

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

by

Anthony Garratt,[†] Kevin Lee^{††} and Kalvinder Shields^{†††}

Abstract

We compare a Global VAR model of actual and expected outputs with alternative models to assess the role of cross-country interdependencies and confidence in forecasting. Forecast performance is judged on point and density forecasts of growth, on probability forecasts of the occurrence of national and global recessionary events and, through a novel ‘fair bet’ exercise, on decision-making using probability forecasts. We find multi-country data and survey data are needed to fully capture the influence of global interactions and expectations in forecasts. We argue that output predictions should avoid simple point forecasts and focus on densities and events 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 May 2015. Corresponding author: Anthony 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 occurring 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 investigates 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 explaining more than 80% of world output movements as well as expected outputs in the G7 countries. The GVAR modelling framework is outlined in Pesaran, Schuermann

¹Analysis 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*.

and Weiner (2004), Garratt *et al.* (2006) and, in the 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, often 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

²A 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).

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

in business cycle fluctuations has been explored recently in Akerlof and Shiller (2009)'s 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 a GVAR model of G7 outputs to examine the role of inter-country interactions and expectations in explaining output growths. A variance-based measure of the persistent effect of shocks shows that, on average, the split between the global and national influences on persistent movements in the countries' outputs is in the ratio 50:50. Further, while most of the persistent movements in output are explained by fundamentals, around 10% are explained by 'sentiment' captured using survey data on expected output alongside the actual output data. GLS demonstrates that there is a considerable role for cross-country interactions and survey data in modelling countries' output growths therefore. The focus of the current paper goes beyond the description of model properties though and considers the use of output growth models for 29 countries 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 into account international links and the information available in survey data. We also find that, focusing on economic criteria, 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

⁴See 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.

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 densities and future output outcomes relevant to decision-makers instead.

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 29 economies over the period 1994q1-2014q2 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

Our modelling framework 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. Survey data provide information on agents' perceptions of the expected contemporaneous output level and of 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 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 will in principle provide 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 $s \geq 1$, the measure is from an official publication (published after the one quarter publication delay). If $s \leq 0$, 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

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

for $t = 1, \dots, T$ 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).⁵ 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 re-written in various alternative forms. For example, the model can also be re-written as a cointegrating VAR in the difference of $\mathbf{y}_{i,t}$ in which the assumed stationarity of the expectational errors built into (1) is reflected in (two) cointegrating vectors that ensure the three output measures move together one-for-one in the long run. This form highlights that the model is consistent with a wide range of expectation formation processes, including cases where actual and expected output are driven by the same fundamentals (e.g. rational expectations models with or without information rigidities), where expectations simply react to actual output movements (e.g. adaptive expectations) and where there is a two way interdependence between actual and expected outputs (with expectations playing a separate role in

⁵Note that this model involves expected growth terms based on the (log) difference between different series not through simple first-differencing.

determining output dynamics via ‘animal spirits’, for example). As elaborated in GLS, the alternative specifications translate into restrictions on the parameters of the VAR which can be tested.⁶

The model can also be written in the levels form

$$\mathbf{y}_{i,t} = \mathbf{A}_{i0} + \sum_{k=1}^{p+1} \mathbf{A}_{ik} \mathbf{y}_{i,t-k} + \boldsymbol{\varepsilon}_{i,t}, \quad t = 1, \dots, T, \quad (2)$$

where $\mathbf{y}_{i,t} = ({}_t y_{i,t-1}, {}_t y_{i,t}^e, {}_t y_{i,t+1}^e)'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{i,at}, \varepsilon_{i,bt}, \varepsilon_{i,ct})' = (\xi_{i,1t}, \xi_{i,1t} + \xi_{i,2t}, \xi_{i,1t} + \xi_{i,3t} + \xi_{i,3t})'$ and the \mathbf{A}_i 's are functions of the original Γ_i 's. This form will again retain the cointegrating properties of (1) and is particularly convenient for producing forecasts.

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 Dees *et al.* (2007), one approach to accommodating these interactions is to proxy the unobservable common factors by global aggregates in country VARs. Here, we can construct global variables, $\mathbf{y}_t^* = \sum_{j=1}^n w_j \mathbf{y}_{j,t}$ using 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

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

In practice, the \mathbf{y}_t^* variable used in model (2) can be defined using country-specific weights, $\mathbf{y}_{i,t}^* = \sum_{j=1}^n w_{ij} \mathbf{y}_{j,t}$ instead, where the weights are chosen so that the foreign variable can

⁶As noted earlier, the evidence in GLS suggests there is indeed a two-way interdependence between actual and expected outputs with 90% of the movements in output explained by fundamentals but 10% explained by ‘sentiment’.

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.⁷

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 $\mathbf{z}_t = (\mathbf{y}'_{1,t}, \dots, \mathbf{y}'_{n,t})'$ 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}_{i_s} and $\mathbf{B}_{i_s}^*$ 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

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

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

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

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.

Model (5) can be written as

$$\mathbf{z}_t = \Phi_0 + \sum_{k=1}^{p+1} \Phi_k \mathbf{z}_{t-k} + \mathbf{v}_t, \quad t = 1, \dots, T, \quad (6)$$

where $\Phi_0 = (\mathbf{I} - \mathbf{B}_0^* \mathbf{W})^{-1} \mathbf{B}_0$, $\Phi_k = (\mathbf{I} - \mathbf{B}_0^* \mathbf{W})^{-1} (\mathbf{B}_k + \mathbf{B}_k^* \mathbf{W})$, $k = 1, \dots, p+1$ and $\mathbf{v}_t = (\mathbf{I} - \mathbf{B}_0^* \mathbf{W})^{-1} \boldsymbol{\epsilon}_t$ with variance-covariance matrix Ω . Its simple linear form makes (6) particularly suitable for forecasting and decision-making using simulation methods.

⁷Note that this is an additional, explicit path of influence for global effects since global influences will effect the survey expectation series that are already included in (2).

Random draws from the estimated variance-covariance matrix Ω can be used to simulate future paths for the \mathbf{z}_t 's ($t = T + 1, \dots, T + h$), taking the estimated Φ_i 's as known. This generates the forecast density $\Pr(\mathbf{Z}_{T+1, T+h} | \mathbf{Z}_{1, T}, M_T^{GVAR})$ showing the likelihood of observing $\mathbf{Z}_{T+1, T+h} = \{\mathbf{z}_T, \mathbf{z}_{T+1}, \dots, \mathbf{z}_{T+h}\}$ given the observed data $\mathbf{Z}_{1, T} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_T\}$ 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 $t = 1, \dots, p+2$ and simulating data for $t = p+3, \dots, T$) each of which can be used to estimate an alternative version of (6) 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 a more detailed discussion.

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

The above GVAR methods focus on the role of expectations and global interactions in growth from a *statistical* modelling perspective. Frequently though, and as emphasised in Granger and Machina (2006), interest might often focus more on the *economic* importance of these effects, emphasising their implications for forecasts used in decision-making. In practice, it is difficult to make general statements on the economic impact of global interactions or confidence on outputs because the decisions of different individuals will be effected by output dynamics in different ways. For this reason, in what follows, we focus on the importance of using survey data and of accommodating global interactions 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, simulation methods can be used to generate the forecast density $\Pr(\mathbf{Z}_{T+1, T+h} | \mathbf{Z}_{1, T}, M_T^{GVAR})$ 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, $\mathbf{z}_{T+1}, \mathbf{z}_{T+2}, \dots$ can be written as $R(\mathbf{Z}_{T+1, T+h})$.

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

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

$$\text{probability of recession} = \int_R \Pr(\mathbf{Z}_{T+1,T+h} | \mathbf{Z}_{1,T}, M_T^{GVAR}) d\mathbf{Z}_{T+1,T+h}. \quad (7)$$

In a simulation exercise, the forecast probability is obtained simply as the proportion of 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 $\nu(r_T, R(\mathbf{Z}_{T+1,T+h}))$ depends on the outcome of a choice variable r_T and the occurrence of the recessionary event, the decision-maker's problem can be written as

$$\max_{r_T} \left\{ \int \nu(r_T, R(\mathbf{Z}_{T+1,T+h})) \Pr(\mathbf{Z}_{T+1,T+h} | \mathbf{Z}_{1,T}, M_T^{GVAR}) d\mathbf{Z}_{T+1,T+h} \right\}. \quad (8)$$

In terms of the simulations, the decision involves simply choosing the value of r_T that maximises the value of the objective function when averaging across the simulations. We can denote the optimal value of the choice variable chosen using model M_T^{GVAR} by r_T^{GVAR} .

Pesaran and Skouras (2000) then suggest using the statistic

$$\bar{\Psi}^{GVAR} = \frac{1}{k} \sum_{\tau=T}^{T+k} \nu(r_{\tau}^{GVAR}, R(\mathbf{Z}_{\tau+1,\tau+h})), \quad (9)$$

calculated over an out-of-sample evaluation period $T, \dots, T+k$ and based around the values of r_{τ}^{GVAR} chosen using model M_{τ}^{GVAR} in each period. Similar statistics can be calculated for any other model M_T^{\dagger} (with associated optimal choice variable r_T^{\dagger}) and these provide the basis of a comparison of the forecast performance of the models on economic grounds.

3.1 Forecast Performance Measured Using a "Fair Bet"

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. (See Johnstone *et. al.*, 2013, for a related analysis.). Framing the decision as a bet allows us to express the outcome of decisions made by an investor using a particular model in terms of the wealth she achieves. The appeal is that a model's probability forecasts are evaluated neither in abstract, nor in isolation, but by whether they would hypothetically have "made money" for the investor compared to bets based on rival models.

Here the investor places a bet on the basis of an estimate of the probability that a recession will occur. Two versions of the 'fair bet' can be considered - symmetric and asymmetric - to highlight different aspects of the forecast event. In the '*symmetric fair bet*' version, it is assumed that an investor gambles every period at a fixed charge of \$1, stating whether she believes the event will take place or not based on a probability forecast π derived from a model. If the investor forecasts the outcome of the event correctly, she receives a payout of $\$(s - 1)$ but otherwise loses the \$1. The payout contingencies are summarised as:

Payout contingencies for outcomes of a symmetric fair bet

Recession Forecast	Recession Occurs	
	<i>Yes</i>	<i>No</i>
<i>Yes</i>	$s - 1$	-1
<i>No</i>	-1	$s - 1$

The bet is fair in the sense that the payout, s , is chosen so that the investor would break even if her bet is based on the unconditional probability, p , that the event occurs. The expected return based on the unconditional probability is given by $(s - 1) [p^2 + (1 - p)^2] - 2p(1 - p)$ and the bet is 'fair' when this return is equal to zero; i.e. where $s = \frac{1}{2p^2 - 2p + 1}$.⁹ The investor's end-of-forecast-period wealth corresponding to $\nu(r_T, R(\mathbf{Z}_{T+1, T+h}))$ in (8)

⁹The payout for a correct prediction is therefore largest in a symmetric bet when $p = 0.5$ and, unconditionally, the event is equally likely to occur or not.

is given by

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

where $I(R)$ is an indicator function taking the value 1 or 0 if recession event R does or does not happen respectively and where r_T is equal to 1 if the investor bets on the recession occurring based on her forecasting model and 0 if she bets against recession. If the model's forecast probability is π and if the investor bets on recession when π exceeds some critical value π^c , then the decision to bet on recession or not ($r_T^\dagger = 1$ or 0) is equivalent to choosing the value for the critical value π^c . Noting that, in this case,

$$E[W_{T+h}] = \begin{cases} (s - 1)\pi - (1 - \pi) & = \frac{\pi}{2p^2 - 2p + 1} - 1 & \text{if } \pi > \pi^c \Leftrightarrow r_T = 1 \\ (s - 1)(1 - \pi) - \pi & = \frac{1 - \pi}{2p^2 - 2p + 1} - 1 & \text{if } \pi < \pi^c \Leftrightarrow r_T = 0 \end{cases}$$

maximum expected wealth is achieved by choosing a critical value of $\pi^c = 0.5$ since $\frac{\pi}{2p^2 - 2p + 1} > \frac{1 - \pi}{2p^2 - 2p + 1}$ if $\pi > 0.5$ and vice versa if $\pi < 0.5$. Here expected wealth rises as $\pi \rightarrow 0$ if $\pi < 0.5$ and as $\pi \rightarrow 1$ if $\pi > 0.5$ reflecting the benefits derived from a model that can make firm predictions that a recession will or will not occur in this symmetric case. Using an estimated model's probability forecasts to predict the occurrence of a recession or not will generate a sequence of financial returns, inserting the chosen r_T^\dagger in place of r_T in (10) for each observation during an evaluation period, and the total \$ outcome can be used to judge the model as in (9).

In the '*asymmetric fair bet*' case, the investor only bets for a stake of \$1 if she believes the event will occur. She wins $\$(s - 1)$ if she is correct and loses \$1 if she is incorrect. The payout contingencies are therefore:

Payout contingencies for outcomes of a asymmetric fair bet

Recession Forecast	Recession Occurs	
	<i>Yes</i>	<i>No</i>
<i>Yes</i> (bet)	$s - 1$	-1
<i>No</i> (no bet)	0	0

In this case, the expected return is given by $(s - 1)p^2 - p(1 - p)$ if the decision is based on the unconditional probability and, to make the bet fair, this is equal to zero when $s = \frac{1}{p}$.¹⁰ The investor's end-of-forecast-period wealth here is given by:

$$W_{T+h} = (s - 1) [r_T \times I(R)] - r_T(1 - I(R)) \quad (11)$$

and, if the investor bets on recession when π exceeds some critical value π^c , then

$$E[W_{T+h}] = \begin{cases} (s - 1)\pi - (1 - \pi) = \frac{\pi}{p} - 1 & \text{if } \pi > \pi^c \Leftrightarrow r_T = 1 \\ 0 & \text{if } \pi < \pi^c \Leftrightarrow r_T = 0 \end{cases}$$

Maximum expected wealth is achieved by choosing a critical value of $\pi^c = p$ since $\frac{\pi}{p} - 1 > 0$ if $\pi > p$. Here, expected wealth rises straightforwardly as $\pi \rightarrow 1$ reflecting the advantage of firm predictions of recession in this asymmetric case. Again, an estimated model's forecasts can be used to predict the occurrence of a recession or not in each observation during an evaluation period and this will generate a sequence of financial returns, inserting the chosen r_T^\dagger in place of r_T in (11) which can again be used to judge the model as in (9).

4 Forecasting Output and Recession in the G7, 1994q1-2014q2

Our forecast evaluation exercise is based on the actual output data for 29 of the world's largest countries and expected output data for the G7 countries over the period 1994q1-2014q2. The actual output data is the real volume GDP index for each country taken from the IMF's *International Financial Statistics* 2015q1. Our intention has been to produce a model of the world's output and we include in our analysis all countries for which data is available in a consistent form over our sample period. The 29 countries in our sample cover the large economies from all continents accounting for 83% of world output in 2013. 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. We include expectations data just for the G7 countries again because these are the only countries that have a long enough time-series for analysis. Full details of the country coverage, data sources and transformations employed are provided in the Data Appendix.

¹⁰In this asymmetric case, the payout for a successful bet increases monotonically as $p \rightarrow 0$.

It is worth noting that, while our focus in this paper is on the interplay between actual and expected outputs, we recognise that the 2015q1 vintage output data used here as our measure of actual output will have incorporated a number of revisions over time and this introduces a further complexity in modelling and forecasting. The VAR modelling framework described above can be readily extended to work with first-release data and subsequent revisions in place of the final vintage measures, as in Garratt *et al.* (2008, 2009). On the other hand, real-time data describing revisions in this way is available only for a small subset of the countries we consider so there is a practical reason for focusing on the final vintage data in this cross-country study. Moreover, an analysis of revisions alongside first-release and expected output measures would substantially increase both the complexity of the modelling (for example, (1) would become a five variable VAR if first-release data and two revisions were included) and the complexity of the evaluation exercise (dealing with the ambiguity over whether the survey responses refer to predictions of the first-release or post-revision outputs, for example). Of course, at least for the G7 countries, two out of the three variables we use for each country (namely the direct measures of expectations) are measured in real time and, as explained in the Data Appendix, we have manipulated the data to ensure that the final vintage measures of actual output we have used are consistent with the real time survey data. We do not model the revisions process explicitly then, interpreting the modelling and forecasting as though undertaken in real time and effectively assuming that revisions to data simply constitute noise.

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.

In our modelling and forecasting exercises, we estimate three equations explaining ${}_ty_{i,t-1}$, ${}_ty_{i,t}^e$ and ${}_ty_{t+1}^e$ for each of the G7 countries plus output equations explaining actual

output for each of the 22 non-G7 countries (making 43 estimated equations in total). The 21 equations for the G7 include global aggregates in the form of weighted averages of actual output growth of the 28 other countries *and* weighted averages of the expectations variables of the other six G7 countries. The 22 equations explaining output growth in the non-G7 countries include the weighted averages of actual output growth of the 28 other countries only. The model therefore captures all the interactions between actual output movements in our 29 countries plus the interactions of the expectations series across the G7 economies. Although our model can produce forecasts of actual output at the global level (i.e. across all 29 countries), our interest here is in judging the relative importance of cross-country interactions versus confidence and expectations for forecasting and so our forecast evaluation exercise focuses on forecasts of output growth and recession in the G7 countries only.

The forecast evaluation exercise is conducted recursively starting with the estimation period 1994q1-2003q3 and extending to 1994q1-2014q1 for exercises involving one-step-ahead forecasts ($h = 1$) and to 1994q1-2013q2 for $h = 4$. The sample choices are relatively arbitrary but are 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 follow a procedure in which the model used for forecasting is updated in each recursion unless the most recent observation of the output series is deemed to be an ‘outlier’. This judgement is based on a Chow test of predictive failure, the observation being considered an outlier if the Chow test is significant at the 1% level. If an outlier is identified, the model is reestimated with the most recent observation overwritten by the forecast from the previous period’s model when survey data are not used or by the previous period’s reported expected value when survey data are used. This procedure effectively dummies out the outlier observation of the series for the purposes of estimation and continues to use the previous period’s forecasting model. The usual recursive updating of models is resumed once the most recent observation on

growth passes the Chow test. 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.¹¹

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 which has a very similar model shows (i) in each of the G7 countries, the actual, nowcast and expected future output series all display similar mean growth rates and are all quite volatile,¹² 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 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, 50% of the persistent movements in countries' output is explained by global factors; and (iv) coefficient restrictions implied by the assumption of rationality in the presence of information rigidities are rejected and, although fundamentals dominate the persistent movements in output, a second variance-decomposition exercise shows that 'sentiment' also contributes around 10% of these persistent movements.

The rejection of the rationality restrictions in (iv) above means that expectational errors found in the survey data have systematic content that cannot be explained by simple information rigidities. This raises the question of whether a reduced form analysis of actual and expected outputs can provide an adequate vehicle for forecasting outputs or whether a more sophisticated structural model has to be used. As a partial answer to this, we undertook a simple recursive one-step-ahead out-of-sample forecasting exercise in which we added some additional macro variables to the output equations and tested whether this additional information helped improve output forecasts. Specifically, for each of the G7 countries, we estimated a second-order VAR explaining $ty_{i,t-1}$, $ty_{i,t}^e$ and $ty_{i,t+1}^e$,

¹¹Outliers were flagged in most models we consider and in most countries at some point during 2008 or 2009 but occurred less frequently - half as often - in models incorporating global interactions compared to those without.

¹²The average of the annualised actual growth rates across the seven countries is 1.8% and the average of the standard deviations is 3.0%.

and then extended this by adding two lags of inflation, short-run and long-run interest rates and exchange rates to the equations. We then computed one-step-ahead forecasts, using the data for 1994q3-2003q3 in the first recursion and for 1994q3-2014q1 in the last, and compared the RMSEs of the basic VAR and extended models. In the event we found the RMSEs from the basic VAR to be lower than those from the extended model for actual outputs in every country apart from Japan while the RMSEs from the extended model were lower for the nowcast and future expectations equations. However, none of the differences in RMSEs in any equation and in any country were statistically significant according to the Giacomini-White (2006) [GW] test of equal forecast performance.¹³

The finding that these core macro variables do not contribute significantly to the point forecasts of output reassured us that, even though the survey measures are not simple rational expectations of output, they do provide a good summary measure of the information relevant to forecasting output available in real time. Taken with the comments of (i)-(iii) above, showing that the direct measures of nowcast and expected future output provide a more sophisticated characterisation of output movements than could be captured by actual output data alone and that international interactions are also extremely important, the finding also reassured us that our reduced form output equations provide a good vehicle for assessing the relative importance of expectations and cross-country interactions in forecasting output.

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

Table 1 reports the root mean squared forecast error (RMSE) for the G7 countries, 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 cross-country linkages and the effects of including the expectations data. The first model variant is a simple univariate autoregressive model (of order two) for each of the G7 countries' own output growth, termed 'AR1'. The second variant uses the expectations data in addition

¹³Details on the macro variables employed and test results are available on request.

to actual output, as in (1), for each of the G7 countries but has no global (trade-weighted) \mathbf{z}_t^* variables. We denote this three-variable model model by ‘VAR3’. The third model explains output growth in all 29 countries and allows for feedbacks through the trade-weighted foreign output variables, but does not include expectations terms from the G7. This is a simple GVAR in actual outputs then and is denoted ‘GVAR1’. The final model is the full GVAR model for the 29 countries as in GVAR1 but including also equations for the expectations series in the G7 countries and feedbacks from the aggregated expectations series in the G7 regressions too as in (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. $_{T+1}y_T - _Ty_{T-1}$. The table shows that the RMSE are less than unity in nearly all cases, so that the models outperform the RW model, but that this improvement is statistically significant only once (the AR1 model for Italy) according to the GW test. The results are similar 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 six out of seven countries (as highlighted by the emboldened statistics) but the improvement over the RW model is statistically significant in just two of these. In terms of providing point forecasts then, only a small subset of the models considered provide statistically significant improvements on a random walk.

Tables 2a and 2b shift the focus from the point forecasts and towards density forecasts. At every recursion, the four estimated models, and the benchmark RW model, are each used to simulate 5000 potential future output paths and, hence, the densities of output growth nowcasts and four-step-ahead growth forecasts.¹⁴ Log predictive scores are then used to judge the models’ performance according to the forecast likelihood of the actual outcome as observed over the forecast evaluation period. The positive (scaled) log scores relative to those of the RW model show that the models outperform the benchmark RW model in nearly every country and nearly every model. Moreover, here, the statistics are often statistically significant. For the nowcasts of Table 2a, the largest value for the log scores is obtained by the GVAR1 or GVAR3 models in every case and these are

¹⁴Note here that we do not account for parameter uncertainty.

significantly better than the RW benchmark in four out of the seven countries. A similar pattern is found with the four-step-ahead forecasts in Table 2b where the GVAR models show the highest log score in six out of seven countries and these are again significantly better than the RW performance in four cases. It is worth noting though that it is not just the global interactions that are improving forecast performance here since the density forecast performance of the VAR3 model is also shown to significantly outperform the RW model in five out of seven cases in Table 2b.

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 models in point forecasting exercises but emphasise their usefulness in probabilistic statements on likely output outcomes. This means that the direct measures of expectations do not substantially improve the point forecasts of actual output which is surprising if the surveys are believed to reflect individuals' own rational real-time point forecasts of future outputs. However, this result matches the finding of GLS that survey responses are not straightforward measures of rationally-formed expectations of actual outputs and suggests instead that the expected output series, which we have seen are less volatile than the actual output series, improve forecast performance by delivering a more realistic range of forecast growth outcomes. Finally, and most clearly, the density forecast results show that interactions that occur across countries have a substantial impact on density forecasting performance, by taking into account other countries' output experiences, as in the GVAR1 and GVAR3 model.

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

We have argued that models' forecasting performance should 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 Germany and the UK grew quite strongly through 2007, Italy and Japan were already experiencing quarters of negative growth and Canada and US were also growing only very slowly.¹⁵ 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.¹⁶ And the recovery from 2009 has been very different, with Canada, Germany and US achieving output levels as high, or higher, than pre-crisis levels relatively quickly while output in France, Italy, Japan and the 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 and, through this variation, they provide a good test-bed for evaluating the four models we have developed for forecasting. In particular, we note the ODR recessions are relatively rare events in most countries and focus on short-run dynamics. In contrast, in terms of simple unconditional probabilities, the BPR recessions are reasonably common events during our sample, being approximately (with some variation over countries) as likely to occur as not, and they take a longer-term perspective influenced by past output levels.

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 $I_{(T+1)}y_T - y_{T-1} < 0$ 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

¹⁵Indeed, US output growth fell in 2008q1 and NBER dated the US recession to have started that quarter. This was well before the Lehman Brothers bankruptcy in 2008q3, for example, when global financial market problems began to dominate the headlines.

¹⁶The maximum year-on-year drop in output averaged at around -7.0% in the five countries outside N. America, dated 2009q2, while the average across Canada and US was -4.2% dated a quarter or so later.

recessionary event - denoted $ODR(0)$ below - occurred in just 10% and 17% of observations between 2003q4-2014q2 in Canada and US respectively, compared to 24%, 21% and 19% in France, Germany and UK, and 45% and 33% 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 $\tilde{y}_T^m = \frac{1}{2m+1} ({}_{T-m+1}y_{T-m} + \dots + {}_{T+1}y_T + \dots + {}_{T+m+1}y_{T+m})$. Hence, for example, we could consider the recessionary event $ODR(4) = I(\tilde{y}_T^4 - \tilde{y}_{T-1}^4 < 0)$, say, describing a drop in the nine-period moving average of output centred on T . This event occurred, on average, 4% 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 17% (majority) and 14% (average) of occasions according to $ODR(0)$ and 24% (majority and 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 based on the nowcast and is defined by $BPR(0) = I\{ {}_{T+1}y_T < \max\{{}_T y_{T-1}, {}_{T-1} y_{T-2}, {}_{T-2} y_{T-3}, \dots\} \}$, while a longer horizon (one-year-ahead) perspective is defined by $BPR(4) = I\{ {}_{T+5}y_{T+4} < \max\{{}_T y_{T-1}, {}_{T-1} y_{T-2}, {}_{T-2} y_{T-3}, \dots\} \}$ 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 only recovered their pre-crisis output levels by the end of our sample, it is not surprising to find that $BPR(0)$ occurs in around half or a little less observations in France, Germany, and UK, and more than half in Italy and Japan. The figure is lower in Canada and US at 21% and 36% 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 40% (majority) and 36% (average) of occasions according to $BPR(0)$ and 29% (majority and 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)¹⁷; and the outcome of two χ^2_1 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 statistics for global recessions. The second part of each of the tables reports the outcome of the symmetric and asymmetric fair bets described in (10) and (11). For the symmetric case, the 0.5 probability threshold is used to in predicting a recession will occur or not when undertaking the bet. For the asymmetric case, the forecast probability is compared to the unconditional probability of the event p to judge whether or not to participate in the bet.

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 on growth are also the models that best predict the *ODR* recessionary events. In Table

¹⁷If ‘YY’ indicates a recession is forecast and it occurs, ‘YN’ indicates a recession is forecast but it is not realised, and so on, then the hit rate is calculated as $HR = (YY + NN)/(YN + NN + NY + YY)$ and the Kuipers Score is defined by $H - F$, where $H = YY/(YY - YN)$ and $F = NY/(NY + NN)$.

3a, where the focus is on the nowcast of $ODR(0)$ recession, the GVAR3 model delivers the highest Kuipers scores in three out of seven countries, and also in predicting global recession. The more demanding dynamic PT test shows the GVAR3 model has significant predictive power in all five cases even taking into account the runs in the events. The dominance of the GVAR3 model is also 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 four of the seven countries and when looking at the G7 average and significantly so in these cases according to the dynamic PT test.

Figures 2 and 3 illustrate the characteristics of the probability forecasts underlying these results. None of the models anticipated the onset of the financial crisis but the two GVAR models and VAR3 model accommodated the news and reacted after one or two quarters, correctly nowcasting recession over the latter part of 2009. For these three models, the forecast probability of global recession drops through 2010 and 2011 but rises again through 2012, anticipating the actual events well. The relative forecasting performance of the three models here are down to small differences in timing although, of course, these can be very important in real world decision-making.

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. Here we report, for each model, the payouts for decisions relating to individual countries' experience of recessions, the average of the countries' payouts, ("country average") and the payouts relating to the global recession events defined earlier. The results show that, according to the fair bet criterion, the VAR3 and GVAR3 models dominate. For the $ODR(0)$ event, the VAR3 model delivers the largest average payout of \$3.34 and \$13.48 (relative to the RW model) for the symmetric and asymmetric cases respectively and also performs best in bets involving the majority of countries. The same applies to the $ODR(4)$ events. The GVAR3 model also performs well and delivers the highest return in bets involving average growth across the G7. Hence, the decision-making here appears to be enhanced primarily through the use of the survey data. As noted previously, this data appears to give more realistic spreads on density forecasts relating to growth and this feature translates to high returns in bets based on probability

forecasts.

Turning to the *BPR* recessions in Table 4, we note that the very high degree of serial correlation in the occurrences of this event means that, in terms of hit rates and Kuipers scores, it is difficult to distinguish between the forecast performance of the various models, with all models performing well relative to the RW benchmark. However, the GVAR models, and especially the GVAR1 model, clearly outperform the others in terms of the fair bet evaluation. In the symmetric case, the GVAR1 model delivers the largest payout for the *BPR*(0) and *BPR*(4) events for nearly all countries and for both global recession events, and the same is true for the *BPR*(4) event in the asymmetric bet scenario.

Figures 4 and 5 illustrate the probability forecasts behind these results looking at the global recession defined by the majority G7 experience. Again the VAR3, GVAR1 and GVAR3 models all reacted reasonably similarly to correctly nowcast the *BPR* events in 2008, to forecast the end of the recession and to predict the turbulence of 2012. However, the figures illustrate that it is the GVAR models (and especially GVAR1 model) which correctly places high probability on the forecast that output would drop below peak through 2012 in many countries, especially in the longer horizon event of *BPR*(4). The differences are again down to small margins then, but the GVAR1 model systematically outperforms the others and this translates into clear benefits according to the returns to the bets. The recessions defined by the *BPR* events involve a longer historical perspective compared to the *ODR* recessions and it appears that the cross-country linkages captured in the GVAR models are best able to capture more prolonged effects of shocks and to predict recessionary events that are defined over a longer timeframe.

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 importantly on the definition of recession. The point is made clearly in the context of the comparison of the models using a 'neutral' fair bet investment scenario. The pay outs associated with the *BPR* recession definitions favour the GVAR models with their sophisticated cross-country linkages but it is the inclusion of survey data in the VAR3 and GVAR3 models that appears most important when we use *ODR* recession definitions.

Third, in terms of a simple 'horse-race' between models and given the clear dominance of GVAR1 in the *BPR* recessions, 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. The GVAR models are able to accommodate more complex dynamics than country-specific models and the ability to capture the propagation of shocks across countries appears important in predicting when countries will regain previous output levels. 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 it is essential to take this into account in density forecasting, event probability forecasting or forecast-based decisions.

And *fourth*, while cross-country effects appear to be 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 the nowcast of output as expressed by survey respondents) contribute more to the models' density forecasts than to their point forecasts. This means

the contribution of expectations data is to generate more realistic spreads in forecasts rather than from the immediacy of including 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 provide straightforward full-information rational expectations of the underlying fundamentals. Rather, the survey measures appear to incorporate 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, delivering a more realistic range of forecasted outputs.

Despite the uncertainties and reservations expressed in the media and by policy makers on predicting growth prospects, the analysis of the paper shows that time series 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 recessionary 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, etc.) are of interest to the commentators since models can only really be judged according to their usefulness in the particular context.

6 Data Appendix

The actual output data used in the analysis is the real volume GDP index taken from the IMF's *International Financial Statistics* 2015q1 for twelve countries in Europe (Austria, Belgium, Finland, France, Germany, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom), six countries in the Americas (Argentina, Brazil, Canada, Chile, Mexico, United States), eight in Asia (China, India, Indonesia, Japan, Korea, Malaysia, Thailand, Turkey), and Australia, New Zealand and South Africa.

Of the 29 countries output series, 14 were seasonally adjusted including the G7. The remaining 15 countries (Argentina, Austria, Belgium, Brazil, Chile, China, Finland, India, Indonesia, Korea, Malaysia, Norway, Sweden, Thailand and Turkey) require seasonal adjustment and, following Dees *et al.* (2007), we use the X12 filter in *Eviews* to accomplish this. The adjustment was performed on the log of the first difference of the variable using the additive option. Using the first observation of the unadjusted log series, the adjusted log changes were then accumulated. The adjusted level series was then obtained by taking the exponential of the adjusted log series. Note also that not all countries have IMF data available from 1990q1, in particular Brazil 1995q1, China 2000q1, India 1997q1, Indonesia 1997q1 and Thailand 1993q1. In these cases, having constructed seasonally adjusted series, we use growth rates from the Dees *et al.* data set to project backwards to 1990q1.

To ensure that the actual and expected series are properly aligned, we treat the measure provided in the IMF's 2015 publication as the 'true' measure of output and assume this is observed after a one quarter delay (so that ${}_{t+s}y_t = {}_T y_t$ for all $s \geq 1$). We then construct, for the G7 countries, 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.

7 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.

Granger, C.W.J and M. Machina (2006), “Forecasting and Decision Theory”, chapter 2 in G. Elliott, C.W.J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 1, 81-98, Elsevier.

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

Johnstone, D.J., S. Jones, V.R.R. Jose and M. Peat, (2013), “Measures of the Economic Value of Probabilities of Bankruptcy”, *Journal of Royal Statistical Society*, 176, 1-19.

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. Faruquee, E. Prasad and B. Turtleboom, (1998), “MULTI-MOD Mark III The Core Dynamic and Steady-State Models”, *IMF Occasional Papers*

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	<i>0.047</i>	0.788	0.655	0.724	0.657
France	<i>0.041</i>	0.834	0.835	0.681	0.803
Germany	<i>0.066</i>	0.961	0.865	0.757	0.823
Italy	<i>0.061</i>	0.885*	0.932	0.688	0.689
Japan	<i>0.084</i>	1.031	0.872	0.985	1.028
UK	<i>0.055</i>	0.786	1.084	0.847	1.123
US	<i>0.050</i>	0.874	0.837	0.893	0.814

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	<i>0.160</i>	0.911*	0.884	0.946	1.050
France	<i>0.154</i>	0.941	0.896	1.051	1.609
Germany	<i>0.211</i>	1.003	0.987	1.159	1.112
Italy	<i>0.244</i>	0.951*	0.807*	1.043	1.488
Japan	<i>0.205</i>	0.993	1.018	1.097	1.538
UK	<i>0.223</i>	0.949*	0.828	0.942*	1.176
US	<i>0.188</i>	0.989*	0.861*	1.045	0.997

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 Giacomini-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	66.967	0.479	0.088	0.947	0.884
France	77.675	0.171	0.248	0.778*	0.773*
Germany	98.767	0.069	0.162	0.521	0.326
Italy	83.669	0.169*	0.281*	0.582*	0.624*
Japan	99.034	-0.052	0.054	0.199	0.097
UK	-76.842	1.072*	0.298	1.932*	1.926*
US	68.753	0.297	0.234	0.934*	0.759*

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	-20.750	4.075*	4.338*	4.886	4.503
France	-13.497	4.524*	3.838*	6.324*	5.656
Germany	-5.659	3.948*	4.885	11.190	10.454
Italy	-33.687	1.330*	2.169*	2.497*	2.282*
Japan	43.852	-0.301	0.053	-0.093	-0.271
UK	-357.163	0.654*	0.671*	0.820*	0.846*
US	-56.573	1.495*	1.665*	2.021*	1.899*

Notes: See notes to Table 1. The scaling takes the form: $\widetilde{LPS}_i = (LPS_i - LPS_{RW}) / |LPS_{RW}|$ for $i = AR1, VAR3, GVAR1$ and $GVAR3$. 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 Giacomini-White (2006) test of equal forecast. performance

Table 3a: Forecasting ‘Output Drop Recessions’ ODR0, 2003q4-14q2

	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	10%	0.90	0.93	0.93	0.93	0.90	na	0.62 (**,**)	0.62 (**,**)	0.57 (**,**)	0.47 (**,**)
France	24%	0.76	0.83	0.83	0.81	0.81	na	0.80 (**,**)	0.64 (**,**)	0.57 (**,**)	0.57 (**,**)
Germany	21%	0.79	0.76	0.74	0.62	0.81	na	-0.22 (-, -)	0.14 (-, -)	0.13 (-, -)	0.45 (**,**)
Italy	45%	0.55	0.57	0.74	0.76	0.81	na	0.23 (-, -)	0.48 (-, -)	0.57 (**,**)	0.64 (**,**)
Japan	33%	0.67	0.67	0.69	0.64	0.57	na	0.00 (-, -)	0.29 (-, **)	0.19 (-, -)	0.04 (-, -)
UK	19%	0.81	0.88	0.83	0.83	0.81	na	0.87 (**,**)	0.47 (**,**)	0.47 (**,**)	0.33 (-, -)
US	17%	0.83	0.83	0.86	0.83	0.88	na	0.35 (-, -)	0.85 (**,**)	0.37 (**, -)	0.88 (**,**)
G7 Majority	17%	0.83	0.88	0.86	0.81	0.90	na	0.88 (-, -)	0.85 (-, -)	0.31 (**, -)	0.90 (**,**)
G7 Average	14%	0.86	0.83	0.93	0.81	0.88	na	-0.07 (-, -)	0.45 (-, -)	0.31 (**, -)	0.64 (**,**)

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	3.91	1.21	1.21	1.21	0.00	0.00	12.50	30.00	20.00	26.00
France	8.82	3.14	4.71	3.14	3.14	0.00	7.60	12.20	10.40	12.60
Germany	7.75	-1.51	-3.02	-10.55	1.51	0.00	8.00	3.67	4.33	2.33
Italy	3.59	1.98	15.86	17.84	21.80	0.00	2.63	11.74	12.74	16.37
Japan	8.40	0.00	1.80	-1.80	-7.20	0.00	-7.00	3.00	1.00	4.00
UK	7.16	4.34	1.45	1.45	0.00	0.00	12.75	19.75	13.50	14.50
US	6.46	0.00	1.38	0.00	2.77	0.00	13.00	14.00	12.00	14.00
Country Average	6.58	1.31	3.34	1.61	3.15	0.00	7.07	13.48	10.57	12.83
G7 Majority	6.46	2.77	1.38	-1.38	4.15	0.00	10.00	5.00	11.00	15.00
G7 Average	5.66	-1.32	3.97	-2.65	1.32	0.00	-1.00	18.00	8.00	11.00

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. 'na' indicates the KS is not calculated where one outcome is always forecast to occur.

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	18%	0.91	0.84	0.87	0.87	0.87	0.33 (-, -)	0.52 (-, -)	0.52 (-, -)	0.63 (-, -)	0.63 (-, -)
France	26%	0.84	0.84	0.87	0.87	0.87	0.82 (**,*)	0.82 (**,*)	0.82 (**,**)	0.85 (**,**)	0.73 (**,**)
Germany	26%	0.82	0.82	0.90	0.79	0.79	0.62 (**, -)	0.62 (**,**)	0.62 (**, -)	0.46 (**, -)	0.46 (**, -)
Italy	53%	0.71	0.76	0.82	0.79	0.79	0.62 (**,**)	0.67 (**,**)	0.67 (**,**)	0.69 (**,**)	0.69 (**,**)
Japan	26%	0.82	0.79	0.84	0.82	0.87	0.55 (**, -)	0.48 (**,**)	0.48 (**,**)	0.53 (**, -)	0.73 (**, -)
UK	24%	0.87	0.87	0.90	0.87	0.87	0.85 (**, -)	0.85 (**, -)	0.85 (**, -)	0.85 (**, -)	0.71 (**,**)
US	24%	0.84	0.87	0.87	0.87	0.90	0.83 (**, -)	0.85 (**,**)	0.85 (**,**)	0.85 (**,**)	0.88 (**,**)
G7 Majority	24%	0.84	0.88	0.89	0.84	0.87	0.82 (**,**)	0.85 (**, -)	0.88 (**, -)	0.59 (**,**)	0.71 (**,**)
G7 Average	24%	0.84	0.87	0.87	0.87	0.89	0.83 (**,**)	0.85 (**, -)	0.85 (**, -)	0.85 (**,**)	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	6.32	1.43	2.86	2.86	2.86	7.86	3.43	8.86	2.43	3.43
France	14.27	0.00	1.63	1.63	1.63	11.20	0.00	1.80	2.60	5.40
Germany	12.64	0.00	4.90	-1.63	-1.63	14.80	0.00	-1.20	-1.40	-0.40
Italy	15.85	3.99	7.98	5.98	5.98	8.10	0.90	3.70	2.70	2.70
Japan	12.64	-1.63	1.63	0.00	3.27	12.60	2.00	0.00	1.80	2.00
UK	13.68	0.00	1.57	0.00	0.00	12.89	0.00	2.22	-8.00	0.22
US	12.12	1.57	1.57	1.57	3.13	12.90	0.00	3.22	2.22	2.22
Country Average	12.50	0.77	3.16	1.49	2.18	12.90	0.00	3.22	2.22	2.22
G7 Majority	12.11	1.57	3.13	0.00	1.57	9.67	3.32	6.44	4.44	5.44
G7 Average	12.11	1.57	1.57	1.57	3.13	9.67	3.32	3.32	3.32	6.44

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	21%	0.93	0.93	0.93	0.93	0.91	0.92 (**,*)	0.92 (**,*)	0.82 (**,**)	0.82 (**,**)	0.72 (-, -)
France	40%	0.79	0.81	0.86	0.76	0.76	0.65 (**, -)	0.68 (**, **)	0.70 (-, -)	0.51 (**, **)	0.51 (**, **)
Germany	48%	0.83	0.83	0.83	0.67	0.83	0.68 (**, -)	0.68 (**, -)	0.67 (**, -)	0.35 (**, **)	0.67 (**, -)
Italy	64%	0.90	0.88	0.86	0.91	0.91	0.79 (**, -)	0.74 (**, -)	0.69 (-, -)	0.79 (**, -)	0.78 (**, -)
Japan	62%	0.81	0.79	0.83	0.86	0.83	0.60 (**, -)	0.55 (**, -)	0.67 (**, **)	0.70 (**, -)	0.67 (**, **)
UK	50%	0.95	0.95	0.93	0.91	0.91	0.91 (**, -)	0.91 (**, **)	0.86 (**, **)	0.91 (**, **)	0.81 (-, -)
US	36%	0.83	0.91	0.91	0.86	0.88	0.79 (**, -)	0.87 (**, **)	0.87 (**, -)	0.82 (**, **)	0.84 (**, -)
G7 Majority	40%	0.88	0.90	0.95	0.88	0.86	0.83 (**, -)	0.86 (-, -)	0.90 (**, **)	0.75 (**, **)	0.70 (**, -)
G7 Average	36%	0.93	0.90	0.90	0.88	0.90	0.90 (**, -)	0.82 (-, -)	0.86 (**, **)	0.74 (**, **)	0.78 (**, -)

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	16.80	0.00	0.00	0.00	-1.51	25.67	-5.00	0.67	-0.33	-1.33
France	21.69	1.93	5.79	-1.93	-1.93	13.71	-0.53	4.35	-1.06	-1.06
Germany	27.84	0.00	0.00	-13.97	0.00	14.5	1.10	0.20	-5.90	0.30
Italy	28.26	-1.85	-3.70	0.00	0.00	12.78	0.00	-2.00	0.00	-1.00
Japan	22.35	-1.89	1.89	3.79	1.89	12.3	0.00	0.00	-1.00	0.00
UK	38.00	0.00	-2.00	0.00	-4.00	19.00	0.00	-1.00	0.00	-2.00
US	22.72	5.55	5.55	5.55	3.70	19.8	1.80	1.80	1.60	0.80
Country Average	25.38	0.53	1.08	-0.94	-0.26	16.82	-0.38	0.57	-0.96	-0.61
G7 Majority	20.52	0.00	0.00	1.69	0.00	37.00	0.00	0.00	1.00	0.00
G7 Average	18.83	0.00	0.00	1.69	1.69	15.00	0.00	0.00	2.50	2.50

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

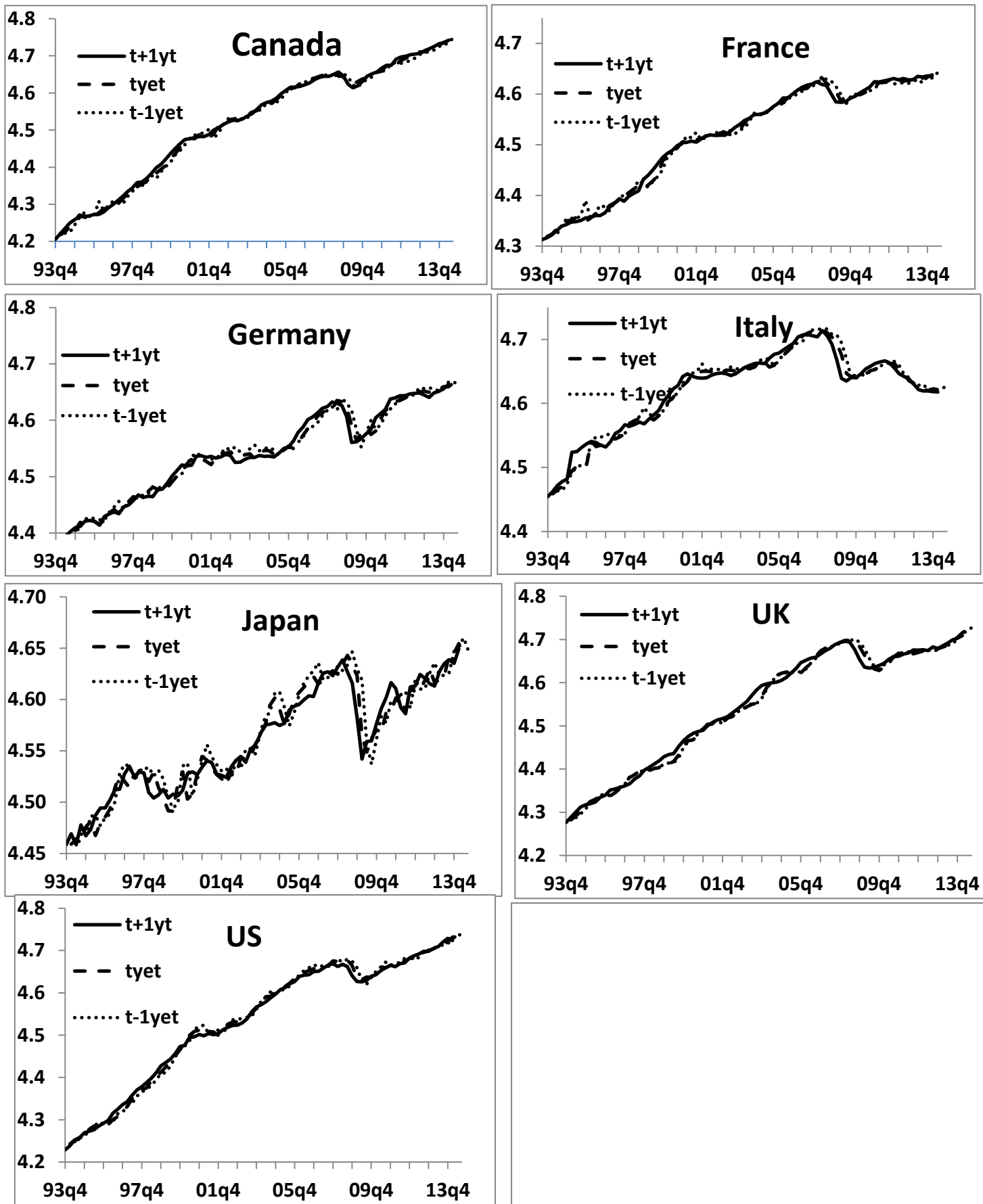
	p	Hit Rates					Kuipers Score				
		RW	AR1	VAR3	GVAR1	GVAR3	RW	AR1	VAR3	GVAR1	GVAR3
Canada	17%	0.86	0.91	0.93	0.93	0.93	0.85 (-, -)	0.90 (**, -)	0.92 (**, **)	0.92 (**, **)	0.92 (**, **)
France	26%	0.86	0.88	0.86	0.95	0.91	0.84 (-, -)	0.86 (**, -)	0.84 (**, -)	0.94 (**, **)	0.89 (**, **)
Germany	38%	0.81	0.79	0.79	0.76	0.76	0.77 (-, -)	0.74 (**, -)	0.74 (**, **)	0.50 (**, **)	0.55 (**, **)
Italy	57%	0.93	0.95	0.95	0.98	0.98	0.86 (-, -)	0.90 (**, -)	0.90 (**, -)	0.95 (**, -)	0.95 (**, -)
Japan	50%	0.91	0.91	0.93	0.95	0.93	0.84 (-, -)	0.84 (**, **)	0.88 (**, **)	0.91 (-, -)	0.88 (**, **)
UK	48%	0.69	0.74	0.83	0.91	0.81	0.63 (-, -)	0.67 (**, -)	0.76 (**, -)	0.85 (**, **)	0.73 (**, -)
US	26%	0.81	0.83	0.83	0.86	0.88	0.80 (-, -)	0.82 (**, -)	0.82 (**, -)	0.84 (**, -)	0.86 (**, **)
G7 Majority	29%	0.88	0.86	0.88	0.86	0.88	0.86 (-, -)	0.86 (**, **)	0.86 (**, -)	0.77 (-, -)	0.86 (**, **)
G7 Average	29%	0.86	0.86	0.86	0.88	0.88	0.83 (-, -)	0.83 (**, **)	0.83 (**, -)	0.86 (-, -)	0.86 (**, **)

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	7.85	2.77	4.15	4.15	4.15	15.00	10.00	9.00	10.00	15.00
France	16.69	1.63	0.00	6.52	3.26	19.72	0.00	0.00	4.45	-4.18
Germany	22.35	-1.89	-1.89	-3.79	-3.79	13.00	0.00	-1.00	-6.75	0.25
Italy	34.44	1.96	1.96	3.92	3.72	15.75	0.75	0.75	1.50	1.50
Japan	34.00	0.00	2.00	4.00	2.00	17.00	0.00	1.00	2.00	1.00
UK	15.87	3.99	11.97	17.96	7.98	7.70	2.20	6.60	11.00	5.50
US	13.43	1.63	1.63	3.26	4.89	8.45	5.64	11.27	16.91	14.09
Country Average	20.66	1.44	2.83	5.15	3.17	13.80	5.64	11.27	16.91	14.09
G7 Majority	20.52	0.00	0.00	1.69	0.00	9.67	0.00	0.00	1.00	0.00
G7 Average	18.83	0.00	0.00	1.69	1.69	15.00	0.00	0.00	2.50	2.50

Figure 1: Actual, Now cast and One-Period-Ahead Expected Outputs



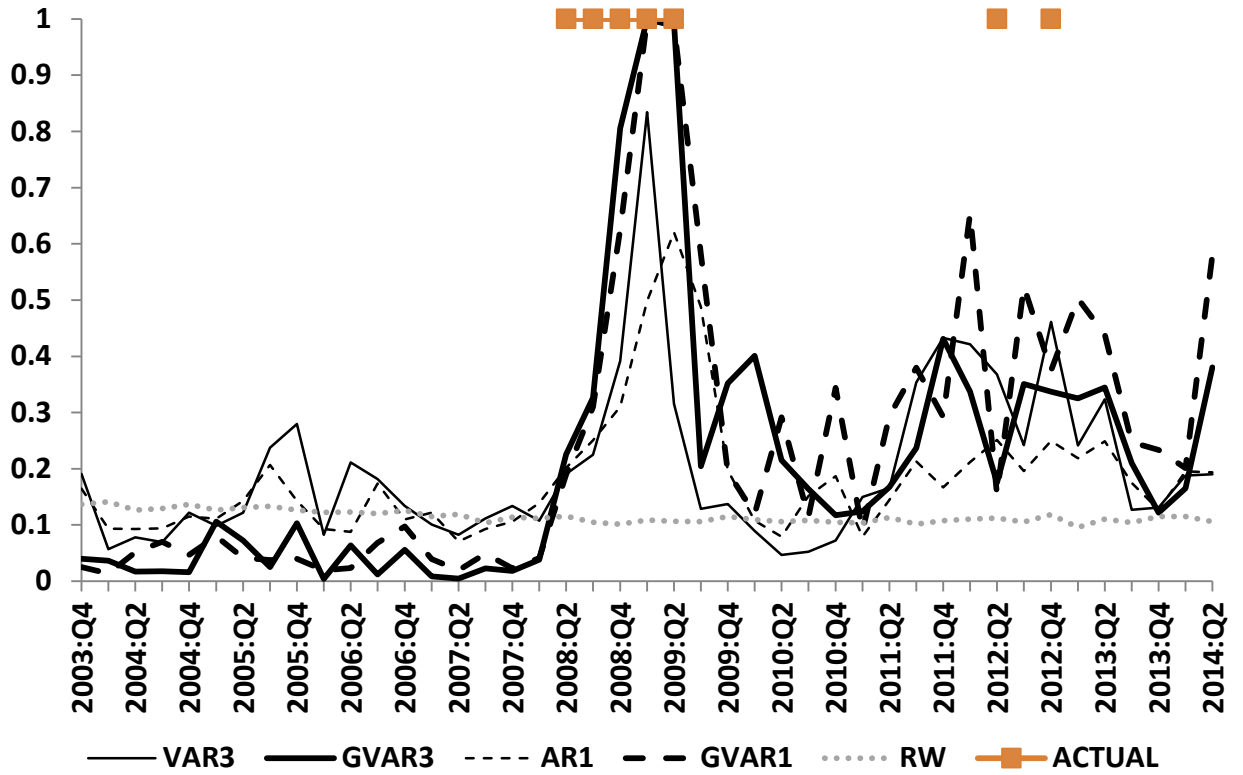


Figure 2: Probability of a Negative Nowcast in 4/7 of the G7.

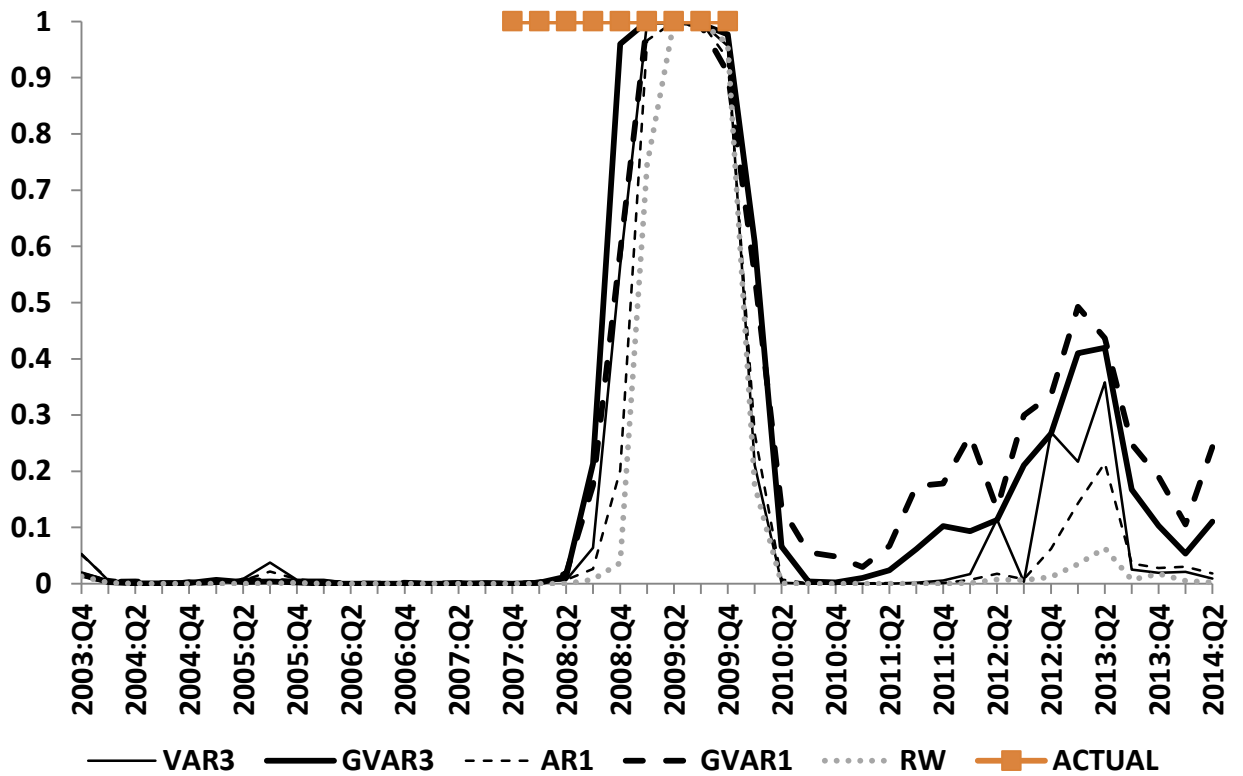


Figure 3: Probability of 9 period moving average growth < 0% in 4/7 of G7.

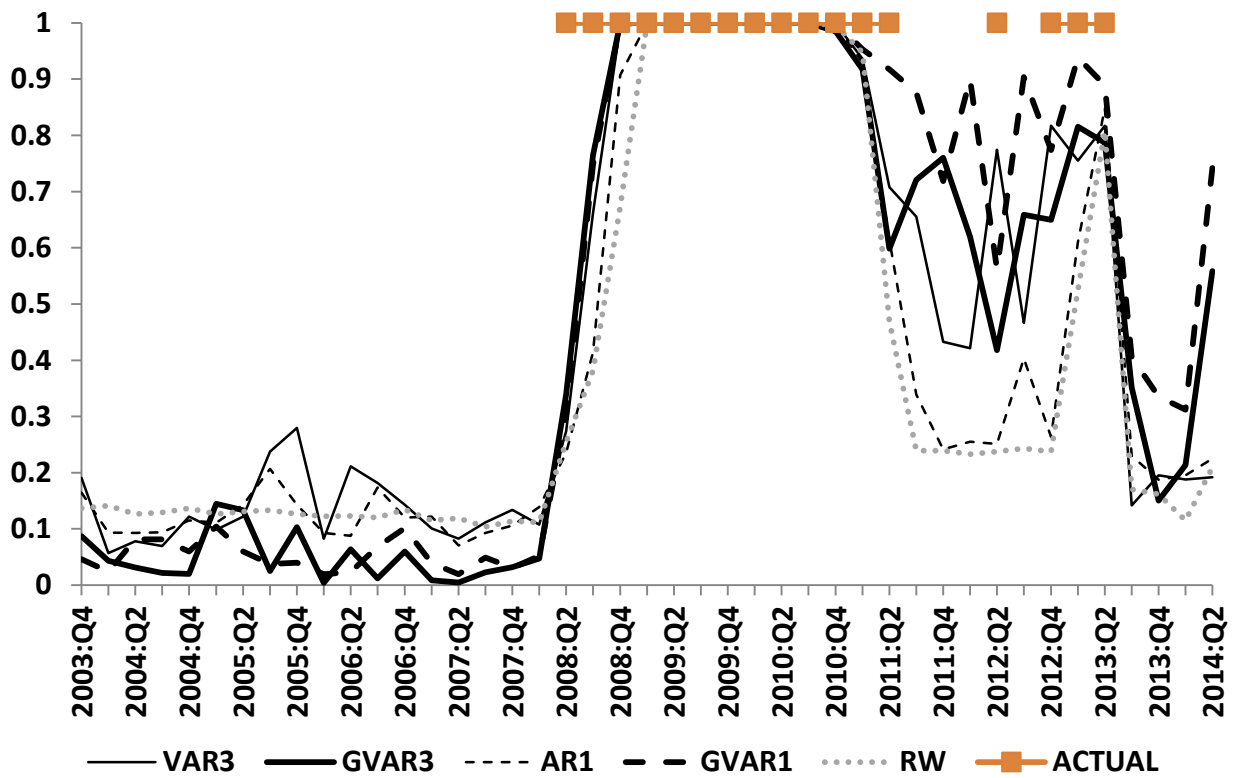


Figure 4: Probability of period T output less than previous Peak in 4/7 of G7.

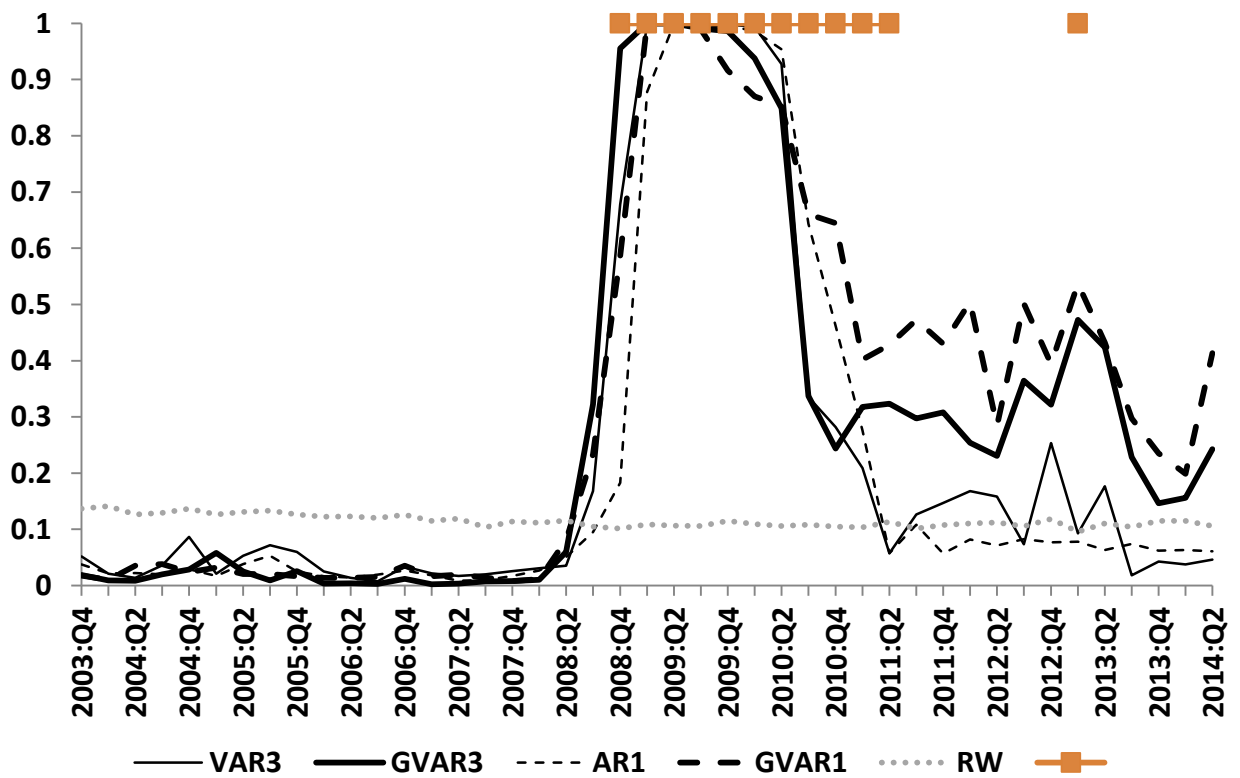


Figure 5: Probability of period T+4 output less than previous peak in 4/7 of G7.