

# **“Warwick Business School Forecasting System”**

## **Summary**

*Ana Galvao, Anthony Garratt and James Mitchell*  
*November, 2014*

The main objective of the *Warwick Business School Forecasting System* is to provide competitive and accurate benchmark (judgement-free) density and probability event forecasts for both UK real GDP (output) growth and CPI inflation –for current and next year GDP growth and CPI inflation. Other forecasts produced, for example, by the Bank of England, OBR, international organisations, commercial entities and averaged consensus forecasts can accordingly be placed in the context of forecasts based on the latest models and techniques used in the academic literature.

Using a range of leading and statistically motivated econometric models – as opposed to a single model - we produce judgement-free point and density forecasts for annualised real UK GDP growth and CPI inflation. Our approach uses model averaging, following well-established methods in statistics, meteorology and, more recently, economics. Model averaging can take a variety of forms but, in the *Warwick Business School Forecasting System*, uses a weighted combination of models’ density forecasts; where higher weights are awarded to models which show the better recent forecasting performance. The emphasis on model averaging reflects the view that any single “correct” model is in practice, if not in reality, unknown. The models used by economists to forecast are therefore likely mis-specified. Individual models also tend to vary in their forecasting performance over time. Hence we apply methods which seek to mitigate these problems, by considering a range of models and allowing the weights on these models to vary over time, in relation to their recent forecasting performance.

An important aspect to the *Warwick Business School Forecasting System* is the explicit quantification, communication and then evaluation of the degree of uncertainty associated with the forecasts. Reporting both the forecast density, and the probability forecasts of economic events of interest extracted from the density, aids decision makers who care about more than just the central (or point) forecast. For example, many decision makers are concerned with risk and the likelihood of tail events, say, occurring. Our focus on forecast uncertainty is shared by the Bank of England who also report density forecasts, again on the assumption that point forecasts, the traditional focus, are of limited interest and value. Indeed, one objective of the *Warwick Business School Forecasting System* is to provide a competitive and more accurate benchmark (judgement-free) alternative to the Bank of England’s own forecast, based on the deliberations and judgement of their Monetary Policy Committee (MPC).

## The Choice of Forecasting Models

We take a combination of density forecasts produced from a range of leading models used in the academic (macro-econometrics) literature. These range from simple univariate models, known to often forecast well, through to multivariate statistical models with time-varying parameters. Importantly, we find that taking such combinations provides robustness to structural breaks; and is insurance against selecting the 'wrong' single model in real-time. The combinations can also be seen to approximate nonlinear specifications.

The five models we consider are:

- (1) A single equation autoregressive model for output growth or inflation.
- (2) A mixed data sampling (MIDAS) approach which consists of time-series regressions that allow for the use of time-series sampled at different frequencies. We use quarterly autoregressive terms for output growth and inflation combined with the following monthly predictors: industrial production, consumer confidence, employment, unemployment rate, consumer confidence, house prices, stock prices, real exchange rate, home sales, 3-month treasury rate, term spread (10 year –3-month), retail prices and oil prices.
- (3) A medium size Bayesian Vector Auto-Regression (BVAR) of 15 variables, comprising the 13 monthly predictor variables used in the MIDAS approach (measured at a quarterly frequency), plus output growth and inflation.
- (4) A second macro variable BVAR of 7 variables, comprising 5 macroeconomic predictors: aggregate consumption, investment, hours, real wages (unit labour costs) and the Bank Rate, plus output growth and inflation.
- (5) A three variable (output growth, inflation and interest rates) time-varying parameter VAR model with stochastic volatility in the errors.

We describe each model in more detail in Appendix 1 below.

The *WBS Forecasting System* then takes a weighted combination of the density forecasts produced from each of these five models. In order both to estimate these weights and to establish the historical performance of the *System* we undertake a series of recursive out-of-sample simulations. These are designed to mimic real-time use of the *System*.

Specifically, we run each of the five models as if in real-time, using (so-called “real-time”) data which would have been the data actually available at the time the forecast was made. The parameters of the models are estimated using a recursive, expanding sample window. Forecasts, up to eight quarters ahead, are then produced once a quarter, towards the end of the second month of each calendar quarter when the ONS publish their “Second Estimate of GDP”. This means that the *System* forecasts are conditioned on these latest official GDP, consumption and investment data.

Combinations of the density forecasts from the five models are produced recursively, each quarter, from 1999Q1-2012Q3 and then evaluated against the subsequently observed outcomes. As the GDP or output growth data are revised over time, there is an issue about how we measure the “outcome”. Consistent with above, we choose to use the second monthly estimate (that is, the value published by the ONS two months after the end of the quarter of interest) on the assumption that this estimate is more reliable than the ONS’s “preliminary” estimate published a month earlier.

To produce the combined density forecasts we use the “linear opinion pool”. This is a weighted linear average of the forecast densities from each of the five models. The weight on each model is either estimated - to reflect the accuracy of that model’s density forecast (we use the logarithmic score, see below, as a measure of “accuracy”) - or is set equal to 1/5 so that each model receives the same weight. The weights calculated using the logarithmic scores have the feature that they evolve over time, allowing for different models to have time-varying effects on the combined density forecast. Equal-weighted combinations have been found, in many applications, to be tough to beat.

Our interest is then not in the density forecasts from the individual models, but, for robustness, in the density forecasts from the combined system. We use the combined density forecasts to make probabilistic statements.

Below we provide an assessment and evaluation of the historical performance of the combined density forecast (labelled WBS), based on these out-of-sample simulations. An accompanying academic paper provides full details.

### **Evaluation of WBS Density Forecasts**

In order to evaluate the density forecasts (or predictive distributions) we compute the logarithmic score, a well-known “scoring rule” used to measure the accuracy of a density forecast. A scoring rule measures the quality of a probability forecast by assigning a numerical score based on the forecast and the subsequently observed “outcome”; where the emphasis is on relative performance, such that competing forecasts can be ranked. Specifically, the logarithmic score is defined as the logarithm of the density forecast of a random variable (in our case output growth or inflation) evaluated at the observed outcome. Intuitively, if the probability of observing the observed outcome is high, then the logarithmic score will also be high. As such the statistic has a positive orientation - higher scores are better, and since (as is the case here) the logarithm of the forecast density is typically negative, scores with smaller absolute values are better.

**Table 1: Logarithmic Scores of the Forecast Densities**

Table 1a: Output Growth

(Observed outcome data: second monthly release of real GDP growth)

	1999Q1-2012Q3			1999-2013 (calendar year forecasts)	
	No Change	BoE	WBS	No Change	WBS
h=1	-1.31	-0.99**	-0.83***	-1.25	-0.42***
h=4	-2.71	-2.32	-2.06**	-2.51	-1.48
h=8	-2.69	-3.07	-2.37	-2.34	-2.24

Table 1b: CPI Inflation

	2004Q1-2012Q3			1999-2013 (calendar year forecasts)	
	No Change	BoE	WBS	No Change	WBS
h=1	-0.95	-0.06***	-0.47**	-0.17	0.51*
h=4	-1.83	-1.62	-1.71*	-1.51	-0.82
h=8	-1.89	-1.95	-1.65	-1.70	-1.43

*Notes:* No change forecasts are computed assuming that expected future values are as current values, and the density is computed by bootstrapping using past forecast errors, varying with the forecast horizon. The Bank of England's density forecasts are two-piece normal densities, assuming market interest rates, as published on the Bank's website. h=1 is the current quarter nowcast/forecast. More accurate forecasters have higher scores. \*, \*\*, \*\*\* means that the forecaster is statistically more accurate than the no change (random walk) forecast at the 10%, 5% and 1% significance levels.

Table 1 reports the logarithmic scores for the no change (so-called random walk), Bank of England and WBS forecast densities. For both output growth and inflation we report two sets of numbers for the logarithmic score. In the first three columns of each table we evaluate the quarterly density forecasts for output growth and inflation (against the quarterly outcomes) and, accordingly, we can compare our forecasts directly with those from the MPC at the Bank of England. In the last two columns we evaluate the calendar year density forecasts of annual output growth or inflation. The accompanying research paper provides a more detailed set of evaluation results using the quarterly forecasts, where we have more data observations.

The Table shows that the WBS forecasts outperform (have a higher score than) the no change forecast (a useful benchmark) for all horizons and both variables, but statistically significant gains are only found for the one quarter (h=1) and one-year-ahead (h=4) forecasts. WBS predictive densities of output growth and inflation score better than the Bank of England's, except when predicting current quarter inflation (h=1).

A second means of evaluating the quality of these forecast densities is via visual inspection of the histogram of the Probability Integral Transforms (PITs). The PIT is the cumulative density function corresponding to the forecast densities evaluated at the observed outcome. As such the sequence of PITs is defined to take values between zero and one, where the histogram is uniform when the forecast density is equal to the true (but unknown) density. The PIT histogram of a “well-calibrated” density forecast should look uniform over time.

For the WBS combined density forecast and the Bank of England, Figures 1 and 2 plot the PIT values for quarterly output growth and inflation - for forecast horizons  $h=1, 4$  and  $8$  ( $h$ =number of quarters ahead).

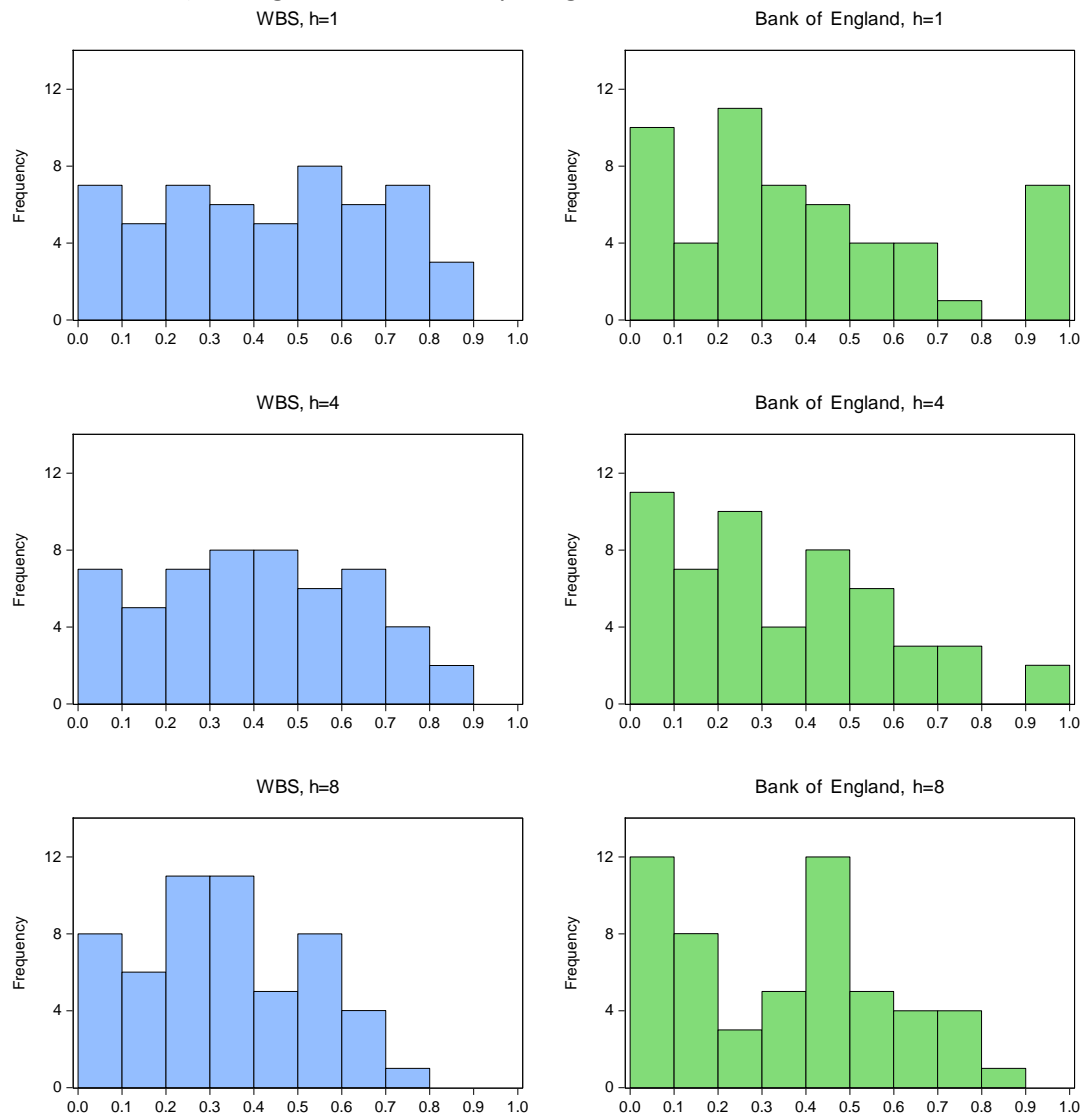
For output growth a well-calibrated density should have uniform PITs, implying that each of the 10 bins or intervals (0 through to 1 in size jumps of 0.1) should contain approximately 5 observations (we have 54 observations in our evaluation period). Visual inspection of Figure 1 suggests that the PITs appear more uniform, for all horizons, for the WBS density than the Bank of England. However, the figures suggest that the WBS densities tend to over-estimate the likelihood of high levels of output growth. The Bank of England was even more optimistic, as evidenced by their PIT plots having even more mass towards the left of the distribution.

For inflation we should expect a well-calibrated density forecast to have approximately 3 observations per bin/interval (as we have only 34 observations in the evaluation period, given the shorter period over which the Bank of England forecasts are available). Looking at Figure 2, the PITs for the *WBS system* appear to be more uniform, for all forecast horizons considered, than the Bank of England’s forecast densities. In particular, the WBS forecasts were less likely to under-predict inflation than the Bank of England at longer horizons.

Formal statistical tests of the properties of the PIT values in Figures 1 and 2 (available in the accompanying academic paper) confirm this visual assessment.

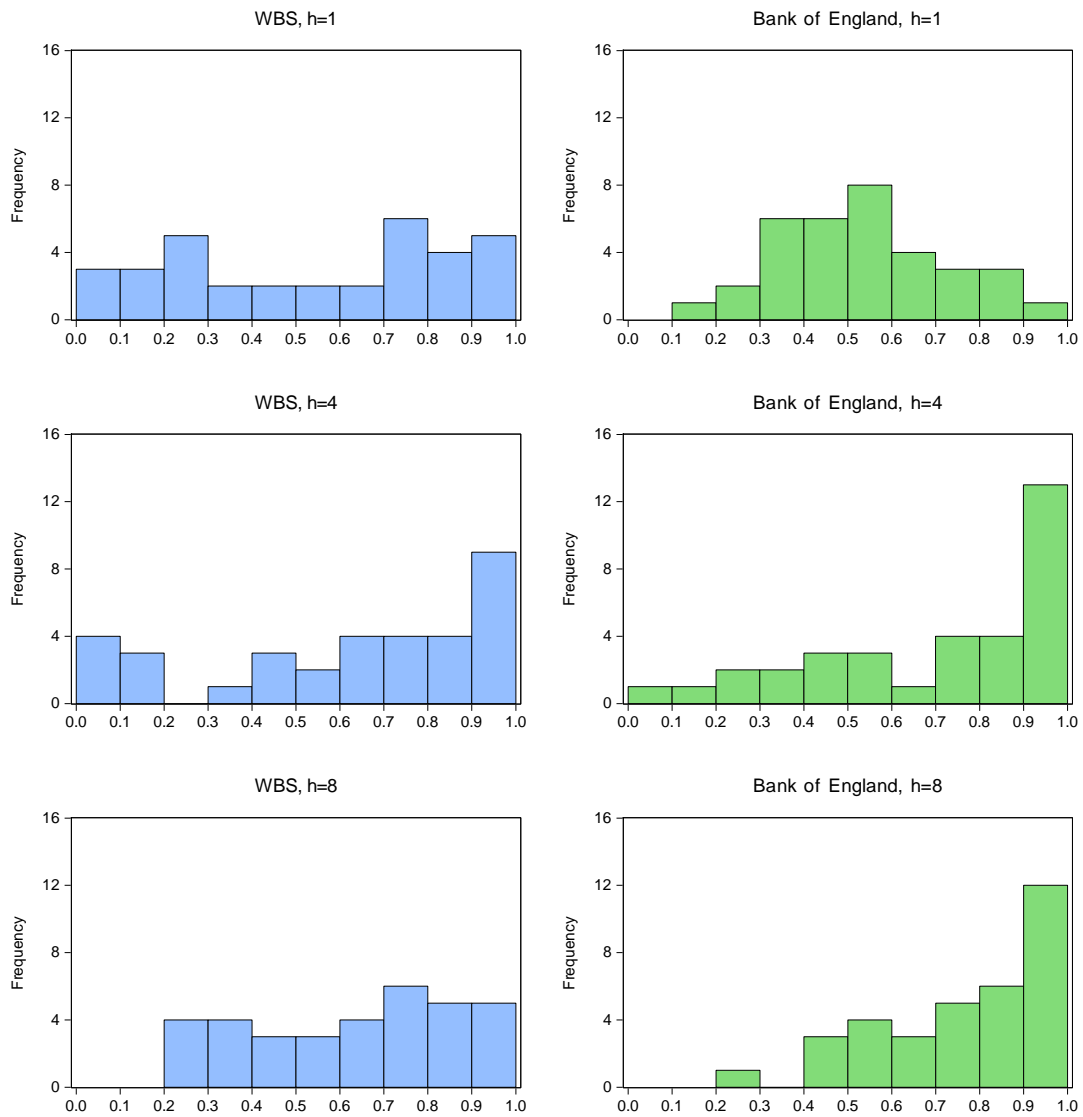
In in this note we summarise the properties and performance of the *Warwick Business School Forecasting System*. The evaluation results support the use of the *System* as a means of producing well-calibrated, but judgement-free, predictive densities. As such we recommend that the *System’s* forecasts for current and next year GDP growth and CPI inflation are used as benchmarks both to draw out the role of judgement in the forecasts of others, whether this judgement is well-founded or otherwise, and assess the risks associated with other forecasts when these do not involve a direct communication of forecast uncertainty.

Figure 1: Evaluating Density Forecast Calibration for Output Growth (against the second monthly release of GDP): Histograms of Probability Integral Transforms



Note: Well calibrated density forecasts have the property that the frequency in each bin should be approximately equal to five. Forecasts computed from 1999Q1 up to 2012Q2.

Figure 2: Evaluating Density Forecast Calibration for Inflation: Histograms of Probability Integral Transforms



Note: Well calibrated density forecasts have the property that the frequency in each bin should be approximately equal to three. Forecasts computed from 2004Q1 up to 2012Q2.

Appendix 1: Forecasting Models included in the *WBS Forecasting System*

	Model	Description	UK Data
1	Autoregressive Model	Lag order fixed at 2; Density forecasts obtained by bootstrap. Estimate separately for quarterly output growth and CPI inflation	Real GDP; CPI seasonally-adjusted with X11 filter.
2	Mixed Data Sampling Model	Single predictor models with autoregressive terms. Estimate separately for output growth and inflation. Predictors are sampled monthly, autoregressive terms are quarterly for output growth and monthly for inflation. Number of monthly lags: 24. Weighting function: beta. We use the values of the first month of the current quarter for a subset of predictors (financial variables) and autoregressive values of inflation. Density forecasts obtained by the bootstrap, assuming the parameters of weighting function are known. Data transformation (log, log-differences) differs across variables.	Uses the same data as model (1) plus the following monthly predictors: industrial production consumer confidence, employment, unemployment rate, consumer confidence, house prices, stock prices, real exchange rate, home sales, 3-month treasury rate, term spread (10 year –3-month), retail prices, oil prices.
3	Bayesian Vector Autoregressive Model: Medium	“Minnesota-type” prior on autoregressive coefficients with overall tightness prior chosen using marginal likelihood. Includes also sum of coefficients prior and “dummy initial observation” prior. Density forecasts obtained by drawing autoregressive parameters and the variance-covariance matrix from their posterior distributions, and by drawing sequentially from disturbances’ posterior distribution. Autoregressive order: 4. Estimated with variables in log-levels.	15 quarterly variables: the 13 predictors of Model (2) plus real GDP and seasonally-adjusted CPI.
4	Bayesian Vector Autoregressive Model: Macro	As model (3). Because of data availability in real time, current quarter forecasts are obtained using a two-quarter-ahead instead of the one-quarter-ahead horizon in model (3).	7 quarterly variables: 5 macroeconomic predictors: aggregate consumption, investment and hours, real wages (unit labour costs), and the bank rate.
5	Time-varying vector autoregressive Model with Stochastic Volatility	Both autoregressive parameters and conditional variances follow a random walk process. Estimated via Gibbs Sampling with a Metropolis step to obtain volatility draws. Autoregressive order=1. Estimated using annual growth rates. Density forecasts obtained using Gibbs draws after excluding burn-in draws, and by drawing sequentially from disturbances’ distribution.	3 quarterly variables: annual growth rate of real GDP and CPI, and the Bank Rate.