(\overline{d})\right]^{-1/2}$.

Under the null of equal accuracy between two forecast series, **S1** is valid for a very wide class of loss functions (not needing to be quadratic, symmetric, or continuous), and for forecast errors that can be non-Gaussian, nonzero mean, serially correlated and contemporaneously correlated. Nonetheless, it can be oversized in small samples and even more so as the forecast horizon increases. To alleviating this problem, Harvey *et al.* (1997) propose an approximately unbiased estimator for the variance of $\overline{d}$, which gives rise to a modified version of the DM test statistic, such that: $\mathbf{S1*} = \mathbf{S1}\left\{\left[n + 1 - 2h + n^{-1}h(h-1)\right]/n\right\}^{1/2}$, where **S1** is the original DM statistic. Harvey *et al.* (1997) also suggest comparing **S1*** with critical values from the T (n-1) distribution, rather than from the N (0; 1) used for the **S1**

statistic. The results for the equal accuracy tests **S1** and **S1**[*] are reported in Table 8. Based on these results, we can state that, for one-step horizon, the predictive accuracies of benchmark model **10** and cointegrated model **21shp** are statistically equivalent.

Table 8: Tests of equal forecast accuracy for **one-step** horizons with models **21shp** and **10**

	Test Statistic	Statistic Distribution	5% critical values	Equal accuracy tests for the forecast series of models **10** and **21shp**	
				STATISTIC	RESULT
Portugal	S1	N(0, 1)	± 1.96	–0.453	Not rejected
	S1[*]	T(11)	|2.20|	–0.434	Not rejected
Spain	S1	N(0, 1)	± 1.96	–0.296	Not rejected
	S1[*]	T(11)	|2.20|	–0.283	Not rejected
France	S1	N(0, 1)	± 1.96	0.141	Not rejected
	S1[*]	T(11)	|2.20|	0.135	Not rejected

Multi-step a-head horizons: Considering now the multi-step forecasts in Tables 6 and 7, we notice a major difference contrasting with the one-step case: the benchmark VAR model **10** (Table 6) is now the worse forecaster with an overall imprecision of 18%, while in the one-step case (Table 4) it was the best predictor, with an overall imprecision of 3.7%. Additionally, we notice that the added dummy variables and exogeneity restriction worsen the forecasting performance of the reduced form VAR, as they also did in the previous case of the one-step horizon. Indeed, the forecast imprecision of 11% (model **00** in Table 6) for the reduced form VAR without the dummy variables and exogeneity restrictions, increases to 16% (model **01** in Table 7) when the dummies and exogeneity restriction are added.

Again, the opposite occurs with the cointegrated VAR models. The forecast imprecision of 11.6% for the cointegrated VAR **20** (Table 6) without the dummy variables and exogeneity restrictions, is reduced to 6.9% in model **21** (Table 7) where the relevant dummy variables and exogeneity restrictions are included. The imprecision is further reduced to 3.9%, when the additional restrictions of homogeneity, symmetry and null cross-price effects are added in model **21shp** (Table 7).

Because model **21shp** presents now the smallest average percentage error (3EQAV), it is considered as the best overall forecaster in the multi-step horizon case. Furthermore, this model also presents remarkably accurate forecasts on an equation by equation basis. Indeed, the MAPE for France in model **21shp** is 4.9% against 16.3% for the differenced (benchmark) model **10**, and 12.3% for the reduced form model **00**; the MAPE for Spain in model **21shp** is 3.1% against 16.5% in model **10** and 10.0% in model **00**; the MAPE for Portugal in model **21shp** is 5.4% against 37.9% of in model **10**, and 12.6% in model **00**.

Consequently, the following conclusions can be extracted at this point. First, the reduce form VAR models are poor predictors independently of including or not the proper structural breaks and exogeneity restrictions, and for whatever forecast horizon. In contrast, the accuracy of the cointegrated VAR models improves with the inclusion of the proper structural breaks and exogeneity restrictions and more so, with the inclusion of the over-identifying restrictions of homogeneity, symmetry and null cross price effects. Once all the restrictions are imposed, these models become excellent predictors for both one- or multi-step horizons. Finally we conclude that benchmark VAR model **10** and cointegrated VAR model **21shp** are the best forecasters in the one-step horizon case but, in the multi-step case, the former becomes the worst forecaster, while the latter maintains its rank of best predictor.

4.2. Second stage: cointegrated, integrated and reduced form VARs

One-step a-head horizons: based on Tables 4 and 5 results, we can also compare the short-run accuracy of models not subject to either integration or cointegration (reduced form VAR models **00** and **01**) with that of models subject to simple integration (differenced VAR **10**) and models subject to cointegration (VAR models **20**, **21** and **21shp**).

Model **10** (Table 4) subject to integration and model **21shp** (Table 5) subject to cointegration, can both be considered as the best forecasters in the one-step case because their imprecision measures, on an equation-by-equation basis, are not statistically different, as showed by the equal accuracy tests displayed in Table 8. Hence, for one-step horizons, integrated and cointegrated VAR models perform equivalently in producing the best forecasts.

The worst forecasters in these circumstances are the models not subject to either integration or cointegration. Their overall imprecision is 11.8% for the reduced form VAR **00**, and 24.9%

for the reduced form VAR **01**. Thus, models not subject to either integration or cointegration are poor short run predictors.

Multi-step a-head horizons: once again, the results in the multi-step case reverse the conclusions extracted from the previous one-step case. Indeed, first differenced model **10** (Table 6) subject to integration, is now the worst forecaster, closely followed by models **00** (Table 6) and **01** (Table 7), not subject to either integration or cointegration. The cointegrated VAR models **21shp** and **21** with an overall imprecision of 3.9% and 6.9% (Table 7) are now, respectively, the best and second best forecasters. Model **21shp** is not only the best overall predictor, but also supplies remarkably precise forecasts on an equation-by-equation basis, for it gives the smallest MAPE for France (4.9%), Spain (3.1%) and Portugal (5.4%) shares.

The precision differences between the best forecaster (model **21shp**, subject to cointegration) and the worse forecaster (model **10**, subject to integration) appear significant; to test the statistical veracity of those differences, we subject these models to the forecast-encompassing tests proposed in Clements and Hendry (1998). These tests are based on the following rules. Given two series of forecasts obtained from, say, model A (MA) and model B (MB), the tests examine whether the forecasts of MB (f_B), can explain the prediction errors of MA (e_A) and vice-versa. Hence, based on the estimation results of $e_{A_t} = \alpha f_{B_t} + \varepsilon_t$, which regresses MA prediction errors on MB forecasts, the test checks out if the null $H_0: \alpha = 0$ is rejected or not. If H_0 is rejected, the forecasts of MB explain the prediction errors of MA and hence, MB forecast-encompasses MA; if H_0 is not rejected, than MB forecasts are not relevant for explaining MA prediction errors and thus, MB does not forecast-encompass MA.

We expect that the tests we are about to carry out show that the forecast precision of models **21shp** and **10** are statistically different, i.e., we expect to find that model **21shp** encompasses model **10** but model **10** does not encompass model **21shp**. If this is the case there will be enough evidence to recognize that cointegrated model **21shp** outperforms integrated (benchmark) model **10**. The tests results are displayed in Table 9.

For all share equations, the results in Table 9 show that model **21shp** always encompasses model **10** while model **10** never encompasses model **21shp**. This means that the forecast accuracy of both models is significantly different and that model **21shp** outperforms model **10**. In view of these results, it is possible to say that cointegration plays a key role in obtaining accurate predictions with VAR models for both short- and long-run horizons scenarios, while

integration (differentiation) is a main factor of forecast precision only for short range horizons. VAR models not subject to either integration or cointegration are poor forecasters for both short- and long-run horizons.

Table 9: Tests of equal forecast accuracy for models **21shp** and **10** and **multi-step** horizons

	model **21shp** versus model **10** and vice-versa	Parameter estimate (NW t-stat)[11]	5% critical value	**RESULTS**
France	$e_{10_t} = \alpha_F f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_F = -0.1435$ $\scriptstyle(-2.90)$		**21shp** encompasses **10**
	$e_{21shp_t} = \beta_F f_{10_t} + \varepsilon'_t$	$\hat{\beta}_F = -0.0306$ $\scriptstyle(-1.34)$		**10** does not encompass **21shp**
Spain	$e_{10_t} = \alpha_S f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_S = 0.1676$ $\scriptstyle(3.64)$	2.20	**21shp** encompasses **10**
	$e_{21shp_t} = \beta_S f_{10_t} + \varepsilon'_t$	$\hat{\beta}_S = 0.0206$ $\scriptstyle(0.96)$		**10** does not encompass **21shp**
Portugal	$e_{10_t} = \alpha_P f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_P = -0.3799$ $\scriptstyle(-5.76)$		**21shp** encompasses **10**
	$e_{21shp_t} = \beta_P f_{10_t} + \varepsilon'_t$	$\hat{\beta}_P = 0.0202$ $\scriptstyle(1.42)$		**10** does not encompass **21shp**

4.3. Third stage: just- and over-identified cointegrated VARs

The final stage of comparison considers the forecasting performance of just-identified cointegrated VAR model **21** with that of the over-identified cointegrated VAR model **21shp** for one-step (Table 5) and multi-step (Table 7) horizons. In both cases, the cointegrated VAR **21shp**, incorporating homogeneity, symmetry and null cross-price effects, presents smaller imprecision measures (both overall, and on an equation-by-equation basis) than the cointegrated VAR **21**, which ignores those theory consistent restrictions. This means that when consumer theory assumptions are included in a well specified cointegrated structural VAR, its forecast accuracy increases. Indeed, the overall imprecision reduces from 9.0% in model **21** to 5.8% in model **21shp** for the one-step horizon (Table 5), and from 6.9% in model **21** to 3.9% in model **21shp**, for the multi-step horizon (Table 7).

---

[11] The significance tests are preformed using Newey and West (1987) consistent covariance matrix to compute the t-statistics (NW t-stat) displayed in brackets.

On an equation-by-equation basis, it can also be established that, for both one- and multi-step horizons, the cointegrated VAR incorporating the over-identifying restrictions improves the forecast precision for all destination shares. Modest precision gains (between 0.5 and 2.5 percentage points) are recorded for the shares of France and Spain and a dramatic increase of precision is achieved for the share of Portugal. In fact, the MAPE for Portugal drops from 29.4% to 8.2%, in the one-step case, and from 22.6% to 5.4% in the multi-step case. Yet, the statistical significance of these gains of precision must also be checked with proper testing. We subject these models to the same forecast-encompassing tests described above. These tests results are displayed in Table 10.

In contrast with the sizable differences between the best and worst forecasters, which statistical significances are reported in Table 9, the differences between the best (**21shp**) and second best (**21**) forecasters do not seem substantial (except for the case of the share of Portugal). Hence, we expect that the encompassing tests show that the precision differences for the shares of Spain and France to be statistically irrelevant. If this is the case, it constitutes evidence of accuracy equivalence between models **21shp** and **21** for these equations forecasts.

Table 10: Tests of equal forecast accuracy for models **21shp** and **21** and **multi-step** horizons

	Model **21shp** versus Model **21** and vice-versa	Parameter Estimate (NW t-stat)	5% critical values	**RESULTS**
France	$e_{21_t} = \alpha_F f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_F = 0.0254$ (1.11)		**21shp** does not encompass **21**
	$e_{21shp_t} = \beta_F f_{21_t} + \varepsilon'_t$	$\hat{\beta}_F = -0.0336$ (-1.26)		**21** does not encompass **21shp**
Spain	$e_{21_t} = \alpha_S f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_S = -0.0551$ (-3.60)	2.20	**21shp** encompass **21**
	$e_{21shp_t} = \beta_S f_{21_t} + \varepsilon'_t$	$\hat{\beta}_S = 0.0188$ (1.08)		**21** does not encompass **21shp**
Portugal	$e_{21_t} = \alpha_P f_{21shp_t} + \varepsilon_t$	$\hat{\alpha}_P = 0.2321$ (9.43)		**21shp** encompass **21**
	$e_{21shp_t} = \beta_P f_{21_t} + \varepsilon'_t$	$\hat{\beta}_P = 0.0361$ (1.48)		**21** does not encompass **21shp**

However, the encompassing tests displayed in table 10 show that, only in the case of the share of France, the forecast accuracy of the models is equivalent, being significantly different in the cases of Spain and Portugal. Consequently, the over-identified cointegrated VAR **21shp**

outperforms the just-identified cointegrated VAR **21** for the shares of Spain and Portugal, being equally precise only in the case of France.

5. CONCLUDING REMARKS

The reported results indicate that the forecast performance of reduced form VAR models is poor, for both short and long-run horizons. Moreover, adding the relevant structural breaks or imposing the appropriate exogenous restrictions does not help the forecast performance of these models. Quite the opposite, the already sizable overall forecast imprecision (about 12%) doubles to around 25% when the dummies and exogeneity restrictions are incorporated. Thus, we conclude that keeping unrestricted reduced form VAR specifications as plain as possible, brings about their best forecasting performance which, in any case, is not commendable. The opposite occurs with the cointegrated VAR models. The lack of the relevant structural breaks and proper exogeneity restrictions in their specifications generates more imprecise forecasts than those obtained with the cointegrated models that include these features. Moreover, these models continue to gain accuracy with the incorporation of further over-identifying theory assumptions. Indeed, the inclusion of homogeneity, symmetry and null cross-price restrictions makes the forecast imprecision of these models to decrease by one half in the one-step horizon case, and by one third in the multi-step horizon case. Hence, for both short- and long-run horizons, well specified cointegrated structural VAR models with all suitable theoretical restrictions, forecast better than unrestricted cointegrated forms. Put another way, the more structured and restricted cointegrated VAR models are, the better forecasters they become. Moreover, their forecast performance is remarkable as settled by their overall accuracy measures of less than 6% for one-step horizons and less than 4% for multi-step horizons.

The forecast precision of the first differenced VAR, used as the benchmark model is praiseworthy only in the one-step horizon case, for its average forecast errors do not exceed 4% of the actual values average. But, in the multi-step horizon case, this model's overall forecast imprecision rises up to 18%. Therefore, the differenced (benchmark) VAR model is an excellent short-run forecaster, but a poor one for longer range horizons.

Once established that the differenced (benchmark) VAR model is an excellent forecaster only for short run horizons; that the reduced form VAR models, although poor predictors in any case, forecast better the simpler they are; and that cointegrated VAR models predict better the

more structured and restricted they are, we can now turn our attention to the debate on the role of cointegration and integration as potential sources of forecast precision.

For one-step horizons, the main conclusion drawn from the evidence gathered in Tables 4 and 5 is that the integrated model **10** is the best predictor, closely followed by the cointegrated model **21shp**. However, the tests reported in Table 8 support the equal accuracy hypothesis for the performance of these two models. Hence, we might only partially agree with Christoffersen and Diebold (1998), in that increased forecast accuracy "*may simply be due to the imposition of integration, irrespective of whether cointegration is imposed*" (p.455) with some reservations on the claimed irrelevance of cointegration. In fact, as established by the equal accuracy tests, both integration and cointegration contribute equivalently to improve the forecast accuracy of VAR models. Yet, due to the form simplicity of the first difference (benchmark) VAR, it becomes the obvious choice, only if forecasting is the sole purpose.

For multi-step horizons, the main conclusion drawn from the evidence gathered in Tables 6 and 7 is that the cointegrated model **21shp** is indisputably and by far the best forecaster, clearly outperforming the differenced (benchmark) model **10**, subject to integration, and the reduced form models **00** and **01**, not subject to either integration or cointegration. Hence, supported by the encompassing tests of Table 9, we must agree with Engle and Yoo (1987) and Clements and Hendry (1998) in that it is cointegration, and not simple integration, that makes all the difference in the forecast performance of VAR models for long run horizons.

In sum, we can say that when forecasting with VAR models, integration is the key factor in short-run precision; cointegration is mandatory in long-run horizons and the absence of either integration or cointegration brings nothing but poor forecasting whatever the horizon. So, models not subject to either integration or cointegration are lousy predictors independently of the horizon, and should not be considered either to explain or forecast tourism demand shares.

Gathering all the evidence together and considering simultaneously the models specifications on one side, and integration, cointegration and reduced forms on the other, we can draw the following broad conclusion: if a VAR system of equations with identifiable structural breaks and exogenous regressors, hosts cointegrated long-run equilibrium relationships and over-identifying theoretical restrictions, its forecasting competence will only be at its best when all these features are incorporated in its equations. Miss one, and the VAR will perform below its

ability; include all and the VAR becomes a prediction device of uncommon accuracy, even for horizons as remote as 8, 10 or 12 steps ahead.

References

Archer, Brian (1976), *Demand Forecasting in Tourism,* Bangor, University of Wales Press.

Banerjee, Anindya, Juan Dolado, John Galbraith and David Hendry (1993), Cointegration, error-correction and the econometric analysis of non-stationary data", Advanced Texts in Econometrics, New York, Oxford University Press.

Clark, Todd and Michael McCracken (2001), "Tests of equal forecast accuracy and encompassing for nested models". *Journal of Econometrics,* 105: 85–110

Clements, Michael and David Hendry (1995), "Forecasting in cointegrated systems", *Journal of Applied Econometrics*, 10: 127–46.

Clements, Michael and David Hendry (1998), "Forecasting economic processes", *International Journal of Forecasting*, 14: 111–31.

Clements, Michael and David Hendry (1999), *Forecasting Non-stationary Economic Time Series*, Cambridge, Massachusetts, MIT Press.

Clements, Michael and David Hendry (2003), "Economic Forecasting: Some Lessons from Recent Research". *Economic Modelling*, 20 (2): 301-29.

Clements, Michael and David Hendry (2005), "Forecasting with breaks" in Elliot G., Granger C. and Timmerman A. (eds*.), Handbook of Economic Forecasting*, Elsevier. Amsterdam.

Christoffersen, Peter F. and Francis X. Diebold (1998), "Cointegration and Long-Horizon Forecasting", *Journal of Business & Economic Statistics*, 16 (4): 450-458.

De Mello, Maria M. (2001), "Theoretical and Empirical Issues in Tourism Demand Analysis", Ph.D. Dissertation, University of Nottingham.

De Mello, Maria M. and Natércia S. Fortuna (2005), "Testing alternative dynamic systems for modelling tourism demand" *Tourism Economics*, 11 (4): 517-37.

De Mello, Maria M. and Kevin S. Nell (2001), "The Forecasting Ability of a Cointegrated VAR Demand System with Endogenous vs. Exogenous Expenditure Variable", Working Paper nº 109 Universidade do Porto*,* Faculdade de Economia.

De Mello, Maria M. and Kevin S. Nell (2005), "The forecasting ability of a cointegrated VAR system of the UK tourism demand for France Spain and Portugal", *Empirical Economics*, 30: 277-308.

De Mello, Maria M., Allan Pack and Thea Sinclair (2002), "A system of equations model of UK tourism demand in neighbouring countries", *Applied Economics*, 34 (4): 509-22.

Diebold, Francis X. and Roberto S. Mariano (1995), "Comparing predictive accuracy", *Journal of Business & Economic Statistics*, 13 (3): 253-63.

Diebold, Francis X. and Glenn D. Rudebusch (1999), *Business Cycles*: *Durations*, *Dynamics and Forecasting,* Princeton University Press, Princeton.

Du Preez, Johann and Stephen Witt (2003), "Univariate versus multivariate time series forecasting: an application to international tourism demand", *International Journal of Forecasting*, 19: 435-51.

Engel, Robert. F. and Clive W. J. Granger (1987), "Cointegration and error correction: representation, estimation and testing", *Econometrica*, 55: 251-76.

Engel, Robert F. and Byung S. Yoo (1987), "Forecasting and testing in cointegrated systems", *Journal of Econometrics*, 35: 143-59.

Robert Fildes, and Keith Ord (2002), "Forecasting competitions: their role in improving forecasting practice and research". In Clements, M. P., and Hendry, D. F. (eds.), *A Companion to Economic Forecasting*. Oxford, Blackwell: 322–253.

Granger, Clive W. J. (1981), "The comparison of time series and econometric forecasting strategies," in *Large Scale Macro-Econometric Models, Theory and Practice*, edited by J. Kmenta and J.B. Ramsey, North Holland: 123-128.

Granger, Clive W. J. (1986), "Developments in the study of cointegrated economic variables", *Oxford Bulletin of Economics and Statistics*, 48: 213-28.

Granger, Clive W. J., and Andrew A. Weiss (1983), "Time series analysis of error-correction models", in Karlin, Samuel, Takeshi Amemiya and Leo A. Goodman (eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics*. New York, Academic Press: 255–278.

Hendry, David (2000), "On detectable and non-detectable structural change". *Structural Change and Economic Dynamics*, 11: 45–65.

Hendry, David (2005), "Robustifying forecasts from equilibrium-correction models", *Journal of Econometrics*, 135 (1-2): 399-426.

Hendry, David (2004), "Unpredictability and the foundations of economic forecasting", Working Paper nº 2004-W15, Oxford University, Nuffield Department of Economics.

Hendry, David and Jurgen A. Doornik, (1997), "The implications for econometric modelling of forecast failure". *Scottish Journal of Political Economy*, 44: 437–61.

Hendry, David and Grayham E. Mizon (1993), "Evaluating dynamic econometric models by encompassing the VAR", in Phillips, Peter C. B. (ed.), *Models, Methods and Applications of Econometrics*: 272-300. Oxford: Basil Blackwell.

Hendry, David and Grayham E. Mizon (2005), "Forecasting in the presence of structural breaks and policy regime shifts", in, Andrews, Donald W. and Stock, James H. (eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg.* Cambridge University Press, Cambridge, UK: 480-502.

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

Harvey, David, Stephen Leybourne, Paul Newbold (1997), "Testing the equality of prediction mean squared errors", *International Journal of Forecasting*, 13: 281-91.

Harvey, David, Stephen Leybourne, Paul Newbold (1998), "Tests for forecast encompassing". *Journal of Business and Economic Statistics*, 16: 254–259.

Johansen, Søren (1988), "A statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, 12: 231-54.

Johansen, Søren (1991), "Estimation and hypotheses testing of cointegrating vectors in Gaussian vector autoregressive models", *Econometrica*, 59: 1551-80.

Johansen, Søren (1995), "Identifying restrictions of linear equations with applications to simultaneous equations and cointegration", *Journal of Econometrics*, 69: 111-33.

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

Johansen, Søren and Katarina Juselius (1992), "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for the UK", *Journal of Econometrics*, 53: 211-44.

Kim, Sang and Haiyan Song (1998), "An empirical analysis of demand for Korean tourism: a cointegration and error-correction approach", *Tourism Analysis*, 3: 25-41.

Kulendran, Nada (1996), "Modelling quarterly tourism flows to Australia using cointegration analysis". *Tourism Economics*, 2: 203-22.

Kulendran, Nada and Maxwell L. King (1997), "Forecasting international quarterly tourism flows using error-correction and time series models", *International Journal of Forecasting*, 13: 319-27.

Kulendran, Nada and Stephen F. Witt (2001), "Cointegration versus least square regression". *Annals of Tourism Research*, 28 (2): 291-311.

Li, Gang, Haiyan Song and Stephan F. Witt (2005), "Recent developments in econometric modelling and forecasting". *Journal of Travel Research*, 44 (2): 82-99.

Li, Gang, Haiyan Song and Stephan F. Witt (2006), "Time varying parameter and fixed parameter linear AIDS: An application to tourism demand forecasting". *International Journal of Forecasting*, 22: 57-71.

Makridakis, Spyros and Michèle Hibon (2000), "The M3-competition: Results, conclusions and implications", *International Journal of Forecasting*, 16: 451-476.

Mills, Terence. C. (ed.) (1999), *Economic Forecasting*, Cheltenham, UK: Edward Elgar Publishing Limited.

Pesaran, M. Hashen and Bahram Pesaran (1997), *Working with microfit 4.0: interactive econometric analysis*, Oxford University Press, New York.

Pesaran, M. Hashem, Yongcheol Shin and Richard J. Smith (2000), "Structural analysis of vector error correction models with exogenous I(1) variables". *Journal of Econometrics*, **97**, 293-343.

Song, Haiyan and Gang Li (2008), "Tourism demand modelling and forecasting - A review of recent research", *Tourism Management*, 29: 203-220.

Song, Haiyan and Stephen F. Witt (2000), *Tourism demand modelling and forecasting: modern econometric approaches*, Pergamon, Amsterdam.

Song, Haiyan and Stephen F. Witt (2006), "Forecasting international tourist flows to Macau". *Tourism Management*, 27 (2): 214-224.

Song, Haiyan, Peter Romilly and Xiaming Liu (2000), "An empirical study of outbound tourism demand in the UK", *Applied Economics*, 32: 611-624.

Song, Haiyan, Stephen F. Witt and Thomas C. Jensen (2003), "Tourism forecasting: accuracy of alternative econometric models". *International Journal of Forecasting*, 19: 123-141.

Witt, S. Stephen and Christine A. Witt (1995), "Forecasting tourism demand: A review of empirical research", *International Journal of Forecasting*, 11 (3): 447-475.

Witt, S. Stephen, Haiyan Song and Panos Louvieris (2003), "Statistical testing in forecasting model selection", *Journal of Travel Research*, 42 (2): 151-158.

Zhou, Allison, Carl Bonham and Byron Gangnes (2007), "Modelling the supply and demand for tourism: a fully identified VECM approach", *University of Hawaii, Department of Economics Working Papers*, 200717.

APPENDIX 1

1.1 Variables definition

The variables in the VAR equations of the UK tourism demand for France, Spain and Portugal are the UK tourism budget shares WP, WS and WF, allocated to the destinations; their tourism effective prices PP, PS, PF; and the UK real per capita tourism budget E. Each destination share is $Wi = \dfrac{EXP_i}{EXP_F + EXP_S + EXP_P}$, where i = P (Portugal); F (France); S (Spain) and EXP_i is the nominal tourism expenditure of UK tourists in destination i. The tourism effective price in i is $Pi = \ln\left(\dfrac{CPI_i/CPI_{UK}}{R_i}\right)$, where CPI_i is the consumer price index of i, CPI_{UK} is the UK consumer price index and R_i is the exchange rate between i and the UK. The UK per capita real tourism expenditure allocated to destinations is $E = \ln\left(\dfrac{\sum_i EXP_i / UKP}{P*}\right)$, where UKP is the UK population and P* is the Stone index.

1.2 Data sources

The data for UK tourism expenditure, disaggregated by destinations and measured in £ million sterling, were obtained from *Business Monitor MA6* (1970-1993), continued as *Travel Trends* (1994-2007). Data on the UK population, price indexes and exchange rates were obtained from the *International Financial Statistics* (IMF) *Yearbooks* (1984, 1990 and 2007).