Rejoinder to Comments on 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, Wharton Financial Institutions Center, and CIMF

L. Vanessa Smith
University of Cambridge, CFAP

July 31, 2009

1 Introduction

We would like to thank all the seven discussants and two anonymous referees for their generous and constructive comments and the associate editor, Keith Ord, for providing us with the opportunity of taking part in this distinguished forum. While there is obviously some overlap between the discussants in the issues they raise, each brings their distinct expertise and experience to bear on our work on forecasting with Global VARs. They also raise a number of important technical and practical issues that are clearly worth further investigation. In this rejoinder we begin with some general remarks regarding the intended purposes and possible uses of the GVAR model and a brief comment about the data. We then organize our comments on the points raised by the seven discussants into three main sections: model parsimony and parameter shrinkage, which collects the myriad of comments around desires to simplify the models; forecast evaluation, where we consider the cross sectional dimension including bias and dispersion; with the third section dealing with a number of remaining issues. We also present some new forecast results extending the evaluation sample to the end of 2006, as well as providing a multivariate benchmark without the foreign ('star') variables, as suggested by Allen, Clements, and Sinclair and Stekler.

The recent financial crisis, which has triggered the first global recession since World War II, has been a humbling experience to any forecaster. It is a crisis of truly global proportion, it is a crisis of risk management, and it may well be a crisis of policy. But if nothing else, it highlights the critical need for explicit and deliberate global modelling of real and financial variables; one example is the GVAR modelling architecture discussed here. How well this or any other global model will fare through and beyond this crisis remains to be seen, but regardless of the outcome, any approach must allow and account for the international transmission of shocks, and the possibility of structural breaks and change.

2 An Overview

The original motivation for developing the GVAR, spelled out in Pesaran, Schuermann and Weiner (2004), and developed more fully in Pesaran, Schuermann, Treutler and Weiner (2006) was as a risk management and policy tool for commercial and central bankers, or other policy makers. The primary objective was to formulate and estimate an econometric model capable of yielding quantitative answers for the following types of questions: what would be the credit risk exposure and the loss distribution of a global loan portfolio in the event of a large real or financial shock in the world economy (such as shocks to oil prices, US equity prices or Japanese output). In order to achieve this objective, it is clear that an explicitly linked global model that allows and accounts for interactions and spillover effects needs to be built. The best point forecast for US GDP or inflation, or for that matter any other particular variable, was not the primary goal or intent. Model (and, in our case, estimation window) averaging, then, is a risk mitigation, portfolio approach to forecasting: in the absence of being able to identify a single best model for the breadth of variables and countries, in a world of structural breaks, it turns out that simple averaging works well.

The intent of the GVAR modelling exercise did not escape our discussants. But as Giannone and Reichlin point out, the evaluation of individual point forecasts is still a natural and frankly necessary exercise, and we agree entirely. Given that we need a somewhat more complicated model to achieve our objective, what then is its forecast performance? Indeed we would have been quite happy if the GVAR model performed as well as other procedures whose only purpose is to generate a good point or density forecast for a given variable or set of variables in a given country.

If the purpose was only to forecast a single variable in isolation, then as Allen and Clements argue, model parsimony might be desirable. However, for policy analysis and risk management where predicting correlations across (often many) variables is also an important consideration, a major focus on parsimony might be misplaced. Moreover, for certain purposes such as identification and estimation of the effects of shocks in the global economy, a relatively large model might be needed, and this ought also to be born in mind in comparative forecast evaluation exercises. In the end, forecasting is best seen as a part of a decision problem and not an end in itself. See, for example, Granger and Pesaran (2000), and Pesaran and Skouras (2002).

Part and parcel of the GVAR architecture is the data set itself. Specifically, the construction of the foreign ("star") variables themselves is the first channel of global interaction. Recall that each country's or region's foreign variable, say output, is the weighted average of all the remaining countries' or regions' variables, where the weights are given by the trade relations. In this way, each set of foreign variables is bespoke to that country. These trade weights come ostensibly from "outside" the GVAR model, and structural breaks in those trading patterns need (and are) themselves to be reflected in the construction of those weights. Secular changes in trade weights can be readily allowed for in the analysis, although results in Dees, di Mauro, Pesaran and Smith (2007) suggest this issue to be of secondary importance.

Over the past decade a more prominent role has been given to the use of large data sets in forecasting and policy analysis. This is perhaps inevitable given the increasing availability of such data sets and the significant increase in computer speeds and size, as noted by Granger. The issue of how best such data sets can be used is an open issue. The GVAR methodology can be viewed as a modest step in that direction.

3 Model Parsimony and Parameter Shrinkage

Perhaps the dominant theme across the set of comments is model parsimony. Clements would like to see a comparison with simpler, single country models, and/or dropping some (or all) of the long-run relationships imposed in the DHPS-GVAR models; Giannone and Reichlin argue for a Bayesian shrinkage approach to reduce the parameter space; and Allen is looking for simplification by dropping foreign variables.

The results clearly show that for almost every variable, there are some countries and some time periods for which the GVAR-AveAve does not outperform the benchmark models. The benchmarks we use in the paper are, of course, simple single equation models (see below for results on single country VAR benchmarks). And as Sinclair and Stekler demonstrate, one need not be confined to just statistical models. As an alternative benchmark they suggest the mean forecast of the (US) Survey of Professional Forecasts (SPF), itself (by definition) an average forecast. They report that, for the two variables which are collected by the SPF, namely US inflation and output, the results are split: compared with the GVAR-AveAve forecast, the SPF does better for output but not for inflation. We are in fact heartened by this outcome considering the acknowledged expertise of the Professional Forecasters in the US. But given the very short evaluation periods at our disposal, forecast evaluation outcomes for any single variable might not be that informative and the results could have gone either way. This is one reason why in our paper we do not consider the forecast performance of individual series, but consider RMSFEs that are averaged across different groups of countries. Similar considerations also apply to our use of a panel version of the Diebold-Mariano test, and not the DM test statistics that could have been computed for individual series.

Several discussants expressed a desire to reduce the model and/or parameter space across both M (models) and W (windows). Lahiri suggests finding the "best" individual model/window combination among the 190 averaged models/windows, but as we show below, there is no such obvious candidate or even set of candidates. Allen postulates that there may be a few GVAR or DGVAR models, estimated over a set of windows, that systematically do better than others, so that the AveAve estimate may actually be "close" to just a few of the underlying 190 models. He suggests helpfully to display the model forecast charts by distinguishing between the GVAR and DGVAR models. We have followed this suggestion and in Figures 1-6 of the paper we now denote the GVAR forecasts with (red) circles, and the DGVAR forecasts with (blue) stars. Recall that the solid line is the AveAve forecast. While there appears a pattern in some instances, there actually is no obvious regularity when all the forecasts are considered across all variables and all countries. For instance, in Figure 1 where we show the one-quarter ahead RMSFE for real output growth, the DGVAR models seem to perform better than the GVAR models for several countries (US, China, Japan in Figure 1a). But this is not the case for the UK or Australia (Figure 1b). For inflation (Figure 2), the reverse is true: the GVAR models do better for the US, China and Japan (Figure 2a), but not for the UK or Australia (Figure 2b). So it is not at all obvious which model and which window one would choose *ex ante*, or even *ex post*.

Similarly, Allen points out that if structural breaks are relatively unimportant, then averaging over all estimation windows is likely to be helpful; see also Pesaran and Timmermann (2007). Yet if structural breaks are important, then excluding the worst performers, most likely forecasts from the widest estimation windows, should improve the composite forecast. A similar point is also raised by Clements. This is true, of course, if the breaks are easily identified *ex ante*. And even if they are identifiable, one needs to determine whether the break is in the mean or in the variances (and covariances) of the model variables. We are not so confident that break points and break sizes can be identified and reliably estimated, and, as we argue in Section 2.3, prefer to make use of an approach that can afford to be agnostic. It is always tempting *ex post* to eliminate the worst or least informative models or windows and keep only the best. But as Figures 1-6 quite clearly show, even *ex post* it is hard to systematically narrow down the set of models and windows.

Allen suggests another avenue for simplification, which is to drop the financial variables for forecasting the real series (e.g. output growth) as we show that these don't seem to help in general. While this is the case broadly, it is not true for every country and for every financial conditioning variable. For instance, output growth forecast performance for the US, Canada, Australia, New Zealand, and Switzerland deteriorates when the long term rate (and real equity prices) are dropped; see Table 14b and the discussion in Section 6.3.

Lahiri makes the point that policy makers may be reluctant to use an average of 190 models rather than a single, 'well specified' GVAR model. Usability of the model has been from the beginning a primary concern of ours. It should be emphasized that for every averaging procedure over models/windows there is a set of average coefficients associated with a single model. One may think of this as a kind of shrinkage approach. Indeed the forecast combination approach denoted in Equation 2 in the paper characterizes a pooled probability density which is of interest in decision making. In this way the AveAve models can equally be used in counter-factual and policy analysis very much in the same way that a single GVAR model is used. An example of the use of a pooled forecast density in a risk management context is given in Pesaran, Schleicher, and Zaffaroni (2009).

The AveAve approach that we use also relates to the Bayesian shrinkage favoured by Bańbura, Giannone and Reichlin (2009), and reviewed in the discussion by Giannone and Reichlin (GR). Unlike these authors we do not see the GVAR methodology as a substitute for parameter shrinkage. Parameter and model uncertainty is equally applicable to GVAR as to any other model. Seen from this perspective forecast averaging (across both model and window dimensions) should be viewed as a way of dealing with parameter instability and model uncertainty, in a manner which complements the GVAR modelling. Comparing forecasts from a single GVAR to forecasts obtained using Bayesian shrinkage, as reported by GR, does not take account of model and parameter uncertainty and instability, and is not what we recommend. A more appropriate exercise would have been to generate Bayesian shrinkage forecasts using our data (which was made publicly available) and compare them with the AveAve forecasts that we report. The recent global recession provides a more effective test of such a comparison and we hope that GR will return to this issue in the coming years.

4 Forecast Evaluation: Bias and Cross Sectional Dispersion

Several discussants made a number of suggestions for alternative or additional ways of evaluating forecast performance. Sinclair and Stekler raised a question of forecast bias; Swanson was interested in evaluating not just point forecasts but also forecast correlations. We have conducted additional analysis to address these important points.

4.1 Bias

We begin the discussion with an examination of forecast bias. Because the RMSFE of the GVAR-AveAve model is typically lower than that of the benchmarks, we normalize the absolute bias by the RMSFE which we call relative bias. In this way, a lower relative bias also indicates a lower absolute bias. For the one quarter ahead forecasts over the evaluation sample of size $n = 8$, the relative bias is computed as

$$\text{Relative Bias} = 100 \times \frac{\left|\sum_{t=T}^{T+8-1} e_t(1)/8\right|}{RMSFE(1,8)}$$

where

$$RMSFE(1,8) = \sqrt{\sum_{t=T}^{T+8-1} e_t^2(1)/8},$$

and $e_t(1) = y_{t+1} - \hat{y}_{t+1}$ is the one quarter ahead forecast error.

Table R1 summarises the results where we present the relative bias of alternative one quarter ahead forecasts for the six variables by the country groupings: Group1, Group2, and Group3 that correspond to the "10 Industrialised Plus China", "All Countries Excluding LA" (Latin America), and "All Countries", respectively. Following the basic format in the paper, we present the bias measures for the DdPS-GVAR-AveW, the DHPS-GVAR-AveW, the AveAve, and the four univariate benchmark models, each of which is estimated over the whole sample. Taking first real output growth, the GVAR model forecasts have the lowest relative bias across all country groupings, with the AveAve model showing the lowest degree of bias. This is the best showing for the GVAR model family among the six variables to be sure, but only for inflation does one of the benchmark models – in this case the two random walk models – have marginally lower relative bias than the GVAR models. The GVAR-AveAve model does best for short term interest rates and real equity prices. For long term rates and real exchange rates, the DdPS-GVAR-AveW and GVAR-AveAve have nearly the same relative bias, and in most cases lower than all the benchmark models.

The lower relative (and absolute) bias of the AveAve forecasts is in line with the theoretical results derived for a few simple models in Pesaran and Pick (2009). The averaging of forecasts across windows and models reduces the bias, but can result in increased forecast error variances. The trade-off depends on the relative importance of parameter instability, particularly when such instabilities occur towards the end of the estimation sample.

4.2 Dispersion

So far all of our forecast evaluations, formal or otherwise, have focussed on the size of forecast errors along the time dimension: how well can the models predict the future values. Yet we stated from the beginning that such point forecasting was not the main intent of the model building exercise. Rather it was a desire to forecast a coherent system of interlinked variables. As such it seems sensible to consider the cross sectional dimension: how well can the models trace the cross sectional dispersion of the forecasts? The set of 134 variables, spread across 26 countries/regions, exhibit a particular dispersion relative to one another. Swanson expressed interest in seeing correlations, and it was in addressing his point that led us to comparisons of the dispersions of realized and forecast values. Dispersion may be thought of as describing the average pair-wise differences of forecasts across a given set of variables, as we will explain below.

This is an issue we have been thinking about for some time. It is quite natural to ask how the forecasts across variables and countries hang together, especially in view of the explicitly global nature of the model to begin with. Swanson suggests evaluating the extent to which the cross correlation of the observations is matched by the cross correlation of the forecasts. Aside from the short sample of the forecasting horizon (just 8 quarters), such a comparison is complicated by the fact that very poor time series forecasts could show a very good match between the cross correlation of actuals and forecasts. For instance, the match could be almost perfect for the random walk model, where the pair-wise cross correlation of n forecasts will differ from the pair-wise correlation of the actuals only by one pair of data points. More specifically, consider the random walk forecasts of π_{it} over the period $t = T+1, T+2, ..., T+n$, given by $\pi^f_{i,T+h} = \pi_{i,T+h-1}$ for $h = 1, 2, ..., n$. Let the pair-wise cross correlation of the realizations of inflation for country i and j be ρ_{ij}, and denote the associated measure computed using the forecasts by $\hat{\rho}_{ij}$. We have

$$\rho_{ij} = \frac{n^{-1}\sum_{h=1}^{n}\pi_{i,T+h}\pi_{j,T+h} - \bar{\pi}_{i,n}\bar{\pi}_{j,n}}{\left[n^{-1}\sum_{h=1}^{n}\pi^2_{i,T+h} - \bar{\pi}^2_{i,n}\right]^{1/2}\left[n^{-1}\sum_{h=1}^{n}\pi^2_{j,T+h} - \bar{\pi}^2_{j,n}\right]^{1/2}},$$

where $\bar{\pi}_{i,n} = n^{-1}\sum_{h=1}^{n}\pi_{i,T+h}$, and

$$\hat{\rho}_{ij} = \frac{n^{-1}\sum_{h=1}^{n}\pi_{i,T+h-1}\pi_{j,T+h-1} - \bar{\pi}^f_{i,h}\bar{\pi}^f_{j,h}}{\left[n^{-1}\sum_{h=1}^{n}\pi^2_{i,T+h-1} - \left(\bar{\pi}^f_{i,n}\right)^2\right]^{1/2}\left[n^{-1}\sum_{h=1}^{n}\pi^2_{j,T+h-1} - \left(\bar{\pi}^f_{j,n}\right)^2\right]^{1/2}},$$

where $\bar{\pi}^f_{i,n} = n^{-1}\sum_{h=1}^{n}\pi_{i,T+h-1}$. But

$$\begin{aligned}
\bar{\pi}_{i,n} - \bar{\pi}^f_{i,n} &= n^{-1}(\pi_{i,T+n} - \pi_{i,T}), \\
n^{-1}\sum_{h=1}^{n}\pi_{i,T+h}\pi_{j,T+h} - n^{-1}\sum_{h=1}^{n}\pi_{i,T+h-1}\pi_{j,T+h-1} &= n^{-1}(\pi_{i,T+n}\pi_{j,T+n} - \pi_{i,T}\pi_{j,T}), \\
n^{-1}\sum_{h=1}^{n}\pi^2_{i,T+h} - n^{-1}\sum_{h=1}^{n}\pi^2_{i,T+h-1} &= n^{-1}(\pi^2_{i,T+n} - \pi^2_{i,T}),
\end{aligned}$$

and it is easily seen that $\hat{\rho}_{ij} - \rho_{ij} \to 0$, as $n \to \infty$. We found very close matches between ρ_{ij} and $\hat{\rho}_{ij}$ for random walk models even for n relatively small.

An alternative measure, which is not subject to the above shortcoming, is to consider how well the cross section dispersion of the forecasts at each point in time match the dispersion of the realizations at the same point in time. Different measures of cross section dispersion can be used for this purpose. One could even examine the dispersion of realizations and forecasts over certain ranges of special interest, such as the tails of the cross section distributions. Some of these issues have been already addressed extensively in the literature on convergence and size distribution of incomes. Here we focus on standard deviations as our measure of dispersion. But other dispersion measures such as the numerator of the Gini coefficient can also be used.[1]

Let $y_{i,T+h}$ and $\hat{y}_{i,T+h}$ be the actual and the associated one quarter ahead forecast, respectively, of a particular variable for a certain country $i = 1, ..., m$, at some quarterly date $h = 1, 2, ..., n$ of the forecast evaluation sample. Then a dispersion ratio measure is defined as the ratio of cross section dispersion of actual to forecast values, computed as

$$\begin{aligned} DR_{T+h} &= \hat{s}_{T+h}/s_{T+h}, \text{ for } h = 1, 2, ..., n, \\ \hat{s}^2_{T+h} &= \textstyle\sum_{i=1}^{m} (\hat{y}_{i,T+h} - \bar{\hat{y}}_{T+h})/m, \ \bar{\hat{y}}_{T+h} = \sum_{i=1}^{m} \hat{y}_{i,T+h}/m, \\ s^2_{T+h} &= \textstyle\sum_{i=1}^{m} (y_{i,T+h} - \bar{y}_{T+h})^2/m, \ \ \bar{y}_{T+h} = \sum_{i=1}^{m} y_{i,T+h}/m. \end{aligned}$$

For a good match between forecasts and realizations we would expect the DR_{T+h} ratio to be close to unity. Large deviations of the dispersion ratio from unity, in either direction, can be viewed as failure of the forecasts to adequately track the cross section dispersion of the realizations. A useful summary measure can then be computed as $\sqrt{\sum_{h=1}^{n} (DR_{T+h} - 1)^2/n}$, which we refer to as the root mean square error of forecast dispersions (RMSEFD).

In view of the trended nature of real output, real equity prices, and the real exchange rate, the cross section dispersions for these variables are computed using their first differences, namely denoting the levels of these variables by $x_{i,T+h}$ and their one-step ahead forecasts by $\hat{x}_{i,T+h}$, respectively, we use $\hat{y}_{i,T+h} = \hat{x}_{i,T+h} - x_{i,T+h-1}$ and $y_{i,T+h} = \Delta x_{i,T+h}$ when computing the dispersion ratios. For inflation, short term, and long term interest rates we set $\hat{y}_{i,T+h} = \hat{x}_{i,T+h}$ and $y_{i,T+h} = x_{i,T+h}$.

The results are summarized in Tables R2a to R2f. A given row shows the dispersion ratios across models for one of the out-of-sample quarters, and a column shows the set of dispersion ratios across quarters for a given model. For real output, short term interest rate, and exchange rate, the AveAve forecasts track the actual dispersions the closest. For real equity prices this is also the case for two of the country sub-groupings (10 industrialised, and all countries excluding Latin America). For

[1] The relationship between alternative measures of dispersions is discussed in Pesaran (2007), where the average pair-wise differences over the cross section units is shown to be proportional to the standard deviation or the numerator of the Gini coefficient, depending on whether the square or absolute values (respectively) are used to represent the pair-wise differences.

long rates, most models had dispersion ratios of one. Note that interest rates are modeled as levels while output growth and inflation are, of course, differences. Not surprisingly it is easier to predict dispersion for industrialised countries, especially for output growth where the AveAve RMSEFD is just 0.31 (as compared to a perfect score of zero). Once we add in emerging markets, forecasting a coherent system of variables becomes harder: the RMSEFD of AveAve forecasts increases to 0.40 when we add in Asian emerging markets and the Middle East (but leave out Latin American countries), and rises further to 0.51 when all countries are considered.

That it is easy to trace out the cross sectional dispersion of the interest rate variables may not be surprising given that they are strongly influenced by monetary policy which tends to be more coordinated across the industrialised countries than elsewhere. So, for instance, the RMSFED is 0.05 or less for all models for the group of 10 industrialised countries plus China, but more than doubles when we move to the second country grouping (Table R2c). A similar result also holds for the long rate (Table 2d). To be sure, it is hard to consistently predict the cross sectional dispersion in every period as is clear from looking down the rows in Tables R2a-R2f, but overall the AveAve forecasts seem to perform relatively well. However, whether the differences of the dispersion ratios across different forecasting models are statistically significant is an important issue which lies outside the scope of the present rejoinder and must be addressed in future research.

4.3 Other Issues

A number of other issues of import have also been raised by several discussants. Lahiri would like us to incorporate parameter estimation uncertainty more explicitly into the forecast evaluation exercise. He also rightly draws attention to the possible cross section dependence of forecast errors - as he puts it, "..the average of the variances of forecast errors from different models is not the same as the variance of the average error." The possible cross section dependence of the forecast errors is clearly a concern for the panel DM test that we have developed, and this is now acknowledged in the published version of the paper (see Section 5.3). Although it is also possible to develop panel DM tests that allow for some forms of cross section dependence of forecast errors, in general such tests would require large sample sizes over time and across the units. In our applications the time dimension is too short for a meaningful use of a panel DM test that allows for cross section dependence. In any case, if we did have sufficiently long time series of forecast errors we could have applied the standard version of DM test (or its various extensions) to individual pairs of forecast errors without any need for a panel version of the test. It is, however, important to bear in mind that so long as the degree of cross section dependence is weak (in the sense defined in Chudik, Tosetti and Pesaran, 2009), AveAve forecasts are not adversely affected by the cross section dependence of the forecasts, since under weak cross section dependence the variance of the forecast error of the average forecast tends to zero as the number of forecasts under consideration becomes sufficiently large.

Sinclair and Stekler ask about data revisions, and real time updating. This is clearly an important consideration in "real time econometrics" and would be a welcome new research agenda in GVAR modelling and forecasting. To do so will be challenging as real time data vintages are currently available only for a few of the 33 countries included in the GVAR.

A number of discussants, such as Granger and Lahiri, commenting on an earlier version of the paper were justly surprised by the lower RMSFE we reported for four step ahead forecasts as compared to one step ahead forecasts. One would have expected the longer horizon forecasts to have a larger RMSFE as compared to the shorter horizon forecasts. This continues to apply to our forecasts. Previously we had divided the four-step ahead forecast errors by 4 to make them comparable to quarterly forecasts. To avoid confusions in the published version we now show the RMSFE of the four quarter ahead forecasts without such scaling, and there are no surprises in this regard.

5 Four More Quarters: Additional Forecast Comparisons

As is often the case in research publication, some time has passed since our paper was first reviewed, and more data has become available since then. We have been able to augment our GVAR data set with an additional year of complete data for all variables in all countries. While it is just one more year, it does increase our forecasting sample by 50%, and perhaps we can somewhat assuage the concern raised by Giannone and Reichlin that the forecasting period is both short and somewhat uneventful. Here we present one quarter ahead RMSFEs as well as panel DM statistics for now 12 quarters out-of-sample. Importantly, the results are based on the same model specification as before. And as before, we update the model parameters recursively every quarter.[2]

The results are presented in Tables R3a and R3b, where R3a shows the one-quarter ahead RMSFEs, and Table R3b shows the panel-DM statistics comparing the GVAR AveAve model to the four univariate benchmarks. Also to address comments by Allen, Clements and Sinclair and Stekler on the value added of the "global" aspect of the GVAR for forecasting, we now include a multivariate benchmark where forecasts are generated from country-specific VARs in domestic variables and oil prices, but without the global interaction terms (or 'star' variables). The main results presented in the paper, based on just 8 quarters out-of-sample, are unchanged by the addition of another 4 quarters. If anything, they are strengthened. For instance, in the case of real output growth, the DM statistic for AveAve model performance is now significant at the 1% level for all countries against all the univariate benchmarks; it was significant at the 1% level for two of the four benchmarks before (AR(1) and RW), at the 5% for one (RW with drift), and insignificant against one (AR(1) with trend). For the remaining variables the performance is largely unchanged.

[2]The extended GVAR data set to the end of 2006, and the related Gauss codes are available upon request.

Turning to the multivariate benchmark, the country-specific VAR models are estimated with linear trends recursively using expanding windows. The order of the VAR is selected using the Akaike criterion with a maximum lag order of 5. The VAR forecasts perform rather poorly for real output, but do reasonably for inflation. However, overall they are dominated by GVAR-AveAve forecasts for almost all variables and country groupings. These results are generally statistically significant as can be seen from the panel DM test statistics reported in Table R3b. Although it is not reported here, the performance of the multivariate benchmark is unaffected if we drop the trend term, but improves by averaging the VAR forecasts across the ten estimation windows. Some marginal improvements also result if we estimate the country-specific VAR models in first differences. But the overall advantage of GVAR-AveAve forecasts over the country-specific VAR models without the global inter-connections remains.

6 A Final Word

We have tried to respond as constructively as possible to many interesting and insightful comments. But we are aware that we have not been able to respond to all the comments, and some of our responses might have missed important points. It is hoped that the present inter-change would be helpful to other researchers interested in multi-country modelling and forecasting. The GVAR framework can also be used in the analysis of other inter-dependencies such as across industries, sectors, and regions. In such applications 'star' variables need to be constructed using input-ouput data or regional transactions. Clearly, there is a great deal of further research to be done.

References

Bańbura, M., Giannone, D. & Reichlin, L. (2009). Large Bayesian VARs. Forthcoming, *Journal of Applied Econometrics*.

Chudik, A., Pesaran, M. H. &Tosetti, E. (2009). Weak and strong cross section dependence and estimation of large panels. CESifo Working Paper Series No. 2689.

Dees, S., di Mauro, F., Pesaran, M. H. & Smith L. V.(2007). Exploring the international linkages of the euro area: A global VAR analysis. *Journal of Applied Econometrics*, 22, 1-38.

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

Granger, C. W. J. & Pesaran M. H. (2000). A decision-based approach to forecast evaluation, in W.S. Chan, W.K. Li and H. Tong (eds.), *Statistics and Finance: An Interface*, Imperial College Press, London.

Pesaran, M. H. (2007). A pair-wise approach to testing for output and growth convergence. *Journal of Econometrics*, 138, 312-355.

Pesaran, M. H. & Pick A. (2009). Forecast combinations across estimation windows. Memo, Cambridge University.

Pesaran M. H., Schleicher, C. & Zaffaroni P. (2009). Model averaging in risk management with an application to futures markets. *Journal of Empirical Finance*, 16, 280-305.

Pesaran, M. H., Schuermann, T., Treutler, B-J. & Weiner S. M. (2006). Macroeconomic dynamics and credit risk: A global perspective. *Journal of Money Credit and Banking*, 38, 1211-1262.

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

Pesaran, M. H. & Skouras S. (2002). Decision-based methods for forecast evaluation, in M.P. Clements and D.F. Hendry (eds.), A Companion to Economic Forecasting, Oxford: Basil Blackwell, Chapter 11, 241-267.

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

Table R1. The Relative Bias of Alternative One Quarter Ahead Forecasts Averaged Across Countries over the Horizon 2004Q1-2005Q4

(In Per Cent)

	Real Output			Inflation		
	Group1	Group2	Group3	Group1	Group2	Group3
DdPS-GVAR-AveW	0.27	0.40	0.41	0.28	0.35	0.39
DHPS-GVAR-AveW	0.32	0.45	0.47	0.66	0.57	0.58
GVAR-AveAve	0.30	0.41	0.44	0.28	0.26	0.29
AR(1)	0.44	0.47	0.51	0.38	0.38	0.47
AR(1) with trend	0.50	0.52	0.54	0.57	0.49	0.50
RW	0.81	0.84	0.84	0.05	0.11	0.12
RW with drift	0.38	0.47	0.50	0.08	0.13	0.15
	Short-Term Interest Rate			Long-Term Interest Rate		
	Group1	Group2	Group3	Group1	Group2†	Group3
DdPS-GVAR-AveW	0.48	0.51	0.54	0.23	0.25	0.25
DHPS-GVAR-AveW	0.55	0.59	0.63	0.31	0.32	0.32
GVAR-AveAve	0.30	0.36	0.36	0.24	0.24	0.24
AR(1)	0.54	0.57	0.65	0.30	0.31	0.31
AR(1) with trend	0.80	0.69	0.69	0.49	0.45	0.45
RW	0.47	0.44	0.42	0.33	0.31	0.31
RW with drift	0.64	0.58	0.59	0.26	0.27	0.27
	Real Equity			Real Exchange Rate		
	Group1	Group2	Group3	Group1	Group2	Group3
DdPS-GVAR-AveW	0.46	0.51	0.50	0.16	0.26	0.27
DHPS-GVAR-AveW	0.46	0.53	0.52	0.39	0.34	0.33
GVAR-AveAve	0.40	0.38	0.36	0.21	0.24	0.28
AR(1)	0.58	0.57	0.56	0.23	0.37	0.41
AR(1) with trend	0.46	0.43	0.44	0.22	0.37	0.38
RW	0.65	0.58	0.56	0.22	0.35	0.41
RW with drift	0.48	0.43	0.41	0.20	0.29	0.34

Notes: Group1, Group2 and Group3 correspond to “10 Industrialised Plus China”, “All Countries Excluding LA” and “All Countries”, respectively. For real output and inflation the grouping “All Countries Excluding LA” comprises 21 countries, while that of “All Countries” comprises all the 26 countries/regions in the GVAR model. For the short-term interest “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. For the long-term interest rate “All Countries” comprises the 10 industralised countries plus South Korea and South Africa. For real equity prices “All Countries Excluding LA” comprises 17 countries, while that of “All Countries” comprises 19 countries. For the real exchange rate “All Countries Excluding LA” 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. The relative forecast bias statistics are computed as the ratio of Abs(bias) to RMSFE averaged across countries in the country group under consideration and are shown in per cent.

†For the case of the long-term interest rate country groupings 2 and 3 are identical.

Table R2a. Ratio of Cross Section Dispersion of Actual to Forecast Values for Real Output

	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
	10 Industrialised Plus China				
2004Q1	0.81	0.81	0.86	0.00	0.79
2004Q2	0.85	0.83	0.87	0.00	0.80
2004Q3	0.87	0.83	0.89	0.00	0.81
2004Q4	0.82	0.82	0.88	0.00	0.79
2005Q1	0.42	0.47	0.50	0.00	0.45
2005Q2	0.46	0.36	0.40	0.00	0.34
2005Q3	1.16	0.58	0.60	0.00	0.56
2005Q4	0.87	0.77	0.82	0.00	0.74
RMSEFD	0.31	0.36	0.33	1.00	0.38
	All Countries Excluding LA				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.59	0.45	0.51	0.00	0.52
2004Q2	0.97	0.68	0.74	0.00	0.77
2004Q3	0.51	0.45	0.49	0.00	0.52
2004Q4	0.72	0.57	0.61	0.00	0.64
2005Q1	0.45	0.32	0.34	0.00	0.36
2005Q2	0.52	0.37	0.39	0.00	0.41
2005Q3	0.95	0.47	0.49	0.00	0.54
2005Q4	0.50	0.41	0.42	0.00	0.46
RMSEFD	0.40	0.55	0.51	1.00	0.49
	All Countries				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.52	0.46	0.48	0.00	0.47
2004Q2	0.59	0.47	0.48	0.00	0.48
2004Q3	0.42	0.39	0.40	0.00	0.40
2004Q4	0.48	0.41	0.44	0.00	0.42
2005Q1	0.36	0.28	0.34	0.00	0.30
2005Q2	0.48	0.36	0.39	0.00	0.38
2005Q3	0.88	0.48	0.55	0.00	0.51
2005Q4	0.38	0.32	0.39	0.00	0.35
RMSEFD	0.51	0.61	0.57	1.00	0.59

Notes: Each quarterly entry in the above table is the dispersion ratio computed as $DR_{T+h} = \hat{s}_{T+h}/s_{T+h}$, for $h = 1, 2, ..., n$ $(n = 8)$, where $\hat{s}^2_{T+h} = \sum_{i=1}^{m}(\hat{y}_{i,T+h} - \bar{\hat{y}}_{T+h})/m$, $\bar{\hat{y}}_{T+h} = \sum_{i=1}^{m} \hat{y}_{i,T+h}/m$, $s^2_{T+h} = \sum_{i=1}^{m}(y_{i,T+h} - \bar{y}_{T+h})^2/m$, and $\bar{y}_{T+h} = \sum_{i=1}^{m} y_{i,T+h}/m$. The variables $y_{i,T+h}$ and $\hat{y}_{i,T+h}$ are the actual and one quarter ahead forecast values respectively, of a particular country $i = 1, ..., m$, where m is the number of countries in the specified country group. RMSEFD is the Root Mean Square Error of Forecast Dispersions computed as $\sqrt{\sum_{h=1}^{8}(DR_{T+h} - 1)^2/8}$.

Table R2b. Ratio of Cross Section Dispersion of Actual to Forecast Values for Inflation

	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
	10 Industrialised Plus China				
2004Q1	1.09	1.12	1.18	0.00	1.28
2004Q2	0.60	0.57	0.55	0.00	0.73
2004Q3	1.23	1.21	1.49	1.50	1.51
2004Q4	0.31	0.35	0.36	0.42	0.42
2005Q1	1.47	1.55	1.37	1.80	1.80
2005Q2	0.58	0.55	0.46	0.70	0.69
2005Q3	1.03	1.02	0.91	1.22	1.22
2005Q4	2.12	1.97	1.74	2.42	2.40
RMSEFD	0.54	0.51	0.48	0.66	0.66
	All Countries Excluding LA				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	1.45	2.38	2.59	1.31	1.30
2004Q2	0.61	1.63	1.71	0.69	0.67
2004Q3	1.04	1.61	1.69	0.91	0.88
2004Q4	0.71	1.70	1.71	1.03	1.00
2005Q1	0.93	1.62	1.61	0.99	0.97
2005Q2	0.77	1.37	1.35	0.97	0.96
2005Q3	1.06	1.69	1.67	1.11	1.09
2005Q4	0.33	0.55	0.54	0.36	0.35
RMSEFD	0.34	0.74	0.80	0.28	0.28
	All Countries				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	1.65	3.05	2.05	1.10	1.08
2004Q2	1.45	2.74	1.78	0.88	0.85
2004Q3	1.62	2.49	1.66	0.91	0.87
2004Q4	1.14	2.43	1.63	0.97	0.93
2005Q1	1.06	2.00	1.33	0.85	0.83
2005Q2	1.95	2.35	1.27	1.12	1.09
2005Q3	1.01	2.53	1.54	1.11	1.07
2005Q4	0.90	0.93	0.57	0.40	0.39
RMSEFD	0.49	1.45	0.63	0.23	0.24

See the notes to Table R2a.

Table R2c. Ratio of Cross Section Dispersion of Actual to Forecast Values for Short-Term Interest Rate

	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
	10 Industrialised Plus China				
2004Q1	1.06	0.99	0.95	0.00	0.95
2004Q2	0.99	0.99	0.94	0.00	0.95
2004Q3	1.03	0.99	0.94	0.97	0.96
2004Q4	1.00	1.02	0.97	1.00	0.99
2005Q1	0.96	1.00	0.95	0.99	0.98
2005Q2	0.97	1.00	0.95	0.98	0.98
2005Q3	1.03	1.02	0.97	1.00	1.00
2005Q4	0.99	0.97	0.92	0.96	0.95
RMSEFD	0.03	0.02	0.05	0.03	0.03
	All Countries Excluding LA				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	1.09	1.19	1.22	1.09	1.11
2004Q2	1.05	1.12	1.13	1.02	1.03
2004Q3	1.02	1.08	1.08	0.99	1.00
2004Q4	0.98	1.14	1.13	1.04	1.05
2005Q1	1.13	1.24	1.22	1.13	1.14
2005Q2	1.07	1.15	1.12	1.03	1.04
2005Q3	0.98	1.10	1.06	0.99	0.99
2005Q4	0.91	1.12	1.08	1.00	1.01
RMSEFD	0.07	0.15	0.14	0.06	0.07
	All Countries				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	1.20	1.99	1.39	1.10	1.11
2004Q2	1.45	1.92	1.30	1.02	1.03
2004Q3	1.29	1.87	1.25	0.99	1.00
2004Q4	0.93	1.88	1.26	1.01	1.02
2005Q1	1.06	1.96	1.30	1.05	1.06
2005Q2	1.87	1.92	1.24	1.00	1.00
2005Q3	0.84	1.91	1.23	0.99	1.00
2005Q4	1.49	1.98	1.27	1.04	1.04
RMSEFD	0.41	0.93	0.28	0.04	0.05

See the notes to Table R2a.

Table R2d. Ratio of Cross Section Dispersion of Actual to Forecast Values for Long-Term Interest Rate

	10 Industrialised				
	AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	1.03	1.02	1.00	0.00	1.01
2004Q2	0.99	1.03	1.01	0.00	1.02
2004Q3	1.02	0.99	0.97	0.99	0.99
2004Q4	0.98	0.99	0.96	0.99	0.98
2005Q1	0.97	0.96	0.93	0.96	0.96
2005Q2	1.00	1.00	0.97	1.00	1.00
2005Q3	1.03	1.05	1.01	1.04	1.04
2005Q4	1.02	1.00	0.96	0.99	0.98
RMSEFD	0.02	0.03	0.04	0.02	0.02
	All Countries (12)				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.96	0.99	0.96	0.95	0.95
2004Q2	0.99	1.01	0.98	0.97	0.97
2004Q3	1.04	1.06	1.04	1.02	1.03
2004Q4	1.06	1.09	1.07	1.06	1.06
2005Q1	1.08	1.10	1.07	1.07	1.07
2005Q2	0.94	0.98	0.95	0.95	0.95
2005Q3	1.05	1.07	1.04	1.03	1.04
2005Q4	1.02	1.05	1.02	1.02	1.02
RMSEFD	0.05	0.06	0.05	0.04	0.05

See the notes to Table R2a.

Table R2e. Ratio of Cross Section Dispersion of Actual to Forecast Values for Real Equity

	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
	10 Industrialised				
2004Q1	0.38	0.15	0.39	0.00	0.10
2004Q2	0.61	0.22	0.49	0.00	0.14
2004Q3	0.42	0.19	0.41	0.00	0.12
2004Q4	0.48	0.21	0.45	0.00	0.13
2005Q1	0.58	0.31	0.62	0.00	0.19
2005Q2	0.66	0.39	0.71	0.00	0.22
2005Q3	0.31	0.17	0.32	0.00	0.11
2005Q4	0.35	0.15	0.24	0.00	0.08
RMSEFD	0.54	0.78	0.56	1.00	0.86
	All Countries Excluding LA (17)				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.60	0.22	0.38	0.00	0.13
2004Q2	0.50	0.22	0.31	0.00	0.11
2004Q3	0.42	0.23	0.35	0.00	0.13
2004Q4	0.43	0.20	0.30	0.00	0.11
2005Q1	0.65	0.37	0.54	0.00	0.19
2005Q2	0.84	0.35	0.49	0.00	0.17
2005Q3	0.26	0.19	0.27	0.00	0.10
2005Q4	0.34	0.20	0.31	0.00	0.09
RMSEFD	0.52	0.75	0.64	1.00	0.87
	All Countries (19)				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.50	0.52	0.70	0.00	0.13
2004Q2	0.39	0.72	0.82	0.00	0.11
2004Q3	0.43	0.69	0.84	0.00	0.16
2004Q4	0.47	0.44	0.53	0.00	0.11
2005Q1	0.82	1.39	1.48	0.00	0.22
2005Q2	0.81	1.55	1.57	0.00	0.21
2005Q3	0.30	0.83	0.83	0.00	0.12
2005Q4	0.41	0.80	0.77	0.00	0.11
RMSEFD	0.52	0.39	0.35	1.00	0.85

See the notes to Table R2a.

Table R2f. Ratio of Cross Section Dispersion of Actual to Forecast Values for Real Exchange Rate

	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
	9 Industrialised Plus China				
2004Q1	0.25	0.08	0.13	0.00	0.12
2004Q2	0.37	0.08	0.14	0.00	0.13
2004Q3	0.56	0.16	0.23	0.00	0.23
2004Q4	0.17	0.07	0.10	0.00	0.10
2005Q1	0.82	0.18	0.25	0.00	0.25
2005Q2	0.34	0.15	0.20	0.00	0.21
2005Q3	0.28	0.09	0.13	0.00	0.14
2005Q4	0.30	0.11	0.16	0.00	0.17
RMSEFD	0.64	0.89	0.84	1.00	0.83
	All Countries Excluding LA				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.40	0.10	0.31	0.00	0.12
2004Q2	0.41	0.11	0.32	0.00	0.12
2004Q3	0.68	0.23	0.57	0.00	0.24
2004Q4	0.19	0.09	0.23	0.00	0.10
2005Q1	0.39	0.15	0.32	0.00	0.14
2005Q2	0.37	0.21	0.42	0.00	0.19
2005Q3	0.39	0.15	0.33	0.00	0.15
2005Q4	0.24	0.12	0.23	0.00	0.11
RMSEFD	0.63	0.86	0.67	1.00	0.85
	All Countries				
	GVAR-AveAve	AR(1)	AR(1) with trend	RW	RW with drift
2004Q1	0.48	0.12	0.38	0.00	0.16
2004Q2	0.49	0.12	0.37	0.00	0.16
2004Q3	0.62	0.21	0.55	0.00	0.26
2004Q4	0.27	0.10	0.27	0.00	0.13
2005Q1	0.47	0.15	0.36	0.00	0.17
2005Q2	0.29	0.13	0.30	0.00	0.14
2005Q3	0.40	0.12	0.31	0.00	0.16
2005Q4	0.39	0.11	0.25	0.00	0.13
RMSEFD	0.58	0.87	0.66	1.00	0.84

Note: 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. Also see the notes to Table R2a.

Table R3a. RMSFEs Over the Extended Horizon 2004Q1-2006Q4 for Different Country Groupings Using the GVAR and Selected Benchmark Models

	Real Output Growth			Inflation		
	Group1	Group2	Group3	Group1	Group2	Group3
DdPS-GVAR-AveW	0.543	0.833	0.933	0.408	0.760	1.214
DHPS-GVAR-AveW	0.517	0.835	0.942	0.476	0.770	1.288
GVAR-AveAve	0.479	0.730	0.818	0.389	0.653	0.809
Univariate Benchmarks						
AR(1)	0.568	0.750	0.869	0.415	0.745	1.158
AR(1) with trend	0.546	0.735	0.850	0.457	0.738	0.828
RW	1.053	1.349	1.424	0.427	0.684	0.684
RW with drift	0.535	0.743	0.846	0.428	0.688	0.690
Multivariate Benchmark						
VAR‡	0.558	0.791	0.947	0.412	0.694	1.022
	Short-Term Interest Rate			Long-Term Interest Rate		
	Group1	Group2	Group3	Group1	Group2†	Group3
DdPS-GVAR-AveW	0.092	0.164	0.661	0.083	0.093	0.093
DHPS-GVAR-AveW	0.080	0.153	0.792	0.076	0.087	0.087
GVAR-AveAve	0.063	0.110	0.284	0.065	0.074	0.074
Univariate Benchmarks						
AR(1)	0.055	0.119	0.622	0.061	0.072	0.072
AR(1) with trend	0.093	0.138	0.288	0.073	0.080	0.080
RW	0.054	0.097	0.109	0.060	0.070	0.070
RW with drift	0.060	0.103	0.121	0.060	0.070	0.070
Multivariate Benchmark						
VAR	0.113	0.198	0.920	0.076	0.084	0.084
	Real Equity			Real Exchange Rate		
	Group1	Group2	Group3	Group1	Group2	Group3
DdPS-GVAR-AveW	6.198	7.560	7.757	3.982	3.659	3.751
DHPS-GVAR-AveW	5.876	7.217	7.289	4.484	4.229	4.252
GVAR-AveAve	5.291	5.891	6.089	3.814	3.487	3.469
Univariate Benchmarks						
AR(1)	5.592	6.271	6.973	3.632	3.485	3.435
AR(1) with trend	5.047	5.660	6.339	3.516	3.475	3.372
RW	5.838	6.229	6.450	3.570	3.405	3.388
RW with drift	5.071	5.546	5.792	3.492	3.264	3.227
Multivariate Benchmark						
VAR	6.304	8.656	8.963	3.824	3.597	3.822

Notes: DdPS-GVAR-AveW and DHPS-GVAR-AveW refer to average forecasts across 10 estimation windows using DdSP-GVAR and DHPS-GVAR models, respectively. DdPS-GVAR denotes the GVAR model with exactly identified 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 stands for the random walk model, and LA for Latin America. Parameters of the benchmark models are estimated recursively using an expanding window. Group1, Group2 and Group3 correspond to "10 Industrialised Plus China", "All Countries Excluding LA" and "All Countries", respectively. See also notes to Table R2. All table entries are simple cross-country averages of RMSFE's in per cent based on one quarter ahead forecasts.

†For the long-term interest rate country groupings 2 and 3 are identical.

‡The VAR models are estimated with linear trends recursively using expanding windows. The lag length is chosen using the Akaike criterion with a maximum lag order of 5. The oil price variable is included as an endogenous variable in all country-specific VAR models.

Table R3b. Panel DM Statistics Over the Extended Horizon 2004Q1-2006Q4 for GVAR-AveAve Forecasts Relative to a Selected Number of Benchmarks

	$z_{ijt} = [e^A_{ijt}(1)]^2 - [e^B_{ijt}(1)]^2$					
	Real Output Growth			Inflation		
Univariate Benchmarks	Group1	Group2	Group3	Group1	Group2	Group3
AR(1)	-1.250	-1.161	-2.514	-1.260	-4.512	-10.406
AR(1) with trend	-2.123	-1.171	-2.528	-2.728	-3.693	-1.225
RW	-5.831	-10.123	-11.873	-2.320	-0.868	2.236
RW with drift	-1.226	-1.354	-2.264	-2.387	-0.931	2.148
Multivariate Benchmark						
VAR‡	-1.378	-1.640	-3.901	-1.160	-0.222	-2.463
	Short-Term Interest Rate			Long-Term Interest Rate		
Univariate Benchmarks	Group1	Group2	Group3	Group1	Group2†	Group3
AR(1)	1.399	-2.864	-17.132	1.177	0.002	0.002
AR(1) with trend	-4.260	-2.503	0.534	-1.689	-1.217	-1.217
RW	1.977	1.645	3.667	1.397	1.057	1.057
RW with drift	0.925	1.129	3.627	1.331	0.891	0.891
Multivariate Benchmark						
VAR	-4.030	-3.137	-7.922	-2.919	-1.900	-1.900
	Real Equity			Real Exchange Rate		
Univariate Benchmarks	Group1	Group2	Group3	Group1	Group2	Group3
AR(1)	-1.495	-2.623	-4.157	1.611	0.455	0.555
AR(1) with trend	0.872	0.563	-2.368	2.195	-0.053	0.589
RW	-2.378	-2.242	-2.213	1.717	1.090	1.046
RW with drift	0.864	1.961	1.027	2.733	2.873	3.336
Multivariate Benchmark						
VAR	-3.527	-7.226	-6.779	-0.205	-1.223	-2.848

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. Group1, Group2 and Group3 correspond to the country groupings "10 Industrialised Plus China", "All Countries Excluding LA" and "All Countries", respectively. See note to Table R2 for further details. For a one sided test, which is appropriate here, the 1% and 5% critical values are -2.326 and -1.645, respectively. A positive value of the panel DM statistic represents evidence against the AveAve forecasts.

†For the long-term interest rate country groupings 2 and 3 are identical.

‡The VAR models are estimated with linear trends recursively using an expanding window. The lag length is chosen using the Akaike criterion with the maximum lag order of 5. The oil price variable is included as an endogenous variable in all individual country VAR models.