

A PREDICTIVE LIKELIHOOD APPROACH TO BAYESIAN AVERAGING

Tomáš Jeřábek¹, Radka Šperková²

¹ Faculty of Economic Studies, University of Finance and Administration, Estonská 500, 101 00 Praha 10, Czech Republic

² Department of Economy and Management, College of Business and Hotel Management, Bosonožská 9, 625 00 Brno, Czech Republic

Abstract

JEŘÁBEK TOMÁŠ, ŠPERKOVÁ RADKA. 2015. A Predictive Likelihood Approach to Bayesian Averaging. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63(4): 1269–1276.

Multivariate time series forecasting is applied in a wide range of economic activities related to regional competitiveness and is the basis of almost all macroeconomic analysis. In this paper we combine multivariate density forecasts of GDP growth, inflation and real interest rates from four various models, two type of Bayesian vector autoregression (BVAR) models, a New Keynesian dynamic stochastic general equilibrium (DSGE) model of small open economy and DSGE-VAR model. The performance of models is identified using historical dates including domestic economy and foreign economy, which is represented by countries of the Eurozone. Because forecast accuracy of observed models are different, the weighting scheme based on the predictive likelihood, the trace of past MSE matrix, model ranks are used to combine the models. The equal-weight scheme is used as a simple combination scheme. The results show that optimally combined densities are comparable to the best individual models.

Keywords: predictive likelihood, density forecasts, Bayesian averaging, Bayesian VAR model, GDP growth, inflation, real interest rates

INTRODUCTION

A combining forecasts from different models to increase their forecast accuracy belongs to the popular macroeconomic research techniques and the first studies in this area have been focused on combining point forecasts; see Timmermann (2006). However, the point forecast evaluation does not take into account a model uncertainty. The solution is the knowledge of the probability distribution of the individual forecasts and therefore density forecasts can be used instead of the point forecasts. As reported in Hall and Mitchell (2007) combining density forecasts provide more accurate description of the true degree of the uncertainty. An empirical question depends on the method used to choose weights for individual models when the combined densities are constructed. A natural way to choose weights in a Bayesian context is to use posterior probability weights based on the marginal likelihood for each model. Unfortunately, there are two problems for the posterior probability weights. The first, the weights are influenced by the number

of estimated time series for each model, the second problem is that over-parameterized models that have a high good in-sample fit, and a low out-of sample forecasting accuracy, have assigned high weights. Andersson and Karlsson (2007) show that an out-of-sample weights, based on the predictive likelihoods, overcome these problems. They illustrated of the predictive likelihood forecast combination procedure at U.S. gross domestic product forecast by using various vector autoregressive (VAR) models. Gerard and Nimark (2008) combine trivariate density forecasts form a dynamic stochastic general equilibrium (DSGE), a Bayesian VAR and a factor-augmented VAR model. They use the predictive likelihood weighting scheme as well as equal weights to three variable forecasts (gross domestic product, inflation and interest rate). Wolters (2012) extends Gerard and Nimark work by combining forecasts form four various DSGE models. The use of the predictive likelihood from the above mentioned models is complicated by the fact that the predictive likelihood does not have a closed form

solution; see Eklund and Karlsson (2005). Therefore, the above authors estimate of the predictive likelihood by using the normal kernel density estimation from the predictive draws. However, Adolfson *et al.* (2007) show that this approach is not practical unless the dimension of density forecast is small and they suggest using a normal approximation of the predictive likelihood based on the mean and the covariance of the marginal predictive distribution. Warne *et al.* (2013) show that predictive likelihood can be computed via missing observations techniques.

This paper proposes to combine trivariate density forecasts from two BVAR, models, a DSGE and a DSGE-VAR model. Several methods to combine forecasts from the set model are considered – predictive likelihood based weights following Adolfson *et al.* (2007) approach as well as Warne *et al.* (2013) approach, the relative performance weights based on the trace of past MSE matrix, and the weights based on model ranks. The equal-weight scheme is used as a simple combination scheme.

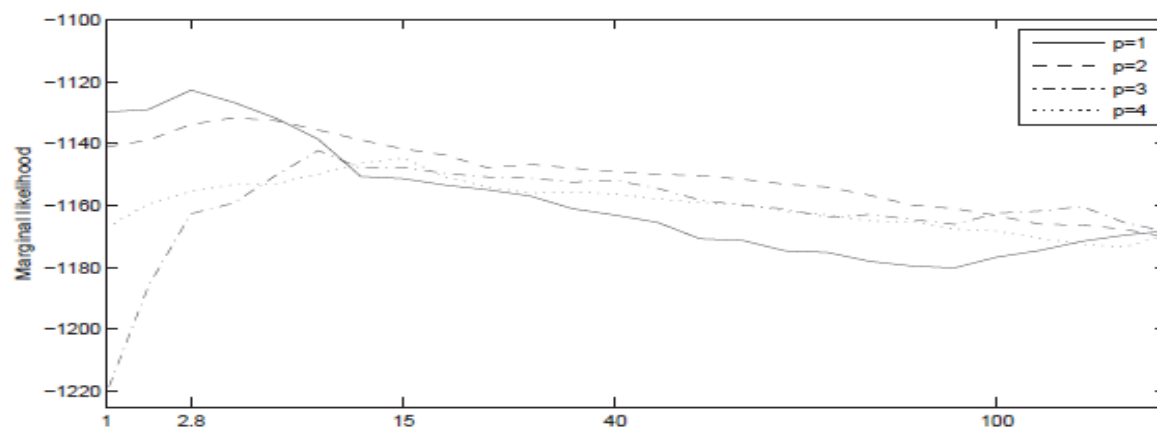
MODELS AND DATA

I consider four models of the Czech Republic economy. Two variants of Bayesian VAR model with 4 lags for endogenous and 4 lags for exogenous variables. The Minnesota prior is used. This prior is based on an approximation which leads to

shrinking the parameters towards zero by assuming that the observed variables of model follow independent random walks; see Doan *et al.* (1984) and Koop and Korobilis (2010). The individual variants of BVAR model differ from one another by setting of the overall tightness parameter which controls the prior standard deviations of all the lag coefficients of the endogenous variables. I set value for BVAR1 and value 0.2 for BVAR2. The values for another hyperparameters were chosen according to Jeřábek *et al.* (2013). I generate 10 000 draws from posterior distributions using Gibbs Sampler with the first 500 draws used as burn-in sample which is removed.

The DSGE model is constructed as a small open economy model and it is borrowed from Seneca (2010). The model parameters are chosen through a combination of calibration and formal estimation, according to Seneca (2010) and Musil (2009). The posteriors of the estimated parameters are simulated using 500 000 draws from the Random-Walk Metropolis-Hastings algorithm where the first 100 000 are removed.

Furthermore, the DSGE model is used to generate a prior distribution for the coefficients of the VAR model, i.e. DSGE-VAR model. As referred by Del Negro and Schorfheide (2004), this prior concentrates most of its probability mass near the restrictions that the DSGE model imposes on the VAR representation and pulls



1: Selection of the optimal λ .
Source: authors

I: Model overview¹

Type	Eq.	Par.	Est. Par.	Observed variables
BVAR with 4 lags	11	59	59	Endogenous: GDP, consumption, investment, government spending, import, export, hours worked, real exchange rate, CPI inflation, wage inflation, interest rate. Exogenous: CPI inflation, GDP, interest rate.
DSGE	62	68	42	
DSGE-VAR with 4 lags	11	59	59	

Source: authors

¹ Eq.: number of equations including shock processes, measurement equations and identities. Par.: total number of parameters. Est. Par.: exact number of estimated parameters including shock variances.

the likelihood estimate of the VAR parameters toward the DSGE model restrictions. The DSGEVAR model is indexed by λ -value which measures the deviation of the DSGE-VAR from the VAR approximation of the DSGE model; see Del Negro and Schorfheide (2012). For the optimal λ -value choice, I use of the maximize marginal likelihood for the estimations of the DSGE-VAR with lags 1 up to 4 over $\lambda \in [0, \infty]$. As seen from Fig. 1, the optimal model choice results in 1 lags and $\lambda = 2.8$.

METHODS

For the formulation of conclusions and recommendations from the methodological point of view use the following methods.

Model Averaging

A combination density forecast p^c is constructed as a weighted linear combination of the model forecasts, i.e.

$$p^c(z_{t+h} | \Omega_t^1, \dots, \Omega_t^M, \omega_{1,h,t}, \dots, \omega_{M,h,t}) = \sum_{m=1}^M \omega_{m,h,t} p(z_{t+h} | \Omega_t^m), \quad (1)$$

where the vector consists three variables GDP growth, CPI inflation and interest rate, and $p(z_{t+h} | \Omega_t^m)$ is trivariate density forecast for horizon $h = 1, \dots, H$ of model $m = 1, \dots, M$ at time t , Ω_t^m is the information set of model including model parameters, variables and equations. Four different weighting schemes are used. The weights $\omega_{m,h,t}$ are based on the predictive likelihood, the trace of mean square errors (MSE) matrix of past forecasts, the relative past forecasting accuracy by assigned ranks from 1 to M according to accuracy measured by the trace of the MSE and the mean forecasts – equal weights.

Predictive Likelihood Weights

The approach based on the predictive likelihood requires to split data into a training sample for estimate the models and hold-out sample used to evaluate forecasting performance of models. As in Gerard and Nimark (2008), the predictive likelihoods are calculated over a recursive scheme where the training sample $Y_t = (y_1, \dots, y_t)'$ relevant to period t is extended by one observe at each recursive step, where vector y consists all observed variables. Let $Z_{t,h} = (z_{t+h}, \dots, z_T)'$ be hold out sample at horizon h . For the predictive likelihood $p(Z_{t,h} | Y_t, m)$ for model m at horizon h holds

$$p(Z_{t,h} | Y_t, m) = \prod_{i=1}^{T-h} p(z_{t+i} | Y_t, m). \quad (2)$$

The predictive likelihood weights are given by

$$\omega_{m,h,t}^{PL} = \frac{p(Z_{t,h} | Y_t, m)}{\sum_{i=1}^M p(Z_{t,h} | Y_t, i)}. \quad (3)$$

I use two approaches to evaluate of the marginal predictive likelihood (MPL) $p(z_{t+h} | Y_t, m)$. The first method consists in MPL approximation by normal MPL which is based on the mean and the covariance of the marginal predictive distribution, see Adolfson *et al.* (2007). Hence, to evaluation equation (2), for each model m the normal distribution is used and I take an average across multiple draws from model m 's marginal predictive distribution. That is

$$p(z_{t+h} | Y_t, m) = \frac{1}{S} \sum_{j=1}^S p_N^j(z_{t+h} | Y_t, m), \quad (4)$$

where $S = 500$ and $p_N(z_{t+h} | Y_t, m)$ is normal MPL.

The second approach – as in Warne *et al.* (2013), I use an importance sampling (IS) estimator for calculate the marginal predictive likelihood in (2), see, e.g., Koop (2003). With Ψ_j being draws from the importance density $q(\Psi_j | m)$, an expression of the IS estimator is

$$p(z_{t+h} | Y_t, m) = \frac{1}{n} \sum_{j=1}^n \frac{p(z_{t+h} | Y_t, \Psi_j, m) p(\Psi_j | Y_t, m)}{q(\Psi_j | m)}, \quad (5)$$

where $n = 10000$. I use the cross-entropy method for select the importance density; see Rubinstein (1997) and Chan and Eisenstat (2012).

MSE Based Weights

This approach is based on the trace of the h -step-ahead MSE matrix $\Sigma_M(h)$ of past forecasts. Weights are built on the relative past forecasting accuracy by ranking the accuracy of the models. Thus, MSE based weights are computed by taking forecasts from previous periods and evaluate of the trace of MSE matrix for each model. For the h -step-ahead MSE matrix holds

$$\Sigma_M(h) = \frac{1}{N_h} \sum_{t=T}^{T+N_h-1} \tilde{\epsilon}_{t+h|t} \tilde{\epsilon}_{t+h|t}', \quad (6)$$

where $\tilde{\epsilon}_{t+h|t} = M^{-1/2} \epsilon_{t+h|t}$, where $\epsilon_{t+h|t}$ stand for the error h – step predictions made during the t and M is a positive definite matrix – I use M as an unit matrix. The weights are then calculated through the inverse relative MSE performance

$$\omega_{m,h,t}^{MSE} = \frac{1}{\sum_{i=1}^M \frac{1}{MSE_{h,t}^i}}, \quad (7)$$

where $MSE_{h,t}^m$ is the trace of $\Sigma_M(h)$ belong to model m and for period t .

RANK Based Weights

Another approach to determinate relative performance weights is based on assigning ranks R from 1 to M pursuant to forecast accuracy obtained

by evaluate the trace of MSE matrix. These weights are computed as

$$\omega_{m,h,t}^R = \frac{\frac{1}{R_{h,t}^m}}{\sum_{i=1}^M \frac{1}{R_{h,t}^i}}. \quad (8)$$

Equal Weights

This is the simplest weighting scheme and consists in putting equal weight on all model in the suite. Thus, in this case holds $\omega_{m,h,t}^{EW} = 1/m$ for each forecast horizon.

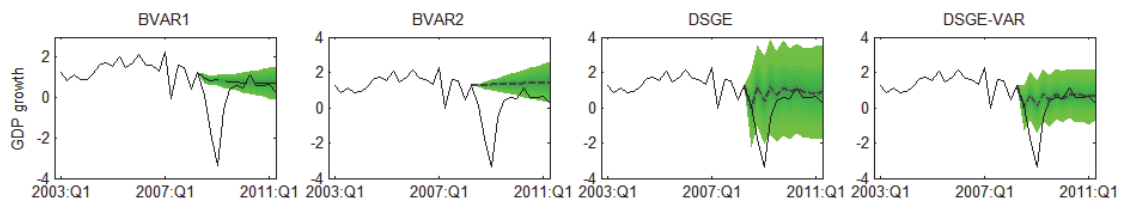
IMPLEMENTATION AND RESULTS

The data sample is given by the period 1999:Q1 until 2012:Q4 and a recursive out-of-sample forecasting scheme was used to evaluate of models and generate weights for each model. The sample from 1999:Q1 to 2006:Q4 was used as initial training sample to estimate of models and to construct density forecast up to twelve quarters ahead, from 2007:Q1 to 2009:Q4. Then, models were estimated for the extended training sample (by one observation), 1999:Q1–2007:Q1, and density forecasts were constructed over the next twelve quarters, i.e. from 2007:Q2 to 2010:Q1, and so on. The final training sample was between 1999:Q1 and 2012:Q3, which allows forecast for period 2012:Q4.

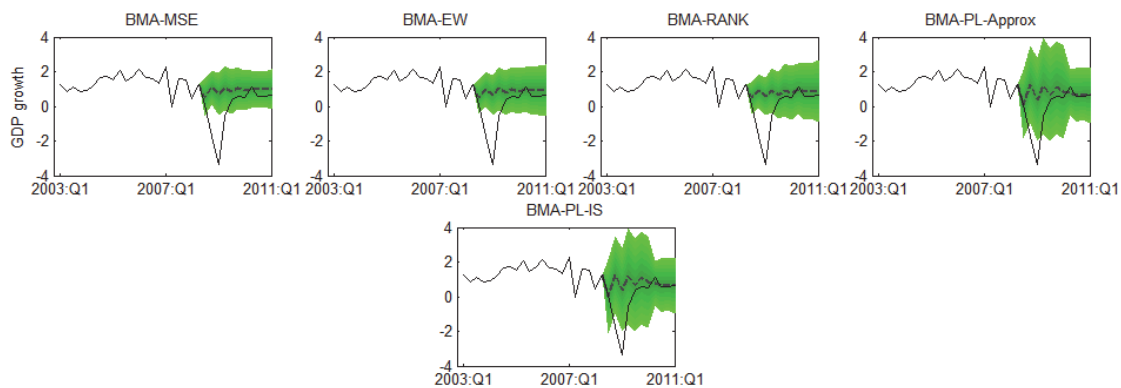
Trivariate density forecasts on the right side of the equation (1) were constructed by taken of multiple draws from posterior parameter

distribution of each model and by iterative making of trivariate density realizations forward up to horizon $h = 12$ for each draw. Complete density forecasts have been built by repeating 2500 times at each forecast horizon; see Adolfson *et al.* (2007) or Christoffel (2010) for algorithm.

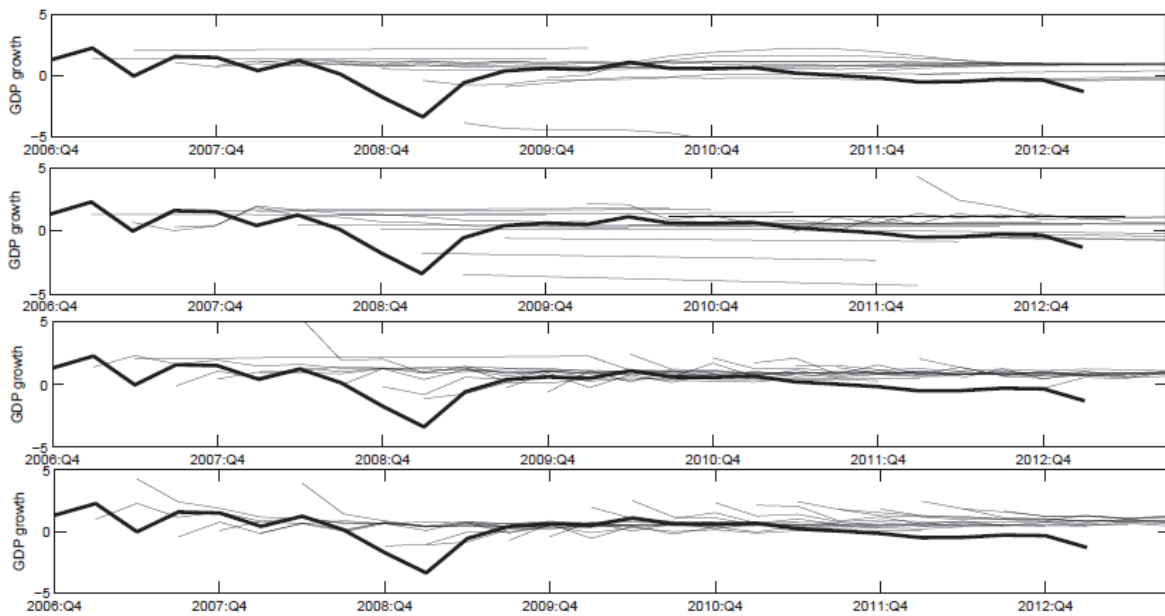
As an example, the univariate density forecasts for GDP growth that have been obtained using data up to 2008:Q2 for each model; see Fig. 1. The black curve shows observed data until forecast start. The different shades show various confidence bands, from 5% to 95%. The line in the middle of the bands shows the mean forecast for each model. I chose the period 2008:Q3, when U.S. bank Lehman Brothers failed and GDP growth turned negative, as start forecasts to compare the forecasting accuracy of models during crisis and recession. Fig. 2 shows that the BVAR models are not able to predict recession. GDP growth is predicted to stay almost constant by BVAR1 and BVAR2 as in 2008:Q2. Very similar results show DSGE and DSGE-VAR models. It may be seen that the BVAR models perform well in terms of density forecast accuracy against DSGE a DSGE-VAR giving higher degree of the forecast uncertainty. Thus, the non-structural models give less uncertainty than structural models and their forecasts do not perform well during the recession. Fig. 2 shows the combined density forecasts for GDP growth derived from the same period as the density forecasts for the individual models. It is seen that the weighting schemes prefer structural model at period of recession, which is in line with the results from the density forecasts of individual model on Fig. 3.



2: GDP growth density forecasts predicted by individual models
Source: author



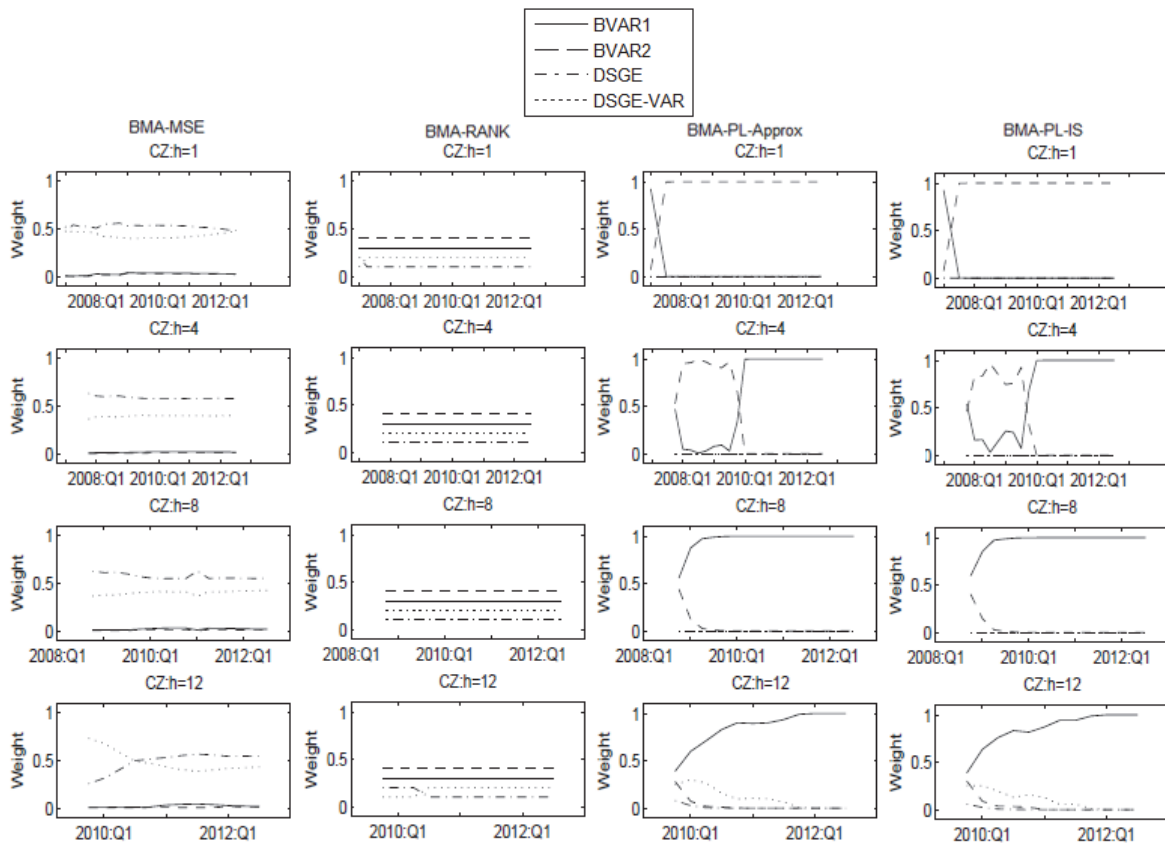
3: GDP growth combined density forecasts
Source: authors



4: GDP growth combined density forecasts
Source: authors

Fig. 4 shows the mean forecast paths from the univariate densities for individual BVAR1, BVAR2, DSGE and DSGE-VAR models together with the data for GDP growth. It is obvious that

the forecast paths of the BVAR models with Minnesota type prior are more or less constant for all observed periods. More accurate mean forecasts provide another two models. The forecast paths



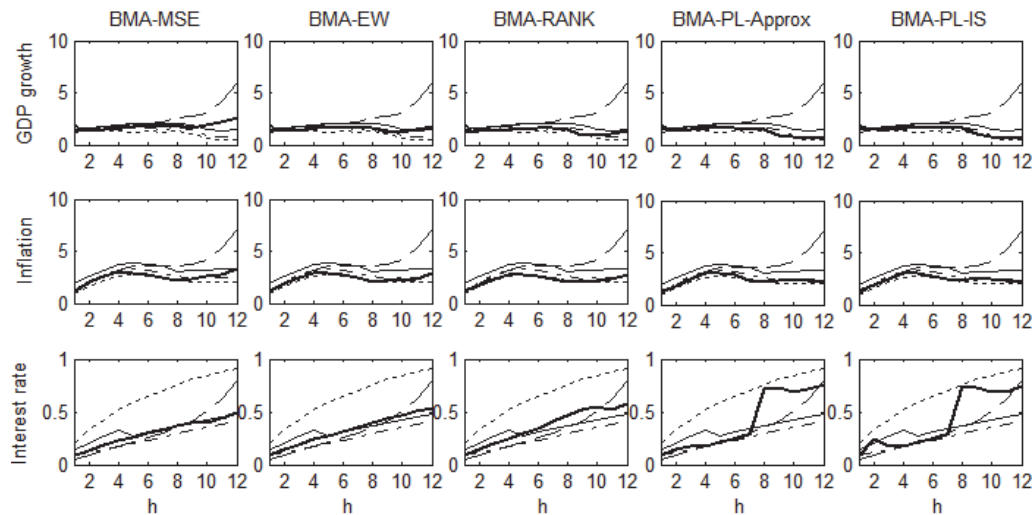
5: Model weights for all use models
Source: authors

for the DSGE-VAR model and the DSGE model are similar to each other. The lowest mean forecast errors are given by DSGE model. This result is confirmed by used forecast error statistics, see Fig. 6 (the first line).

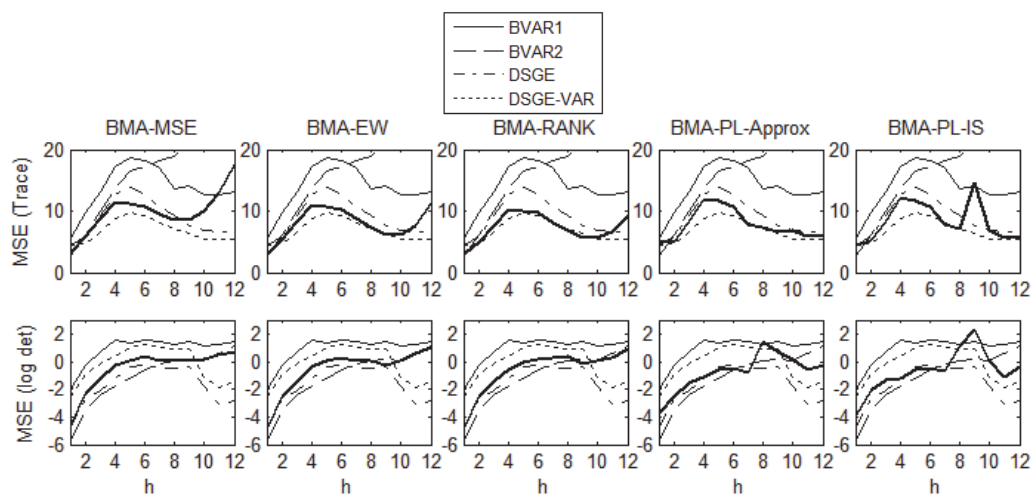
How the weights are evolved through the hold-out sample is shown in Fig. 5 for one, four, eight and twelve – step – ahead – forecasts. Five different model weights are used – the weights based on the MSE matrix, model rank and two approaches to the predictive likelihood evaluation. The equal weights are used, as well. It may be seen that the weights based on the trace of the h -step-ahead MSE matrix prefer DSGE and DSGE-VAR model with very similar weights. BVAR model weights were assigned value equal to zero. This is consistent with the statement according to the Fig. 4 which shows that DSGE and DSGE-VAR models provide the most accurate point forecasts. Rank based weights assigned the highest rank to the BVAR2

model and the lowest weights to DSGE model over all observed forecast horizons. The rank weights are as well as MSE weight computed based on the trace of the h -step-ahead MSE matrix. The differences between MSE weight and rank weight are caused by more cautious approach to the rank based weight algorithm – none of the models can be exclude and extremely preferred. Both approaches to the predictive likelihood provide very similar results. The weights assigned by both the approximation approach and importance sampling approach are almost identical. The models with lower degree of the forecast uncertainty are preferred.

The combined point forecast of models m for horizon h are determined as mean of combined density forecast (1). The forecast comparison covers both the univariate measures based on RMSE and the multivariate MSE-based measures with log determinant and trace of the MSE matrix (6); see Warne *et al.* (2013). Figs. 6 and 7 show point forecasts



6: RMSE of models
Source: authors



7: MSE of models
Source: authors

comparison of individual and combined forecasts where the bold line represents of combined forecasts. It may be seen that a model performing well in terms of density forecast accuracy does not necessarily make the most point forecasts accuracy. As indicated above, two type of the BVAR models used in the paper have higher density accurate against DSGE and DSGE-VAR models. But it is clear that the combined forecasts have in general an accuracy higher than forecasts from most single models.

From individual models, the results for observed variables are different from each other. The RMSE statistic shows that the DSGE-VAR model provides the best result for the inflation as well as GDP growth, as already mentioned. Conversely, this model fails for the real interest rate. We find that the forecasting accuracy of the DSGE and the DSGEVAR model is very similar from both the univariate RMSE and the multivariate trace MSE view. When using the multivariate log determinant MSE statistic, the DSGE model outperforms DSGE-VAR model. The contradiction for forecast evaluation between two use multivariate statistics is described by Herbst and Schorfheide (2012).

There is no much difference between the accuracy of the combined models each other evaluated by the RMSE statistic except interest rate estimated by predictive likelihood weighting scheme for the longer horizons – from eight quarter ahead horizon forecasts. MSE statistic provides an overall view of forecasts. As reported by Timmerman (2006), the equal weighted mean forecast combinations are difficult to overcome in terms of accuracy. The trace of the MSE matrix, which is influenced by the least predictable variables, shows that the rank based weight yields more accurate forecasts against other models. But, an increase of the forecast accuracy is difficult to excuse by increasing of the computational effort equated to the equal weighting forecasts. Thus, weighting scheme that

give significant weight to more models are preferred over weighting scheme that assign high weight to one or two models. The MSE log determinant statistic is influenced by both the most and the least predictable variables.

DISCUSSIONS

In this paper, we dealt with only little explored part of the economy forecast area, namely combining multivariate density forecasts. At the first, we have compared the accuracy of mean trivariate density forecasts of two BVAR models, DSGE model and DSGE-VAR model. The mean forecasts were observed from both the univariate and multivariate view. Forecast accuracy of DSGE model is comparable to the forecasting accuracy DSGE-VAR model. Especially, DSGE-VAR and DSGE models yield relatively precise GDP growth and inflation, they overcome non-structural BVAR models. However, DSGE-VAR fails for interest rate forecasts, when DSGE model provide the higher forecast accuracy. Thus, forecast accuracy used models are different. For overcoming this problem the combining forecasts from several models were used. We have employed five different weighting scheme based on the trace of the MSE matrix, the model rank, two approaches to the predictive likelihood compute and equal weight. The weighting schemes based on predictive likelihood assign almost zero weights to DSGE and DSGE-VAR model in the combined density forecast. This is due to high uncertainty of the model density forecast. There is not much of a difference between the accuracy of the other combination schemes. Weight forecasts increase the forecasting accuracy. The rank based weight and equal weigh model yield more accurate forecasts against other models. Results show that weighting scheme that give significant weight to more models are preferred over weighting scheme that tend to consider one best model.

CONCLUSION

A combining forecasts from different models to increase their forecast accuracy belongs to the popular macroeconomic research techniques and the first studies in this area have been focused on combining point forecasts. This paper proposed to combine trivariate density forecasts from two BVAR, models, a DSGE and a DSGE-VAR model. Several methods to combine forecasts from the set model were considered – predictive likelihood based weights following Adolfson *et al.* (2007) approach as well as Warne *et al.* (2013) approach, the relative performance weights based on the trace of past MSE matrix, and the weights based on model ranks. The equal-weight scheme was used as a simple combination scheme. There is no much difference between the accuracy of the combined models each other evaluated by the RMSE statistic except interest rate estimated by predictive likelihood weighting scheme for the longer horizons – from eight quarter ahead horizon forecasts. MSE statistic provides an overall view of forecasts. As reported by Timmerman (2006), the equal weighted mean forecast combinations are difficult to overcome in terms of accuracy. The trace of the MSE matrix, which is influenced by the least predictable variables, shows that the rank based weight yields more accurate forecasts against other models. But, an increase of the forecast accuracy is difficult to excuse by increasing of the computational effort equated to the equal weighting forecasts. Thus, weighting scheme that give significant weight to more models are preferred over weighting scheme that assign high weight to one or two models.

Acknowledgement

This work is supported by funding of specific research at University of Finance and Administration, Faculty of Economic Studies.

REFERENCES

- ADOLFSON, M., LASÉEN, S., LINDÉ, J. et al. 2008. Evaluating an estimated new Keynesian small open economy model. *Journal of Economic Dynamic & Control*, 32(8): 2690–2721.
- ANDERSSON, M. K. and KARLSSON, S. 2007. Bayesian Forecast Combination for VAR Models. *Sveriges Riksbank Working Paper*, No. 216.
- CHAN, J. C. and EISENSTAT, E. 2012. Marginal Likelihood Estimation with the Cross Entropy Method. *Centre for Applied Macroeconomic Analysis Working Papers*, No. 2012-18
- CHRISTOFFEL, K., COENEN, G. and WARNE, A. 2010. Forecasting with DSGE models. *European Central Bank Working Papers*, No. 1185.
- DEL NEGRO, M. and SCHORFHEIDE, F. 2004. Priors from general equilibrium models for VARs. *International Economic Review*, 45(2): 643–673.
- DOAN, T., LITTERMAN, R. B. and SIMS, C. A. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, (3)1: 1–100.
- GERARD, H. and NIMARK, K. 2008. Combining multivariate density forecasts using predictive criteria. *Reserve Bank of Australia. Research Discussion Paper*, No. 2008-02.
- HALL, S. G. and MITCHELL, J. 2004. Density Forecast Combination. *National Institute of Economic and Social Research Discussion Paper*, No. 249.
- HERBST, E. and SCHORFHEIDE, F. 2012. Evaluating DSGE Model Forecasts of Comovements. *Journal of Econometrics*, 171(2): 156–166.
- JEŘÁBEK, T., TROJAN, J. and ŠPERKOVÁ, R. 2013. Predictive performance of DSGE model for small open economy - the case study of Czech Republic. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 61(7): 2229–2238.
- JORE, A. S., MITCHELL, J. and VAHEY, S. P. 2010. Combining forecast densities from VARs with uncertain instabilities. *Journal of Applied Econometrics*, 25(4): 621–634.
- KOOP, G. 2003. *Bayesian Econometrics*. John Wiley & Sons Ltd.
- KOOP, G. and KOROBILIS, D. 2010. Bayesian Multivariate Time Series Methods for Empirical Macroeconomics. *Foundations and Trends in Econometrics*, 3(4): 267–358.
- MUSIL, K. 2009. *International Growth Rule Model: New Approach to the Foreign Sector of the Open Economy*. Dissertation Thesis. Brno: Masaryk University.
- RUBINSTEIN, R. Y. and KROESE, D. P. 2007. *Simulation and the Monte Carlo Method*. Wiley.
- SENECA, M. 2010. A DSGE model for Iceland. *The Central Bank of Iceland Working Paper*, No. 50.
- TIMMERMANN, A. 2006. Forecast combinations. In: ELLIOT, G., GRANGER, C., TIMMERMANN, A. (eds.), *Handbook of Economic Forecasting*, 1: 135–196.
- WARNE, A., COENEN, G. and CHRISTFFEL, K. 2013. Predictive Likelihood Comparisons with DSGE and DSGE-VAR Models. *European Central Bank Working Paper Series*, No. 1536.
- WOLTERS, M. H. 2012. Evaluating point and density forecasts of DSGE models. *Goethe University Frankfurt MPRA paper*, No. 36147.

Contact information

Tomáš Jeřábek: jerabek@mail.vsfs.cz
 Radka Šperková: radka.sperkova@gmail.com