Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
International Journal of Forecasting 26 (2010) 211–215
international journal of forecasting
www.elsevier.com/locate/ijforecast

Editorial

Bayesian forecasting in economics

The Bayesian paradigm uses probabilities to express uncertainty about all unknowns. This defining characteristic of the paradigm renders it the natural choice for use in forecasting, with uncertainty about the unknown future being automatically quantified in probabilistic terms. Moreover, the simple rules of probability that underlie the Bayesian approach yield probabilistic forecasts that reflect the uncertainty about *all* of the unknowns that characterize an empirical problem, and condition *only* on what is known. As such, parameter (and, if desired, model) uncertainty is automatically factored into Bayesian forecasts, rather than forecasts being conditioned on what is inherently unknowable.

Bayesian forecasts may incorporate informative prior beliefs about the parameters of an assumed model, if this is deemed appropriate, or may be primarily influenced (via the likelihood function) by the observed historical data. A forecast model may be characterized only in terms of intrinsically fixed parameters, or may also be dependent on (a hierarchy of) stochastic parameters (or latent variables). A model may be taken as given, or treated as only one of many alternatives from which forecasts can be produced. Whatever the specific approach adopted, all components of the statistical problem – future values, historical values, parameters, latent variables, and models – are viewed as arguments of a joint probability distribution to which standard probability manipulations are applied to produce the distribution of future values conditional on past data. Hence, up to a point, all Bayesian forecast problems are tackled in a manner that is broadly common, and consistent with the well-known calculus of probability distributions.

The eleven papers in this special issue demonstrate the application of the Bayesian paradigm to forecasting problems that arise in economics and its allied spheres. As such, the papers illustrate the way in which the Bayesian approach deals with the particular empirical challenges that arise therein. These challenges include modelling unobservable dynamic random variables (such as financial volatility), dealing with multiple sources of non-linearity (regime shifts, random coefficients, limited dependent variables), and producing accurate forecasts from large, highly parameterized models.

The articles range across all fields of economic activity, including macroeconomics, finance and firm-level economic decision making. As is consistent with the Bayesian empirical literature in general, Gibbs-based Markov chain Monte Carlo (MCMC) schemes predominate, although some variation in simulation schemes is still visible—evidence of the healthy state of Bayesian computational developments. Several of the papers highlight the importance of prior specification to the production of accurate forecasts from complex, highly parameterized models. Due attention is also given in the issue to forecast evaluation, with the usual type of mean squared error criteria (for point forecasts) supplemented, in many cases, with assessments that are relevant to the distributional forecasts produced via the Bayesian method.

The issue begins with two papers that are concerned with an issue which is of paramount importance in financial economics: quantifying the distribution of future stock market returns, including the probability of the extreme returns that would lead to a large capital loss. Geweke and Amisano ('*Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns*') produce Bayesian predictive distributions for

doi:10.1016/j.ijforecast.2010.02.003

returns on the S&P500 market index, over multiple future horizons, based on five alternative models. The authors highlight the distinction (often blurred in the literature) between *comparing* the predictive accuracy of alternative models and *evaluating* their performance against an absolute standard. The comparative technique is, in itself, inherently Bayesian, based as it is on the decomposition of the Bayes factor (or the ratio of marginal likelihoods) into a sequence of one-step-ahead predictive likelihoods, evaluated at returns observed *ex-post*. Importantly, the technique is used to show how individual observations contribute to one model being favoured over another, and how the dominance of a model (within a set) can alter over the sample period. The evaluation technique – based on the probability integral transform (PIT) – is inherently frequentist in style. That is, the PIT method benchmarks the sequence of cumulative predictive distributions (evaluated at *ex-post* returns) against the i.i.d. uniform distribution which *would* result *if* the data *were* generated by the assumed model. The evaluation results reported by the authors complement those of the comparative method. Specifically, the PIT goodness-of-fit test results are consistent with the predictive likelihood results, in highlighting the overall inferiority of the generalized autoregressive conditional heteroscedastic (GARCH) model with Gaussian innovations. However, assessments based on various transformations of the PIT values indicate deficiencies in all five models in capturing the behaviour of observed returns—something that the comparative analysis fails to pick up.

Hoogerheide and van Dijk ('*Bayesian Forecasting of Value at Risk and Expected Shortfall Using Adaptive Importance Sampling*') also focus on the predictive distribution of asset returns. Their particular focus, however, is on the lower tail of the (associated) loss distribution, and on producing accurate Bayesian forecasts of the value at risk (VaR) quantile and expected shortfall (ES). The authors extend the concept of the 'optimal importance density' (Geweke, 1989) to the VaR/ES context, in which the joint distribution of both parameters and future returns is relevant, and where accurate estimation in one (small) part of the support is critical. In brief, the optimal importance distribution is specified as a bivariate mixture that assigns disproportionately high probabilities to draws from the tail of the loss distribution. The components of this mixture are, in turn, obtained as mixtures of Student-t distributions, via the iterative procedure of Hoogerheide, Kaashoek, and van Dijk (2007). Using two standard returns models as the basis for illustration, the new method is shown to yield a substantial increase in estimation accuracy (as measured by the numerical standard error) relative to a standard MCMC sampling algorithm. The method thus shows great promise in reducing the computation time required to produce accurate estimates of these important risk measures. Further exploration of the (relative) computational performance of the IS-based method, both for multipstep horizons and in multivariate settings, will be an important area of future research.

The next two papers in the issue shift the focus from financial to macroeconomic forecasting, and examine the role that prior beliefs play in determining predictive accuracy. Beechey and Osterholm ('*Forecasting Inflation in an Inflation Targeting Regime: A Role for Informative Steady-State Priors*') investigate whether the forecasting performance of autogressive models can be improved by incorporating information about the numerical inflation targets set by many central banks under inflation targeting regimes. The authors hypothesize that when a central bank's commitment to such a monetary policy is strong and credible, the inflation target is likely to serve as a valuable informative prior for the steady state of the inflation process. This is because a credible target acts as an anchor for inflationary expectations and, hence, as a stabilizing influence on inflationary outcomes. Using data from various countries over the last two decades, the authors do indeed find that informative priors typically generate a substantial improvement in inflation forecasting performance relative to alternative specifications.

Lahiri and Sheng ('*Learning and Heterogeneity in GDP and Inflation Forecasts*') use a Bayesian information processing framework to identify the relative importance of alternative pathways through which forecasters of GDP growth and inflation adapt to new information and determine the term structure of forecasts and the associated patterns of forecast accuracy. In their model, forecast disagreement assumes a special role, and arises from two sources: differences in the initial prior beliefs of forecasters, and differences in their subsequent interpretations of public information. The authors use a panel of monthly forecasts made by a number of professional forecasters for seven industrial countries over the period 1991–2007 to illustrate these points. In addition, case studies on (i) forecast disagreement around the 9/11 terrorist attack in the US, and (ii) the inflation targeting experience of Italy after 1997, provide additional support for the analysis.

Maintaining the focus on the macroeconomy, the next five papers all adopt models in which latent variables play some role, thus increasing the complexity of the MCMC algorithms used to estimate the relevant predictive distributions. In the first two of these papers (Rodriguez and Puggioni, and Giordani and Villani), the Kalman-filter based forward-filtering-backward-sampling (FFBS) algorithm (Carter & Kohn, 1994; Frühwirth-Schnatter, 1994) is used to produce multi-move draws of the latent states, as is typical in the treatment of such models. Rodriguez and Puggioni ('*Mixed Frequency Models: Bayesian Approaches to Estimation and Prediction*') present a Bayesian approach to mixed frequency models, whereby a dependent variable is regressed on (several lags of) explanatory variables observed at a higher frequency. In essence, the paper provides a Bayesian alternative to the frequentist MIDAS (mixed data sampling) regression of Ghysels, Santa-Clara, and Valkanov (2002) and Ghysels, Sinko, and Valkanov (2007), with the deterministic weighting structure which is imposed on the regression coefficients in the MIDAS approach being replaced by more flexible stochastic priors. The authors develop the approach within a dynamic linear model (DLM) framework, thus enabling the regression structure to be supplemented by random dynamic coefficients. The issue of variable (or model) selection within this context is explored in some detail, with attention given both to sensible prior specification on the model space, and to the application of an MCMC scheme that enables the closed-form computation of the marginal likelihood for each model. The approach is used to produce a range of alternative predictions of US quarterly GDP, based on a term structure of monthly interest rates, with the model-averaged (point) predictions being the most accurate of all of the alternatives considered.

Giordani and Villani ('*Forecasting Macroeconomic Time Series with Locally Adaptive Signal Extraction*') also adopt a state space representation for forecasting key macroeconomic variables. Unlike in the work of Rodriguez and Puggioni, however, where standard Gaussian autogressive processes are adopted for the mean parameters in the model, these authors also allow for random variation in both the mean and conditional variance, including large infrequent shifts therein, in addition to mixed normal innovations. The primary aim of adopting such a specification is to combine flexibility with robustness. That is, the mean parameters are allowed to shift, potentially by a large amount on occasion. However, the flexible innovation (and conditional variance) specifications serve to absorb the impact of some extreme observations, such that overly volatile predictions are not produced. The authors document the performance of point, interval and density forecasts of various macroeconomic time series for several variants of their model, and draw the general conclusion that a good forecast performance is indeed aided by the particular combination of regime shifts and non-Gaussian distributional assumptions that they propose.

Structural breaks also feature in the paper by Jochmann, Koop and Strachan ('*Bayesian Forecasting using Stochastic Search Variable Selection in a VAR subject to Breaks*'). A standard VAR model is augmented to cater for an unknown number of breaks, with a break being allowed to occur with constant probability in each time period (including over the forecast horizon). The forecasts are produced as a mixture distribution, with the weights of the mixture being the probabilities assigned to the events 'break' and 'no break' occurring during the forecast horizon. The issue of overparameterization, which is typically managed by 'shrinkage priors' in VAR models, is handled by applying the stochastic search variable selection (SSVS) method of George, Sun, and Ni (2008). SSVS is a hierarchical 'spike and slab' prior, whereby each VAR coefficient is randomly assigned a zero mean prior distribution with either a small or a large prior variance. Importantly, in the recursive forecasting exercise conducted by the authors, the weights assigned to different sub-models (as defined by the SSVS method) vary over time, so that the model-averaged predictive reflects the changing importance of different variables in the full model. Whilst the SSVS prior serves to produce more accurate forecasts (overall) – for a selection of US macroeconomic data – the parsimony that it enforces is not sufficient to counteract the large forecast inaccuracy associated with the most general break model. In results that are somewhat consistent with those of Giordani and Villani, the authors find that a model that restrict breaks to occur *only* in the error covariance matrix, when combined with the SSVS prior, performs best according to a range of forecast accuracy measures.

The article by Shorfheide, Sill and Kryshko (‘*DSGE Model-based Forecasting of Non-modelled Variables*’) develops a framework for generating forecasts for certain macroeconomic variables that are not explicitly modelled in a dynamic stochastic general equilibrium (DSGE) model, but are of interest to a policy maker. They build a hybrid empirical model that augments the core DSGE model with auxiliary equations that link the non-modelled (or ‘non-core’) variables with the state variables of the DSGE model. The advantage of this approach is that it de-couples the estimation of the DSGE model and the analysis of the auxiliary regressions, making it more suitable for real time forecasting of the non-core variables of interest. In the first step, the authors use Bayesian methods to estimate the DSGE model comprising seven core variables. Conditional on the DSGE model parameter estimates, they apply the Kalman filter to obtain estimates of the latent state variables given the most recent information. The filtered state variables are then inserted as regressors in linear measurement equations for the non-core variables. These regressions are, in turn, estimated by Bayesian methods and used to produce forecasts of the non-core variables—supplementing the forecasts of the core variables produced directly from the DSGE model. In terms of pseudo out-of-sample predictions, the approach compares somewhat favourably to a collection of first-order autoregressive models. The advantage of this approach is that, in addition to delivering an interpretation of the predicted trajectories that is consistent with economic theory, forecasts can be produced in real time.

Still maintaining a state-space structure, but now with a shift of focus to firm-level demand forecasting, Yelland (‘*Bayesian Forecasting of Parts Demand*’) provides an inventive Bayesian solution to the problem of forecasting the demand for computer parts with little or no demand history. A hierachical prior structure is used to assign *common* prior information to parts that are used in the same final product, in addition to part-specific prior information. It is also used to impose the type of ‘life cycle’ curve that characterizes the short lives of parts in the rapidly changing technology industry. For parts whose demand history is simply *too* short to yield accurate inferences (and forecasts), informative priors based on the history of similar but ‘superannuated’ parts are specified. With the model supplemented with a standard autoregressive latent state, an MCMC algorithm (with an FFBS step) is used to estimate the model and produce (point) forecasts. The accuracy of the latter, compared to various benchmark models, is documented by the author via various metrics. The results auger well for the application of this type of approach to the challenging problem of forecasting the demand for manufacturing parts.

The issue is rounded off by two multivariate applications in which the more traditional regression structure is adopted, obviating the need for any filtering-based component in the simulation algorithm. Griffiths, Newton and O’Donnell (‘*Predictive Densities for Models with Stochastic Regressor and Inequality Constraints: Forecasting Local-Area Wheat Yield*’) demonstrate a technique for estimating the predictive probability density function for both a linear and a non-linear regression model with stochastic regressors. The approach is novel in that the predictive density incorporates both inequality information about the coefficients and uncertainty from the stochastic regressors. The authors illustrate the approach by forecasting the wheat yield in local areas in the state of Western Australia. The wheat yield is heavily dependent on rainfall during three periods: germination, growing and flowering, with the rainfalls associated with each period being included as regressors in the model. The timing of the forecast thus determines which of these regressors are treated as stochastic and which are taken as given, with the shape of the predictive density changing accordingly. The Bayesian approach provides a natural and unified approach for accommodating the many sources of uncertainty that are present in this forecasting example.

Finally, Zellner and Ando (‘*Bayesian and Non-Bayesian Analysis of the Seemingly Unrelated Regression Model with Student-t Errors and its Application for Forecasting*’) develop a computationally efficient method for applying Bayesian inference in seemingly unrelated regression (SUR) models with heavy-tailed error distributions. They suggest a new algorithm that combines direct Monte Carlo (DMC) with importance sampling in order to compute various quantities of interest to a forecaster. A DMC algorithm for the Bayesian Method of Moments (BMOM) approach (in which maximum entropy methods are used to construct 'postdata' densities from moment conditions) is also presented. The authors provide numerical illustrations, including the production of sales forecasts for ten industrial sectors in Japan. In his *Comment* on the Zellner-Ando paper, Geweke agrees that allowing for excess kurtosis in the disturbances of the SUR model is a worthwhile endeavour, but contends that existing MCMC algorithms are more appropriate for the purpose of Bayesian inference than the proposed DMC method. Geweke also raises important issues regarding parameterization and identification, as pertaining to the Zellner-Ando analysis. In their *Rejoinder*, Zellner and Ando continue to emphasize the attractiveness of the DMC method in many applications, most notably in on-line forecasting situations, and provide further elaboration of their method.

To conclude, we quote from the excellent summary of Bayesian forecasting by Geweke and Whiteman (2006, chap. 1):

'Twas not always so easy..... In the beginning there was diffuseness, conjugacy and analytical work.'

Before the advent of fast, accessible computers and finely honed numerical techniques, this was indeed the state of affairs in Bayesian statistics! However, with the aid of these recent developments, Bayesian analysis, including forecasting, now knows no limits, and the scope of the papers included in this issue is clear evidence of that. We thank the authors for their fine contributions to the issue and for the committed and professional manner in which they have responded to comments from referees and editors, within quite tight dead lines. We also extend our thanks to the large number of referees who have contributed to this exercise, providing constructive, detailed and timely reports at all stages in the process. Their efforts are warmly acknowledged.

References

Carter, C., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, *81*, 541–553.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, *15*, 183–202.

George, E., Sun, D., & Ni, S. (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, *142*, 553–580.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, *57*, 1317–1339.

Geweke, J., & Whiteman, C. (2006). Bayesian forecasting. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 4–80). Amsterdam: Elsevier.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2002). *The MIDAS touch: Mixed data sampling regression models*. Technical report, University of North Carolina, Chapel Hill and University of California, Los Angeles.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, *26*, 53–90.

Hoogerheide, L. F., Kaashoek, J. F., & van Dijk, H. K. (2007). On the shape of posterior densities and credible sets in instrumental variable regression models with reduced rank: An application of flexible sampling methods using neural networks. *Journal of Econometrics*, *139*, 154–180.

Kajal Lahiri
Department of Economics,
University at Albany: SUNY, United States
E-mail address: KLahiri@albany.edu.

Gael Martin*
Department of Econometrics and Business Statistics,
Monash University, Australia
E-mail address: Gael.Martin@buseco.monash.edu.

* Corresponding author.