CFCM

CENTRE FOR FINANCE, CREDIT AND MACROECONOMICS

Working Paper 14/06

Forecasting Global Recessions in a GVAR Model of Actual and Expected Output in the G7

Anthony Garratt, Kevin Lee and Kalvinder Shields

Produced By:

Centre for Finance, Credit and Macroeconomics
School of Economics
Sir Clive Granger Building
University of Nottingham
University Park
Nottingham
NG7 2RD

Tel: +44(0) 115 951 5619
Fax: +44(0) 115 951 4159
enquiries@cfcm.org.uk

Forecasting Global Recessions in a GVAR Model of Actual and Expected Output in the G7*

by

Anthony Garratt,† Kevin Lee†† and Kalvinder Shields†††

Abstract

The forecasting performance of a Global VAR model of actual and expected outputs in the G7 economies is compared with that of alternative models to judge the usefulness of modelling cross-country interdependencies and employing survey data. Both effects are found to be important in calculating density forecasts, in forecasting the occurrence of recessionary events defined at the national and G7-wide levels and, through a novel 'fair bet' exercise, in decision-making based on forecasts. The analysis argues for a nuanced approach to presenting output predictions, avoiding simple point forecasts and focusing on features of future growth dynamics relevant to decision-makers.

Keywords: Cross-country interactions, Survey expectations, Probability Forecasts, Global and National Recession, Forecast evaluation

JEL Classification: C53, E32, E37

*†University of Warwick, UK; †University of Nottingham, UK; ††University of Melbourne, Australia. Version dated April 2014. Corresponding author: Tony Garratt, Warwick Business School, The University of Warwick, Coventry, CV4 7AL, UK. E-mail: Anthony.Garratt@wbs.ac.uk . Tel: (+44) 024 761 50052.

1 Introduction

This paper investigates the importance of two aspects of forecasting business cycles highlighted during the slowdown in economic growth that followed the financial crisis of 2007. *First,* given the global reach of the slowdown, we investigate the importance of incorporating cross-country interactions in a forecasting model. *Second*, we acknowledge the potential role of confidence and pessimism in initiating and propagating business cycle dynamics and the contribution of expectations data, obtained directly from surveys, to the calculation of output forecasts. The paper considers the importance of these factors for point forecasts and density forecasts of output outcomes in standard statistical terms. But it also focuses on the extent to which different models are able to forecast the likelihood of particular recessionary events and to make predictions that would be useful in decision-making. This is important given that little of the popular discussion of growth predictions focuses on the exact forecast of output at a specific time in the future; rather, the discussion concentrates on the likely occurrence of 'global recession, 'double-dip recessions', 'signs of green shoots', and so on; i.e. more broadly defined events occuring over some future interval.[1] Our forecast evaluation exercise considers the contribution of international interactions and expectations data in helping to forecast the likelihood of this sort of event and also considers a decision-making situation involving a bet that the events occur. This provides a novel perspective, and one which matches the popular view on forecasting output outcomes, on the usefulness of this data and of output modelling in general.

The paper investigate the importance of the two factors through a comparison of the forecasting performance of a range of models aimed at isolating the separate contributions of taking into account cross-country interdependencies and survey expectations. The most general model considered is a multi-country Global Vector Autoregressive (GVAR) model of actual and expected outputs in each country. The GVAR modelling framework is outlined in Pesaran, Schuermann and Weiner (2004), Garratt *et al.* (2006) and, in the

[1] Analysis of the nature and timing of business cycle events has also been thoroughly explored in the literature; see, for example, van Dijk *et al.*'s (2005) special issue of Journal of Applied Econometrics.

context of forecasting, in Pesaran, Schuermann and Smith (2009), *inter alia.* It uses trade-weighted averages of foreign variables to capture the effect of external influences in an otherwise-unconstrained VAR model of separate national models. The further inclusion of direct measures of expected outputs (at home and abroad) allows the model to accommodate the complex dynamic interactions that arise when decisions are made by forward-looking agents influenced by confidence and pessimism on current and future growth prospects at the national and international levels. The individual country models are then brought together in a single coherent GVAR system which accommodates the complexity of cross-country interactions while at the same time allowing for the sophisticated short-run dynamics found in the data.

Of course, there is no shortage of papers in the academic literature concerned with investigating cross-country interactions in the global business cycle, including the large-scale structural econometric systems of the United Nations' Project LINK, or the IMF's multi-regional model MULTIMOD, for example (see Laxton *et al.*, 1998).[2] There have also been many modelling exercises aiming to provide a statistical characterisation of macroeconomic variables across many countries, typically estimating dynamic factor models to identify global, nation-specific and idiosyncratic components in different series and across different sets of countries; see, for example Lumsdaine and Prasad (2003), Kose *et al.* (2003, 2008), del Negro and Otrok (2008), or Cruccini *et al.* (2011) These latter models are not typically used in forecasting, however, and while global interactions are at the heart of the forecasts delivered by the large structural models, it is not easy to isolate the contribution of the global effects in these models. We believe our GVAR analysis is well suited to this exercise. At the same time, our model can capture the influence of expectations effects at both the national and global levels through the inclusion of the direct measures of expectations.[3] The potential role of confidence and expectation formation in business cycle fluctuations has been explored recently in Akerlof and Shiller (2009)'s

[2]A recent IMF characterisation of periods of recession in advanced economies since the 1960's shows that recessions are deeper and last longer when they are synchronized across countries compared to those that are more localised; see Kannan *et al.* (2009).

[3]The usefulness of output forecasts obtained directly from surveys is explored in a multi-country context in Isiklar and Lahiri (2007).

discussion of ‘animal spirits’, for example, and in the analyses of information rigidities in Barsky and Sims (2012), Blanchard *et al* (2013), and Coibion and Gorodnichenko (2012), *inter alia.* This aspect of business cycle dynamics is potentially crucial in any forecasting exercise and can be readily incorporated using the GVAR methods we use in our models.

In a complementary paper, Garratt, Lee and Shields (2013) [GLS] use the same GVAR model of G7 outputs to examine the role of inter-country interactions and expectations in explaining output growths in the G7 over the period 1994q1-2013q1. In that paper, a variance-based measure of the persistent effect of shocks to the actual and expected outputs is applied to the estimated GVAR model to characterise countries' output dynamics, showing that the effects of the shocks are very complex and prolonged mainly because of the cross-country interactions that exist within the G7. A decomposition method is then used to evaluate the importance of global factors and the role of sentiment (as identified using the survey expectations data) in business cycle fluctuations at the infinite horizon. This shows that, on average, the split between the global and national influences on persistent movements in the countries' outputs is in the ratio 60:40 respectively, while around 30% of the permanent effects of shocks to output is explained in terms of sentiment compared to 70% related to fundamentals.

GLS demonstrates that there is a considerable role for cross-country interactions and survey data in modelling countries' output growths. The focus of the current paper goes beyond the description of model properties and considers the use of output growth models in producing forecasts and, in particular, in producing probability forecasts of recessionary events. The use of models in producing forecasts of the likely occurrence of different types of recessionary events, in addition to density forecasts and point forecasts of output growths, focuses attention on decision-making and the economic significance of forecasts to complement the more usual statistical assessment of their forecasting performance.[4]

As we shall see, the results show that, judged by statistical criteria, the performance of models that nowcast and forecast countries' outputs is considerably enhanced by taking

[4]See Garratt *et al.* (2003, 2006) for discussion of the estimation of event probability forecasts; and Clements (2006), Lahiri and Wang (2007, 2013), and Garratt and Lee (2010) for a discussion of their use in forecast evaluation.

into account international links and the information available in survey data. We also find that, relating to economic significance, both the expectations data and the international interactions are important in calculating density forecasts, in forecasting the occurrence of recessionary events defined at the national and G7-wide levels and, through a 'fair bet' exercise, in decision-making based on forecasts. Ultimately, the analysis argues for a nuanced approach to representing and evaluating output predictions, avoiding simple point forecasts and focusing on features of future growth dynamics relevant to decision-makers.

The layout of the remainder of the paper is as follows. Section 2 describes our modelling framework, explaining how our national models of actual and expected output growths are developed and brought together in the GVAR. Section 3 explains the use of density and probability forecasts in model evaluation using statistical and economic criteria, introducing a novel, generally-applicable approach to making an economic evaluation of forecasts over a range of different recessionary events based around a fair bet. Section 4 describes the GVAR model obtained for the G7 economies over the period 1994q1-2013q1 and the details of our forecasting exercise. Section 5 concludes with a brief summary of the findings.

2 Modelling and Forecasting Global Economic Outputs

2.1 Actual and expected outputs

The modelling framework of GLS uses measures of actual output that take into account that these data are typically published with a lag of one quarter so that agents are always unsure of the current state of the economy.[5] Survey data provide information on agents' perceptions of current output levels and expected future output levels. Surveys will certainly reflect the economic 'fundamentals' driving activity in the economy, contemporaneously and in the future, but these data could also incorporate the effects of

[5]The first-release data is also often revised introducing a further complexity in decision-making. As we explain below, in what follows, we ignore the revisions process, effectively assuming that subsequent revisions simply constitute noise. See Garratt *et al.* (2008, 2009) and Lee *et al.* (2012) for more discussion of the analysis of revisions data.

consumer or business confidence, uncertainty, learning or imitative/herding behaviour, and so on. These latter effects might in turn have an impact on actual and expected output levels at different phases of the business cycle. In any case, an analysis of the first-release actual output data and survey data provides a more accurate reflection of the innovations in output than can be obtained by looking at actual output data alone and this is potentially important in a study of output forecasts.

We denote (the logarithm of) output in country i at time t by $y_{i,t}$ and the measure of $y_{i,t}$ published in time $t+s$ by ${}_{t+s}y_{i,t}$. If $s \geqslant 1$, the measure is from an official publication (published after the one quarter publication delay). If $s \leqslant 0$, the measure is a direct measure of expectations on $y_{i,t}$ as published in $t+s$ (and the point is emphasised by a superscript 'e'). Focusing for expositional purposes on a single country i for the time being, a modelling framework that can accommodate the publication delays and the role of surveys on contemporaneous and future outputs is given by

$$\begin{bmatrix} {}_{t}y_{i,t-1} - {}_{t-1}y_{i,t-2} \\ {}_{t}y^{e}_{i,t} - {}_{t}y_{i,t-1} \\ {}_{t}y^{e}_{i,t+1} - {}_{t}y^{e}_{i,t} \end{bmatrix} = \Gamma_{i0} + \sum_{k=1}^{p} \Gamma_{ik} \begin{bmatrix} {}_{t-k}y_{i,t-1-k} - {}_{t-1-k}y_{i,t-2-k} \\ {}_{t-k}y^{e}_{i,t-k} - {}_{t-k}y_{i,t-1-k} \\ {}_{t-k}y^{e}_{i,t+1-k} - {}_{t-k}y^{e}_{i,t-k} \end{bmatrix} + \begin{bmatrix} \xi_{i,1t} \\ \xi_{i,2t} \\ \xi_{i,3t} \end{bmatrix} \quad (1)$$

for $t = 1, ..., T$ where the Γ_i's are (country-specific) matrices of parameters and the ξ_i's are mean zero innovations with variance-covariance Ω_i and where, as an illustration, we focus here on the case where only contemporaneous and one-period-ahead forecasts are used. This model explains, in the order of the variables in (1), the growth in actual output at time $t-1$ (published in time t following the one-quarter publication delay), the expected contemporaneous growth in output (published as a nowcast in the survey dated at time t), and the expected one-period ahead growth in output (also published in the survey dated at time t).[6] The model assumes that actual output growth is stationary and that expectational errors are stationary but is quite general otherwise.

The model is reasonably estimated in the form in (1) but it can be written in levels

[6]Note that this model involves expected growth terms obtained not through simple first-differencing of the (logs of) series but based on the (log) difference of different series.

form

$$\mathbf{y}_{i,t} = \mathbf{A}_{i0} + \sum_{k=1}^{p+1} \mathbf{A}_{ik}\ \mathbf{y}_{i,t-k} + \boldsymbol{\varepsilon}_{i,t}, \qquad t = 1, ..., T, \tag{2}$$

where $\mathbf{y}_{i,t} = (\ _t y_{i,t-1},\ _t y^e_{i,t},\ \ _t y^e_{i,t+1})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{i,at},\ \ \varepsilon_{i,bt},\ \ \varepsilon_{i,ct})' = (\xi_{i,1t},\ \ \xi_{i,1t} + \xi_{i,2t},\ \ \xi_{i,1t} + \xi_{i,3t} + \xi_{i,3t})'$ and the $\mathbf{A}_i$'s are functions of the original Γ_i's.[7] Its simple linear form makes (2) particularly suitable for forecasting and decision-making using simulation methods. Random draws from the estimated variance-covariance matrix Ω_i can be used to simulate future paths for the $\mathbf{y}_{i,t}$'s ($t = T+1, ..., T+h$), taking the estimated $\mathbf{A}_i$'s as being true and known. This generates the entire forecast density $\Pr(\mathbf{Y}_{i,T+1,T+h} \mid \mathbf{Y}_{i,1,T})$ showing the likelihood of observing $\mathbf{Y}_{i,T+1,T+h} = \{\mathbf{y}_{i,T},\ \mathbf{y}_{i,T+1},...,\mathbf{y}_{i,T+h}\}$ given the observed data $\mathbf{Y}_{i,1,T} = \{\mathbf{y}_{i,1},\ \mathbf{y}_{i,2},...,\mathbf{y}_{i,T}\}$ and taking into account the stochastic uncertainty surrounding the model. Alternatively, the estimated model can be used to generate artificial histories (using actual data for $t = 1, .., p+2$ and simulating data for $t = p+3, .., T$) each of which can be used to estimate an alternative version of (2) and to generate simulated future paths. The resultant density obtained across all simulated futures takes into account both the stochastic uncertainty and parameter uncertainty associated with the model. (see Garratt *et al.* (2003) and Garratt *et al.* (2006) for further detail and discussion).

2.2 Global interactions

The national model of output growth described in (2) can be readily extended to accommodate global interactions that arise because of the potential effects of common factors that drive output in many countries simultaneously. These could be justified through common productivity shocks (i.e. common fundamentals), for example, or through self-reinforcing outcomes of bouts of global pessimism or optimism which drive changes in risk premia across all countries, say (i.e. common drivers of sentiment). As elaborated in GLS, Dees *et al.* (2007) argue that unobservable common factors can be reasonably proxied by the inclusion of global aggregates in the country VARs. We can construct

[7] The model can also be written as a cointegrating VAR in the difference of $\mathbf{y}_{i,t}$ in which the assumed stationarity of the expectational errors is reflected in (two) cointegrating vectors that ensure the three output measures move together one-for-one in the long run.

global variables, $\mathbf{y}_t^* = \sum_{j=1}^{n} w_j \mathbf{y}_{j,t}$ using fixed weights w_j, and the corresponding growth series can then be used to supplement the original VAR in (1) in estimation. The national model in (2) can then be written

$$\mathbf{y}_{i,t} = \mathbf{B}_{i0} + \sum_{k=1}^{p+1} \mathbf{B}_{ik}\ \mathbf{y}_{i,t-k} + \sum_{k=0}^{p+1} \mathbf{B}_{ik}^*\ \mathbf{y}_{t-k}^* + \boldsymbol{\varepsilon}_{i,t}, \qquad i = 1, ..., n \quad \text{and} \quad t = 1, ..., T. \tag{3}$$

Here, the effects of the common factor are accommodated through the inclusion of the current and lagged values of the global variable in the models explaining individual countries' output growths. In practice, the $\mathbf{y}_t^*$ variable used in model (2) can be defined using country-specific weights, $\mathbf{y}_{i,t}^* = \sum_{j=1}^{n} w_{ij} \mathbf{y}_{j,t}$, where the weights are chosen so that the foreign variable can best capture the influence of different countries on country i (using trade volumes or some other metric, for example). Similarly, the order of the lags of $\mathbf{y}_{i,t}$ and $\mathbf{y}_{i,t}^*$ do not have to be the same. But in any case, the national model in (3) provides a straightforward means of incorporating global influences on a country's output, either exerted alongside the other macroeconomic influences captured by the direct measures of expectations included in $\mathbf{y}_{it}$'s or through the common global factors proxied by the inclusion of the weighted cross-sectional averages.

The final stage in the construction of a global VAR (GVAR) explaining actual and expected outputs across the n countries is motivated by noting that we can arrange the country-specific series into a single $3n \times 1$ vector $\mathbf{z}_t = (\mathbf{y}_{1,t}', .., \mathbf{y}_{n,t}')'$ and that we can write $\mathbf{y}_{i,t}^* = \mathbf{w}_i \mathbf{z}_t$ where $\mathbf{w}_i$ is the $1 \times 3n$ vector containing country i's weights. Arranging the individual vectors of parameters $\mathbf{B}_{is}$ and $\mathbf{B}_{is}^*$ in $\mathbf{B}_s$ and $\mathbf{B}_s^*$ and arranging individual vectors of weights in $\mathbf{W}$, the n country-specific models in (3) can be stacked to write

$$\mathbf{z}_t = \mathbf{B}_0 + \sum_{s=1}^{p+1} \mathbf{B}_s \mathbf{z}_{t-s} + \sum_{s=0}^{p+1} \mathbf{B}_s^* \mathbf{W} \mathbf{z}_{t-s} + \boldsymbol{\epsilon}_t, \qquad t = 1, ..., T, \tag{4}$$

where $\boldsymbol{\epsilon}_t = (\boldsymbol{\varepsilon}_{1t}', .., \boldsymbol{\varepsilon}_{nt}')'$ and so

$$\mathbf{z}_t = (\mathbf{I} - \mathbf{B}_0^* \mathbf{W})^{-1} \left(\mathbf{B}_0 + \sum_{s=1}^{p+1} (\mathbf{B}_s + \mathbf{B}_s^* \mathbf{W}) \mathbf{z}_{t-s} + \boldsymbol{\epsilon}_t \right), \qquad t = 1, ..., T, \tag{5}$$

providing a GVAR model that explicitly models all the interdependencies that exist between actual and expected outputs in all n countries. The errors $\boldsymbol{\epsilon}_t$ abstract from the influences on $\mathbf{z}_t$ arising from the global measures and, while in practice there might be cross-country correlations in these innovations, the variance-covariance matrix Σ will be close to diagonal and these shocks can be thought of as nation-specific ones; global shocks are represented through the $(\mathbf{I}-\mathbf{B}_0^*\mathbf{W})^{-1}\boldsymbol{\epsilon}_t$ term.

3 Recessions, Decision-Making and the Economic Evaluation of Forecasts

The above GVAR methods focus on the importance of the effects of expectations and global interactions in growth from a modelling perspective. Frequently though, interest focuses more on the *economic* importance of these effects, emphasising their implications for forecasts used in decision-making. The economic consequences of the effects are difficult to measure in practice because different individuals will be effected by output dynamics in different ways. For this reason, in what follows, we judge the economic importance of using survey data and of accommodating global interactions according to their importance in forecasting the likely occurrence of a range of recessionary events.[8]

The GVAR model described above can be readily used to produce forecasts of the probability of specified events taking place and to make decisions that depend on the events. As noted earlier, random draws from the estimated variance-covariance matrix Ω can be used to simulate future paths for the $\mathbf{z}_t$'s $(t = T+1, ..., T+h)$, taking the estimated parameters as known or taking into account parameter uncertainty. This generates the forecast density $\Pr(\mathbf{Z}_{T+1,T+h} \mid \mathbf{Z}_{1,T},\ M_T^{GVAR})$ given the data to date, $\mathbf{Z}_{1,T}$, and based on the the GVAR model available at time T, denoted M_T^{GVAR}. Further, any recessionary event defined as a set of outcomes involving nowcast and future actual outputs, $\mathbf{z}_{T+1}$, $\mathbf{z}_{T+2}$,... can be written as $R(\mathbf{Z}_{T+1,T+h})$. This event could focus on a particular country's output experiences or could look at all countries together to consider 'global recession'. The probability that the event occurs is

$$\text{probability of recession}= \int_R \Pr(\mathbf{Z}_{T+1,T+h} |\ \mathbf{Z}_{1,T},\ M_T^{GVAR})\ \ \partial\mathbf{Z}_{T+1,T+h}. \tag{6}$$

In a simulation exercise, the forecast probability is obtained simply as the proportion of

[8]The link between the economic importance of downturns in growth, individual decision-making and forecasting recessionary event probabilities is elaborated in Lee and Shields (2011).

the simulations in which the event is observed to occur. One natural criterion with which to judge the economic significance of the global interactions, for example, is to compare the performance of the GVAR model - which takes these interactions into account - in correctly forecasting the occurrence of events of interest compared to that of a model that ignores global interactions. As we discuss below, there are a range of statistics that can be calculated based on hit rates (i.e. how often the model successfully predicts the event will or will not happen) to provide an assessment of the economic value of model features following this approach.

The economic significance of expectations data and of global interactions in decision-making can be considered even more directly if we know the nature of the decision to be made. In a decision-making context, where an individual's objective function $\nu(r, R(\mathbf{Z}_{T+1,T+h}))$ depends on the outcome of a choice variable r and the occurrence of the recessionary event, the decision-maker's problem can be written as

$$\max_{r}\left\{ \int \nu(r, R(\mathbf{Z}_{T+1,T+h}))\operatorname{Pr}\left(\mathbf{Z}_{T+1,H} \mid \mathbf{Z}_{1,T},\ M_{T}^{GVAR}\right) \ \ d\mathbf{Z}_{T+1,T+h} \right\}. \qquad (7)$$

In terms of the simulations, the decision involves evaluating the objective function for different values of r in each simulation and simply choosing the value of r that maximises the value of the objective function when averaging across the simulations; we might denote the chosen optimal value of by $r_T^{\dagger}$. Pesaran and Skouras (2000) suggest a decision-based criterion function for the evaluation of the predictive density function from the model given by $\Psi_T = \mathrm{E}_P\left[\nu(r_T^{\dagger},\ R(\mathbf{Z}_{T+1,T+h}))\right]$ based around the value $r_T^{\dagger}$ advised by the GVAR model and where $\mathrm{E}_P[]$ is the expectations operator with respect to the "true" probability density function of $\mathbf{Z}_{T+1,h}$ conditional on information at time T. In practice, an absolute standard for the forecast evaluation is not available because the true probability density function of the forecast variable is not known. But an economic evaluation of a model can be based on the sample counterpart of the criterion function; namely

$$\overline{\Psi}_T = \frac{1}{k}\sum_{\tau=T-k}^{T} \nu(r_{\tau}^{\dagger},\ R(\mathbf{Z}_{\tau+1,\tau+h})), \qquad (8)$$

calculated over an out-of-sample evaluation period $T-k, ...T$. A similar statistic can be calculated for any other model (with associated optimal choice variable $r_T^{\dagger}$).

This discussion makes it clear that we generally need to specify an explicit objective function to measure a model's economic value in forecasting. An objective function that can be used in a very wide variety of circumstances, and which allows comparison of forecast performance to be made across models and across different events, is one based on the returns to a 'fair bet' on whether an event takes place. Two versions of the 'fair bet' might be considered. In the 'symmetric' version, it is assumed that an investor gambles every period with a \$1 stake, stating whether she believes the event will take place or not. If she forecasts the outcome of the event correctly, she receives a payout of \$$s$ but otherwise loses the \$1. In the 'asymmetric' case, the investor only bets for a stake of \$1 if she believes the event will occur and wins \$$s$ if she is correct. The bet is fair in each case because the payout is chosen so that the investor would break even if she made the choice randomly on the basis of the unconditional probability that the event occurs.

With the unconditional probability denoted by p, the payout in the *symmetric fair bet* is obtained by noting that the expected return is given by

$$[p^2 + (1-p)^2](s-1) - 2p(1-p) = 0$$

so that

$$s = \frac{1}{2p^2 - 2p + 1}.$$

The investor is assumed to want to maximise the end-of-forecast-period wealth which, in this case, is given by

$$W_{T+h} = [\ \ (r_T^\dagger \times I(R)) + (1 - r_T^\dagger)(1 - I(R))\ \ ](s-1) - r_T^\dagger(1 - I(R)) - I(R)(1 - r_T^\dagger) \tag{9}$$

where $I(.)$ is an indicator function taking the value 1 or 0 if the event does or does not happen and where $r_T^\dagger = 1$ or 0 depending on whether the decision-maker believes the event R will occur or not, based on their forecasting model. The forecast probability can be converted in to the binary event by setting $r_T^\dagger = 1$ when the forecast probability is greater than 0.5 (i.e. more likely to occur than not). In the *asymmetric fair bet* case, the expected return is given by

$$(s-1) \times p^2 - p(1-p) = 0$$

so that

$$s = \frac{1}{p},$$

and end-of-forecast-period wealth is given by

$$W_{T+h} = [\ \ r_T^\dagger \times I(R)\ \ ](s-1) - r_T^\dagger(1 - I(R)). \quad (10)$$

The expressions in (9) or (10) can then be used as $\nu(r_\tau^\dagger,\ R(\mathbf{Z}_{\tau+1,\tau+h}))$ in the criterion function (8) to judge the forecasting performance of the models in terms of decision-making.

4 Forecasting Output and Recession in the G7, 1994q1-2013q1

The forecast evaluation exercise is based on actual and expected output data for the G7 economies (Canada, France, Germany, Italy, Japan, United Kingdom, and United States) available over the period 1994q1-2013q1. The quarterly expectations data are provided in the surveys published by *Consensus Forecasts: A Digest of International Economic Forecasts* in March, June, September and December. The actual output data is the real volume GDP index for each country taken from the IMF's *International Financial Statistics* 2013q3. We treat the measure provided in the 2013 publication as the 'true' measure of output and assume this is observed after a one quarter delay (so that ${}_{t+s}y_t = {}_Ty_t$ for all $s \geqslant 1$). We then construct the corresponding series of nowcast and expected output levels at t using the final vintage series up to $t-1$ and the *Consensus Forecasts* of nowcast and expected growth published at t. Our data manipulation effectively assumes that the 'true' actual output series is released with a one quarter delay and is not subsequently revised, and that individuals know the true value of output up to one quarter previously and that it is their expectations of growth in the true output series that is reported in the surveys.

The actual output series for each of the G7 countries, ${}_{t+1}y_{i,t}$, is plotted in Figure 1 alongside the expected contemporaneous output series, ${}_ty_{i,t}^e$, and the expected future output series, ${}_{t-1}y_t^e$, obtained from the surveys. The series are aligned to show the output value at time t with the series published and dated at $t+1$, t, and $t-1$ respectively. The plots show that the three series move together relatively closely at most times but that there are periods of quite large divergence. For example, the very rapid decline in output at the end of 2007 and early 2008 was not anticipated in the expectations data and large expectational errors continued for two or three further quarters in most countries.

GLS describe in some detail the statistical properties of the series and the properties of the estimated regressions used to construct the GVAR model. The regression exercise was conducted recursively starting with the estimation period 1994q1-2003q3 and extending to 1994q1-2013q1. These sample choices are relatively arbitrary but made to maximise the estimation period while maintaining a reasonably long period for forecast evaluation. The period is, of course, one in which many countries experienced slowdowns in growth following the global financial crisis of 2007/8. This makes the period a good one for the purpose of evaluating forecasts of the probability of recession, but it also creates difficulties in forecasting if the modelling takes no account of possible structural breaks. To deal with this, we followed a procedure in which the model used for forecasting was updated in each recursion unless the most recent observation of the actual output series was deemed to be an 'outlier'. This judgement was based on a Chow test of structural stability where the observation was considered to be an outlier if the Chow test is significant at 1% level. In this case, the model obtained in the previous period was retained for the purpose of forecasting, effectively dummying out the most recent observation of the series for estimation purposes. The recursive updating of models is resumed once the observed growth data no longer appear as an outlier. This procedure is applied equally to all the models so that the forecast comparison is a fair one and it reflects how probability forecasts and decisions might have been made in real time.

Our interest in this paper is on the forecasts from the models rather than the properties of the series or models themselves. However, it is worth noting that the analysis of GLS shows (i) in each country, the actual, nowcast and expected future output series all display similar mean growth rates and are all quite volatile,[9] although the survey-based measures show less volatility than the actual series; (ii) output dynamics are very complicated in the estimated VAR models, with statistically significant feedbacks to actual, nowcast and

[9]The average of the annualised actual growth rates across the seven countries is 1.8% and the average of the standard deviations is 3.0%.

expected future outputs growths from lags of all these variables in most countries; (iii) cross-country interactions are very important in the model so that, applying a variance decomposition measure to investigate output dynamics, 60% of the persistent movements in countries' output is explained by global factors; and (iv) coefficient restrictions that can be imposed on the VAR to reflect the hypothesis that expectations on output are formed rationally are strongly rejected by the data, and a second variance-decomposition exercise shows that only 70% of the persistent movements in output can be explained by a model incorporating rationally-formed expectations, leaving 30% due to 'sentiment'. In short, the use of direct measures of nowcast and expected future output provides a more sophisticated characterisation of output movements than could be captured by actual output data alone, and international interactions are also found to be extremely important. The question now is whether the importance of these factors in modelling carries over to forecasting and economic decision-making.

4.1 The Forecast Performance of Alternative Models: Point Forecasts and Density Forecasts

Table 1 reports the root mean squared forecast error (RMSE) based on forecasts from four alternative models, each expressed relative to the RMSE from a standard random walk [RW] benchmark model. As well as the GVAR model in equation (5), we also consider three additional variants to assess the role of the GVAR model's cross-country linkages and also the effects of including the expectations data. The first model variant is a simple univariate autoregressive model (of order two), in each countries' own output growth, termed 'AR1'. The second variant includes the two expectations terms in (1) in addition to actual output but has no global (trade-weighted) $\mathbf{z}_t^*$ variables. We denote this three-variable model model by 'VAR3'. The third model includes the trade-weighted foreign output variables in each countries' growth equation but excludes the expectations terms; it is a simple GVAR in actual outputs only and is denoted 'GVAR1'. The final model is the full GVAR model including lags of actual and expected variables as well as the global (trade-weighted) variable, as in equation (5). This model is denoted 'GVAR3'.

In Table 1a, the models are judged according to their ability to nowcast current output growth at time T as it is revealed in $T+1$; i.e. ${}_{T+1}y_T - {}_{T}y_{T-1}$. The tables shows that the RMSE are less than unity in nearly every case, so that the models outperform the RW model, but that this improvement is not statistically significant anywhere according to the Giacomini-White (2006) test of equal forecast performance. The results are slightly stronger in Table 1b, where the models are judged by their ability to forecast, at time T, the four-step-ahead output growth as revealed in $T+5$; i.e. ${}_{T+5}y_{T+4} - {}_{T+4}y_{T+3}$. Here the VAR3 model is shown to have the best forecasting performance in 5/7 countries (as highlighted by the emboldened statistics) and the improvement over the RW model is statistically significant in three of these cases. In terms of providing point forecasts, though, none of the models appears to perform particularly well.

Tables 2a and 2b shift the focus from the point forecasts and towards density forecasts. At each recursion, the four estimated models, and the benchmark RW model, were each used to simulate 10000 potential future output paths and, hence, the densities of output growth nowcasts and four-step-ahead growth forecasts. The log predictive scores judge the models' performance according to the forecast likelihood of the actual outcome as observed over the forecast evaluation period. In this case, a positive log score relative to that of the RW model shows the model outperforms the benchmark and, again, the results show positive signs in nearly every country and nearly every model. In this case, though, the statistics are more often statistically significant. For the nowcasts, the largest value for the log scores was obtained by the GVAR1 or GVAR3 models in every case and these are significant in 4/7 countries. This pattern is repeated with the four-step-ahead forecasts.

These results provide some useful insights on output growth forecasting and the role of the survey data and international interdependencies. Specifically, the results downplay the usefulness of the forecasting models in point forecasting exercises - the focus of most popularist coverage - but emphasise their usefulness in probabilistic statements on likely output outcomes. Further, the results show the direct measures of expectations do not substantially improve the nowcasts of the (yet-to-be-published) actual output levels either in point forecasts or in density forecasts. This is counter-intuitive if the surveys reflect individuals' best guess of what measured output will turn out to be based on the information available to them in real time. But the result matches the finding of GLS that the survey expectations are not straightforward measures of rationally-formed expectations of underlying fundamentals. On the other hand, the survey data does appear to help in forecasting within a country, in that, when used with the actual output data, the expected current and expected future series provide useful input into the models' longer horizon forecasts through their long-run cointegrating properties. Finally, and most clearly, the results show that it is the interactions that occur across countries that has the greatest impact on forecasting performance, by taking into account other countries' output experiences, as in the GVAR1 model, and by taking into account the direct measures of expected output changes overseas, as in the GVAR3 model.

4.2 The Forecast Performance of Alternative Models: Forecasting the Probability of Recessions

We have argued that models' forecasting performance might be judged using economic as well as statistical criteria and that the economic significance of countries' growth series is often expressed in terms of recessionary events. Every country in the G7 experienced a reduction in output at some point during 2008 and 2009 but the context of the reductions and their size and timing were quite distinct so that the economic significance of the downturns, and the value of forecasting them correctly, differed from country to country. For example, while some countries grew quite strongly through 2007 (e.g. Germany, UK), others such as Italy and Japan were already experiencing quarters of negative growth and Canada and US were also growing only very slowly.[10] Nearly all countries experienced negative quarterly growth through the "crisis quarters" 2008q3-2009q1 although the output reductions were much larger in Europe and Japan than in North America.[11] And the recovery from 2009 has been very different, with some countries achieving output levels as high, or higher, than pre-crisis levels relatively quickly (e.g. Canada, Germany, US),

[10]Indeed, US ouput growth fell in 2008q1 and NBER dated the US recession to have started that quarter. This was well before the Leyman Brothers bankrupcy in 2008q3, for example, when global financial market problems began to dominate the headlines.

[11]The maximum year-on-year drop in output averaged at around -14% in the five countries outside N. America, dated 2009q2, while the average across Canada and US was -9% dated a quarter or so earlier.

while output in France, Italy, Japan and UK remained below pre-recession levels even in 2013.

The differences in countries' output growth experiences translate into different episodes of recession depending on the definition employed. In what follows, we focus on four types of recessionary event: two 'output drop recessions' (ODR), where recession is defined to occur when output growth is negative; and two 'below peak recessions' (BPR), where recession is defined to occur if output is lower than its previous peak level. The four recessionary events illustrate the variety of events that might be considered relevant to decision makers, both relating to individual countries and aggregating in different ways across countries to define a global recession. Moreover, they provide a very good test-bed for evaluating the four models we have developed for forecasting. The ODR recessions are relatively rare events in most countries so that the models will be exposed if they rely too heavily on the status quo. In contrast, in terms of simple unconditional probabilities, the BPR recessions are reasonably common events during our sample, being approximately as likely to occur as not. On the other hand, by their nature, these events occur in runs and so a forecast evaluation exercise needs to take into account this serial correlation in occurrences.

Recessions defined by output drops are 'ahistoric' in the sense that the event is based on the rate of change of output only. One ODR event of interest at T is whether time-T output growth is negative; i.e. nowcasting $I({}_{T+1}y_T - {}_Ty_{T-1} < 0)$ where output growth in T is measured after a one period delay at $T+1$. The relatively modest slowdowns in growth observed in N. America, and the relatively strong recoveries observed there, mean that this recessionary event - denoted $ODR(0)$ below - occurred in just 14% of observations between 2003q4-2013q1 in Canada and US, compared to 16%, 21% and 27% in France, Germany and UK, and 35% and 32% in Italy and Japan. The time frame of the event can be extended if we work with the moving average series centred on the time T observation $\widetilde{y}_T^m = \frac{1}{m+1}({}_{T-m+1}y_{T-m} + ... + {}_{T+1}y_T + ... + {}_{T+m+1}y_{T+m})$. Hence, for example, we could consider the recessionary event $ODR(4) = I(\widetilde{y}_T^4 - \widetilde{y}_{T-1}^4 < 0)$, say, describing a drop in the nine-period moving average of output centred on T. This event occurred, on average, 5% more often than $ODR(0)$ during our evaluation period as the smoothing arising from taking moving averages spreads out the effects of the periods of rapid contraction.

Each of the definitions can be aggregated to define 'global recession' which might be defined as occurring when the majority of countries (i.e. at least four out of seven) are individually experiencing recession, say, or when the average output growth across the seven countries meets the recession criterion for example. As illustrated in Figures 2 and 3, the evaluation period saw global recession in 19% (majority) and 14% (average) of occasions according to $ODR(0)$ and 27% (majority) and 24% (average) of occasions according to $ODR(4)$ reflecting the severity of the impact of the financial crisis on growth rates across all countries over the last decade.

Recessions defined by output relative to its previous peak are more sensitive to history. A short-horizon 'below peak recession' is defined by $BPR(0) = I\{\ _{T+1}y_T < \max\{_Ty_{T-1},\ _{T-1}y_{T-2},\ _{T-2}y_{T-3}, ...\}\}$ while a longer horizon (one-year-ahead) perspective is defined by $BPR(4) = I\{\ _{T+5}y_{T+4} < \max\{_Ty_{T-1},_{T-1}y_{T-2},_{T-2}y_{T-3}, ...\}\}$ comparing the outcome four periods ahead with the most recent peak. Given that the financial crisis takes place roughly half-way through our sample evaluation period, and given that many European countries have not recovered their pre-crisis output levels, it is not surprising to find that $BPR(0)$ occurs in around half the observations in France, Germany, and UK, and more than half in Italy and Japan. The figure is lower in Canada and US at 27% and 41% respectively. The $BPR(4)$ definition requires output to remain below peak for longer than $BPR(0)$ and occurs less frequently therefore. According to the BPR definitions, and as shown in Figures 4 and 5, global recession was experienced in 51% (majority) and 43% (average) of occasions according to $BPR(0)$ and 49% (majority) and 38% (average) of occasions according to $BPR(4)$.

4.2.1 The Comparative Performance

The performance of the alternative models in forecasting the various recessionary events is described formally in Tables 3 and 4. The first part of each of the tables shows the proportion of times the event occurred in each country during the evaluation period (i.e. the unconditional probability of the event p) and then, for each model: the hit-rate (i.e. the proportion of accurate predictions) where the event is predicted to occur when the forecast probability exceeds 0.5; the Kuipers Score (a statistic that takes values between -1 and 1 and summarises the degree of correspondence between predictions and outcomes rather like a correlation coefficient); and the outcome of two χ_1^2 tests of the null that there is no relationship between the outcome and the predictions. The two tests are the 'reduced rank regression' and 'dynamically-augmented reduced rank regression' tests described in Pesaran and Timmermann [PT] (2009). The first of these tests is a standard contingency-table test of the null that the model is no more successful in predicting outcomes than would be achieved based only on the unconditional probability. The second test takes into account the possibility that there are predictable runs in the data. The table also shows the corresponding global statistics. The second part of each of the tables reports the outcome of the symmetric and asymmetric fair bets described in (9) and (10) using the same 0.5 probability threshold to choose whether or not to bet as when calculating the hit rates and Kuipers score.

Broadly speaking, the statistics in the first part of Tables 3(a) and 3(b) echo the log score results in Table 2 in that the models that deliver the best density forecasts are also the models that best predict the ODR recessionary events. In Table 3a, where the focus is on the nowcast of $ODR(0)$ recession, the GVAR1 and GVAR3 models deliver the highest Kuipers scores in 5/7 countries, and also in predicting global recession. The more demanding dynamic PT test shows the GVAR3 model has significant predictive power in four countries even taking into account the runs in the events. The dominance of the GVAR3 model is even more apparent in Table 3b concerned with predicting the longer horizon recessions defined by $ODR(4)$. The model has the largest Kuipers score in predicting recession in five of the seven countries and at the global level, and significantly so in these cases according to the dynamic PT test.

Figures 2 and 3 illustrate the characteristics of the forecasts underlying these results. None of the models anticipated the onset of the financial crisis but the two GVAR models accommodated the news and reacted after one or two quarters, correctly nowcasting recession over the latter part of 2009. The inclusion of the survey data in the models improves the predictive power but it is the interactions across countries that raises the forecast probabilities above 0.5 in the nowcast plots of Figure 2. The importance of the survey data is more apparent in Figure 3 where the GVAR3 model correctly forecasts the longer horizon $ODR(4)$ recession to start a little earlier and end a little later than the GVAR1 model so that, as noted earlier, the contribution of the survey data appears to come through the feedbacks across time rather than as an early warning of a downturn.

The figures in the second part of Tables 3(a) and 3(b) describe the returns to the fair \$1 bet and provide more insights on the relative forecast performance of the models in terms of decision making. Averaging over the countries, the payout for correctly nowcasting the event (s in (9) and (10)) is \$1.53 for the $ODR(0)$ recession in the symmetric case and \$5.00 in the asymmetric case. The difference in payout reflects the fact that, although the investor gambles every period in the symmetric case, she can win the bet if she correctly predicts no recession (which happens reasonably frequently), while she has to choose to make the bet and correctly anticipate the recession in the asymmetric case. In these circumstances, it is clear that none of the models are very helpful in the symmetric case with most models performing no better, or even losing money, relative the RW model which makes \$6.32 on average by simply nowcasting 'no recession' over the 38 quarters of our evaluation period. However, the nowcasting advantages of the GVAR models translate into reasonably large gains in the asymmetric case: a profit of \$13.82 is made on average using the GVAR3 model across the seven countries which represents an infinitely higher rate of return over that delivered by the RW model (which advises never to take the bet and delivers \$0 as a consequence). For the longer horizon recession $ODR(4)$, the payout for the symmetric and asymmetric bets are \$1.62 and \$4.00 respectively, reflecting the higher likelihood of this type of recession occurring. Again, the investor would obtain relatively little using the models to forecast the recession here in the symmetric case (improving the rate of return over RW model by 15%) but using the GVAR3 model would improve the rate of return by 74% over the RW model in the asymmetric case. It is also perhaps worth noting that the GVAR3 model outperforms all models in forecasting global recession, however defined, in both Tables 3a and 3b.

Figures 4 and 5 and Table 4 describes the corresponding results for nowcasting and forecasting the below peak recessions $BPR(0)$ and $BPR(4)$. All the models reacted reasonably quickly to correctly nowcast this event and their relative performances are better distinguished by their attempt to forecast the end of the recession. The figures illustrate that it is the GVAR models (and especially GVAR3) which correctly anticipates that output will remain below peak through 2012 in many countries whether judged by the one-step-ahead nowcast in $BPR(0)$ or forecasting one year ahead as in $BPR(4)$. The very high degree of serial correlation in the occurrences of this event mean that, in terms of hit rates and Kuipers scores, it is difficult to distinguish between the forecast performance of the various models according to the BPR(0) definition (with all models performing well relative to the 'no model' benchmark). This is also true of the symmetric fair bet outcomes where the roughly 50-50 unconditional probability translates into a payout of $1.90 for correctly predicting whether a recession occurs or not so that even the RW model delivers $24.36 over the 38 quarters and there is no substantial gain made using the other models. The GVAR models' advantage in terms of forecasting four periods ahead translates to improved performance in predicting BRP(4) at the end of the evaluation period though, with the GVAR3 model generating a rate of return some 34% higher than the RW model in the symmetric fair bet and 63% higher in the asymmetric case. As in the ODR results, the occurrence of global BPR recession, whether defined by the majority of countries or the average growth across countries, is again predicted most accurately by GVAR3.

These results provide useful insights on evaluating the forecasting performance of the models. The economic value of the models is much weaker in the symmetric case where the economic context means the payout for correctly predicting recession is relatively small. But the payouts, and the gains from correctly predicting recession, rise in the more demanding asymmetric case. Although the GVAR3 model was not able to anticipate the effects of the financial crisis, and its advantage in terms of differences in hit rates and Kuipers scores across models seems modest, its better forecasting performance translates into systematic and quite substantial financial advantage in the context of the asymmetric fair bet.

5 Concluding remarks

The empirical exercise provides some important insights on forecasting output dynamics in particular and illustrates some useful features of forecast evaluation exercises in general.

First the exercise shows that none of the time series models of G7 outputs significantly outperform a simple random walk model if judged according to the countries' point forecasts but they all perform well when judged according to their density forecasts. This is true for nowcasts as well as four-quarter-ahead forecasts. The improvement in forecasting provided by the time series models in the more sophisticated density forecasts also translates into improved forecasts on event probabilities and, depending on the loss functions involved, improved decision-making. The evaluation of models of output based solely on point forecast performance could be severely misleading and the exercise illustrates well the general point that forecast evaluation should be multifaceted.

Second, there is a good correspondence between models' relative performance in density forecasting and their performance in predicting recession when using criteria linked to the hit rates, but the extent to which this translates to effective decision-making depends crucially on the decision-making context. While this may be obvious, the point is made very clearly in the context of the comparison of the models in a 'neutral' fair bet investment scenario. The low payouts associated with bets in the symmetric case means that decisions based on the random walk model are as effective as those based on any of the time series models. However, the more demanding decision context of the asymmetric setup translates into very different returns from the various models and provide good discriminatory power in evaluating the models with GVAR3 performing much better than the other models.

Third, in terms of a simple 'horse-race' between models, it appears that it is more important to include the cross-country interactions between the G7 countries' outputs than to include the countries' survey-based expectations data when making output forecasts. This is true both considering each country in turn as well as when forecasting the occurrence of global recession. The importance of the cross-country effects in nowcasting and one-year-ahead forecasting matches the finding of GLS that the majority of the (infinite-horizon) persistent effect of shocks to output relate to common international influences. The results demonstrate the inter-relatedness of countries' output dynamics and the global nature of recession and confirm that is essential to take this into account in density forecasting, event probability forecasting or forecast-based decisions.

And *fourth*, while cross-country effects are the most important, in practice it is best to include both cross-country and expectational effects when forecasting densities or recessionary event probabilities. Perhaps unexpectedly, the direct measures of expectations (which include people's nowcast of output as expressed through the survey) contribute more to the models' longer term forecasts than to the models' nowcasts. The contribution appears to arise from their long-run influence, tying forecasted actual and expected outputs together in the long run through the cointegrating relations in the model, rather than from the immediacy of including any up-to-date news content contained in the survey nowcast. This result is also in line with the findings of GLS which cast doubt on the idea that the survey data is a straightforward prediction of outputs based on a full-information rational expectation of the underlying fundamentals. Rather, the survey measures appear to incorporate information rigidities, learning behaviour and other potential 'sentiment' effects which help to predict (and indeed define) future paths of output over and above the effects of fundamentals.

Despite the uncertainties and reservations expressed in the media and by policy makers on predicting growth prospects, the analysis of the paper shows that economic models can provide good insights on future output dynamics and recessionary events. The analysis confirms that it is important to explicitly incorporate into the model cross-country output interactions and information contained in surveys, matching the widely-expressed belief that the recent experiences are global in nature and that confidence and pessimism about output prospects can play an important role beyond that played by simple projections of fundamentals. But the analysis also emphasises the need for a more nuanced approach to representing predictions on output, providing forecasts of the entire range of possible outcomes and the likelihood of recesssionary events, rather than just point forecasts. And the analysis emphasises the need for a clearer statement on which features of the slowdown (recession length, depth and welfare measures) are of interest to the commentators since models can only really be judged according to their usefulness in the particular context.

6 Bibliography

Akerlof, G.A. and R. J. Shiller (2009), *Animal Spirits: How Human Psychology Drives the Economy, and Why it Matters for Global Capitalism*, Princeton University Press.

Barsky R.B. and E.R. Sims (2012), "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence", *American Economic Review*, 102(4), 1343-77.

Blanchard, O.J., G. Lorenzoni and J-P L'Huillier, (2013), "News, Noise, and Fluctuations: An Empirical Exploration", *American Economic Review*, forthcoming.

Clements, M. (2006), "Evaluating the Survey of Professional Forecasters probability distributions of expected inflation based on derived event probability forecasts", *Empirical Economics*, 31, 49–64.

Coibion, O. & Y. Gorodnichenko (2012). "What Can Survey Forecasts Tell Us about Information Rigidities?", *Journal of Political Economy*, 120(1), 116-159.

Crucini, M.,A. Kose and C. Otrok, (2011), "What are the Driving Forces of International Business Cycles?", *Review of Economic Dynamics*, 14(1), 156-175.

Dees, S., F. di Mauro, M.H. Pesaran and L.V. Smith (2007), "Exploring the International Linkages of the Euro Area: A Global VAR Analysis", *Journal of Applied Econometrics*, 22, 1-38.

Del Negro, M. and C. Otrok, (2008), "Dynamic factor models with time-varying parameters: measuring changes in international business cycles", *Staff Reports*, 326, Federal Reserve Bank of New York.

van Dijk, D., H. van Dijk and P.H. Frances, (2005), "On the Dynamics of Business Cycle Analysis: Editors' Introduction", *Journal of Applied Econometrics*, 20, 2, 147-150.

Garratt, A. and K. Lee, (2010), "Investing Under Model Uncertainty: Decision Based Evaluation of Exchange Rate Forecasts in the US, UK and Japan", *Journal of International Money and Finance*, 29, 3, 403-422.

Garratt, A., K. Lee, E. Mise and K. Shields (2008), "Real Time Representations of the Output Gap", *Review of Economics and Statistics*, 90, 4, 792-804.

Garratt, A., K. Lee, E. Mise and K. Shields (2009), "Real Time Representations of the UK Output Gap in the Presence of Model Uncertainty", *International Journal of*

Forecasting, 25, 81-102.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin (2003), "Forecast Uncertainty in Macroeconometric Modelling: An Application to the UK Economy", *Journal of the American Statistical Association*, 98, 464, 829-838.

Garratt, A., K. Lee, M.H. Pesaran and Y. Shin (2006), *National and Global Macroeconometric Modelling: A Long-Run Structural Modelling Approach*, Oxford University Press.

Garratt, A., K. Lee and K. Shields (2013), "Economic Sentiment, International Interdependence and Output Dynamics in the G7", *mimeo*, University of Nottingham, October.

Giacomini, R. and H. White (2006), "Tests of Conditional Predictive Ability", *Econometrica*, 74, 1545-1578.

Isiklar, G. and K. Lahiri, (2007), "How far ahead can we forecast? Evidence from cross-country surveys", *International Journal of Forecasting*, 23, 167–187.

Kannan, P., A. Scott, and M.E. Terrones, (2009), "From Recession to Recovery: How Soon and How Strong?", chapter 3 in *IMF World Economic Outlook*, 103-138.

Kose A., C. Otrok and C.H. Whiteman, (2003), "International Business Cycles: World, Region, and Country-Specific Factors", *American Economic Review*, 93(4), 1216-1239.

Kose, A., C. Otrok, and C.H. Whiteman, (2008), "Understanding the evolution of world business cycles", *Journal of International Economics*, 75(1), 110-130.

Lahiri, K. and J.G. Wang, (2007), "The value of probability forecasts as predictors of cyclical downturns", *Applied Economics Letters*, 14, 11–14

Lahiri, K. and J.G. Wang, (2013), "Evaluating probability forecasts for GDP declines using alternative methodologies", *International Journal of Forecasting*, 29, 175–190

Laxton, D., P. Isard, H. Faruqee, E. Prasad and B. Turtleboom, (1998), "MULTIMOD Mark III The Core Dynamic and Steady-State Models", *IMF Occassional Papers* 164

Lee, K., N. Olekalns, K. Shields and Z. Wang (2012), "The Australian Real-Time Database: An Overview and an Illustration of its Use in Business Cycle Analysis", *Economic Record*, 88, 283, 495-516.

Lee, K. and K. Shields, (2011), "Decision-Making in Hard Times: What is a Recession, Why Do We Care and When Do We Know We Are in One?", *North American Journal of Economics and Finance,* 22, 1, 43-60.

Lumsdaine, R.L. and E.S. Prasad, (2003), "Identifying the common component of international economic fluctuations: a new approach", *Economic Journal*, 113, 101–127.

Pesaran, M.H., T. Schuermann and L.V. Smith (2009) "Forecasting Economic and Financial Variables with Global VARs", *International Journal of Forecasting*, 25, 4, 642-675.

Pesaran, M.H., T. Schuermann and S.M. Weiner (2004). "Modeling Regional Interdependencies using a Global Error-Correcting Macroeconometric Model", *Journal of Business & Economic Statistics* 22, 129-162 and 175-181

Pesaran, M.H., and S. Skouras (2002), "Decision-Based methods for Forecast Evaluation", in M.P. Clements and D.F. Hendry (eds.) *A Companion to Economic Forecasting*, 241-267, Basil Blackwell, Oxford.

Pesaran, M.H. and A. Timmermann (2009) "Testing Dependence Among Serially Correlated Multicategory Variables", *Journal of the American Statistical Association*, vol 104, 485, 325-337.

Table 1a: RMSE for Output Growth Nowcasts

(Actual RMSE for RW, Ratio relative to RW for other models*)*

	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*0.047*	0.802	0.645	0.711	**0.599**
France	*0.040*	0.834	1.011	**0.731**	0.738
Germany	*0.063*	0.961	0.930	**0.827**	0.916
Italy	*0.061*	0.904	0.894	0.773	**0.719**
Japan	*0.082*	0.969	0.911	0.902	**0.880**
UK	*0.061*	**0.881**	1.147	0.906	1.233
US	*0.049*	**0.856**	0.973	0.993	0.901

Table 1b: RMSE for Four-Step-Ahead Output Growth Forecasts

(Actual RMSE for RW, Ratio relative to RW for other models*)*

	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*0.163*	0.918	**0.835**	1.016	1.090
France	*0.157*	**0.942**	0.951	0.970	1.191
Germany	*0.198*	1.008	**0.971**	1.144	1.047
Italy	*0.227*	0.961*	**0.856***	0.945	1.049
Japan	*0.204*	**0.992**	1.042	1.060	1.066
UK	*0.253*	0.972*	**0.749***	0.884*	0.869
US	*0.192*	0.900*	**0.893***	1.021	1.093

Notes: RW denotes the random walk model for actual output growth in each country; AR1 denotes a univariate autoregressive (order 2) model of actual output growth in each country; VAR3 denotes a 3-variable VAR (order 2) model of actual output growth and current and one-period ahead survey expectations in each country; GVAR1 is the global version of AR; and GVAR3 is the global version of VAR. The * denotes that the RMSE is significantly lower than that from the random walk model, working at the 10% level of significance, and applying the Giocomini-White (2006) test of equal forecast performance

Table 2a: Average Log Predictive Scores for Output Growth Nowcasts

(Average Log scores for RW, Scaled difference of log score from RW for other models*)*

	RW	**AR1**	**VAR3**	**GVAR1**	**GVAR3**
Canada	*42.118*	0.181	1.130	1.309	**1.594**
France	*45.906*	0.494*	-0.181	1.471*	**1.615***
Germany	*71.005*	0.067	0.053	**0.724**	0.475
Italy	*90.829*	0.105*	0.202*	0.299	**0.401***
Japan	*55.964*	0.051	-0.010	**0.717**	0.536
UK	*-155.057*	0.404*	-0.416	**1.240***	0.980*
US	*51.503*	0.352*	-0.234	**0.863***	0.630*

Table 2b: Average Log Predictive Scores for Four-Step-Ahead Output Growth Forecasts

(Actual Log scores for RW, Scaled difference of log score from RW for other models*)*

	RW	**AR1**	**VAR3**	**GVAR1**	**GVAR3**
Canada	*-33.520*	2.528*	2.874*	**2.926**	2.515
France	*-46.580*	1.347*	0.956*	1.786*	**1.787**
Germany	*-15.356*	1.529	-1.253	**3.843**	1.699
Italy	*-13.734*	2.218*	2.833*	3.699*	**4.269***
Japan	*20.734*	0.331	0.268	0.725	**0.867**
UK	*-527.329*	0.325*	0.716*	0.812*	**0.836***
US	*-67.997*	1.163*	1.239*	**1.499***	1.368*

Notes: See notes to Table 1. The * denotes that the log predictive score is significantly larger than that from the random walk model, working at the 10% level of significance, and applying the Giocomini-White (2006) test of equal forecast. performance

Table 3a: Forecasting 'Output Drop Recessions' ODR0, 2003q4-13q1

	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*14%*	0.86	0.89	**0.92**	0.89	**0.92**	0 (–,–)	0.58 (**,**)	**0.69** (**,**)	0.58 (**,**)	**0.69** (**,**)
France	*16%*	0.81	0.78	0.78	**0.84**	**0.84**	0 (–,–)	−0.19 (–,**)	−0.19 (–,**)	**0.52** (**,–)	0.48 (**,–)
Germany	*21%*	**0.78**	0.73	0.76	0.68	0.76	0 (–,–)	−0.23 (**,**)	0.30 (*,–)	0.17 (–,–)	**0.42** (**,**)
Italy	*35%*	0.62	0.62	0.78	0.76	**0.84**	0 (–,–)	0.13 (–,–)	0.55 (**,**)	0.59 (**,–)	**0.65** (**,**)
Japan	*32%*	0.62	0.62	**0.68**	0.65	0.59	0 (–,–)	0.13 (–,–)	**0.34** (–,–)	0.23 (–,–)	0.06 (–,–)
UK	*27%*	**0.73**	**0.73**	**0.73**	0.70	0.70	0 (–,–)	0.00 (–,–)	**0.28** (–,–)	−0.28 (–,–)	0.07 (–,–)
US	*14%*	0.86	0.86	0.89	**0.92**	**0.92**	0 (–,–)	0.39 (–,–)	0.89 (**,–)	**0.91** (**,**)	**0.91** (**,**)
Majority	*19%*	0.81	0.81	0.86	0.86	**0.89**	0 (–,–)	0.33 (–,–)	0.86 (**,**)	0.86 (**,**)	**0.88** (**,**)
Average	*16%*	**0.84**	0.81	**0.84**	**0.84**	**0.84**	0 (–,–)	−0.17 (–,–)	0.36 (**,–)	0.36 (–,–)	**0.38** (*,–)

Table 3a (cont.): Forecasting 'Output Drop Recessions' ODR0, 2003q4-13q1

(*Actual Return for RW,* Improvement over RW for other models)

	Returns to Fair Bet (Symmetric)					Returns to Fair Bet (Asymmetric)				
	RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*4.76*	1.31	**2.61**	**2.61**	**2.61**	*0.00*	12.80	**19.20**	12.80	**19.20**
France	*5.57*	-1.37	-1.37	-1.37	-1.37	*0.00*	0.00	0.00	5.17	**10.33**
Germany	*6.87*	-3.03	-1.51	-1.51	-3.03	*0.00*	0.00	14.50	14.50	**25.38**
Italy	*7.10*	0.00	**11.03**	7.35	**11.03**	*0.00*	1.85	16.62	9.23	**20.31**
Japan	*7.51*	-3.56	**0.00**	-5.34	-3.56	*0.00*	0.00	6.25	**8.33**	6.25
UK	*7.59*	**0.00**	**0.00**	**0.00**	-1.65	*0.00*	0.00	**8.10**	0.00	2.70
US	*4.76*	0.00	1.31	**2.61**	**2.61**	*0.00*	6.40	6.40	**12.80**	**12.80**
Majority	*6.28*	0.00	2.89	**4.33**	**4.33**	*0.00*	4.29	8.57	**12.86**	**12.86**
Average	*14.38*	1.66	1.66	1.66	**3.32**	*0.00*	0.00	5.17	**10.33**	**10.33**

Note: p is the unconditional probability of the event 2003q4-2013q1. The figures in parentheses (., .) below the Kuipers Scores show, respectively, the outcome of the static and dynamic versions of the Pesaran and Timmerman (2009) tests of no additional predictive power beyond that of the unconditional probability; a '**' indicates significance at 5% level, '*' indicates significance at 10% level, and '-' indicates no significance at 10% level.

Table 3b: Forecasting 'Output Drop Recessions' ODR4, 2003q4-13q1

	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*15%*	0.85	0.76	**0.79**	**0.79**	**0.92**	0.00 (–,–)	−0.17 (–,–)	**0.11** (–.–)	**0.11** (–,–)	**0.11** (–.–)
France	*24%*	0.85	0.88	0.88	0.88	**0.91**	0.83 (**,*)	0.86 (**,*)	0.86 (**,**)	0.86 (**,**)	**0.89** (**,**)
Germany	*24%*	0.82	**0.88**	0.85	**0.88**	**0.88**	0.58 (**,–)	**0.86** (**,**)	0.66 (**,–)	0.72 (**,–)	0.72 (**,–)
Italy	*39%*	0.76	0.74	**0.82**	0.79	**0.82**	0.71 (**,**)	0.74 (**,**)	**0.77** (**,**)	0.74 (**,**)	**0.77** (**,**)
Japan	*27%*	0.82	**0.85**	**0.85**	**0.85**	0.82	0.62 (**,–)	**0.83** (**,**)	**0.83** (**,**)	**0.83** (**,**)	0.56 (**,–)
UK	*27%*	0.82	0.82	0.85	0.82	**0.88**	0.80 (**,–)	0.80 (**,–)	0.83 (**,–)	0.80 (**,–)	**0.86** (**,**)
US	*27%*	0.79	0.85	0.82	0.82	**0.85**	0.77 (**,–)	**0.83** (**,**)	0.80 (**,–)	0.80 (**,–)	**0.83** (**,**)
Majority	*27%*	0.82	0.85	0.85	0.85	**0.88**	0.80 (**,**)	0.83 (**,**)	0.83 (**,**)	0.83 (**,**)	**0.86** (**,**)
Average	*24%*	0.85	0.88	0.88	0.88	**0.91**	0.83 (**,**)	0.86 (**,**)	0.86 (**,**)	0.86 (**,**)	**0.89** (**,**)

Table 3b (cont.): Forecasting 'Output Drop Recessions' ODR4, 2003q4-13q1

(Actual Return for RW, Improvement over RW for other models*)*

	Returns to Fair Bet (Symmetric)					Returns to Fair Bet (Asymmetric)				
	RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*6.08*	-4.04	-2.61	-2.61	-2.61	*0.00*	0.00	**5.60**	**5.60**	**5.60**
France	*13.58*	1.58	1.58	1.58	**3.16**	*9.38*	3.13	3.13	3.13	**6.25**
Germany	*12.00*	**3.16**	1.58	**3.16**	**3.16**	*9.38*	3.13	3.13	**6.25**	**6.25**
Italy	*14.67*	**0.00**	**0.00**	-1.91	**0.00**	*7.69*	1.54	**3.08**	1.54	**3.08**
Japan	*14.38*	**1.66**	**1.66**	0.00	0.00	*10.67*	0.00	0.00	**2.67**	**2.67**
UK	*4.38*	0.00	1.66	0.00	**3.32**	*8.00*	0.00	2.67	0.00	**5.33**
US	*12.73*	3.32	3.32	3.32	**4.97**	*5.33*	5.33	2.67	5.33	**8.00**
Majority	*5.57*	-1.37	0.00	0.00	**1.37**	*8.00*	1.66	1.66	1.66	**3.32**
Average	*13.58*	1.58	1.58	1.58	**3.16**	*9.38*	1.58	1.58	1.58	**3.16**

Table 4a: Forecasting 'Below Peak Recessions' BPR0, 2003q4-13q1

	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*27%*	**0.89**	**0.89**	**0.89**	0.84	0.92	**0.87** (**,*)	**0.87** (**,*)	0.73 (**,–)	0.61 (**,–)	0.82 (**,**)
France	*51%*	0.86	0.95	0.95	0.95	**0.97**	0.78 (**,–)	0.90 (**,–)	0.90 (**,–)	0.90 (**,–)	**0.95** (**,–)
Germany	*49%*	0.78	0.78	**0.81**	0.54	0.78	**0.64** (**,–)	**0.64** (–,–)	0.62 (**,**)	0.14 (–,–)	0.63 (**,**)
Italy	*62%*	**0.89**	**0.89**	**0.89**	**0.89**	**0.89**	**0.78** (**,–)	**0.78** (**,–)	0.77 (–,–)	**0.78** (**,–)	0.77 (–,–)
Japan	*65%*	0.78	0.78	0.78	0.78	**0.81**	0.53 (**,–)	0.53 (**,–)	**0.55** (**,–)	0.53 (**,–)	0.54 (**,**)
UK	*51%*	**0.97**	**0.97**	0.89	0.92	0.84	**0.95** (**,–)	**0.95** (**,–)	0.78 (–,–)	0.84 (**,–)	0.78 (–,–)
US	*41%*	0.89	0.89	**0.92**	**0.92**	**0.92**	0.85 (**,–)	0.85 (**,–)	**0.88** (**,**)	**0.88** (**,**)	**0.88** (**,**)
Majority	*51%*	0.86	0.95	0.95	**0.97**	**0.97**	0.78 (**,–)	0.90 (**,–)	0.90 (**,–)	**0.95** (**,–)	**0.95** (**,–)
Average	*43%*	**0.95**	0.92	0.86	0.84	0.92	**0.91** (**,**)	0.84 (–,–)	0.72 (**,–)	0.67 (**,–)	0.83 (**,**)

Table 4a (cont.): Forecasting 'Below Peak Recessions' BPR0, 2003q4-13q1

(Actual Return for RW, Improvement over RW for other models*)*

	Returns to Fair Bet (Symmetric)					Returns to Fair Bet (Asymmetric)				
	RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*17.50*	0.00	0.00	-1.65	**1.65**	*16.20*	0.00	**5.40**	0.00	**5.40**
France	*26.95*	6.00	6.00	**7.99**	**7.99**	*13.26*	2.84	2.84	**3.79**	**3.79**
Germany	*20.96*	0.00	**2.00**	-11.99	-2.00	*11.61*	0.00	4.22	0.00	**6.63**
Italy	*25.31*	**0.00**	**0.00**	**0.00**	**0.00**	*11.57*	0.00	**1.22**	0.00	**1.22**
Japan	*18.13*	-1.84	-1.84	-3.68	**0.00**	*10.29*	**0.00**	-0.54	**0.00**	**0.00**
UK	*34.95*	**0.00**	-6.00	**0.00**	-6.00	*17.05*	**0.00**	-0.95	**0.00**	-0.95
US	*26.72*	0.00	**1.93**	**1.93**	**1.93**	*16.13*	0.00	**1.47**	**1.47**	**1.47**
Majority	*26.95*	6.00	6.00	**7.99**	**7.99**	*13.26*	2.84	2.84	**3.79**	**3.79**
Average	*31.74*	**-1.96**	-5.89	-3.93	**-1.96**	*18.38*	0.00	0.00	**1.31**	**1.31**

Table 4b: Forecasting 'Below Peak Recessions' BPR4, 2003q4-13q1

	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*19%*	0.86	0.89	**0.92**	0.89	**0.92**	0.0 $(-,*)$	0.58 $(**,-)$	0.68 $(**,**)$	0.58 $(**,**)$	**0.69** $(**,**)$
France	*49%*	0.81	0.78	0.78	**0.84**	**0.84**	0.0 $(-,-)$	−0.19 $(**,-)$	−0.19 $(**,-)$	**0.52** $(**,-)$	0.48 $(**,-)$
Germany	*35%*	**0.78**	0.73	0.76	0.68	0.76	0.0 $(-,-)$	−0.22 $(**,-)$	0.30 $(**,*)$	0.17 $(**,*)$	**0.42** $(**,**)$
Italy	*51%*	0.62	0.62	0.78	0.76	**0.84**	0.0 $(-,-)$	0.13 $(**,-)$	0.55 $(**,-)$	0.59 $(**,-)$	**0.65** $(-,-)$
Japan	*49%*	0.62	0.62	**0.68**	0.65	0.59	0.0 $(-,-)$	0.13 $(**,-)$	**0.34** $(**,-)$	0.23 $(**,-)$	0.06 $(**,**)$
UK	*49%*	**0.73**	**0.73**	**0.73**	0.70	0.70	0.0 $(-,-)$	0.0 $(**,-)$	**0.27** $(**,-)$	−0.28 $(**,-)$	0.07 $(**,-)$
US	*32%*	0.86	0.86	0.89	**0.92**	**0.92**	0.0 $(-,-)$	0.39 $(**,-)$	0.89 $(**,-)$	**0.91** $(**,**)$	**0.91** $(**,**)$
Majority	*49%*	**0.78**	0.70	0.70	0.76	**0.78**	0.63 $(**,-)$	0.63 $(**,-)$	0.63 $(**,-)$	0.68 $(**,-)$	**0.70** $(**,-)$
Average	*38%*	0.78	0.81	0.84	**0.86**	**0.86**	0.74 $(**,-)$	0.77 $(**,-)$	0.79 $(**,-)$	**0.82** $(**,-)$	**0.82** $(**,**)$

Table 4b (cont.): Forecasting 'Below Peak Recessions' BPR4, 2003q4-13q1

(Actual Return for RW, Improvement over RW for other models*)*

	Returns to Fair Bet (Symmetric)					Returns to Fair Bet (Asymmetric)				
	RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	*7.72*	2.89	4.33	4.33	**5.77**	*4.29*	8.57	12.86	12.86	**17.14**
France	*12.96*	2.00	0.00	**6.00**	**6.00**	*6.33*	1.06	0.00	**3.17**	**3.17**
Germany	*19.97*	-1.84	0.00	**1.84**	0.00	*12.92*	-1.85	0.00	**3.69**	1.85
Italy	*30.95*	**2.00**	**2.00**	**2.00**	**2.00**	*15.16*	0.95	0.95	0.95	**1.89**
Japan	*32.95*	0.00	-2.00	**2.00**	**2.00**	*16.89*	0.00	-1.06	1.06	**2.11**
UK	*16.96*	0.00	13.99	13.99	**17.99**	*8.44*	0.00	7.39	7.39	**9.50**
US	*12.85*	1.78	1.78	1.78	**7.12**	*6.25*	2.08	2.08	2.08	**8.33**
Majority	*14.96*	0.00	0.00	**4.00**	**4.00**	*7.39*	0.00	0.00	**2.11**	**2.11**
Average	*17.76*	1.89	3.78	**5.66**	**5.66**	*9.86*	1.64	3.29	**4.93**	**4.93**

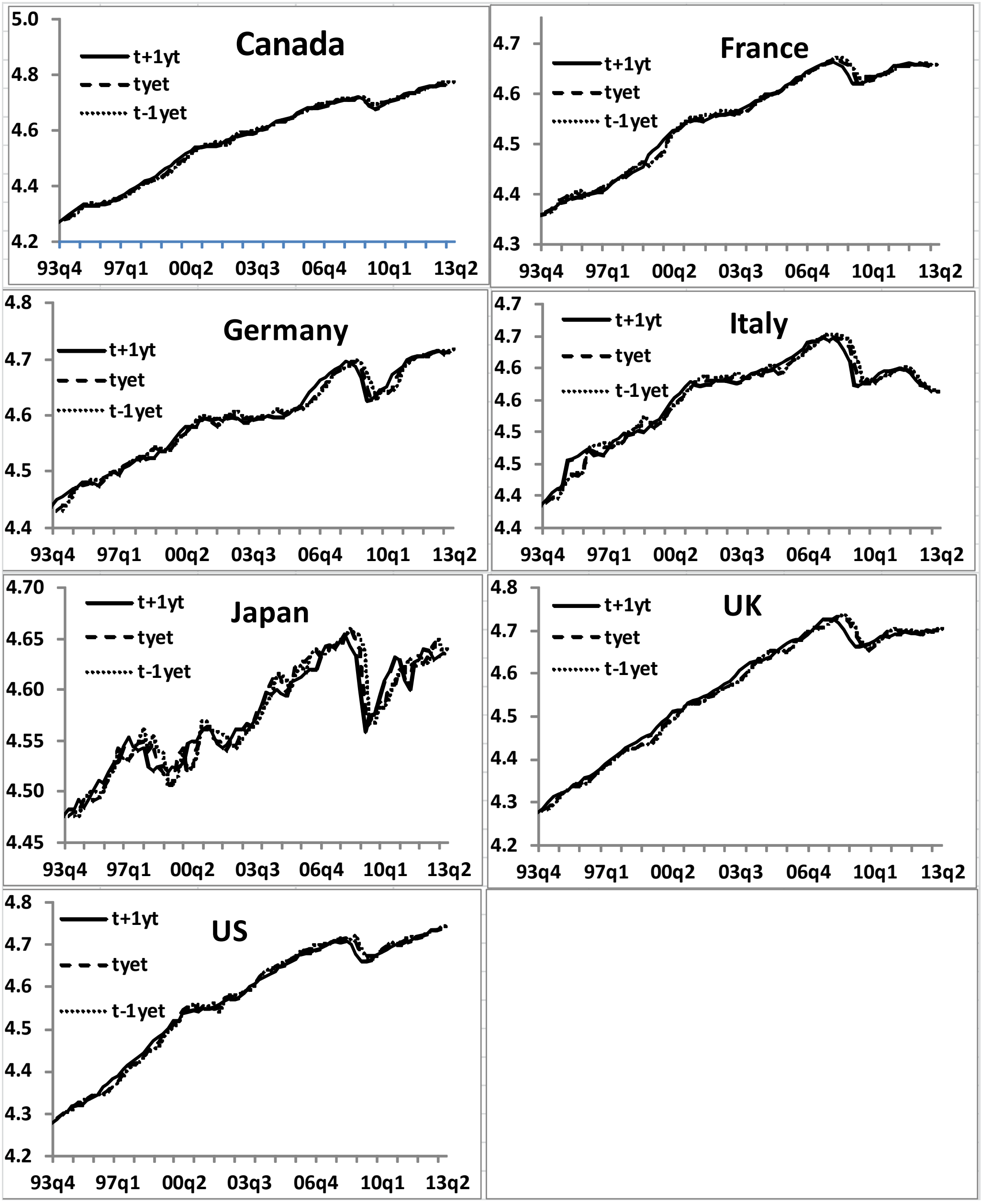

Figure 1: Actual, Nowcast and One-Period-Ahead Expected Output

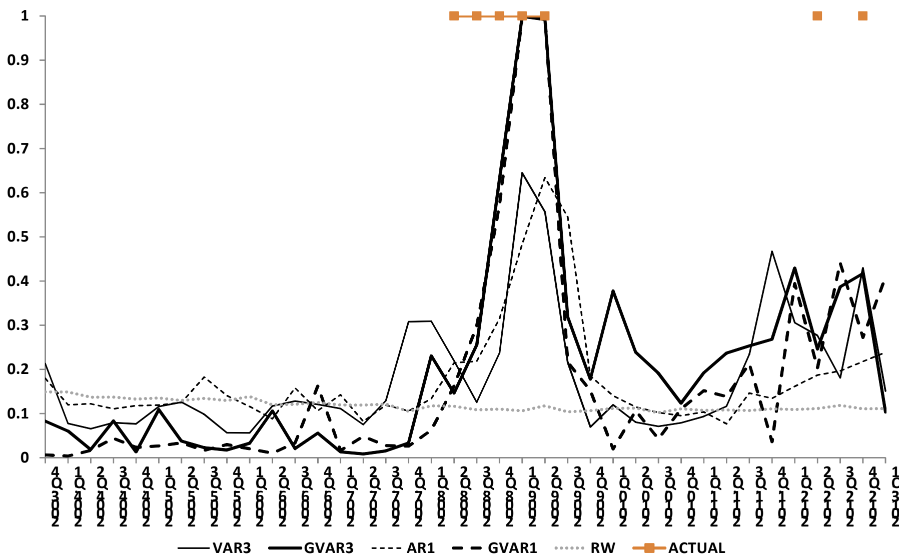

Figure 2: Probability of a Negative Nowcast in 4/7 of the G7

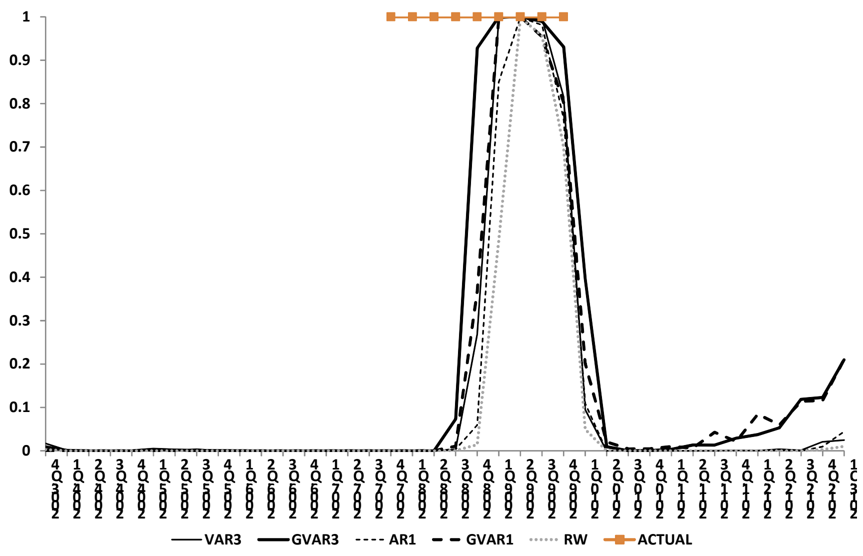

Figure 3: Probability of 9 period moving average growth $< 0\%$ in 4/7 of G7

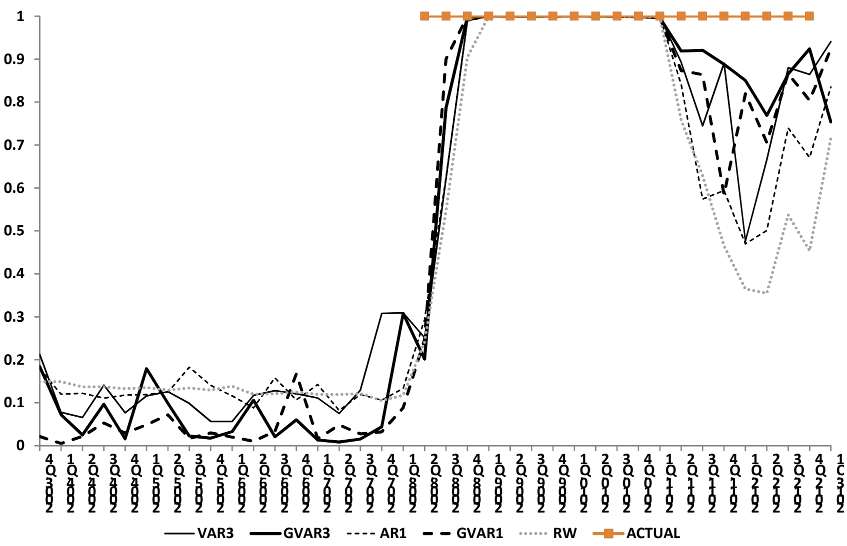

Figure 4: Probability of period T output less than previous peak in 4/7 of G7

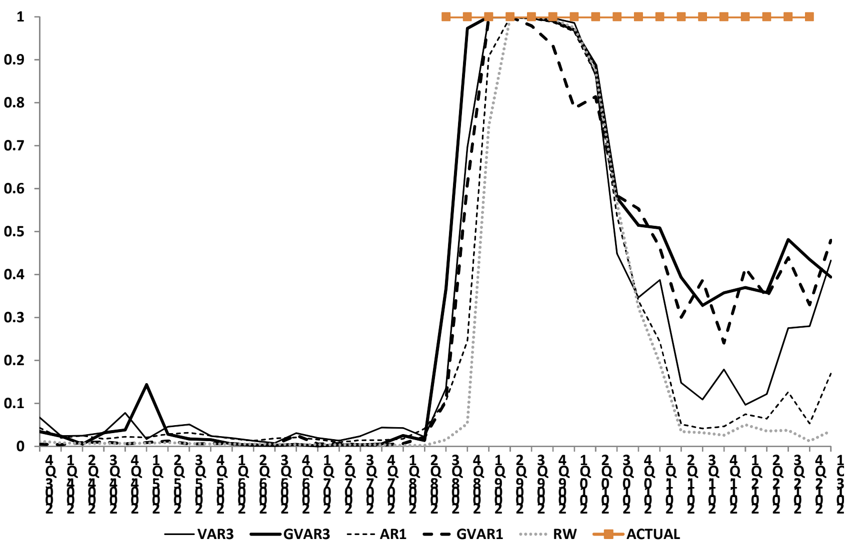

Figure 5: Probability of period T+4 output less than previous peak in 4/7 of G7.