Forecasting Economic and Financial Variables with Global VARs*

M. Hashem Pesaran
Faculty of Economics and CIMF, University of Cambridge, and USC

Til Schuermann†
Federal Reserve Bank of New York and Wharton Financial Institutions Center

L. Vanessa Smith
CFAP, Judge Business School, University of Cambridge

November 26, 2007

Abstract

In this paper the GVAR model, previously estimated over the 1979Q1-2003Q4 period by Dees, de Mauro, Pesaran, and Smith (2007), is used to generate out-of-sample one quarter and four quarters ahead quarterly forecasts of real output, inflation and interest rates over the period 2004Q1-2005Q4. The forecasts are compared to typical benchmarks: univariate autoregressive and random walk models. Following the theoretical contributions of Pesaran and Timmermann (2007), the effects of model and estimation uncertainty on forecast outcomes are examined by pooling of forecasts obtained from different GVAR models estimated over alternative estimation periods. Given the size of the modeling problem – 134 variables from 26 regions made up of 33 countries covering about 90% of world output – and the heterogeneity of economies considered – industrialised, emerging, and less developed countries – as well as the very real likelihood of possibly multiple structural breaks, averaging forecasts across both models and windows makes a significant difference.

JEL Classifications: C32, C51, C53

Key Words: Forecasting using GVAR, structural breaks and forecasting, average forecasts across models and windows, financial and macroeconomic forecasts

---

*Part of this work was presented at a plenary session of the 27th International Symposium on Forecasting, New York City, June 24-27, 2007 and at the Bank of England Research Workshop on Dynamic Factor Models held at the Bank of England, 8-10 October 2007. We are grateful for comments by the discussant James Stock, as well as to Stephane Dees, Emanuil Manos Halicioglu and particulalrly to Donna Harris and M.Tugrul Vehbi for their excellent assistance in extending the GVAR data.

†Any views expressed represent those of the authors only and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

Suppose one were interested in forecasting output growth and inflation across a number of different countries; how would one go about it? What additional variables might help in such forecasting (the oil price comes to mind), and should one also consider adding financial variables such as equity returns and the long term interest rate? Should they be treated separately (two isolated equations) or together, say in a VAR? Should one consider only domestic or also foreign variables? If foreign variables are included, should they be endogenised as well? How important are cointegrating relationships, either across variables within a country or even across countries (PPP relationships come to mind)? And how should one address the ever-present problem of structural breaks which may happen multiple times in any one or several of the relations in the forecasting model under consideration?

In this paper we employ the Global Vector Autoregressive (GVAR) model, originally introduced in Pesaran, Schuermann and Weiner, PSW, (2004), and further developed in Dees, de Mauro, Pesaran and Smith, DdPS, (2007), for answering some of these questions. We do so with the recognition that macroeconomic policy analysis and risk management require taking account of the increasing interdependencies that exist across markets and countries. Indeed there are major differences in cross country correlation of output growths, inflation, and interest rates. For instance, equity returns and long term interest rates are much more closely correlated across countries as compared to output growth and inflation. This invariably means that many different channels of transmissions must be taken into account. The GVAR approach directly models the interlinkages using trade-weighted observable macroeconomic aggregates and financial variables. It allows for interdependence at a variety of levels in a transparent manner that can be empirically evaluated, including long run relationships consistent with the theory and short run relationships consistent with the data.

Nonetheless, with a modeling task of this size, it would be surprising if a single model would be universally preferred over any other. Recognising that a broader set of models might be needed to tackle the problem, we turn to the model averaging literature to arrive at better overall forecasts; Bayesian model averaging is a prominent example; see Timmermann (2006) for a recent survey on forecast combination. But simply averaging across models does not address the structural break problem. Indeed as we show, the standard Bayesian model averaging approach implicitly assumes that the underlying data generating process and the models remain stable. We solve this problem by using recent developments in the forecast pooling literature that propose to estimate the model over different sample windows (Pesaran and Timmermann, 2007). In this way parameter estimates are automatically allowed to vary over time. This strategy is especially useful when not only the nature but also the number of breaks is unknown. Finally we combine the two averaging approaches – across models and across sampling windows – to arrive at an average-average (AveAve) forecast which turns out to outperform forecasts from any single model or estimation window.

In this paper the GVAR model, previously estimated over the 1979Q1-2003Q4 period by DdPS, is used to generate out-of-sample one quarter and four quarters ahead quarterly forecasts of real output, inflation and interest rates across 26 countries/regions over the following two years, 2004Q1-2005Q4. The forecasts are compared to typical benchmarks: univariate autoregressive and random walk models. Following the theoretical contributions of Pesaran and Timmermann (2007), the effects of model and estimation uncertainty on forecast outcomes are examined by pooling of forecasts obtained from different GVAR models estimated over alternative estimation periods. All modeling exercises face the trade-off between bias and efficiency, and model averaging serves to increase the latter.

Given the size of the modeling problem – 134 variables from 26 regions made up of 33 countries covering about 90% of world output, and the heterogeneity of economies considered – industrialised, emerging, and less developed countries – as well as the very real likelihood of possibly multiple structural breaks, averaging across both models and windows makes a significant difference. The "AveAve" forecasts from the GVAR computed as the double averages of forecasts from different models estimated over different observation windows are in general better than forecasts from a single GVAR model estimated over a single observation window. The AveAve forecasts also tend to perform better than the AveM forecasts computed as averages of forecasts from different models all estimated on a single window, or AveW forecasts computed as averages of the forecasts obtained from the same model estimated across different windows. The GVAR based AveAve forecasts beat the benchmarks in the case of output, inflation and real equity prices. The results are mixed for other variables such as interest rates where in general the AveAve does as well as though in some cases worse than the benchmark forecasts. We go on to consider the effect of excluding real equity prices and long run interest rates from the GVAR model, and this has only marginal effects on forecast performance. Broadly the same also holds when only real equity prices are excluded. After dropping these two financial variables from the GVAR model, the differences in the forecast performances are not statistically significant. It is, however, clear that real variables and long run interest rates are important in forecasting real equity prices.

The plan of the paper is as follows. In Section 2 we describe the range of issues faced by the modeler in the course of building a global forecasting model, including model averaging and forecast pooling. Section 3 introduces the GVAR model and the data set used for estimation. In Section 5 we introduce the benchmark models against which the GVAR will be compared and present results. Section 7 provides some concluding remarks.

2 Model Building, Evaluation and Testing: Issues and Trade-offs

In the course of developing a model one typically goes through three stages: building, which is done entirely on an in-sample basis, evaluation which may involve some form of cross-validation, and final testing. Broadly, the objective in the initial "build" stage is to focus on statistical significance and goodness-of-fit: which functional form to use, which variables to include, possible relationships among the conditioning variables (captured, for instance, through cointegration), and so on. During evaluation one may test for the presence of structural breaks that might have occurred during the sample used for the "build" stage. Structural breaks can occur in a host of different ways such as breaks in a trend or a cointegrating relationship, and these are discussed in more detail in Section 2.2. Ideally evaluation is done with a separate sample, though that can often be prohibitively costly, one reason why techniques like cross-validation have considerable appeal. Essentially these first two stages can be considered as trading off bias (build) and efficiency (evaluate). Finally the model is put to the test: genuine out-of-sample forecast evaluation.[1] At each stage the researcher is faced with a plethora of choices, some of which we shall consider in this paper.

The GVAR framework allows for a rich structure which, if correct – and relatively stable – should yield better forecasts over short and long horizon than simpler competitors. The structure may include trends with co-trending restrictions, across country cointegration, weak cross-country dependence of shocks (innovations), trade relations and so on. Structural change could occur in any and all of these relations.

At the other extreme are a set of very simple models, the simplest of which may well be the random walk without a drift that uses the current values as forecasts for all horizons. Modest variations on the random walk theme are the random walk with drift and the univariate first or second order autoregressive (AR) models. These forecasting procedures, while deceptively simple, nonetheless are often tough to beat out-of-sample. The empirical macro and finance literature is littered with such examples.

2.1 Building a Global Model

In this section we provide a brief discussion of some of the issues we face when constructing the basic GVAR. They are, of course, the same issues faced by the simpler models. When it comes to forecast evaluation, it is natural to look at the model building stage for culprits of success or failure of the different models. For more detailed discussions, we direct the reader to PSW and DdPS.

The first and perhaps most obvious decision is which set of variables to choose to adequately capture the real and financial dynamics of the global economy. Although it is typically easier to forecast the former than the latter, it does not necessarily follow that one should choose mostly real variables for the modeling exercise. Currently the GVAR makes use of seven variables, described

[1] For a discussion of these steps and the desire to have three separate datasets, see Weiss and Kulikowski (1991).

in more detail in Section 3. The first version of the GVAR (PSW) included real output, real money supply, a price index, exchange rates, a short-term interest rate, and a stock market index. The seventh variable, common to all countries, is the price of crude oil. In the second version of the GVAR (DdPS) the money supply variable was dropped due to lack of a consistent measure across all countries, and a long-term interest rate was added to allow for simple yield curve relationships. Indeed there is strong evidence that yield curves forecast recessions (see, for instance, Estrella and Mishkin, 1998).

Several other choices need to be made; here are a few, with the final choice in parentheses. How to measure foreign variables (use trade shares); which countries to aggregate into regions (depends on the application; for instance, shared geography, e.g. Latin America, or shared currency, e.g. Eurozone); how to aggregate countries into regions (use PPP output weights); how many lags to include, domestic and foreign, by country/region (depends, but largely one, possibly two lags have been typical). More structure through over-identifying long-run restrictions can also be added (Dees, Holly, Pesaran, and Smith, 2007; DHPS).

2.2 Structural Breaks and Forecast Combinations

There is now considerable evidence that autoregressive models used in economic and financial forecasting are often unstable and subject to structural breaks, despite their success relative to other alternatives. In an extensive study of a wide variety of economic time series, Stock and Watson (1996) find that the majority of these relations are subject to structural breaks. Other studies that document instability of autoregressive models include Alogoskoufis and Smith (1991) and Garcia and Perron (1996). Structural instability is identified by Clements and Hendry (1998, 2006) as a key factor in poor forecast performance. It is also important to note that even if conditional models (e.g. country specific models in the GVAR) are structurally stable (as it is found to be the case for many of the country models in DdPS), the unconditional model which is used to generate forecasts could be subject to structural breaks. For example, consider a capital asset pricing model (which is a conditional model) where the individual firm returns are regressed on the market returns. Suppose that the parameters of these regressions are stable. But imagine that there has been a bubble in the market with a break in the univariate process of the market return which is fully reflected in the individual asset returns. It is clear that in this case the forecasting of the market return and individual asset returns will be subject to structural breaks, although the underlying CAPM models might be structurally stable.

Structural breaks can arise from institutional changes, large macroeconomic shocks, changes in economic policy, to name a few. Structural breaks can occur in a number of places in the model, from changes in the coefficients to trend breaks to changes in the error variance. Moreover, these changes could occur in one or more relations or in one or more countries, not to mention the possibility of multiple breaks. Even when the point of the break is known, depending on the size of the break there is a trade off between bias and efficiency – forecasts that use only post break data are unbiased but could be inefficient as compared to biased forecasts that also include part of the available pre-break observations. The choice of the optimal observation window depends on the full knowledge of the break point as well as the size of the break. These issues are considered in some detail in Pesaran and Timmermann (2007) who also consider a number of alternative procedures that can be used to exploit information on the break points and the sizes of breaks in forecasting.

In general, however, information about breaks is limited, particularly as far as the size of the breaks are concerned. The question then arises as to whether the optimal window size can be estimated in practice. For this to be the case we need reliable estimates of the point of the break(s) as well as the size of the breaks in the parameters. This is possible at best in the case of very simple models. In view of these difficulties rolling windows of a fixed size are often used in practice, but this comes with its own problems; if one is close to the break, the optimal window size would be short, but if one is far from the break, the optimal window would be long. It is also not clear that the same rolling window size would be appropriate over the full sample period. Whether one uses an expanding or a rolling window in estimation, the resultant forecasts will be based on a single estimation window, which need not be appropriate given that the choice of the estimation window (whether expanding or rolling) has been made in an *ad hoc* manner.

One possibility would be to extend the idea of pooling of forecasts obtained from different models (but based on the same given estimation window) to pooling of forecasts based on the same model but computed across alternative estimation windows. The rationale behind this approach is very similar – when unsure about the optimal window size use many different window sizes and then pool the results. This idea was suggested in Pesaran and Timmermann (2007) and has been recently shown by Pesaran and Pick (2007) to possess some optimality properties in forecasting the mean of a process. It is shown that the average forecast across different windows dominates (in the root mean squared error sense) forecasts from a single window when forecasting the mean of a given process subject to a break so long as the break point is not too close to the end of the sample. This is shown to be true irrespective of the size of the break.

In what follows we provide a formal Bayesian account that aims at integrating the uncertainties that prevail across models and across estimation windows.

2.3 Bayesian Model Averaging in the Presence of Model Instability

Model averaging and forecast combination have a rich history in statistics and forecasting. An early survey of the literature on forecast combination is provided by Bates and Granger (1969), with Timmermann (2006) providing a more recent survey. Here too there is a wide range of choices faced by the econometrician: what is the set of admissible models; what weighting scheme should be used to combine the forecasts from each model, and so on.

To fix ideas, suppose that we have available up to T observations of the variable of interest, $Z_{T,T} = (\mathbf{z}_1, \ldots \mathbf{z}_T)$, but that the estimation window is just of length w, $Z_{w,T} = (\mathbf{z}_{T-w+1}, \ldots \mathbf{z}_T)$. The future variables to be forecast are denoted $Z_{T+1,h} = (z_{T+1}, \ldots z_h)$. We can describe the forecasting problem as estimating the forecast probability density function, namely $\Pr(Z_{T+1,h}|Z_{w,T})$. To do so we need a model $\mathfrak{M}_m$ which in turn needs to be estimated over the estimation window of size w from the end of estimation sample at T, to obtain an estimate, $\widehat{\Pr}(Z_{T+1,h}|Z_{w,T}, \mathfrak{M}_m)$. In the face of model uncertainty we may want to pool over a total of, say, M models. Using Bayes rule we arrive at the familiar Bayesian Model Averaging expression:

$$\widehat{\Pr}(Z_{T+1,h}|Z_{w,T}) = \sum_{m=1}^{M} \widehat{\Pr}\left(\mathfrak{M}_m|Z_{w,T}\right) \widehat{\Pr}(Z_{T+1,h}|Z_{w,T}, \mathfrak{M}_m), \tag{1}$$

where $\widehat{\Pr}(Z_{T+1,h}|Z_{w,T}, \mathfrak{M}_m)$ is the predictive density of $Z_{T+1,h}$ conditional on model $\mathfrak{M}_m$ and $\widehat{\Pr}\left(\mathfrak{M}_m|Z_{w,T}\right)$ is the posterior probability of model $\mathfrak{M}_m$, both estimated over the observation window w.

If a particular model $\mathfrak{M}_m$ is stable over time, then obviously it would be desirable to use the longest sample window possible for estimation, namely $Z_{T,T}$ in our notation. In reality, however, this is unlikely to be the case, but unfortunately the Bayesian Model Averaging expression given by (1) implicitly makes the assumption of model stability.

In reality some or all of the models under consideration could be subject to structural breaks and different choices of estimation samples might be warranted. With this in mind, a more pragmatic approach would be to also average each model over different sampling windows, starting from a minimum window size to the largest permitted by the available data set. Allowing for both model and estimation window uncertainty yields

$$\widehat{\Pr}(Z_{T+1,h}|Z_{T,T}) = \sum_{m=1}^{M} \sum_{w=T}^{T-W+1} \widehat{\Pr}\left(\mathfrak{M}_m|Z_{w,T}\right) \widehat{\Pr}(Z_{T+1,h}|Z_{w,T}, \mathfrak{M}_m), \tag{2}$$

where $\widehat{\Pr}\left(\mathfrak{M}_m|Z_{w,T}\right)$ may be thought of as the weight attached to model $\mathfrak{M}_m, m = 1, \ldots M$, estimated over the sample window $w = T, T-1, \ldots, T-W+1$. The windows are arranged from the longest window of size T to the shortest window of size $T-W+1$. See Assenmacher-Wesche and Pesaran (2007) for an application of this approach to forecasting the Swiss economy.

Bayesian model averaging requires the specification of model weights, here the prior probability of model $\mathfrak{M}_m$ and a prior probability of the model's coefficients, collected in θ_m, conditional on $\mathfrak{M}_m$, for $m = 1, \ldots M$. When there is little certainty about which model is the right one, and if in addition the models are subject to structural breaks, the simplicity of equal weights is quite appealing. To be sure, this choice entails risks as one may consider some bad models that should perhaps have been better left out. It is worth noting, however, that in his Handbook survey, Timmermann (2006) reports that across many different empirical applications, the equal weighting scheme is tough to beat.

3 The GVAR Model

The GVAR is composed of individual country vector error correcting models in which the core domestic variables are related to country-specific foreign variables. The model covers 33 countries that account for about 90% of world output with the euro area considered as a single economy (eight economies are grouped into one). In total there are 26 country/region specific models that are linked within a unified GVAR framework including Europe, the Anglo-Saxon world, Latin America, South East Asia, China, Korea, India, Saudi Arabia, Turkey and South Africa. For a more detailed list of countries included in the GVAR model along with the trade weights used to construct the foreign (star) variables, see DdPS (2007). The individual country models are formulated and estimated over the period 1979Q2 - 2003Q4.

Most country specific models include the following core variables:

$$\left.\begin{array}{c} y_{it} = \ln\left(GDP_{it}/CPI_{it}\right), \quad p_{it} = \ln(CPI_{it}), \\ q_{it} = \ln(EQ_{it}/CPI_{it}), \quad e_{it} = \ln(E_{it}), \\ \rho_{it}^{S} = 0.25\ln(1+R_{it}^{S}/100), \quad \rho_{it}^{L} = 0.25\ln(1+R_{it}^{L}/100), \\ p_t^o = \ln(P_t^o) \end{array}\right\} \tag{3}$$

where

$$\begin{array}{rcl} GDP_{it} & = & \text{Nominal Gross Domestic Product of country } i \\ & & \text{during period } t, \text{ in domestic currency,} \\ CPI_{it} & = & \text{Consumer Price Index in country } i \text{ at time } t, \\ & & \text{equal to 1.0 in a base year (say 1995),} \\ EQ_{it} & = & \text{Nominal Equity Price Index,} \\ E_{it} & = & \text{Exchange rate of country } i \text{ at time } t \text{ in terms of US dollars,} \\ R_{it}^{S} & = & \text{Nominal short term rate of interest per annum, in percent,} \\ R_{it}^{L} & = & \text{Nominal long term rate of interest per annum, in percent,} \\ P_t^o & = & \text{Price of oil (in USD).} \end{array}$$

The domestic and foreign variables included in the country-specific models are summarised in the table below. Note that the endogeneity of oil prices reflects the large size of the US economy, while the inclusion of only three foreign variables in the US, as resulting from the weak exogeneity tests, reflects the importance of the US financial markets within the global financial system.

Table 1. Domestic and foreign variables included in the country-specific models

	All Countries Excluding US		US	
Variables	Endogenous	Foreign	Endogenous	Foreign
Real Output	y_{it}	y_{it}^*	$y_{us,t}$	$y_{us,t}^*$
Inflation	Δp_{it}	Δp_{it}^*	$\Delta p_{us,t}$	$\Delta p_{us,t}^*$
Real Exchange Rate	$e_{it} - p_{it}$	-	-	$e_{us,t}^* - p_{us,t}^*$
Real Equity Price	q_{it}	q_{it}^*	$q_{us,t}$	-
Short-Term Interest Rate	ρ_{it}^S	ρ_{it}^{*S}	$\rho_{us,t}^S$	-
Long-Term Interest Rate	ρ_{it}^L	ρ_{it}^{*L}	$\rho_{us,t}^L$	-
Oil Price	-	p_t^o	p_t^o	-

It is also worth mentioning that due to data availability, and the fact that not all countries have well developed capital markets, not all countries contain the same number of domestic variables. Table 2 below shows how the total number of 134 domestic variables in the world economy used in DdPS (2007) are distributed across each variable. That is, it summarises the number of countries, out of a total of 26 (recall that the 8-country Euro area is treated as a single country in the model), for which each variable is available.

Table 2. Country Composition of Endogeneous Variables in the GVAR model

Variables	# Countries	
Real Output	26	
Inflation	26	
Real Equity Price	19	Excluding: China, Brazil, Mexico, Indonesia, Turkey, Saudi Arabia, Peru
Real Exchange Rate	25	Excluding: US
Short-Term Interest Rate	25	Excluding: Saudi Arabia
Long-Term Interest Rate	12	Including: USA, Euro Area, Japan, UK, Canada, South Korea, Australia, South Africa, Norway, Sweden, Switzerland, New Zealand
Oil Prices	1	Included only in the US model as endogenous

3.1 Country-Specific VARX* Models

The variables given in Table 1 are modeled for each economy using a VARX* structure as described below. Suppose there are a set of $N+1$ countries indexed by $i = 0, 1, 2, ..., N$, with country 0, say the US, as the reference country. For country i, consider the VARX*$(2,1)$ specification

$$\mathbf{x}_{it} = \mathbf{h}_{i0} + \mathbf{h}_{i1}t + \mathbf{\Phi}_{i1}\mathbf{x}_{i,t-1} + \mathbf{\Phi}_{i2}\mathbf{x}_{i,t-2} + \mathbf{\Psi}_{i0}\mathbf{x}^*_{it} + \mathbf{\Psi}_{i1}\mathbf{x}^*_{i,t-1} + \mathbf{u}_{it}$$

where

$$\begin{aligned} \mathbf{x}_{it} &\quad : \quad k_i \times 1 \text{ vector of domestic variables} \\ \mathbf{x}^*_{it} &\quad : \quad k^*_i \times 1 \text{ vector of foreign variables} \end{aligned}$$

and $\mathbf{u}_{it}$ is a serially uncorrelated and cross sectionally weakly dependent process such that for each t and i, and the set of granular weights, w_{ij}, we have[2]

$$\bar{\mathbf{u}}_{it} = \sum_{j=0}^{N} w_{ij}\mathbf{u}_{jt} \xrightarrow{p} \mathbf{0}, \text{ as } N \to \infty.$$

The error correction form of the VARX*$(2,1)$ specification may be written as

$$\Delta\mathbf{x}_{it} = \mathbf{c}_{i0} - \boldsymbol{\alpha}_i\boldsymbol{\beta}'_i[\mathbf{z}_{i,t-1} - \boldsymbol{\gamma}_i(t-1)] + \mathbf{\Psi}_{i0}\Delta\mathbf{x}^*_{it} + \mathbf{\Gamma}_i\Delta\mathbf{z}_{i,t-1} + \mathbf{u}_{it},$$

where $\mathbf{z}_{it} = (\mathbf{x}'_{it}, \mathbf{x}^{*\prime}_{it})'$, $\boldsymbol{\alpha}_i$ is a $k_i \times r_i$ matrix of rank r_i and $\boldsymbol{\beta}_i$ is a $(k_i + k^*_i) \times r_i$ matrix of rank r_i. By partitioning $\boldsymbol{\beta}_i$ as $\boldsymbol{\beta}_i = (\boldsymbol{\beta}'_{ix}, \boldsymbol{\beta}'_{ix*})'$ conformable to $\mathbf{z}_{it}$, the r_i error correction terms defined by the above equation can be written as

$$\boldsymbol{\beta}'_i(\mathbf{z}_{it} - \boldsymbol{\gamma}_i t) = \boldsymbol{\beta}'_{ix}\mathbf{x}_{it} + \boldsymbol{\beta}'_{ix*}\mathbf{x}^*_{it} + (\boldsymbol{\beta}'_i\boldsymbol{\gamma}_i)\,t,$$

which clearly allows for the possibility of cointegration both within $\mathbf{x}_{it}$ and between $\mathbf{x}_{it}$ and $\mathbf{x}^*_{it}$ and consequently across $\mathbf{x}_{it}$ and $\mathbf{x}_{jt}$ for $i \neq j$.

Conditional on r_i cointegrating relations, the co-trending restrictions, $\boldsymbol{\beta}'_i\boldsymbol{\gamma}_i = \mathbf{0}$, and long-run restrictions on $\boldsymbol{\beta}_i$ can be tested. For estimation, $\mathbf{x}^*_{it}$ are treated as "long-run forcing" or $I(1)$ weakly exogenous with respect to the parameters of the conditional model, an assumption found acceptable when tested.[3] The VARX* model is estimated separately for each country conditional on $\mathbf{x}^*_{it}$, taking into account the possibility of cointegration both within $\mathbf{x}_{it}$ and across $\mathbf{x}_{it}$ and $\mathbf{x}^*_{it}$.

3.2 Solution and Properties of the GVAR model

Although estimation is done on a country by country basis, the GVAR model needs to be solved for all the endogenous variables of the global economy simultaneously. Let $\mathbf{x}_t = (\mathbf{x}'_{0t}, \mathbf{x}'_{1t}, ..., \mathbf{x}'_{Nt})'$

[2] For country i, the weigts w_{ij}, $j = 0, 1, ..., N$ with $w_{ii} = 0$ is granular if $\lim_{N\to\infty}\left(\sum_{j=0}^{N} w_{ij}^2\right) = 0$.

[3] Conditions under which $\mathbf{x}^*_{it}$ can be treated as weakly exogenous are discussed in Chudik and Pesaran (2007) in the context of an infinite dimensional VARs. It is shown that starting with a high dimensional global VAR it can be decomposed into country-specific VARX* models if there is a finite number of dominant countries and/or common factors.

be the $k \times 1$ global vector of endogenous variables with $k = \sum_{i=0}^{N} k_i$. The key to solving the model is to note that the link between $\mathbf{x}_t$ and the variables in the i^{th} country model, which can expressed in terms of $\mathbf{z}_{it} = (\mathbf{x}'_{it}, \mathbf{x}^{*\prime}_{it})'$, is given by the identity

$$\mathbf{z}_{it} = \mathbf{W}_i \mathbf{x}_t, \tag{4}$$

where $\mathbf{W}_i$ is a $(k_i + k_i^*) \times k$ 'link' matrix defined by the trade weights.

Using the identity (4) and stacking the $N + 1$ individual country models yields the Global VAR model obtained as

$$\mathbf{x}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{F}_1 \mathbf{x}_{t-1} + \mathbf{F}_2 \mathbf{x}_{t-2} + \boldsymbol{\varepsilon}_t, \tag{5}$$

where the coefficients of (5) embody the global interdependencies and are determined by the parameter matrices of the underlying country specific models. There are no restrictions on the covariance matrix $\boldsymbol{\Sigma} = \mathbf{E}(\boldsymbol{\varepsilon}_\mathbf{t} \boldsymbol{\varepsilon}'_\mathbf{t}) = (\boldsymbol{\Sigma}_{ij})$. For each country there is a $k_i \times 1$ vector of estimated residuals $\widehat{\mathbf{u}}_{it}$ from which can be calculated $\widehat{\boldsymbol{\varepsilon}}_{it}$, and hence $\widehat{\boldsymbol{\Sigma}}_{ij} = \sum_{t=1}^{T} \widehat{\boldsymbol{\varepsilon}}_{it} \widehat{\boldsymbol{\varepsilon}}'_{jt} / T$. For further details see PSW and DdPS.

The GVAR model entertained by DdPS (2007) has 134 endogenous variables, 71 stochastic trends and 63 long-run (cointegrating) relations. It is globally stable in that all its roots lie either on or inside the unit circle. Although log-linear with a simple overall structure, the GVAR is a large and complicated model which allows for a high degree of interdependence and dynamics. It has two routes for between country interdependence: through the impact of the $\mathbf{x}^*_{it}$ variables and through the error covariances. Shocks to one country can have marked effects on other countries, depending on their size and the patterns of their trade. DdPS find that the long run forcing assumption is rejected only in 5 out of 153 cases, while evidence of structural instability is found primarily in the error variances (47% of the equations – clustered in the period 1985-1992). Overall DdPS demonstrate that the GVAR model is quite effective in dealing with the common factor interdependencies and international co-movements of business cycles.

4 GVAR Models and Estimation Windows

Modeling a complex system as the global economy is naturally subject to considerable uncertainties. There are many choices to be made at the level of individual country models – the variables to be included in the country-specific models, the lag orders, the number of cointegrating, and whether to impose long and short run theory restrictions on the parameters, just to mention a few of the choices to be made. The number of possible GVAR models that could be considered as a result of such combinations of choices is enormous. Considering only the uncertainty regarding the number of cointegrating relations and fixing the lag orders p_i and q_i in the individual country VARX*(p_i, q_i) models, using the information provided in Table 3, we would end up with $6^{12} \times 5^7 \times 4^6 \times 3 \approx 2.1 \times 10^{15}$

number of GVAR models, a very large number indeed![4] Even if one fixes the number of cointegrating relations for each country to its estimated value, $\hat{r}_i$, allowing only for uncertainty with respect to p_i and q_i, with $p_{\max i} = q_{\max i} \leq 2$, this would amount to $2^{26} = 67,108,864$ possible GVAR models.[5] Allowing both for uncertainty with respect to p_i and q_i and the number of cointegrating relations would result in an even larger number of GVAR models that would be clearly infeasible to deal with in practice. In what follows we shall focus only on a limited number of GVAR type models in order to make the analysis feasible and to illustrate our approach.

Table 3. Number of Endogenous Variables Included per Country Models in GVAR

# Endogenous Variables (k_i)	List of Countries
6	USA, Euro Area, Japan, UK, Canada, Korea, Australia, South Africa, Norway, Sweden, Switzerland, New Zealand (12 countries)
5	Argentina, Chile, Malaysia, Philippines, Singapore, Thailand, India (7 countries)
4	China, Brazil, Mexico, Peru, Indonesia, Turkey (6 countries)
3	Saudi Arabia

While it is certainly desirable to consider a large number of models, one needs to be cautious about the models selected so as not to include too many that are a priori obvious not to perform well. The literature on forecasting is typically silent on this issue. Economic theory, if available, could provide some guidance as to a reasonable choice of models. In any case this is an issue that deserves considerable attention.

In constructing the model space we begin with the GVAR specification estimated in DdPS based on data ending in the last quarter of 2003. This seems a sensible starting point since the DdPS-GVAR specification was developed prior to the forecast evaluation period, 2004Q1-2005Q4. For the purpose of the forecasting exercise, the data used in DdPS is further extended from 2004Q1 to 2005Q4 along the lines described in the Appendix.

Other GVAR type models can now be developed from the DdPS-GVAR specification. Given the uncertainty regarding the true number of cointegrating relations, one possibility would be to

[4] Note that in the case of a country model with k_i endogenous variables we could have k_i different models, one with 0 number of long run relations, another model with 1, a third model with 2 long run relations etc. Therefore, for a global economy composed of $N+1$ countries each with k_i endogenous variables we would have $\Pi_{i=0}^{N} k_i$ different GVAR models.

[5] This restriction on the maximum lag orders is considered in DdPS (2007) given the limited data availability.

set the number of cointegrating relations for all country specific models to zero, and thus consider a GVAR model in first differences, to be denoted as DdPS-DGVAR model, without changing the lag orders of the individual country models. For this model, we can then allow for uncertainty with respect to the true lag orders of the country-specific models by considering all possible combinations of lag orders for the DGVAR model with $p_{\max}, q_{\max}$ not exceeding 1, given the limited availability of data. This yields the additional models, DdPS-DGVAR(p_i, q_i), for p_i, $q_i = 0$ and 1.

Additional GVAR models can be specified by dividing the countries into two groups, as shown in Table 4 below, with Group A consisting of 10 industrialised countries plus China, with the remaining 15 countries placed in Group B. For Group A, the more developed economies, we set the lag orders and number of cointegrating relations to those of the DdPS-GVAR model, while for the remaining less developed economies, we impose zero cointegrating restrictions reflecting our greater uncertainty regarding the true number of long run relations for these countries. For Group B we also allow for uncertainty with respect to the lag order of the individual country/regions as above. We denote these models by DdPS-GVAR$_{ab}(p_i,q_i)$, for p_i, $q_i = 0$ and 1.

Table 4. Country Groups

Group A 10 Industrialised Countries Plus China	Group B Remaining 15 Countries	
US	India	Malaysia
Euro Area[6]	Brazil	Chile
China	Mexico	Peru
Japan	Korea	Singapore
UK	Indonesia	
Canada	South Africa	
Australia	Argentina	
Sweden	Turkey	
Switzerland	Thailand	
Norway	Philippines	
New Zealand	Saudi Arabia	

Another class of GVAR models can be developed from the long-run restricted specification in DHPS. The DHPS-GVAR model incorporates long-run structural relationships, suggested by economic theory, in an otherwise unrestricted GVAR model. DHPS show how the GVAR model needs to be modified in order for long-run relations such as Purchasing Power Parity (PPP) to be imposed on the country specific models, which include the effective exchange rate amongst the domestic variables rather than the real exchange rate as in the GVAR. The long term properties of

[6]Euro Area includes Austria, Belgium, Finland, France, Germany, Italy, Netherlands and Spain.

this model are based on market arbitrage and stock-flow equilibrium conditions, while the short run dynamics are left unconstrained. Using this specification we then spin off other long-run restricted models, yielding overall 19 GVAR models which we use in the forecasting exercise.

4.1 Choice of Observation Windows

The next issue to consider is the choice of the window size and the frequency of window updates. These choices are to some extent restricted by the availability of data. For this reason we select ten quarterly estimation windows, with the first window, $W1$ starting in 1979Q1 ending up with window $W10$ that starts in 1981Q1. We further experimented by increasing the space of models as well as selecting ten bi-quarterly estimation windows beginning from 1979Q1, and the results were qualitatively similar. The space of models and estimation windows considered are set out in Table 5 below.

Table 5. Space of GVAR Models and Estimation Windows

Space of Models (19)	
DdPS-GVAR	DHPS-GVAR
DdPS-DGVAR	-
DdPS-DGVAR(0,0)	DHPS-DGVAR(0,0)
DdPS-DGVAR(0,1)	DHPS-DGVAR(0,1)
DdPS-DGVAR(1,0)	DHPS-DGVAR(1,0)
DdPS-DGVAR(1,1)	DHPS-DGVAR(1,1)
DdPS-GVAR$_{ab}$(0,0)	DHPS-GVAR$_{ab}$(0,0)
DdPS-GVAR$_{ab}$(0,1)	DHPS-GVAR$_{ab}$(0,1)
DdPS-GVAR$_{ab}$(1,0)	DHPS-GVAR$_{ab}$(1,0)
DdPS-GVAR$_{ab}$(1,1)	DHPS-GVAR$_{ab}$(1,1)
Beginning of Estimation Windows (10)	
W1:1979Q1, W2:1979Q2, W3:1979Q3, W4:1979Q4, W5:1980Q1, W6:1980Q2, W7:1980Q3, W8: 1980Q4, W9:1981Q1, W10:1981Q2	

Note: The DHPS-DGVAR model is excluded from the above table since it coincides with the DHPS-GVAR$_{ab}$(1,0) model, given that the specification across all underlying individual country models in the case of DHPS-GVAR is a VARX*$(2,1)$. See DHPS for further details.

4.2 Trade Weights

For in-sample estimation over the period 1979Q1-2003Q4 we use fixed trade weights averaged over the three year window 1999-2001. For out-of-sample recursive forecasting we use the trade weighted average over 2001-2003 to compute 2004 forecasts and the trade weighted average over 2002-2004 to compute 2005 forecasts. All country specific models were estimated for the case of an unrestricted intercept and no trend. Only 6 out of 26 countries rejected the null of co-trending, namely China, Japan, Argentina, New Zealand, India and Turkey, at the 1% significance level.

5 Forecast Evaluation: Methodological Considerations

Before presenting the forecast results we first consider a number of standard benchmarks used in the forecast evaluation literature. We also develop a panel version of the Diebold and Mariano, DM, (1995) test which allows us to statistically test the GVAR forecasts against each of the benchmarks for a given variable across different country groupings. Note that we only have eight one-quarter ahead forecasts (obtained over the period 2004-2005) for each of the variables per country, which is not sufficient for statistical testing. However, by pooling forecast errors for the same variable across different countries, we are able to carry out the panel DM test so long as it is appropriately adapted to take account of the panel nature of the pooled series.

5.1 Benchmark Models

We compare the forecast performance of the GVAR model to forecasts from random walk and AR(1) models, with and without drifts. The specifications of the four benchmark models are

$$\begin{aligned} \text{Random walk } (RW) &: \quad y_t = y_{t-1} + \varepsilon_t, \\ \text{Forecast} &: \quad y_{t+h} = y_t. \end{aligned}$$

$$\begin{aligned} \text{Random walk } (RW) \text{ plus drift } \mu &: \quad y_t = \mu + y_{t-1} + \varepsilon_t, \\ \text{Forecast} &: \quad y_{t+h|t} = h\hat{\mu} + y_t, \\ \hat{\mu} \text{ is obtained by estimation of} &: \quad \Delta y_t = \hat{\mu} + \hat{\varepsilon}_t. \end{aligned}$$

$$\begin{aligned} AR(1) &: \quad y_t = a + \gamma y_{t-1} + \varepsilon_t, \\ \text{Forecast} &: \quad y_{t+h|t} = \hat{a} + \hat{\gamma} y_{t+h-1|t}. \end{aligned}$$

$$\begin{aligned} AR(1) \text{ plus trend} &: \quad y_t = a + \beta t + \gamma y_{t-1} + \varepsilon_t, \\ \text{Forecast} &: \quad y_{t+h|t} = \hat{a} + \hat{\beta} t + \hat{\gamma} y_{t+h-1|t}. \end{aligned}$$

The drift parameter, μ, and the parameters of the AR models, α, β, and γ, can be estimated recursively using the full estimation window starting in 1979Q1, shorter windows, or averages across different windows. Following the standard in the literature, in what follows the parameters of the benchmark models are estimated using an expanding window, if applicable. Clearly, the issue of parameter update does not arise for the RW model.

Although admittedly simple, the endurance of the above models as benchmarks in the empirical macro and finance literature, aside from theoretical motivations (e.g. market efficiency), stems from the simple fact that they have been surprisingly hard to beat.

5.2 Pooling GVAR forecasts

In Section 4 we presented 19 different models within the GVAR family estimated over 10 different sample windows. Recall that there are two sets of GVAR models: five based on DHPS impose overidentifying restrictions, and six based on DdPS that do not. The remaining specifications are simply variants of these models. These models are summarised in Table 5.

The estimation sample spans 1979Q1 to 2003Q4 for a total of 100 quarters (or 25 years). This is the same sample used to estimate the GVAR model presented in DdPS; indeed we use precisely that fitted model here for our forecast evaluation to avoid being subject to data snooping. Our new data sample goes through 2005Q4, which gives us 8 quarters for out-of-sample forecast evaluation. For a given model, each out-of-sample quarter is forecast with the maximum amount of data available. Specifically, 2004Q1 is the out-of-sample forecast with data from 1979Q1 to 2003Q4. Next, 2004Q2 is forecast by adding the realised first quarter (2004Q1), and so on.

To allow adaptation to structural breaks the estimation window is changed. Specifically, the start date of the estimation sample is moved forward by one quarter, and the process of out-of-sample forecasting is repeated as before. This denotes a new sample or estimation window. A total of 10 such samples from the longest to the shortest, in one quarter increments, are used so that the last estimation start date is at 1981Q3; see Table 5 above. This estimation process is repeated for each of the 19 models. Thus for each out-of-sample forecast period, say 2004Q1, there are 10 windows and 19 models yielding a total of 190 forecasts for each variable to be pooled or averaged. We denote AveM the average forecast over models for a particular estimation window, AveW the average forecasts from a particular model estimated over different estimation windows, and AveAve the average forecast over both models and estimation windows.

We allow for averaging across sample windows for the benchmark models as well, although the RW model is invariant to the estimation window as no parameters per se are estimated. To be sure, when a drift is added to the RW model, that drift estimate could change as the estimation window changes. In the tables and figures below we present only the AveW results for the benchmarks. Results for the benchmark forecasts based on other windows are very similar and are available from the authors on request.

5.3 A Panel Version of the Diebold and Mariano Test of Forecasting Skill

Before proceeding to the results we need a way of determining whether the forecasts from GVAR can be statistically distinguished from a benchmark forecast at conventional significance levels. To begin, for any given model we are interested in computing the root mean-squared forecast error (RMSFE) of a given model or set of models, as in

$$RMSFE(h,n) = 100\sqrt{n^{-1}\sum_{t=T}^{T+n-1} e_t^2(h)},\ h = 1,2,3,4,$$

where $e_t(h) = (y_{t+h} - \hat{y}_{t+h|t})/h$ is the h-quarter ahead forecast error, with y_{t+h} being the actual value and $\hat{y}_{t+h|t}$ the corresponding forecast. The forecast horizon is denoted by h, and n is the size of the forecast sample. In our analysis we consider up to $h = 4$ (four-quarters ahead). We report results for one-quarter ahead ($h = 1$) and one year ahead ($h = 4$), but for statistical testing we focus on the one-quarter ahead forecasts which yields $n = 8$ (2004Q1-2005Q4).

Consider now the loss differential of forecasting the variable j in country i, using method A relative to method B :

$$\begin{aligned} z_{ijt} &= \left[e_{ijt}^A(1)\right]^2 - \left[e_{ijt}^B(1)\right]^2, \\ A &\equiv AveAve \text{ Forecast, } B \equiv \text{Benchmark Forecast,} \end{aligned}$$

for $i = 1,2,...,m$; $j = 1,2,...,k; t = 1,2,...,n$; where $e_{ijt}(1)$ is the one-quarter ahead forecast error, m is the number of countries, k is the number of variables, and n is the forecast sample.[7] The panel DM statistic is developed as follows: for a given variable j (say real output growth), consider

$$\begin{aligned} z_{it} &= \alpha_i + \varepsilon_{it} \\ H_0 &: \alpha_i = 0 \\ H_1 &: \alpha_i < 0 \text{ for some } i, \end{aligned}$$

suppressing the variable index j for simplicity. Under the null, and assuming that $\varepsilon_{it} \sim iid(0, \sigma_i^2)$,

$$\overline{DM} = \frac{\bar{z}}{\sqrt{V(\bar{z})}} \sim N(0,1)$$

where

$$\bar{z} = m^{-1}\sum_{i=1}^{m} z_{it}, \quad \bar{z}_i = \frac{1}{n}\sum_{i=1}^{n} z_{it}$$

and

$$V(\bar{z}) = \left(\frac{1}{mn}\right)\left(m^{-1}\sum_{i=1}^{m} s_i^2\right),\ s_i^2 = \frac{\sum_{i=1}^{n}(z_{it} - \bar{z}_i)^2}{n-1}.$$

[7] z_{ijt} may also be expressed in terms of absolute rather than squared loss.

For one-quarter ahead forecasts no adjustment for serial correlation is needed, since it is reasonable to assume that the loss differentials are serially uncorrelated. The same, of course, will not be true of forecast comparisons that involve forecasts of two or more quarters ahead. To handle such cases, the panel DM statistic can be readily modified to deal with the serial correlation of h-quarter ahead forecasts by using a Newey-West type estimator of $Var(\bar{z}_i)$. We do not pursue this extension here since for $h > 1$ we do not have sufficient data to reliably carry out the panel DM tests.

6 Forecast Evaluation Results

Given how many models, sample windows and combinations thereof are considered in our forecast evaluation exercise, one is hard pressed to present the results in simple summary form. We begin the discussion with a series of figures, tables and a chart. Our aim is to shed light on three main issues of particular interest. First to evaluate the performance of forecast averaging strategies in the GVAR context. Second, to see if the AveAve forecasts from the GVAR can beat the forecasts from the benchmark models. Finally, to assess the extent to which using financial variables such as long term interest rates and real equity prices are likely to be helpful in forecasting real output and inflation in the world economy.

6.1 Performance of AveAve Forecasts in the GVAR Context

Figures 1 to 6 are intended to address the first question and display RMSFE (in per cent) for the six core variables of the GVAR model in the case of seven main industrialised economies plus China (parts a and b of the figures), and the average of RMSFE across all the 10 industrialised countries plus China (part c of the figures). Similar figures are also available for the remaining countries. In these figures the horizontal axis shows the 10 windows from longest to shortest. The circles on the vertical lines associated with a particular window shows the RMSFE of the 19 models estimated using that window. The AveAve forecast is given by the horizontal line, and dominates all other forecasts when all the circles lie above this line. This does happen, for example, in the case of forecasting US output, but it is rather rare.

The AveAve forecasts seem to do particularly well in the case of output, inflation and the short term interest rate. This can be seen clearly from Figures 1c, 2c and 3c where the average RMSFEs across the 10 industrialised economies plus China are shown. In the case of all these three variables only a few of models/windows do better than the AveAve forecasts across these 11 countries. Also the models and windows that do better are not the same across the countries. The results for other variables are less clear cut but overall favour the use of the AveAve strategy. It is also worth noting that the AveAve forecasts do better than the GVAR models specified in DdPS(2007) and DHPS(2007), whether estimated over the longest window or their forecasts averaged across the 10 windows. The latter are denoted by DdPS-GVAR-AveW and DHPS-GVAR-AveW and their RMSFEs (for one quarter ahead and four quarters ahead) are summarised in Tables 6a-11a for the six core variables under three different country groupings, namely the ten industrialised countries plus China ('10+China'), all countries excluding Latin America ('All Excluding LA'), and all countries/region ('All Countries'). The RMSFEs of AveAve one quarter ahead forecasts lie below those of DdPS-GVAR-AveW and DHPS-GVAR-AveW in the case of all variables and all country groupings. The same holds for the four quarter ahead forecasts except for inflation in the case of 10+China country group (Table 7a), and real exchange rate for 10+China and all countries excluding LA.

Finally, the above results continue to hold when the AveAve forecasts are compared to the AveM forecasts (averages across models for a given window). These results are available from the authors upon request.

6.2 Performance of AveAve Forecasts Relative to the Benchmarks

Turning to the second question concerning the performance of the AveAve forecasts relative to the benchmark, we have also summarised the RMSFEs of all the four benchmarks for all the six variables averaged across the same three country groupings. Recall that the selected benchmarks are RW (random walk), RW with drift, a univariate AR(1) model with and without a drift.[8] In addition to the RMSFEs for one and four quarter ahead forecasts, in part b of Tables 6-11 we also provide the panel DM statistics for the one-quarter ahead AveAve forecasts relative to the four benchmark.

But before proceeding to some of the details, a summary taken from part b of Tables 6–11, is displayed in Figure 7. For each of the three country groupings we show the proportion of forecasts where the AveAve model beats the benchmarks at the 95% confidence level or better.[9] Since there are four benchmarks, the best the GVAR-AveAve forecasts can do is 4 out of 4, with the overall performance index set at 100%. If the GVAR-AveAve forecasts were beaten by all the benchmarks the performance index would take the value of -100%. Nothing will be recorded in the figure if the differences between the AveAve and the benchmark forecasts are not statistically significant. For each variable, when the AveAve model does better (in the panel DM sense), the bars are in the positive region; when a benchmark competitor does better, the bars are in the negative region. Of course it is possible that the AveAve can beat some models and lose to others.

Beginning with the first set of bars on the left of Figure 7 (and Table 6b), we first note that none of the benchmark forecasts do better than the GVAR-AveAve in forecasting output growth, and

[8] As noted earlier the parameters of the benchmark models have been estimated using an expanding window, when applicable. We also tried an AveW version of the benchmark models and overall obtained very similar results.

[9] Note that the alternative of forecast superiority is one sided and hence the appropriate critical value at the 5% significance level is -1.675, for testing the loss of GVAR-AveAve forecasts relative to the benchmark forecasts, and +1.675 for testing the alternative that the GVAR-AveAve forecasts are worse than the benchmarks.

based on the panel DM test the GVAR-AveAve forecasts significantly beat half of the benchmark models for 10+China country group, and beat all the benchmark models when we consider all countries. When Latin America is excluded from the panel, the AveAve output growth forecasts significantly beat three of the four benchmark forecasts.

Moving on to the next variable, inflation, the AveAve performance is as good overall and significantly better in the case of the industrialised economies, beating three out of four benchmarks for the 10+China country group (See Table 7b). Similar results are obtained for All Excluding LA country group. In this case the GVAR-AveAve inflation forecasts significantly beat two of the benchmark forecasts without being beaten by any of the benchmarks. These findings are particularly interesting considering a stream of research documenting the difficulty of beating simple models like the random walk in forecasting inflation, at least for the US; see inter alia Atkeson and Ohanian (2001), and Stock and Watson (2007).

The situation is, however, mixed when the Latin American countries are included in the comparison. The AveAve forecasts continue to do better than two of the benchmarks, but are significantly beaten by the two random walk benchmarks. Several of Latin American countries experienced a period of hyperinflation during our estimation sample, and so perhaps it is not surprising that the GVAR cannot forecast inflation in these countries over the 2003-205 period with a more normal inflationary experience. Unlike the random walk models that adjusts to new inflationary circumstances very quickly, the GVAR adapts more slowly and cannot cope when the change is too large and too abrupt.

Interest rates turn out to be harder to forecast. Indeed the random walk benchmark does better than the AveAve forecasts for all country groupings (Table 8b). Taking first the short term rate of interest, for the panel of industrialised countries plus China, the AveAve forecast does worse than the benchmarks, but when we look at the All Excluding LA group, the AveAve forecast does better than half of the benchmark models (it beats the AR(1) models). When all the countries are included, the GVAR-AveAve forecast only beats the AR(1) benchmark, while the three other benchmarks all do statistically better.

Considering the forecasts of long term interest rates, first recall that long term interest rates are included only in 12 out of the 26 country-specific models, namely the 10 industrialised countries/region plus South Korea and South Africa. Moreover, the RW model always beats the AveAve model, though this is statistically significant in the case of 10+China group. On average across all region groupings the AveAve does better than the benchmark models about 8% of the time, while 17% of the time a benchmark outperforms the AveAve forecasts of long rate.

Not surprisingly it is much harder to accurately forecast real equity prices and exchange rates as compared to forecasting output, inflation and interest rates. This is clearly seen in the large magnitudes of RMSFEs reported for these variables in Tables 10a and 11a. Nevertheless, it is interesting that the AveAve-GVAR forecasts of real equity prices perform significantly better than some of the benchmark forecasts, and are not beaten significantly by any of the four benchmarks including the RW ones. See Table 10 and Figure 7. The same is not true of the AveAve forecasts of real FX, which are generally worse than the benchmark forecasts, but not by much and not significantly either. Only the RW with a drift manages to statistically beat the GVAR-AveAve forecasts of real FX.

We also calculated RMSFEs where the errors were weighted by country PPP-GDP. The results turn out to be qualitatively similar to the equal-weighted results described above and are available upon request.

6.3 The Relevance of Capital Markets in Forecasting Output, Inflation and the Short-Term Interest Rate

In what follows we seek to assess the role that capital markets play in forecasting output, inflation and the short-term interest rate. To this end, we carry out the same forecasting exercise as described above, entertaining the following two modified sets of DdPS-GVAR and DHPS-GVAR models. The first set consists of the GVAR models with the equity variable dropped from all country specific models, while a second set excludes both the real equity and long-term interest rate variables from all the country-specific models. In carrying out this exercise we further exclude the 5 Latin American countries from the GVAR models in order to avoid any predominant effects related the distinctive behaviour of these economies over the period under investigation that could potentially overshadow the aim of our exercise. Thus, for all results relating to this exercise the DdPS-GVAR and DHPS-GVAR models and their variants comprise of 21 country/regions.

The space of models and selection of estimation windows are the same as in the previous experiments. Similarly, in obtaining the quarterly forecasts for 2004 and 2005, the individual country models are estimated for the case of an unrestricted intercept and no trend, following results of the co-trending tests, with the trade-weights adjusted as described above. However, dropping countries or variables from the GVAR gives rise to a new model, which means that the lag orders and number of cointegrating relations for the individual country models need to be re-estimated. We begin by focussing on the GVAR models that exclude the 5 Latin American countries but include all variables as our "benchmark" models. For these models, the lag orders of the corresponding individual country models are selected by using the Akaike Criterion with $pmax_i = 2$ and $q\max_i = 1$. The rank of the cointegrating space for each country/region is computed using Johansen's trace and maximal eigenvalue statistics as set out in Pesaran, Shin and Smith (2000) for models with weakly exogenous I(1) regressors, in the case where unrestricted constants and restricted trend coefficients are included in the individual country error correction models, at the 5% significance level.[10]

[10] The critical values used are those reported in MacKinnon, Haug and Michelis (1999).

The number of cointegrating relations is subsequently adjusted by inspection of their persistence profiles, which are calculated based on the solution of the GVAR model, and the eigenvalues of the system. That is, if the persistence profiles do not converge to zero for any number of countries, the number of cointegrating relations for those countries are reduced by one until all persistence profiles of the GVAR long run relations converge to zero, and none of its eigenvalues are outside the unit circle. The same strategy is followed when selecting the lag orders and rank of the cointegrating space for the individual country/regions models of the GVAR without equity, and without equity and long-term interest rates.

The forecasting results using the AveAve approach are presented in Tables 12a - 12d. Tables 12a - 12c show the 1-quarter and 4-quarters ahead RMSFE by country for the three versions of the GVAR model for output, inflation and the short rate respectively: first including both equity prices and the long rate, then dropping equity prices, and finally also dropping the long rate. At the bottom of each table we also present results for the 11 country (10 industrialised plus China) and All Countries groupings.[11] Table 12d shows the panel DM statistics for the 1-quarter ahead forecasts to test whether the apparent differences between the three specifications are in fact statistically distinguishable.

The results for all three variables do not show overwhelming evidence that including the financial variables (equity prices and the long rate) contributes substantively to the forecast accuracy of output, inflation or the short rate. In fact, the results of the panel DM statistics in Table 12d show that none of the differences for the 1-quarter ahead forecast in Tables 12a - 12c are statistically significant. Taking, for instance, output for the 11 countries group, the RMSFE is unchanged when dropping equity prices at 0.517, and it rises marginally to 0.522 when dropping the long rate (Table 12a). When considering all 21 countries the results are quite similar. RMSFE increases from 0.766 to 0.771 when dropping equity prices but increases no further when also dropping the long rate. This basic pattern carries over to the 4-quarters ahead RMSFEs.

Clearly, these results are rather disappointing particularly as far as the relevance of long rates for forecasting output is concerned. It is generally believed that the term premium, measured as the spread of the long term over the short term rate, are helpful in forecasting output growth. However, this evidence is predominantly reported for the US, and to our knowledge there are no systematic studies of this issue in the case of other economies.[12] In fact if we consider country-by-country results given in Table 12a, we see that we do get larger RMSFEs for output growth in the case of US (by 11%), Canada (by 17%), Australia (by 13%), Sweden (by 0.7%), Switzerland (by

[11] Note that for these tables where the Latin American countries are dropped from the GVAR, the groupings "All Countries Ex. LA" and "All Countries" coincide, and are replaced by the grouping "All Countries".

[12] Most of the evidence on the term spread and output growth refer to the US and recently have been subject to a number of re-examinations. See, for example, Hamilton and Kim (2002), Favero, Kaminska, and Söderström (2005), and Rudebusch, Sack and Swanson (2007).

12%) and New Zealand (by 7%) when the long term rate and real equity prices are excluded from the model. However, the results are reversed in the case of other countries with a balanced overall outcome for the industrialised countries plus China, namely an average RMSFE of 0.522 when long term interest rate and real equity prices are excluded from the GVAR as compared to 0.552 when they are included – only a very modest overall improvement which turns out not to be statistically significant.

Similar mixed results also hold for forecasting inflation.

7 Conclusions

The GVAR model estimated over the 1979Q1-2003Q4 period was used to generate out-of-sample one quarter and four quarters ahead forecasts of real output, inflation and a number of financial variables across 26 countries over the period 2004Q1-2005Q4. The forecasts were compared with univariate autoregressive and random walk models. The effects of model and estimation uncertainty on forecast outcomes was examined by pooling of forecasts obtained from different GVAR models estimated over alternative estimation periods. The AveAve forecasts based on GVAR models seem to perform well in a number of dimensions.

8 Appendix

A.9 Data Sources

A.1. Real GDP

IFS data is used for all countries except for Singapore for which Datatstream data is used. For cases where the IFS data was either too volatile relative to the DdPS data or not available, the DdPS data was used and we extrapolated forward using the growth rate of the latest IFS data. This was the case for the following countries: Brazil DdPS data (1990Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Germany DdPS data (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Indonesia (1983Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Italy (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Malaysia (1988Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Netherlands (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), New Zealand (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Peru (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4), Spain (1979Q1-2003Q4) extrapolated forward using the growth rate of IFS data (2003Q4-2005Q4). For Belgium DdPS data (1980Q1-2003Q4) extrapolated forward using the growth rate of OECD data (2003Q4-2005Q4). For Switzerland we use the data as described in the Appendix of Assenmacher-Wesche and Pesaran (2007). For the rest of the countries not mentioned above the latest IFS BVPZF GDP series in 2000 constant prices or the B.PVF.volume series was used.

Seasonal adjustment was performed for the following countries: Argentina, Austria, Brazil, Chile, Finland, India, Indonesia, Korea, Malaysia, Mexico, Norway, Peru, Philippines, Sweden, Thailand, Turkey.

Interpolation from annual figures was conducted for the following countries using the procedure described in Supplement A of Dees, di Mauro, Pesaran and Smith (2007): Argentina (1979Q1-1989Q4), Belgium (1979Q1-1979Q4), Brazil (1979Q1-1989Q4), Chile (1979Q1-1979Q4), China (1979Q1-2005Q4), India (1979Q1-1996Q3), Indonesia (1979Q1-1982Q4), Malaysia (1979Q1-1987Q4), Mexico (1979Q1-1979Q4), Philippines (1979Q1-1980Q4), Saudi Arabia (1979Q1-2005Q4) and Thailand (1979Q1-1986Q4).

A.2. Consumer Price Indices

The IFS CPI 64zf series is used for all countries except for: Brazil, IFS 64zf data was available for the period 1980Q1-2005Q4 with the average growth rate of prices for 1980 used to backfill to 1979Q1, China (IFS 64 xzf), Germany Datastream data[13], and Switzerland, the CPI data as described in the Appendix of Assenmacher-Wesche and Pesaran (2007) is used.

[13] The CPI code for datastream is based on west germany data only pre-unification (code: BDCONPRCX) and it is exactly the same as the IFS CPI 64zf series post-unification.

Seasonal adjustment was performed for the following series: Austria, Finland, Germany, India, Japan, Korea, Malaysia, Netherlands, Norway, Turkey and the UK.

A.3. Equity Price Indices

For Argentina DdPS Datatstream data (1988Q1-2001Q4) is used which is based on quarterly averages of weekly data points as opposed to daily observations, while for 2002Q1-2006Q4 we extrapolated forward using IFS Industrial Share Index 62zf. Datastream Total Market Index (TOTMK) is used for Chile (1989Q3-2005Q4), Finland (1988Q1-2005Q4), India (1990Q1-2005Q4), Korea (1987Q3-2005Q4), Malaysia (1986Q1-2005Q4), New Zealand (1988Q1-2005Q4), Norway (1980Q1-2005Q4), Philippines (1987Q3-2005Q4), Spain (1987Q1-2005Q4), Sweden (1982Q1-2005Q4), Thailand (1987Q1-2005Q4) with the growth rate of IFS Industrial Share Index 62zf series used to backfill, except for Malaysia where Bloomberg data of DdPS is used to backfill. For the rest of the countries, where an equity price series is available, Datastream Total Market Index (TOTMK) is used throughtout (1979Q1-2005Q4).

The Datastream Total Market Index (TOTMKT) calculation includes the most important companies based on market value. The precise number of constituents varies from market to market, according to the size of the market capitalisation, and changes to reflect current market conditions. The quarterly averages were calculated initially extracting the last Wednesday of each month from Datastream daily values. Quarterly averages were then computed by averaging the last Wednesday of each month within a quarter.

A.4. Exchange Rates

The GTIS US $ exchange rate is used for Brazil (1994Q1-2005Q4), Chile (1994Q1-2005Q4), Peru (1991Q1-2005Q4) and for the rest of the countries (1986Q1-2005Q4) and backfilled with the growth rate of the IFS rf series.

A.5. Short-Term Interest Rates

IFS is used as the main source for short term interest rates. The IFS Deposit Rate 60Lzf series is used for Argentina, Chile, China and Turkey. The IFS Discount Rate 60zf series is used for New Zealand and Peru. The IFS Treasury Bill Rate IFS 60Czf series is used for Canada, Malaysia, Mexico, Philippines, South Africa, Sweden, UK and US. The IFS Money Market Rate 60Bzf series is used throughout the whole sample period for Australia, Brazil, Finland, Germany, Inodnesia, Italy, Japan, Korea, Norway, Singapore, Spain, Switzerland, Thailand and throughout 1979Q1-1998Q4 for Austria, Belgium, France, Netherlands. For the latter group, the IFS Money Market Rate 60Bzf series for Germany was used 1999Q1-2005Q4. For India the average of Datastream's 90-180 day Bank Deposit Middle Rate (1991Q1-2005Q4), 91 Day T-Bill Primary Middle Rate (1997Q2-2005Q4), 91 Day T-Bill Secondary Middle Rate (1993Q1-2005Q4) and IFS Money Market Rate 60Bzfseries (1979Q1-1998Q1) is used.

A.6. Long-Term Interest Rates

The IFS Government Bond Yield 61zf series is used for all 18 countries for which long term interest rate data is available, namely Australia, Belgium, Canada, France, Germany, Italy, Japan, Korea, Netherlands, New Zealand, Norway, South Africa, Spain, Sweden, Switzerland, UK and USA. For Austria, the IFS 61zf is used for the period 1979Q1-2000Q3 and the series is completed with the OECD 10 Year Federal Government Benchmark bond series (AUT.IRLTLT01.ST.).

A.7. Oil Price Index

Monthly averages of the Brent Crude series from Datastream

Note that for real GDP when DdPS IFS data is used and interpolation is required this is done on the annual figures provided in the code corresponding to the DdPS paper. Furthermore, the latest IFS data refers to the updated real GDP data collected from the IFS at the end of 2006.

A.9.1 Assessing the Joint Significance of Seasonal Components

To asses the joint significance of the seasonal components for real output and the price level we used the following procedure:

1. Let S_1, S_2, S_3 and S_4 be the usual seasonal dummies, such that S_i, $i = 1, 2, 3, 4$, takes the value of 1 in the i^{th} quarter and zero in the other three quarters.

2. Construct $S_{14} = S_1 - S_4$, $S_{24} = S_2 - S_4$, $S_{34} = S_3 - S_4$.

3. Run a regression of DY (DP) on an intercept and S_{14}, S_{24}, S_{34}. Denote the OLS estimates of S_{14}, S_{24} and S_{34} by a_1, a_2 and a_3.

4. Asses the joint significance of the seasonal components by testing the hypothesis that $a_1 = a_2 = a_3 = 0$ using the F-statistic.

In summary, 16 out of 26 countries were seasonally adjusted for real output and 11 out of 26 for inflation. For cases where the seasonal components were found significant, seasonal adjustment was performed on the log of the corresponding variable in level, that is, Y (P) using the X11 procedure in Eviews under the additive option.

References

[1] Alogoskoufis, G. and R. Smith (1991). On error correction models: specification, interpretation, estimation. *Journal of Economic Surveys* 5, 97–128.

[2] Assenmacher-Wesche, K. and M.H. Pesaran (2007). Assessing forecast uncertainties in a VECX model for Switzerland: an exercise in forecast combination across models and observation windows. IZA discussion paper 3071.

[3] Atkeson, A. and L.E. Ohanian (2001). Are Phillips curves useful for forecasting inflation? Federal Reserve Bank of Minneapolis, *Quarterly Review* 25, 2-11.

[4] Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47-78.

[5] Bates, J. M. and C. W. J. Granger (1969). The combination of forecasts. *Operational Research Quarterly* 20, 451-468.

[6] Chib, S. (1998). Estimation and comparison of multiple change-point models. *Journal of Econometrics* 86, 221–241.

[7] Chudik, A. and M. H. Pesaran (2007). Infinite dimensional VARs and factor models. Unpublished manuscript, Cambridge University.

[8] Clements, M.P. and D.F. Hendry (1998). Forecasting economic time series. Cambridge University Press, Cambridge, UK.

[9] Clements, M.P. and D.F. Hendry (2006). Forecasting with structural breaks. In: G. Elliot, C. W. J. Granger, and A. Timmermann (eds), *Handbook of Economic Forecasting*, vol 1., Elsevier, Amsterdam, 605- 658.

[10] Dees, S., F. di Mauro, Pesaran, M.H. and L.V. Smith (2006). Exploring the international linkages of the euro area: a global VAR analysis. *Journal of Applied Econometrics* 22, 1-38.

[11] Dees, S., S. Holly, Pesaran, M.H. and L.V. Smith (2006). Long run macroeconomic relations in the global economy. *Economics - The Open-Access, Open-Assessment E-Journal*, 2007-3.

[12] Diebold, F.X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253-265.

[13] Estrella, A. and F.S. Mishkin (1998). Predicting U.S. recessions: financial variables as leading indicators. *Review of Economics and Statistics* 80, 45-61.

[14] Favero, C. A., I. Kaminska, and U. Söderström (2005). The Predictive Power of the Yield Spread: Further Evidence and a Structural Interpretation. Unpublished manuscript, Università Bocconi.

[15] Garcia, R. and P. Perron (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics* 78, 111–125.

[16] Geweke, J. and C. H. Whiteman (2006). Bayesian forecasting. In: G. Elliot, C. W. J. Granger, and A. Timmermann (eds), *Handbook of Economic Forecasting*, Elsevier, Amsterdam, 3-80.

[17] Granger, C. (1989). Combining forecasts—twenty years later. *Journal of Forecasting* 8, 167–173.

[18] Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 357-384.

[19] Hamilton, J. D. and K.D. Heon (2002). A Re-examination of the Predictability of Economic Activity Using the Yield Spread. *Journal of Money, Credit, and Banking*, 34, 340-60.

[20] MacKinnon, J.G., Haug, A.A., and L. Michelis (1999). Numerical distribution functions of likelihood ratio tests for cointegration. *Journal of Applied Econometrics* 14, 563-577.

[21] Min, C. and A. Zellner (1993). Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *Journal of Econometrics* 56, 89–118.

[22] Palm, F.C. and A. Zellner (1992). To combine or not to combine? Issues of combining forecasts. *Journal of Forecasting* 11, 687–701.

[23] Pesaran M.H., Pettenuzzo, D. and A. Timmermann, (2006). Forecasting time series subject to multiple structural breaks. *Review of Economic Studies* 73, 1057-1084.

[24] Pesaran M.H. and A. Pick (2007). Optimality properties of forecast averaging when forecasting the mean. Mimeo.

[25] Pesaran M.H., Schuermann, T. and S.M. Weiner (2004). Modeling regional interdependencies using a global error correcting macroeconometric model. *Journal of Business and Economic Statistics* 22, 129-162.

[26] Pesaran, M.H., Shin, Y. and R. Smith (2000). Structural analysis of vector error correction models with exogenous I(1) variables. *Journal of Econometrics* 97, 293-343.

[27] Pesaran M.H. and A. Timmermann (2004). How costly is it to ignore breaks when forecasting the direction of a time series? *International Journal of Forecasting* 20, 411-425.

[28] Pesaran M.H. and A. Timmermann (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics* 137, 134-161.

[29] Rudebusch, G. D., B. P. Sack, and E. T. Swanson (2007). Macroeconomic Implications of Changes in the Term Premium, *Federal Reserve Bank of St. Louis Review*, 89, 241-69.

[30] Stock, J.H. and M.W. Watson (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* 14, 11-30.

[31] Stock, J.H. and M.W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 3-33.

[32] Timmermann, A. (2006). Forecast combinations. In: G. Elliot, C. W. J. Granger, and A. Timmermann (eds), *Handbook of Economic Forecasting*, vol 1., Elsevier, Amsterdam, 135-196.

[33] Weiss, S.M. and C.A Kulikowski (1991). *Computer Systems that Learn.* Palo Alto, CA: Morgan Kaufmann.

Table 6a. Forecasts of Real Output Growth for Different Country Groupings Using the GVAR and Selected Benchmark Models. Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
GVAR_AveW	0.571	0.864	1.051
GVARLR_AveW	0.543	0.864	1.058
AveAve	0.514	0.769	0.941
AR(1)	0.613	0.813	1.018
AR(1) with trend	0.584	0.795	0.987
RW	1.075	1.400	1.568
RW with drift	0.573	0.803	0.985

Models	Four Quarters Ahead		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	0.233	0.424	0.584
DHPS-GVAR-AveW	0.271	0.504	0.659
GVAR-AveAve	0.206	0.399	0.530
AR(1)	0.274	0.405	0.582
AR(1) with trend	0.269	0.401	0.538
RW	0.792	1.093	1.222
RW with drift	0.209	0.393	0.530

Notes: DdPS-GVAR-AveW and DHPS-GVAR-AveW denote average forecasts across 10 estimation windows using DdSP-GVAR and DHPS-GVAR models, respectivel. TDdPS-GVAR denotes the GVAR model with exactlyidentified long-run relations developed in Dees, di Mauro, Pesaran and Smith (2007), and DHPS-GVAR denotes the GVAR model with the long run structural relationships imposed, as in Dees, Holly, Pesaran and Smith (2007). GVAR-AveAve denotes the average forecast across 19 models and 10 estimation windows. RW denotes the random walk benchmark model. LA denotes Latin America. For this set of results the grouping All Country Excluding LA comprises 21 countries, while that of All Countries comprises all the 26 countries/region in the GVAR model. Parameters of the benchmark models are estimated recursively over the longest window.

Table 6b. Panel DM Statistics for GVAR-AveAve Forecasts of Real Output Growth Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e^A_{ijt}(1)]^2 - [e^B_{ijt}(1)]^2$ 10 Industrialised Plus China	All Countries Excluding LA	All Countries
AR(1)	-1.274	-1.684	-2.606
AR(1) with trend	-1.988	-1.747	-1.403
RW	-4.108	-7.529	-8.376
RW with drift	-1.236	-1.829	-1.800

Notes: $e^A_{ijt}(1)$ denotes the forecast error corresponding to the one-quarter ahead AveAve forecast of the GVAR model; $e^B_{ijt}(1)$ denotes the forecast error of the corresponding benchmark model's one-quarter ahead forecast over the longest window. Clearly no estimation is needed for the random walk denoted by RW. For this set of results the grouping All Countries Excluding Latin America (LA) comprises 21 countries, while that of All Countries comprises 26 countries.

Table 7a. Forecasts of Inflation for Different Country Groupings Using the GVAR and Selected Benchmark Models. Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	0.464	0.821	1.229
DHPS-GVAR-AveW	0.553	0.844	1.373
GVAR-AveAve	0.438	0.681	0.886
AR(1)	0.486	0.786	1.214
AR(1) with trend	0.521	0.780	0.865
RW	0.508	0.709	0.720
RW with drift	0.512	0.715	0.730
Models	**Four Quarters Ahead**		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	0.116	0.201	0.449
DHPS-GVAR-AveW	0.189	0.226	0.593
GVAR-AveAve	0.128	0.181	0.247
AR(1)	0.150	0.289	0.625
AR(1) with trend	0.179	0.286	0.341
RW	0.118	0.178	0.182
RW with drift	0.126	0.187	0.198

See notes to Table 6a.

Table 7b. Panel DM Statistics for GVAR-AveAve Forecasts of Inflation Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e^A_{ijt}(1)]^2 - [e^B_{ijt}(1)]^2$		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
AR(1)	-1.456	-3.719	-7.128
AR(1) with trend	-3.638	-3.206	-0.212
RW	-3.994	-0.420	3.415
RW with drift	-4.262	-0.880	3.306

See notes to Table 6b.

Table 8a. Forecasts of Short-Term Interest Rates for Different Country Groupings Using the GVAR and Selected Benchmark Models. Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	0.096	0.173	0.788
DHPS-GVAR-AveW	0.088	0.161	0.950
GVAR-AveAve	0.063	0.103	0.355
AR(1)	0.053	0.116	0.635
AR(1) with trend	0.080	0.126	0.260
RW	0.047	0.081	0.096
RW with drift	0.054	0.088	0.109

Models	Four Quarters Ahead		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	0.055	0.091	0.327
DHPS-GVAR-AveW	0.054	0.090	0.440
GVAR-AveAve	0.041	0.064	0.143
AR(1)	0.036	0.086	0.423
AR(1) with trend	0.061	0.095	0.173
RW	0.031	0.050	0.060
RW with drift	0.039	0.059	0.081

Notes: For this set of results the grouping All Country Excluding LA comprises 20 countries, while that of All Countries comprises 25 countries as there is no domestic short-term interest rate available for Saudi-Arabia. See also notes to Table 6a.

Table 8b. Panel DM Statistics for GVAR-AveAve Forecasts of Short-Term Interest Rates Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e^A_{ijt}(1)]^2 - [e^B_{ijt}(1)]^2$		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
AR(1)	0.909	-3.599	-6.491
AR(1) with trend	-1.762	-2.472	2.052
RW	2.138	2.928	2.935
RW with drift	1.374	1.811	2.921

See notes to Tables 6b and 8a.

Table 9a. Forecasts of Long -Term Interest Rates for Different Country Groupings Using the GVAR and Selected Benchmark Models. Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead: 10 Industrialised	All Countries (12)
DdPS-GVAR-AveW	0.087	0.098
DHPS-GVAR-AveW	0.083	0.095
GVAR-AveAve	0.068	0.078
AR(1)	0.070	0.082
AR(1) with trend	0.066	0.076
RW	0.059	0.070
RW with drift	0.057	0.069

Models	Four Quarters Ahead: 10 Industrialised	All Countries (12)
DdPS-GVAR-AveW	0.036	0.042
DHPS-GVAR-AveW	0.043	0.049
GVAR-AveAve	0.028	0.036
AR(1)	0.041	0.052
AR(1) with trend	0.034	0.040
RW	0.031	0.038
RW with drift	0.026	0.034

Notes:The grouping "All Countries" in this table comprises the 10 industralised countries plus South Korea and South Africa. Also see Table 2 , and the notes to Table 6a.

Table 9b. Panel DM Statistics for GVAR-AveAve Forecasts of Long-Term Interest Rates Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e_{ijt}^{A}(1)]^2 - [e_{ijt}^{B}(1)]^2$: 10 Industrialised	All Countries (12)
AR(1)	-2.286	-1.870
AR(1) with trend	0.373	0.438
RW	2.101	1.641
RW with drift	2.553	1.880

See notes to Tables 6b and 9a.

Table 10a. Forecasts of Real Equity Prices for Different Country Groupings Using the GVAR and Selected Benchmark Models.
Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead 10 Industrialised	 All Countries Excluding LA (17)	 All Countries (19)
DdPS-GVAR-AveW	5.942	7.413	7.792
DHPS-GVAR-AveW	5.110	6.451	6.725
GVAR-AveAve	4.750	5.356	5.676
AR(1)	5.458	6.052	6.741
AR(1) with trend	4.829	5.347	6.075
RW	5.655	5.998	6.260
RW with drift	4.885	5.347	5.650
Models	**Four Quarters Ahead 10 Industrialised**	** All Countries Excluding LA (17)**	** All Countries (19)**
DdPS-GVAR-AveW	4.719	5.798	5.900
DHPS-GVAR-AveW	2.981	3.764	3.909
GVAR-AveAve	2.908	3.120	3.183
AR(1)	3.648	3.967	4.314
AR(1) with trend	2.751	2.744	3.184
RW	3.918	3.885	3.933
RW with drift	2.834	2.900	2.917

Notes: The grouping "All Countries Excluding LA" here comprises 17 countries, while that of "All Countries" comprises 19 countries. See Table 2 for the list of the countries and the notes to Table 6a for further details.

Table 10b. Panel DM Statistics for GVAR-AveAve Forecasts of Real Equity Prices Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e^{A}_{ijt}(1)]^2 - [e^{B}_{ijt}(1)]^2$ 10 Industrialised	 All Countries Excluding LA	 All Countries
AR(1)	-2.506	-3.441	-3.217
AR(1) with trend	-0.453	-0.189	-2.009
RW	-3.069	-3.130	-2.486
RW with drift	-0.718	-0.170	-0.506

See notes to Tables 6b and 10a.

Table 11a. Forecasts of Real Exchange Rates for Different Country Groupings Using the GVAR and Selected Benchmark Models. Simple Cross Country Averages of RMSFE's in Percent.

Models	One Quarter Ahead		
	9 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	4.177	3.636	3.694
DHPS-GVAR-AveW	4.743	4.078	4.136
GVAR-AveAve	4.097	3.478	3.507
AR(1)	3.834	3.413	3.423
AR(1) with trend	3.820	3.452	3.405
RW	3.801	3.373	3.410
RW with drift	3.793	3.285	3.294

Models	Four Quarters Ahead		
	9 Industrialised Plus China	All Countries Excluding LA	All Countries
DdPS-GVAR-AveW	1.400	1.490	1.746
DHPS-GVAR-AveW	1.794	1.722	1.943
GVAR-AveAve	1.529	1.491	1.688
AR(1)	1.860	1.817	1.998
AR(1) with trend	1.751	1.839	1.913
RW	1.860	1.780	2.015
RW with drift	1.589	1.480	1.683

Notes: The grouping "All Countries Excluding LA" here comprises 20 countries, while that of "All Countries" comprises 25 countries, as there is no domestic exchange rate in the model for the US. For the same reason there are 9 industrialised countries instead of 10 in this set of results. See also notes to Table 6a.

Table 11b. Panel DM Statistics for GVAR-AveAve Forecasts of Real Exchange Rates Relative to a Selected Number of Benchmarks

Benchmark Models	$z_{ijt} = [e^{A}_{ijt}(1)]^2 - [e^{B}_{ijt}(1)]^2$		
	10 Industrialised Plus China	All Countries Excluding LA	All Countries
AR(1)	1.454	0.619	0.541
AR(1) with trend	1.507	0.179	0.553
RW	1.521	0.930	0.699
RW with drift	1.901	1.875	1.944

See notes to Tables 6b and 11a.

Table 12a. Forecasts of Real Output Growth for Individual Country and Different Country Groupings Using Variants of the GVAR Model Excluding Latin America. Simple Cross Country Averages of RMSFE's in Percent.

	One Quarter Ahead GVAR Models Excluding Latin America		
Country/Group	With EQ & LR	Without EQ	Without EQ & LR
US	0.181	0.182	0.201
EA	0.245	0.198	0.207
China	0.138	0.127	0.112
Japan	0.604	0.644	0.599
UK	0.197	0.196	0.198
Canada	0.225	0.252	0.264
Australia	0.360	0.355	0.407
Sweden	1.034	1.027	1.041
Switzerland	0.317	0.361	0.357
Norway	1.833	1.781	1.768
New Zealand	0.552	0.569	0.591
10 Industrialised Plus China	0.517	0.517	0.522
All Countries	0.766	0.771	0.771
	Four Quarters Ahead GVAR Models Excluding Latin America		
Country/Group	With EQ & LR	Without EQ	Without EQ & LR
US	0.089	0.096	0.119
EA	0.184	0.145	0.159
China	0.189	0.164	0.148
Japan	0.390	0.440	0.341
UK	0.144	0.144	0.135
Canada	0.115	0.138	0.179
Australia	0.084	0.134	0.189
Sweden	0.270	0.285	0.298
Switzerland	0.196	0.212	0.191
Norway	0.556	0.546	0.501
New Zealand	0.165	0.185	0.173
10 Industrialised Plus China	0.217	0.226	0.221
All Countries	0.409	0.419	0.406

Notes: The group "All Countries" for this set of results comprises 21 countries (Latin America countries are excluded from this specification of the GVAR model).

Table 12b. Forecasts of Inflation for Individual Country and Different Country Groupings Using Variants of the GVAR Model Excluding Latin America. Simple Cross Country Averages of RMSFE's in Percent.

	One Quarter Ahead		
	GVAR Models Excluding Latin America		
Country/Group	With EQ & LR	Without EQ	Without EQ &LR
US	0.524	0.541	0.527
EA	0.246	0.245	0.241
China	1.318	1.313	1.317
Japan	0.302	0.308	0.299
UK	0.120	0.141	0.135
Canada	0.510	0.533	0.501
Australia	0.348	0.345	0.324
Sweden	0.354	0.320	0.329
Switzerland	0.224	0.225	0.199
Norway	0.492	0.553	0.568
New Zealand	0.406	0.395	0.399
10 Industrialised Plus China	0.440	0.447	0.440
All Countries	0.671	0.669	0.675

	Four Quarters Ahead		
	GVAR Models Excluding Latin America		
Country/Group	With EQ & LR	Without EQ	Without EQ & LR
US	0.114	0.118	0.107
EA	0.061	0.063	0.061
China	0.449	0.440	0.451
Japan	0.091	0.093	0.094
UK	0.054	0.067	0.057
Canada	0.110	0.111	0.108
Australia	0.105	0.104	0.068
Sweden	0.109	0.107	0.124
Switzerland	0.070	0.077	0.072
Norway	0.112	0.113	0.115
New Zealand	0.155	0.177	0.143
10 Industrialised Plus China	0.130	0.134	0.127
All Countries	0.186	0.186	0.184

See notes to Table 12a.

Table 12c. Forecasts of Short-Term Interest Rates for Individual Country and Different Country Groupings Using Variants of the GVAR Model Excluding Latin America. Simple Cross Country Averages of RMSFE's in Percent.

	One Quarter Ahead GVAR Models Excluding Latin America		
Country/Group	With EQ and LR	Without EQ	Without EQ and LR
US	0.114	0.114	0.106
EA	0.019	0.028	0.025
China	0.034	0.035	0.032
Japan	0.017	0.013	0.017
UK	0.048	0.058	0.062
Canada	0.087	0.085	0.080
Australia	0.037	0.033	0.031
Sweden	0.059	0.053	0.047
Switzerland	0.046	0.047	0.040
Norway	0.130	0.124	0.116
New Zealand	0.099	0.089	0.086
10 Industrialised Plus China	0.063	0.062	0.058
All Countries	0.102	0.103	0.097

	Four Quarters Ahead GVAR Models Excluding Latin America		
Country/Group	With EQ and LR	Without EQ	Without EQ and LR
US	0.119	0.126	0.119
EA	0.024	0.036	0.032
China	0.005	0.008	0.010
Japan	0.017	0.022	0.018
UK	0.047	0.058	0.051
Canada	0.046	0.047	0.034
Australia	0.026	0.028	0.019
Sweden	0.024	0.016	0.014
Switzerland	0.019	0.027	0.024
Norway	0.049	0.043	0.031
New Zealand	0.065	0.085	0.082
10 Industrialised Plus China	0.040	0.045	0.039
All Countries	0.067	0.069	0.063

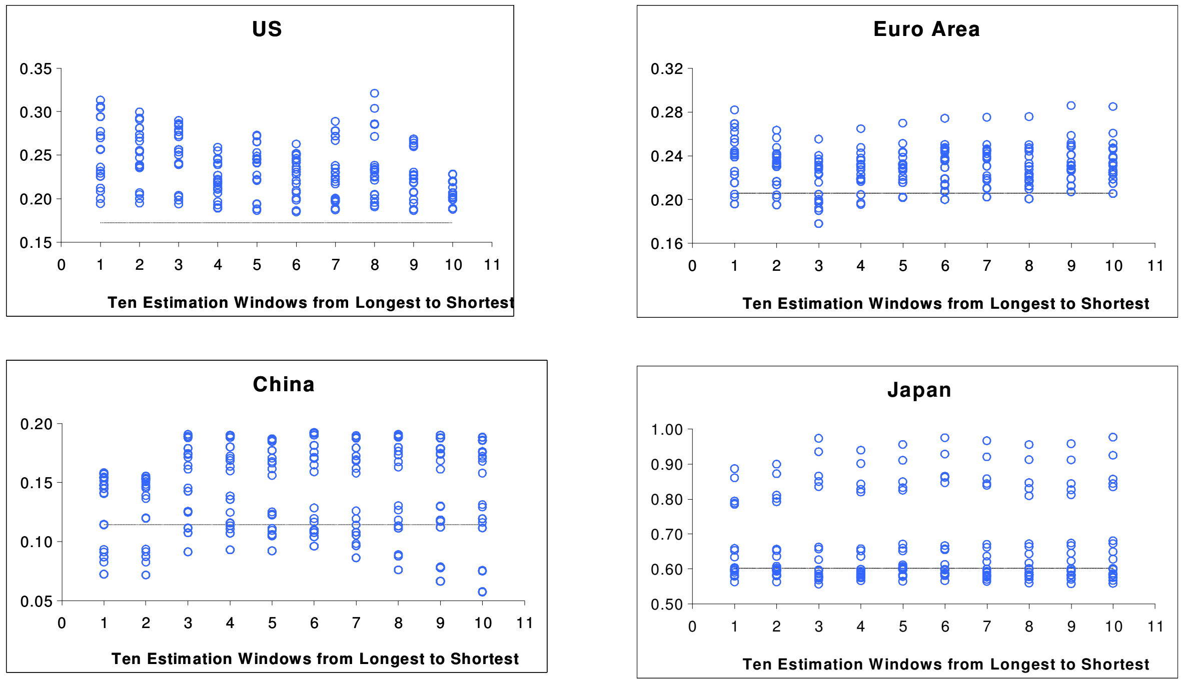
Notes: The group All Countries for this set of results comprises 20 countries (Latin America countries are excluded from the GVAR model and Saudi Arabia does not have a domestic short-term interest rate).

Table 12d. Panel DM Statistics for AveAve Forecasts of the GVAR Model Excluding Latin America

Models	$z_{ijt} = [e^{A}_{ijt}(1)]^2 - [e^{B}_{ijt}(1)]^2$ 10 Industrialised Plus China	All Countries
	Real Output	
Without EQ	0.488	-0.856
Without EQ & LR	0.377	-0.563
	Inflation	
Without EQ	-1.509	1.200
Without EQ & LR	-0.153	-0.061
	Short-Term Interest Rate	
Without EQ	0.506	-0.781
Without EQ & LR	0.748	0.085

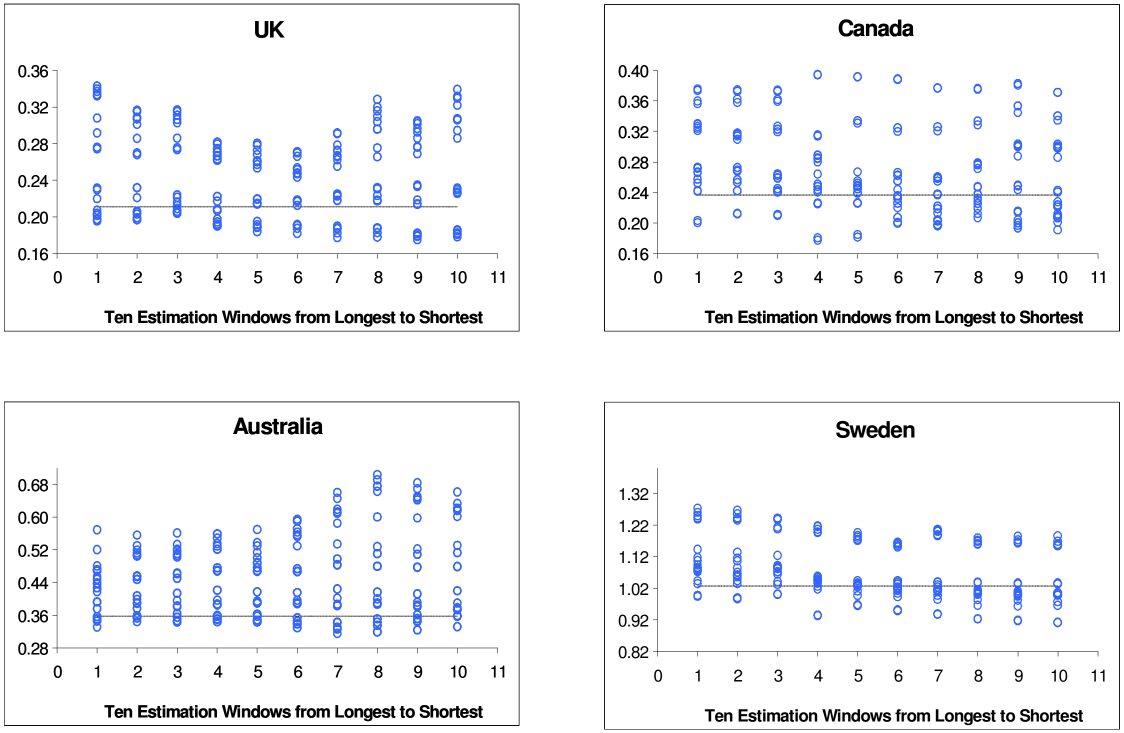
Notes: $e^{A}_{ijt}(1)$ denotes the forecast error corresponding to the one-quarter ahead AveAve forecast of the GVAR model that includes equity and the long-term interest rate; $e^{B}_{ijt}(1)$ denotes the forecast error corresponding to the one-quarter ahead AveAve forecast of the GVAR model excluding equity (without EQ) or excluding equity and the long-term interest rate (without EQ & LR). The group All Countries for this set of results comprises 21 countries (Latin America countries are excluded from the GVAR model) for real output and inflation and 20 countries for the short-term interest rate due to the non-availablility of data for this variable for Saudi Arabia.

Figure 1a. RMSFEs of One-Quarter Ahead Forecasts for Real Output Growth



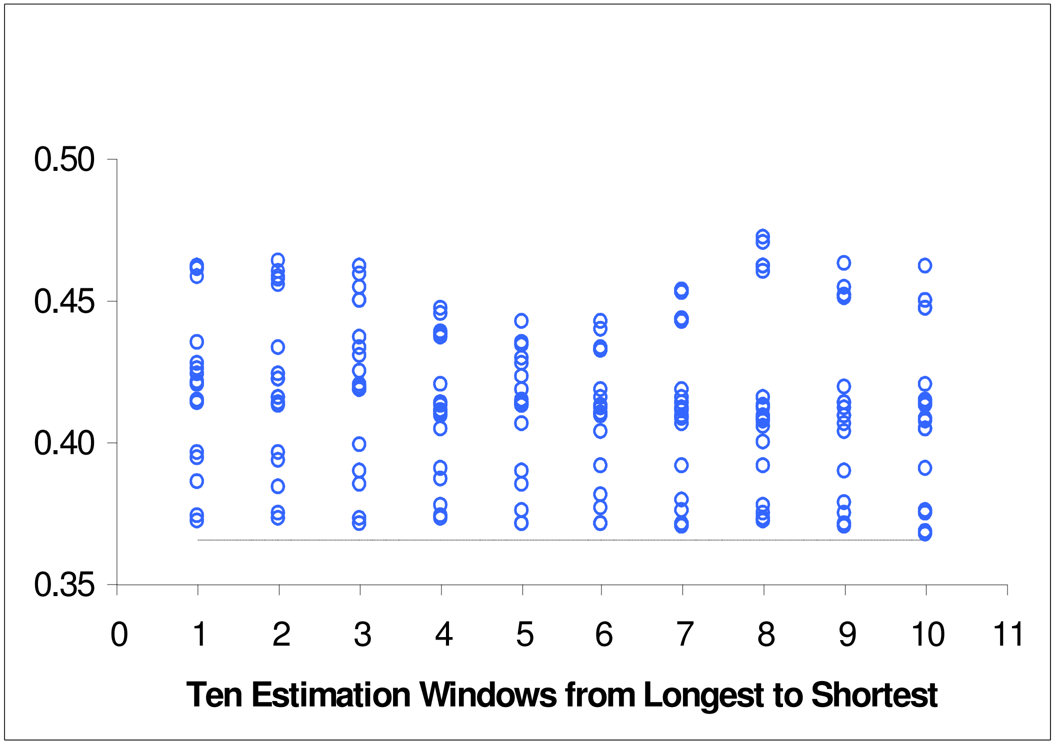
Note: Each circle denotes the RMSFE of a particular model estimated on a particular window. The straight line in the figures refers to the GVAR-AveAve forecasts across all the 19 models and 10 estimation windows.

Figure 1b. RMSFEs of One-Quarter Ahead Forecasts for Real Output Growth



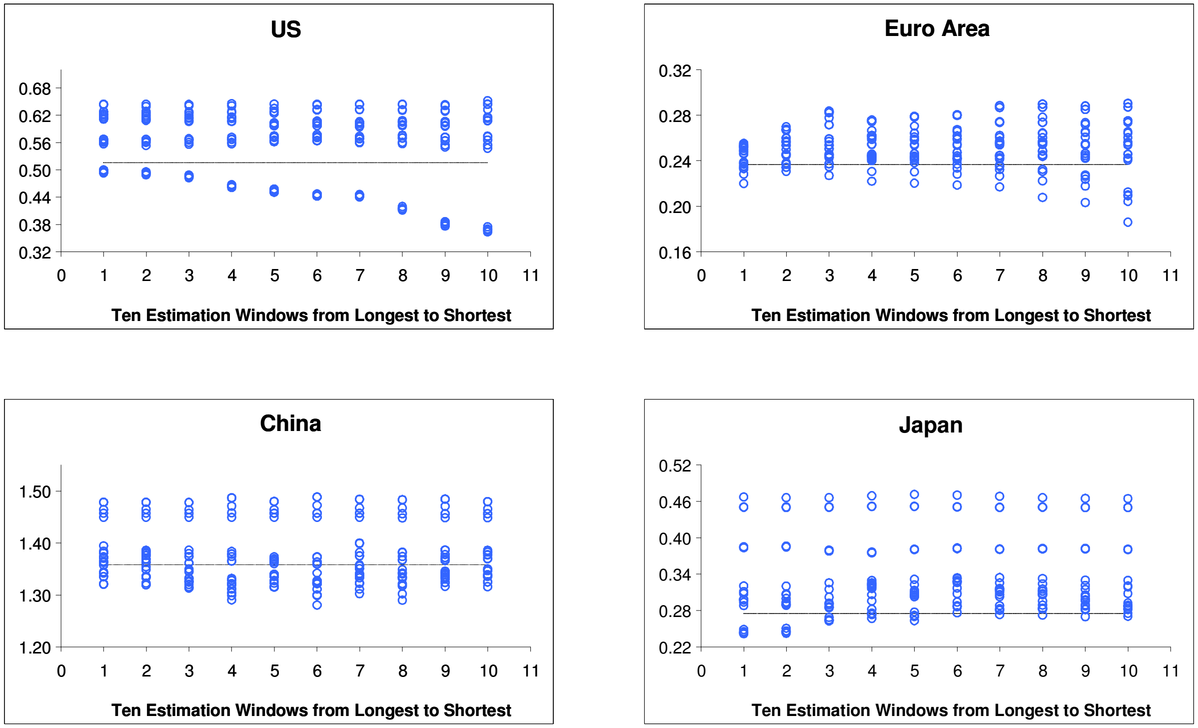
See note to Figure 1a.

Figure 1c. Average RMSFEs of One-Quarter Ahead Forecasts for Real Output Growth



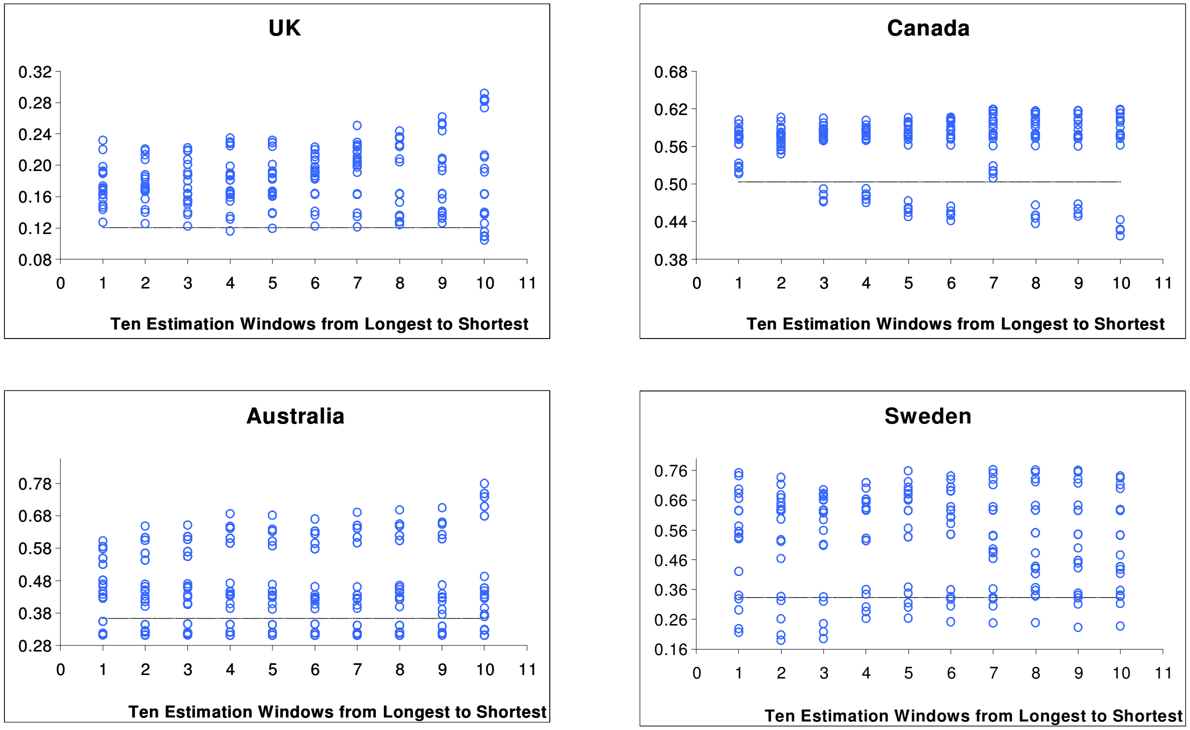
Note: The average is for the 10 industrialised countries plus China. See Table 2 for the list of countries.

Figure 2a. RMSFEs of One-Quarter Ahead Forecasts for Inflation



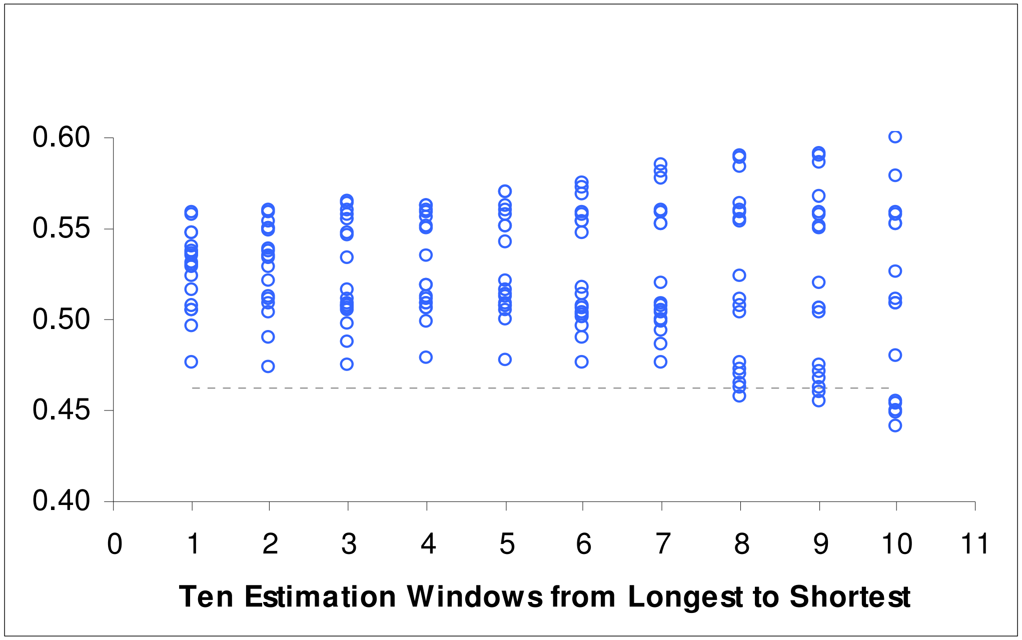
See note to Figure 1a.

Figure 2b. RMSFEs of One-Quarter Ahead Forecasts for Inflation



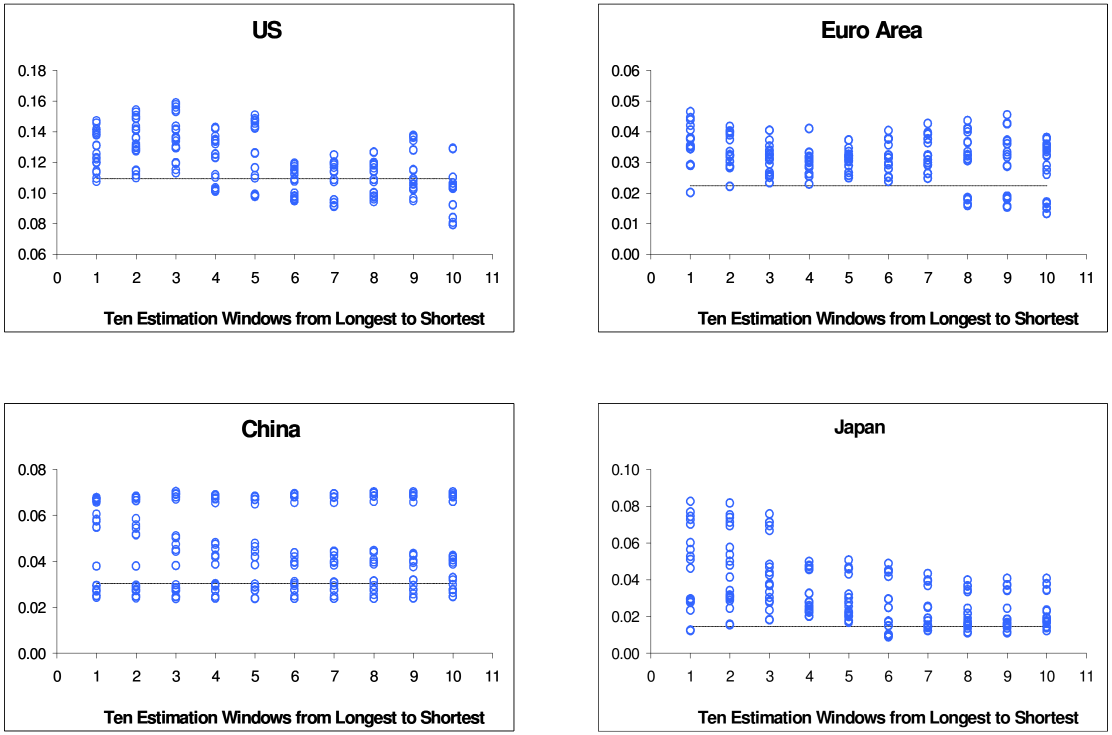
See note to Figure 1a.

Figure 2c. Average RMSFEs of One-Quarter Ahead Forecasts for Inflation



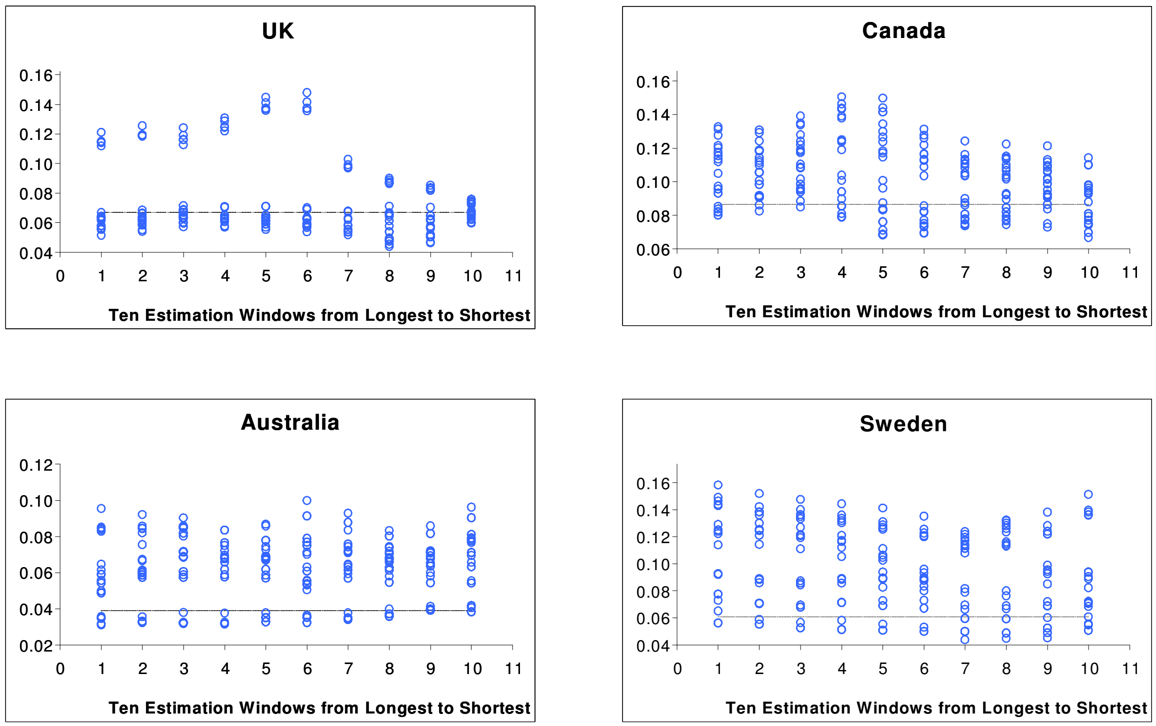
See note to Figure 1c.

Figure 3a. RMSFEs of One-Quarter Ahead Forecasts for Short-Term Interest Rate



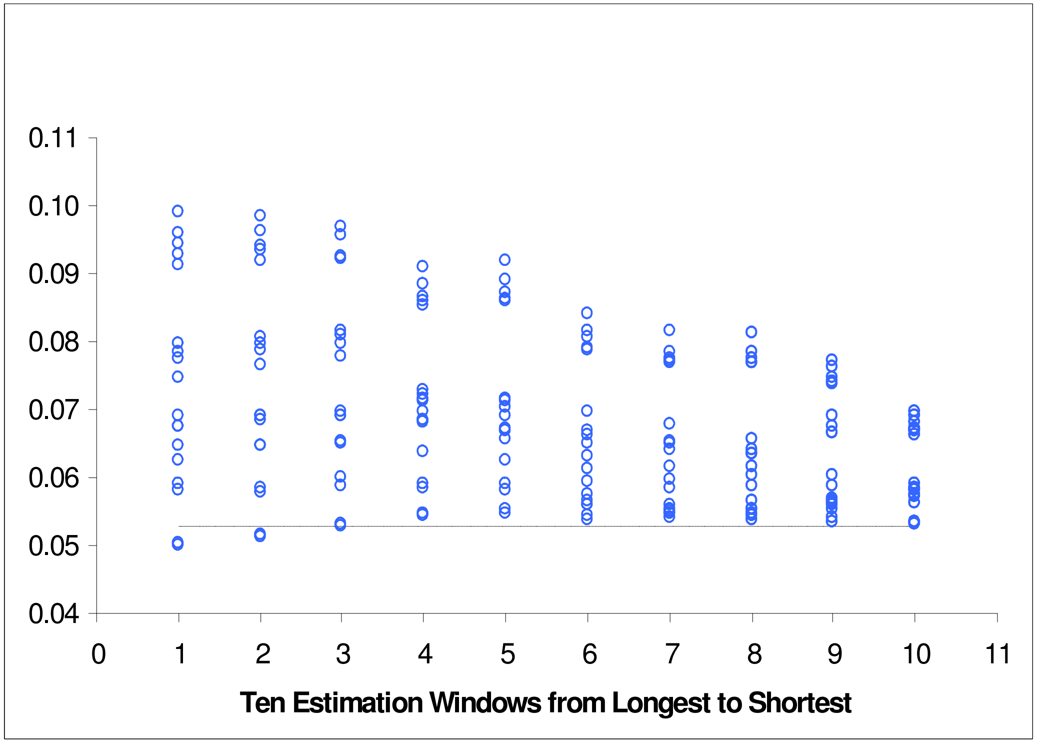
See note to Figure 1a.

Figure 3b. RMSFEs of One-Quarter Ahead Forecasts for Short-Term Interest Rate



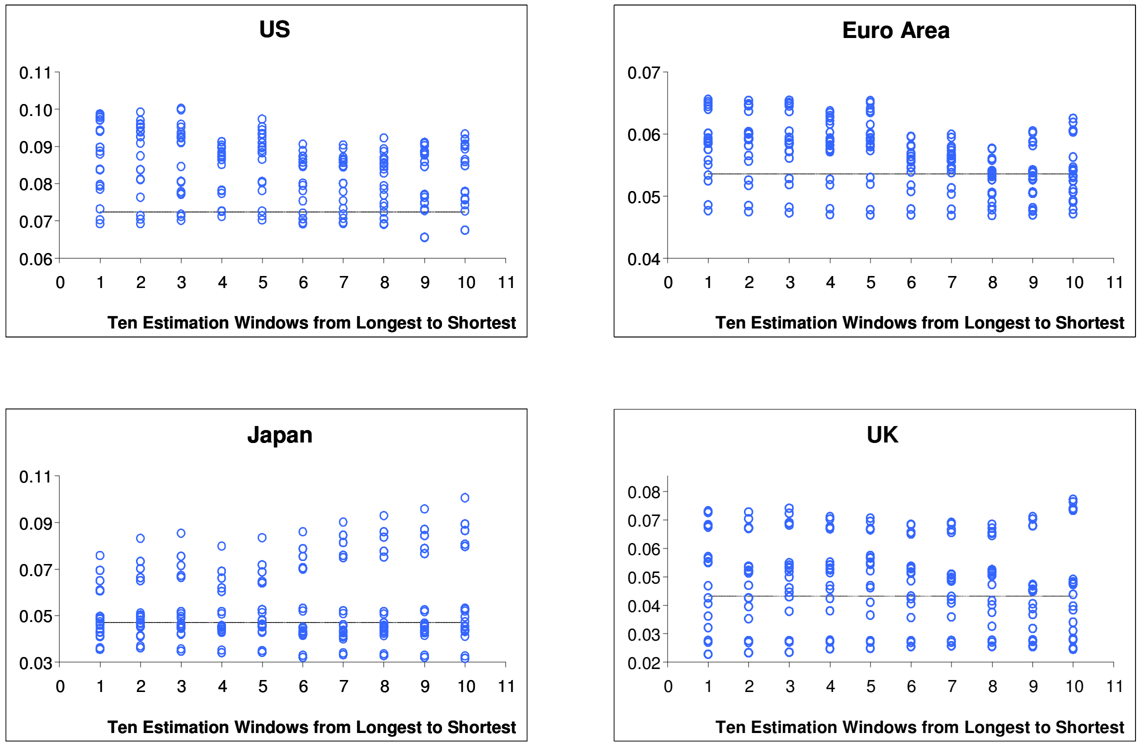
See note to Figure 1a.

Figure 3c. Average RMSFEs of One-Quarter Ahead Forecasts for Short-Term Interest Rate



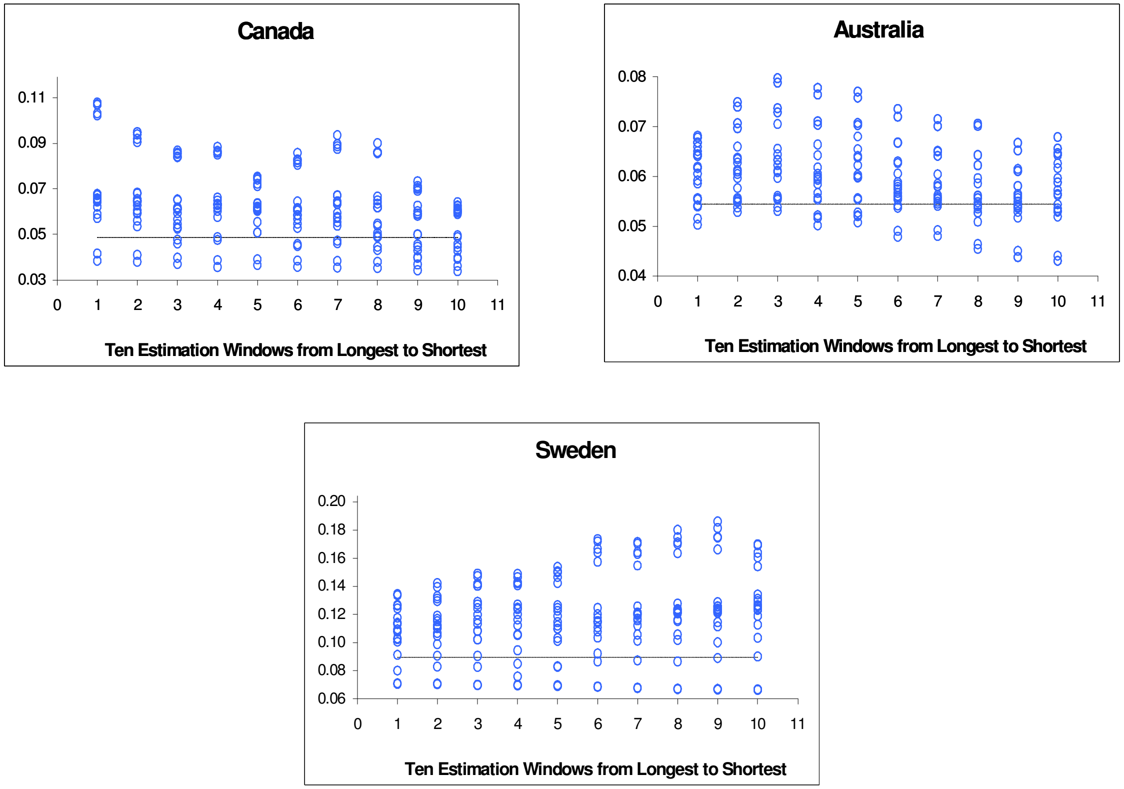
See note to Figure 1c.

Figure 4a. RMSFEs of One-Quarter Ahead Forecasts for Long-Term Interest Rate



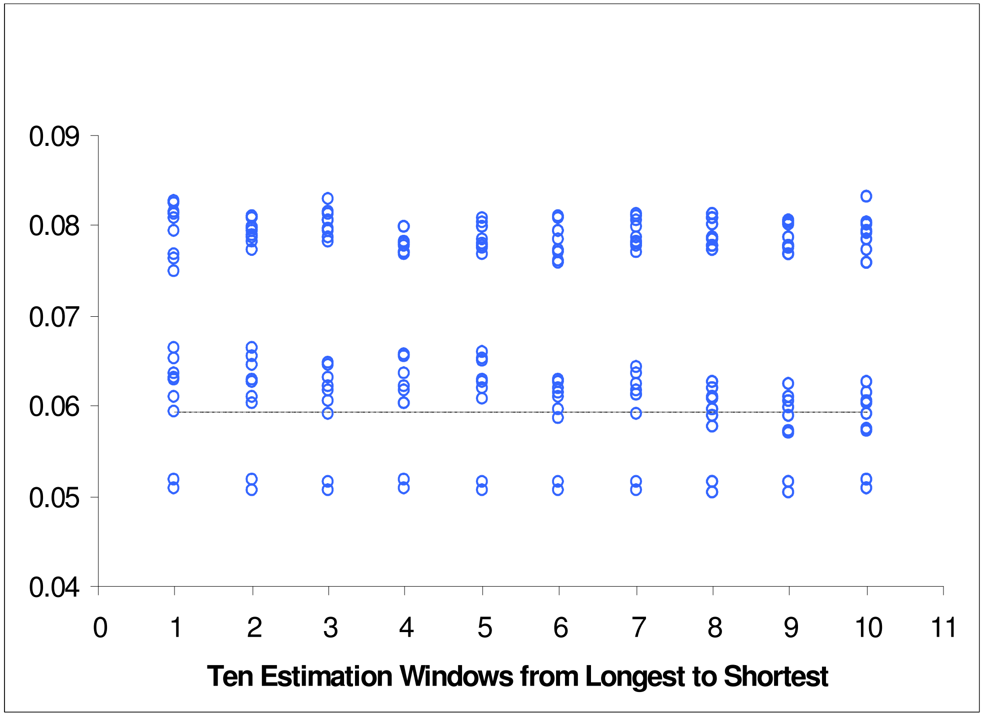
See note to Figure 1a.

Figure 4b. RMSFEs of One-Quarter Ahead Forecasts for Long-Term Interest Rate



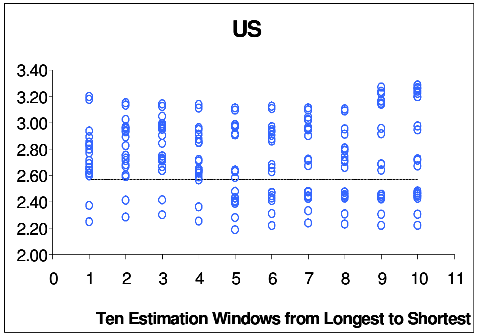
See note to Figure 1a.

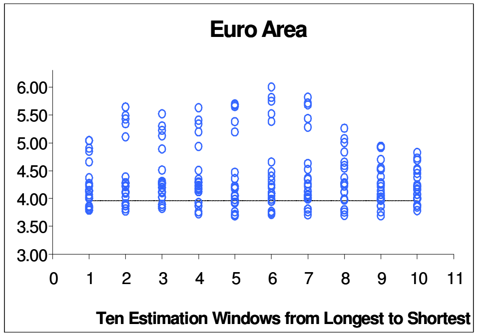
Figure 4c. Average RMSFEs of One-Quarter Ahead Forecasts for Long-Term Interest Rate

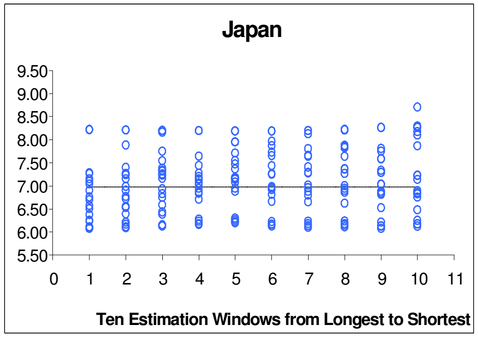


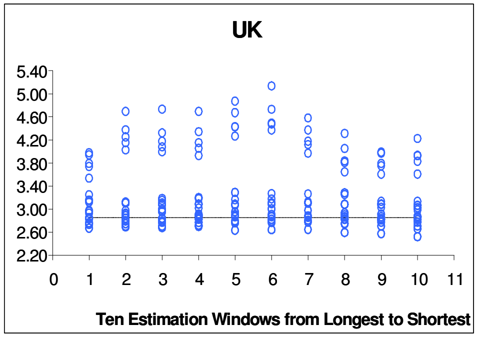
Note: The average is for the 10 industrialised countries. See Table 2 for the list of countries.

Figure 5a. RMSFEs of One-Quarter Ahead Forecasts for Real Equity



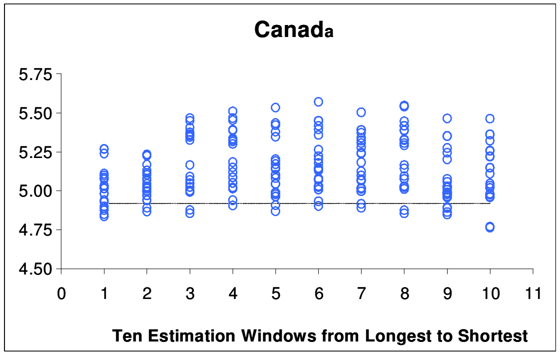


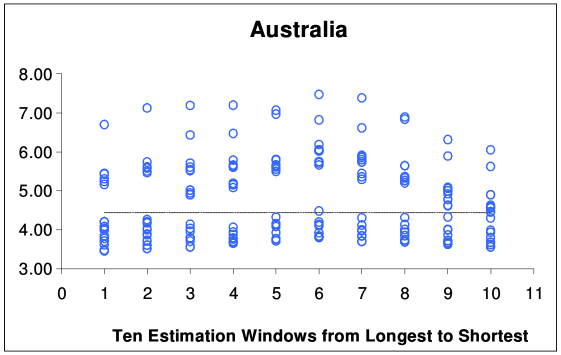


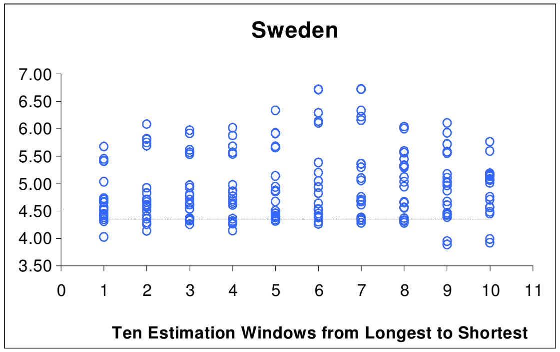


See note to Figure 1a.

Figure 5b. RMSFEs of One-Quarter Ahead Forecasts for Real Equity





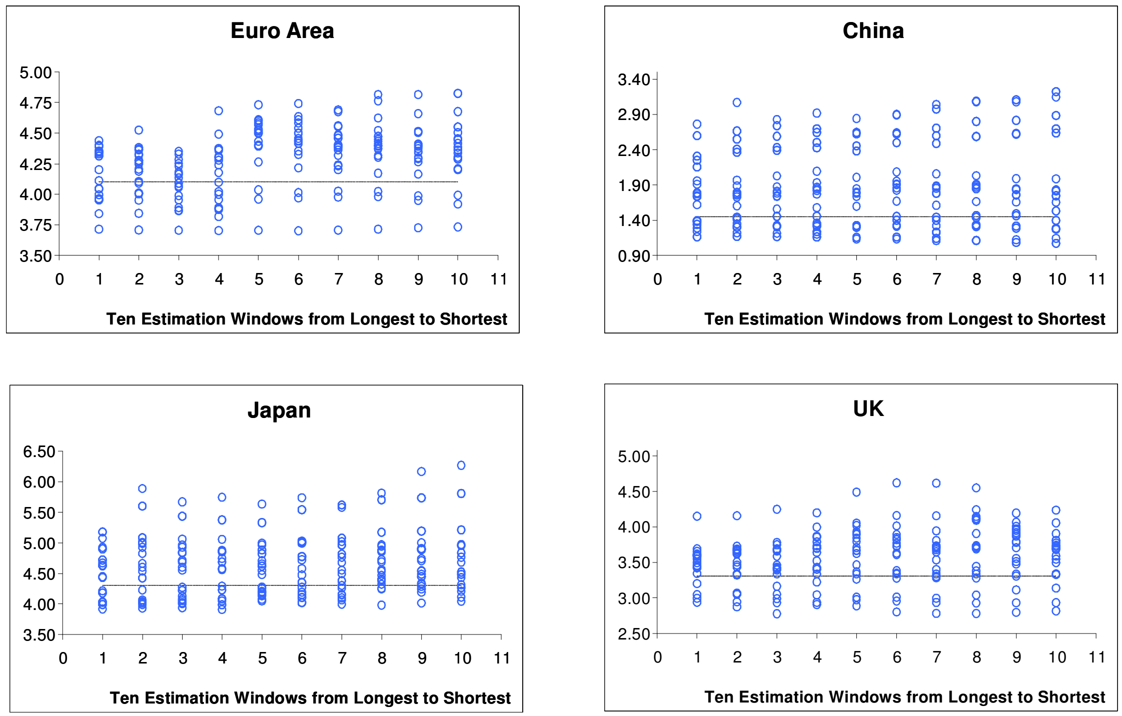


See note to Figure 1a.

Figure 5c. Average RMSFEs of One-Quarter Ahead Forecasts for Real Equity

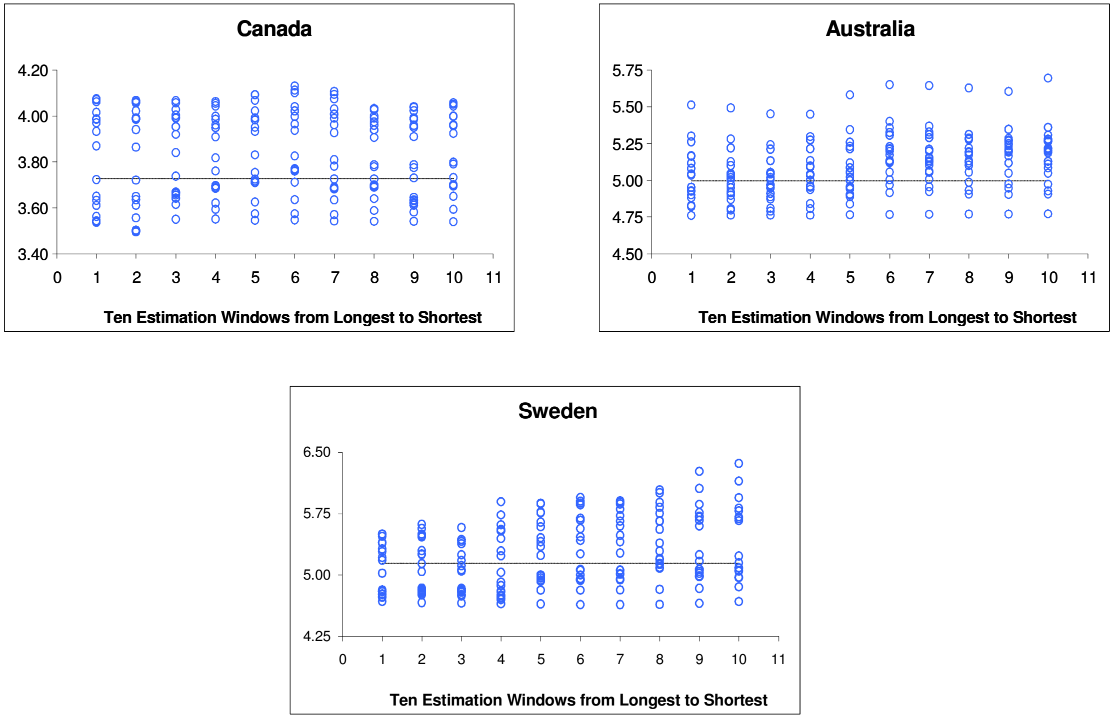
See not to Figure 4c.

Figure 6a. RMSFEs of One-Quarter Ahead Forecasts for Real Exchange Rate



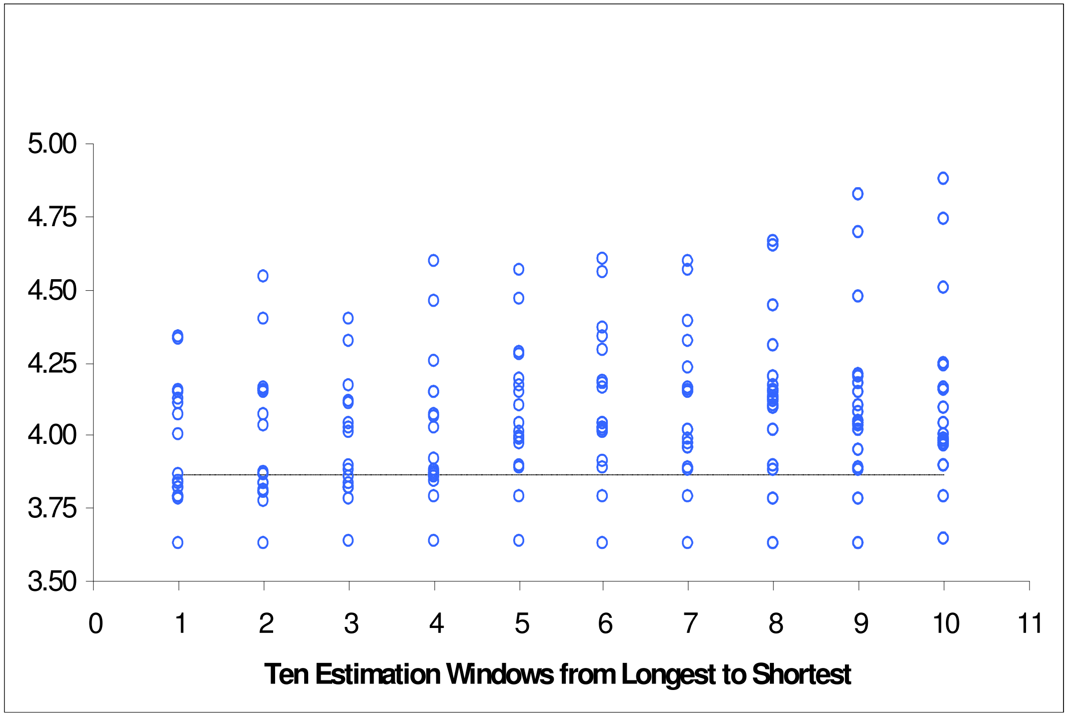
See note to Figure 1a.

Figure 6b. RMSFEs of One-Quarter Ahead Forecasts for Real Exchange Rate



See note to Figure 1a.

Figure 6c. Average RMSFEs of One-Quarter Ahead Forecasts for Real Exchange Rate



Note: The average is for the 9 industrialised countries, excluding the US, plus China. See Table 2 for the list of countries.

Figure 7. Performance of AveAve forecasts based on GVAR models versus the forecasts from the four benchmarks. % of Forecast where GVARAveAve beats Benchmark at 95% CI or better