DOI: 10.20472/ES.2018.7.1.004

FORECASTING CZECH GDP USING BAYESIAN DYNAMIC MODEL AVERAGING

TOMÁŠ KAREL, PETR HEBÁK

Abstract:

Forecasting future path of macroeconomic aggregates has become crucial for monetary and fiscal policymakers. Using Czech data, the aim of this paper is to demonstrate the benefits of the Bayesian dynamic averaging and Bayesian Vector Autoregressive Models (BVAR) in forecasting real GDP growth. Estimation of richly parameterized VARs often leads to unstable estimates and inaccurate forecasts in models with many variables. Bayesian inference and proper choice of informative priors offers an effective solution to this problem by shrinking the variance of model parameters. Bayesian dynamic model averaging (DMA) then makes it possible to account for model uncertainty by combining predictive abilities of many competing VAR models considered by a researcher. Since forecasting performance of individual models may vary over time, the DMA can adapt their weights in dynamic and optimal way. It is shown that the application of DMA leads to substantial forecasting gains in forecasting Czech real GDP.

Keywords:

Bayesian dynamic model averaging, Minnesota prior, Bayesian Vector Autoregressive model, GDP forecasting

JEL Classification: E17, C10, C11

Authors:

TOMÁŠ KAREL, University of Economics Prague, Czech Republic, Email: tomas.karel@vse.cz
PETR HEBÁK, Metropolitan University Prague, Czech Republic, Email: hebak@centrum.cz

Citation:

TOMÁŠ KAREL, PETR HEBÁK (2018). Forecasting Czech GDP using Bayesian dynamic model averaging. International Journal of Economic Sciences, Vol. VII(1), pp. 65-81., 10.20472/ES.2018.7.1.004

INTRODUCTION

Fair forecasts of future path of macroeconomic aggregates is a crucial issue for economic researchers and a key tool for monetary and fiscal policy makers. However, forecasting the growth in Gross Domestic Product (GDP) is generally considered a very complicated and complex exercise. Many different approaches have been suggested in the literature see Canova (2007) and Koop and Korobilis (2010) for a review. In this paper, we use a Vector autoregression model (VAR) which is a flexible time series method able to capture complex dynamic interrelationships among macroeconomic variables. However, a rich parametrization of these models often leads to inaccurate estimates of model parameters and to unstable out-of-sample forecasts. The total number of estimated parameters is given by a relation $(p \times k \times k) + [k \times (k+1)]/2 + k$ where k represents the number of variables in the given model and p number of lags. For a VAR model with 10 variables and 4 lags, it is thus necessary to estimate a total of 465 unknown parameters including the vector of constants. Economic time series, especially those in the Czech Republic, are not long enough to provide ground for reasonably-behaved estimates of such richly parameterized models. Obviously, in situations where the model has more parameters than observations, it is not possible to make any estimate of its parameters (Hawkins, 2004).

Litterman (1979) and Litterman et al. (1986) introduced an effective Bayesian way how to tackle the issue of a large number of parameters. Their approach relies on the combination of the likelihood function with the informative prior distribution. This approach has become successful as it allows effective reduction of variance of parameters and stabilization of the forecasts. Using classical (frequentist) terminology, it substantially reduces the estimation error by introducing only relatively small biases in the parameter estimates. Experience with the performance of the BVAR-based forecasts covering their five-year use in the Federal reserve bank of Minneappolis was subsequently summed up one year later in the publication Litterman et al. (1986). The Bayesian VAR model proved to be effective and a flexible tool for prediction macroeconomic variables (Bikker, 1998) especially in situations where a large set of time series is available and a very high number of parameters needs to be estimated (Banbura, Giannone, & Reichlin, 2010).

It is also important to note that the “best” model for forecasting can potentially change over time. For instance, optimal lag length or hyperparameter setting can differ across the business cycle. We address this issue using a strategy proposed by Raftery et al. (1997) which they refer to as Bayesian dynamic model averaging (DMA). Dynamic model averaging allows to choose the most appropriate model for forecasting at each point of time. Koop and Korobilis (2010) were first to implement this methodology to econometric problem of inflation forecasting.

1 LITERATURE OVERVIEW

One of the first works on forecasting macroeconomic aggregates with Bayesian vector autoregression models is by Litterman et al (1986). A comparison of the BVAR forecasts of macroeconomic aggregates using EU-7 and EU-14 data is offered by Bikker (1998). Comparison of classical VAR and BVAR models for a group of selected macroeconomic indicators of the Eurozone was carried out by Félix and Nunes (2003). The BVAR models generally showed better predictive performance compared to their classic VAR counterparts, in particular for short-term horizon up to one year. Another article with significant impact on this work is by Canova (2007). It compares different approaches to the prediction of inflation in G7 countries. Forecasts are based on ARIMA, VAR and BVAR models and are compared using Theil-U statistics. The study shows that, BVAR models generally provide more accurate forecasts than traditional VAR models on quarterly, one-year and two-year time horizons. This paper also draws on the work by Giannone et al., (2015) who suggested to weight the forecasts over the grid of hyperpameters within the Minnesota-prior setting. However, unlike Giannone et al (2015), we use dynamic form of Bayesian models averaging (DMA), which allows dynamic change of the weights of individual models over time, depending on the forecasting performance of individual models in the recent past.

2 METHODOLOGY

Bayesian inference is based on combining a priori knowledge (external information - Prior) and a a function of the parameters of a statistical model given data (Likelihood) derived from the data.

$$POSTERIOR = PRIOR \times LIKELIHOOD \tag{1}$$

External, undated information enters into the model in the form of prior distribution of the individual VAR model parameters. The shape of the priori distribution is defined by parameters (hyperparameter) for more details see (Koop, 2003). In this paper we study a choice of hyperparameters of Litterman´s Minnesota Prior (Litterman et al, 1978) within the (B)VAR model of the form

$$\boldsymbol{y}_t = \boldsymbol{a} + \sum_{t=1}^{p} \boldsymbol{A}_t \boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{2}$$

where $\boldsymbol{y}_t$ for $t = \{1,\ 2,\ ...,\ T\}$ is an k x 1 vector of endogenous variables, $\boldsymbol{\varepsilon}_t \sim N(\boldsymbol{0};\boldsymbol{\Sigma})$ is an vector of exogenous shocks, and $\boldsymbol{a}, \boldsymbol{A}_t, \boldsymbol{\Sigma}$ are matrices of suitable dimensions containing unknown parameters of the model.

In the following subsections we briefly describe main concepts used in this paper. First, we provide rationale behind Minnesota prior and its forms. We pay particular attention to the issue of how its hyperparameters are chosen. Then we present the DMA method and show how the forecasts obtained from different models can be efficiently weighted.

2.1 PRIORS FOR BVAR

The Bayesian VAR model contains (similarly to any regression model) two types of unknown parameters for which it is necessary to formulate prior views. The first group consists of autoregression parameters contained in parametric matrices $\boldsymbol{A}_t$, the second group is formed by the elements in the covariance matrix of the exogenous shocks. terms. Although the prior distributions are, in principle, fully in the researchers' hands and should reflect their prior views as close as possible, for computational reasons the parameters in the matrices $\boldsymbol{A}_t$ usually follow multivariate normal distribution. For the same reason, prior views on the covariance matrix of error terms is captured by some form of inverse Wishart distribution. The joint prior distribution for both groups of parameters can then be formulated in two ways, depending on whether the prior dependence or independence between the two parameter-sets is assumed. If one assumes that probability distributions of the parameters in $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ are dependent, the joint prior density $p(\boldsymbol{\alpha},\boldsymbol{\Sigma})$ can be factored as $p(\boldsymbol{\alpha},\boldsymbol{\Sigma}) = p(\boldsymbol{\Sigma})\,p(\boldsymbol{\alpha}|\boldsymbol{\Sigma})$. While conditional distribution $p(\boldsymbol{\alpha}|\boldsymbol{\Sigma})$ may complicate elicitation of some kinds of economically reasonable prior views, the form above bears considerable computational advantages. In particular, in this case posterior density can be obtained in a closed-form and no posterior sampling is needed. (Gibbs Sampler or Metropolis Hastings algorithm). The second way is when dependence between prior parameters $p(\boldsymbol{\alpha},\boldsymbol{\Sigma}) = p(\boldsymbol{\alpha})\,p(\boldsymbol{\Sigma})$ is assumed and the posterior density doesn´t have a closed analytical form and posterior simulators are needed (Koop, 2003).

2.2 LITTERMAN´S MINNESOTA PRIOR

Researchers from the University of Minnesota and the Federal Reserve Bank of Minneapolis, Doan, Litterman, and Sims, trying to reduce parameter shrinking using the BVAR model first introduced an prior distributrion that later became known as the Litterman Prior distribution (Doan, Litterman, & Sims, 1984) and (Litterman et al, 1986) or more commonly also as so-called Minnesota prior. Their approach to the choice of prior is based on a considerable simplification of the functioning of the economic system and simple calculation of posterior densities. It also takes into account certain features that can be expected in time series of an economic nature. One of these features is the assumption of the persistent behavior of economic variables close to the Random Walk process. Litterman's prior distribution is based on the replacement of the covariance matrix $\mathbf{\Sigma}$ by its estimate $\hat{\mathbf{\Sigma}}$.. The original Litterman version assumes that $\mathbf{\Sigma}$ it is a diagonal matrix (Litterman, 1979). The parameters of each VAR model equation are then estimated separately at the set $\sigma_{ii}^2 = s_i^2$ where s_i^2 is the variance estimation of the random cpmponent, is determined using the OLS method in the *i*-th equation and σ_{ii}^2 is the *ii* element of $\hat{\mathbf{\Sigma}}$. In the case that the covariance matrix $\mathbf{\Sigma}$ is not assumed to be diagonal, it is possible To replace it with an estimate $\hat{\mathbf{\Sigma}}$ (Koop, 2003).

As noted Koop and Korobilis (2012) the disadvantage of this approach is that we do not deduce the unknown matrix of parameters $\mathbf{\Sigma}$ by the Bayesian method, but rather replace it by its estimation $\hat{\mathbf{\Sigma}}$. An advantage, however, is a considerable simplification of the posterior density calculation and even the possibility of its expression in analytical form. Litterman's prior distribution also offers high flexibility in the choice of prior density (Lütkepohl, 2005). After replacing the covariance matrix $\mathbf{\Sigma}$, it is necessary to set the prior densities (*hyperparameters*) for the parameter vector $\boldsymbol{\alpha}$. For Litterman 's Minnesota prior density, the assumption is

$$\boldsymbol{\alpha} \sim N\left(\underline{\boldsymbol{\alpha}}, \underline{\boldsymbol{V}}\right), \tag{3}$$

where $\underline{\boldsymbol{\alpha}}$ and $\underline{\boldsymbol{V}}$ are hyperparmeters of normal distribution. Most of (in some cases all) values of the prior mean values are set to zero, which reduces the standard error

of model parameter estimation. However, if the model is working with non-stationary time series, Litterman's prior division uses a mean value $\underline{\boldsymbol{\alpha}}$ such that individual variables behave similarly to random walk processes. The priori mean value hyperparametr $\underline{\boldsymbol{\alpha}}$ is then set as in the previous case, except for the elements corresponding to its first delay dependent variable in each VAR model equation. For these own lags, the value 1 is set, respectively. A value close to 1 e. g. in this paper 0.95, taking into account the above-mentioned property of high persistence of economic time series Koop and Korobilis (2010) and Karlsson (2012). The values of the other prior hyperparameters are then set to zero. Koop and Korobilis (2010) further state that if $\underline{\boldsymbol{V}}_i$ denotes the block of the prior covariance matrix $\underline{\boldsymbol{V}}_i$ associated with k coefficients in the i-th equation, then its diagonal elements are determined as:

$$\underline{V}_{i,jj} = \begin{cases} \dfrac{\underline{a}_1}{p^2} & \text{for lag } p \text{ variable } j = i \text{ where } p \in \langle 1; P \rangle, \\ \dfrac{\underline{a}_2 \sigma_{ii}^2}{p^2 \sigma_{jj}} & \text{for lag } p \text{ variable } j \neq i \text{ where } p \in \langle 1; P \rangle, \\ \underline{a}_3 \sigma_{ii}^2 & \text{for other endogenous variables..} \end{cases}$$

For simplicity and user-friendliness is the setting of all elements of the prior covariance matrix simplified here only for the choice of 3 scalars $\underline{a}_1$, $\underline{a}_2$ and $\underline{a}_3$ instead of having to set prior mean and prior variance for all $(p \times k \times k) + [k \times (k+1)]/2 + k$ parameters of the VAR model. The more the values of the individual parameters of VAR model are reduced to zero, the lower is the effect of the delay. When choosing values $\underline{a}_1 < \underline{a}_2$, the weighting of a particular variable is degraded to the lags of other variables. This reflects the prior assumption that the process is more closely influenced by the closer than the distant past. Elements of covariance matrix σ_{ii}^2 are usually replaced by values s_i^2. Given the advantageous properties of Litterman's prior density, the posterior density has a normal distribution with hyperparameters $\overline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{V}}$

$$\boldsymbol{\alpha} | \boldsymbol{y} \sim N\left(\overline{\boldsymbol{\alpha}}, \overline{\boldsymbol{V}}\right). \tag{4}$$

The posterior mean value of $\bar{\boldsymbol{\alpha}}$ has the form

$$\overline{\boldsymbol{\alpha}} = \overline{\boldsymbol{V}}\left[\underline{\boldsymbol{V}}^{-1}\boldsymbol{\alpha} + \left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \boldsymbol{X}\right)' \boldsymbol{y}\right]. \tag{5}$$

The variance of normal distribution is obtained as

$$\overline{\boldsymbol{V}} = \left[\underline{\boldsymbol{V}}^{-1} + \left(\hat{\boldsymbol{\Sigma}}^{-1} \otimes \left(\boldsymbol{X}'\boldsymbol{X}\right)\right)\right]^{-1}. \tag{6}$$

In the classical (frequency) approach, the number of parameters and their lags are often reduced by so-called hard restrictions based on the ad-hoc results and t-tests of zero value or similar techniques, ensuring prior hard restriction of a specific variable, its delay in the model under consideration. Restrictions on the number of parameters executed using the Litterman´s Minnesota prior distribution are called soft restrictions as they allocate a certain probability distribution to the VAR model, giving the researcher a more flexible view of the indefiniteness he encounters when constructing the model (Canova, 2007). Due to the fact that time series of economic nature also contain a number of non-systematic noise, which can burden the estimation of parameters with unintended inaccuracies, Litterman's prior density is able to reduce these non-significant parameters to zero.

A more general version of Litterman's prior parameter setting was introduced, for example, by Canova (2007), where the diagonal elements $\underline{V}_{i,jj}$ of the covariance matrix are intended as

$$\underline{V}_{i,jj} = \begin{cases} \dfrac{\underline{a}_1}{d(p)} & \text{for lag } p \text{ variable } j = i \text{ where } p \in \langle 1; P \rangle, \\ \dfrac{\underline{a}_1 \underline{a}_2 \sigma_{jj}^2}{d(p)\sigma_{ii}^2} & \text{for lag } p \text{ variable } j \neq i \text{ where } p \in \langle 1; P \rangle, \\ \underline{a}_1 \underline{a}_3 & \text{for other endogenous variables.} \end{cases}$$

As d(*p*) there is a so-called decay function, which in Litterman's previous version was set to a fixed value in the form of p^2 In this work decay function is exponential $d(p) = \underline{a}_4^{-p+1}$ where $\underline{a}_4 > 0$.

2.3 BAYESIAN DYNAMIC MODEL AVERAGING

Bayesian Dynamic Modeling method was proposed by Raftery et al. (1997) and Raftery (1999) as an effective tool for combining the predictive abilities of a set of competing models over time. The concept of dynamic model averaging reflects the fact that forecasting performance of individual models may vary over time, depending on the explanatory variables considered, the number of lags used or the prior distribution of parameters employed (e. g.models can exhibit different forecasting performance across the phases of the business cycle). The weights (prior probabilities) of the models at the initial period are usually assumed to be equal for all models, i.e.

$$\pi_{t=0} = 1/R \tag{7}$$

Prior probabilities $\pi_{t=0}$ are then updated depending on previous forecasting abilities. Probability of model *r*, where *r* = 1,2, … *R* in time *t* is

$$\pi_{t|t-1,r} = \left(\pi_{t|t-1,r}^{\alpha} + c\right)\left(\sum_{r=1}^{R} \pi_{t|t-1,r}^{\alpha} + c\right)^{-1} \tag{8}$$

Where α is called the forgetting factor, $0 < \alpha < 1$, which most often takes values close to one. E.g., for quarterly data choice $\alpha = 0,99$ suggests that observations five years ago receive approximately 80% as much weight as last period´s observation. The choice of this value consistent with fairly stable models where parameters change is gradual. In case of choice $\alpha = 0,95$, observations five years ago receive only about 35% as much weight as observation in the last period (Koop and Korobilis, 2010). The *c* symbol indicates a small positive constant, providing a non-zero probability value for the model. Raftery et al. (1997) recommend to set the constant to 0.001/ R. Each of subsequent predictive step updates the posteriori probabilities of all competing models depending on the quality of last prediction quality. In this work Kullback-Leibler information criterion (KLIC) is used to evaluate the quality of predictions and balance assignments (Kullback and Leibler, 1953). The KLIC criterion takes into account the whole prediction density of s of individual models, unlike point characteristics such as Mean Square Error (MSE), Root Mean Square Error (RMSE) or Mean Absolute Error (MAE).

3 APPLICATION

In our application Predictions were constructed using 18 models with different choice of hyperparameters and number of lags (2, 3 and 4). Six different types of Litterman´s Minnesota priors setting and two types of decay function were considered . The first decay function by Koop and Korobilis (2010) settings with linear decay form and the second one by Cannova (2007) considering geometric decay function. These two types of setting hyperparameters were set in Tight, Middle and Non-informative way. On Figure 1 we can see the list of all considered models with different settings of decay function (Koop and Korobilis 2010 and Cannova 2007), different settings of tightness (Tight, Middle and Non-Inf) and different number of lags (2, 3 and 4).

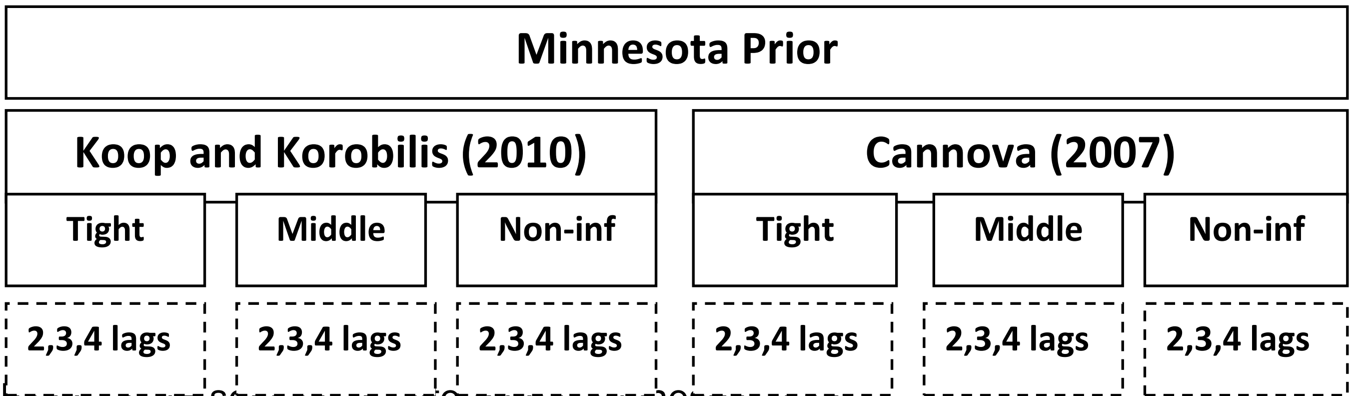

Figure 1: 18 settings of Litterman´s Minnesota prior

4 DATA AND VARIABLES

The data used in the empirical part of this work were obtained from three sources. The first source is the publicly accessible database of the Czech National Bank (CNB), the database of the Czech Statistical Office (CZSO) and the database of the European Statistical Office (EUROSTAT). The variables entering the models were chosen on the basis of economic theory and experience from previous studies (see a.g. Adam and Plasil (2014). Additionaly, statistical criteria such as marginal likelihood (Bayesian approach to model selection) or AIC were taken into account, see Koop (2003)

Note that the choice of time series used in the candidate models was also influenced by the availability and data quality of relevant economic indicators. Therefore, only quarterly time series covering the full period from 1998 to 2016 with minimal changes in the methodology were considered. E.g. Methodology of construction of the industrial production index has been changed and after the consultation of the impact of methodological changes with the head of the Department of Industry Statistics of Czech Statistical office we decided to not include this indicator into our model.

The starting year 1998 is seed a compromise between the sufficient length of time series analyzed, the availability of time series data, and the methodological consistency . The change of the monetary policy regime, approved in December 1997 by the Czech National Bank, also supports this choice. It is also possible to say that some time has already passed since the significant structural changes of the Czech economy caused by the transition and its transformation from the central planning regime to the market economy regime during the first half of the nineties. From the year 1998, the Czech economy can already be considered (with a certain degree of simplification) structurally similar to the current one.

For selecting the number of lags in VAR model, we used combination of economic and statistics criterions. In this work we consider VAR models with 2, 3 and 4 lags because of posterior probabilities based on Bayes factor (Koop 2003, Koop and Korobilis, 2012) Bayes factor takes into account the combination of prior probability and marginal likelihood of each considered model. For the choice of the number of lags we used upgraded version of Bayes factor designed by Karlsson (2012) as well. Karlsson (2012) replaced marginal likelihood in Bayes factor with predictive marginal likelihood of training sample. Results based on both versions of Bayes factor recommended the choice of 2,3 and 4 lags. To predict the time series considered, it is often appropriate to transform the used time series of individual economic and financial indicators. Literature focusing on empirical analysis of time series offers two basic approaches (Lütkepohl, 2005). The VAR model can be estimated at either levels or differences. Typically, non-stationary macroeconomic data is entering the level model. In this work the models are estimated in the second variant, in year-on-year differences comparing the relative development of the selected quarter with the same quarter of the previous year. The choice of transformation using year-to-year changes has been chosen with regard to its use in the vast majority of authors of well-known foreign publications that deal with macroaggregates predictions using BVAR models. Furthermore, the recommendations of Koop and Korobilis (2010), Cannova (2007) and Blake and Mumtaz (2012) were taken into account and the fact that the year-on-year growth rate belongs to the most frequently monitored quantities in the professional economic public and industry.

All time series entering the model were seasonally adjusted, with the HP (Hodrick-Prescott) filter being applied to the selected time series.The time series HP filter application allows to eliminate the oscillations relatively efficiently longer than those that correspond to the length of the business cycle. As Cannova (2007) states, the Bayesian approach to VAR model estimation allows unlike the classic approach to work directly with non-stationary time series. Although the classical and Bayesian approaches are similar in many ways, when the unit root is present in the time series, both approaches are fundamentally different. While in the classical approach the asymptotic distribution of parameter estimates in the case of a unit root is non-standard, the posterior division in Bayesian estimation is not affected by this fact. Thus, in the case of using the Bayesian approach to estimating model parameters, unlike the classical approach, the failure to fulfill the assumption of stationarity is not a major problem (Canova, 2007).

Variable	**Transformation,**
Czech GDP	Year-to-year difference, quarterly
Germany GDP	Year-to-year difference, quarterly
Inflation	Year-to-year difference of CPI (Customer Price Index), quarterly
Nominal effective exchange rate	Year-to-year difference (CZK to weighted basket of foreign currencies), quarterly
PRIBOR - Prague interbank offered rate	Year-to-year difference, quarterly

Table 1: List of variables

5 RESULTS

Dynamic model averaging methodology rapidly helped to improve quality of forecast. On Table 2 we can see how DMA and Minnesota prior improved forecast statistics in comparison to non-informative prior.[1] The first line in each column is the value of characteristics for DMA forecast, the second line is the value of characteristics if the Non-informative prior (classical VAR) is used and the last bolt line is the percent improvement of DMA in comparison to Non-Informative Prior.

Table 1: Comparison of Dynamic Model Averaging and Classical OLS estimate (Non-Informative prior) and percentual improvement

DMA (Non-informative) ***Percentual improvement**	**T+1**	**T+2**	**T+4**	**T+8**
MSE	0,804 (1.080) **74.5%**	2,907 (4.140) **70.2%**	8,136 (12.347) **65.9%**	8,620 (19.244) **44.8%**
RMSE	0,897 (1.039) **86.3%**	1,705 (2.035) **83.8%**	2,852 (3.514) **81.2%**	2,936 (4.387) **66.9%**
MAE	0,656 (0.803) **81.8%**	1,179 (1.414) **83.4%**	2,190 (2.647) **82.8%**	2,427 (3.273) **74.2%**
THEILU1	0,154 (0.169) **90.9%**	0,282 (0.321) **87.7%**	0,459 (0.505) **90.9%**	0,473 (0.543) **87.2%**
THEILU2	0,658 (0.762) **86.3%**	0,668 (0.797) **87.7%**	0,684 (0.842) **90.9%**	0,632 (0.688) **87.2%**

Data Source: Authors´ computations

[1] The forecast using non-informative prior is comparable with classical OLS VAR forecast (Koop et al, 2007).

At the T+1 horizon, the MSE decreased by 26.5%, at the T + 2 horizon by 29.8%, at the T + 4 horizon by 34.1% and at the highest T + 8 horizon by 55.2%. Using Dynamic model averaging methodology has led to improvement of all considered characteristics. The largest improvement was achieved at two years (T+8) horizon.

On following figures we can see the GDP forecast of all models (blue dashed lines), the final dynamic model averaging forecast (dark blue line) and the real GDP growth (red line).

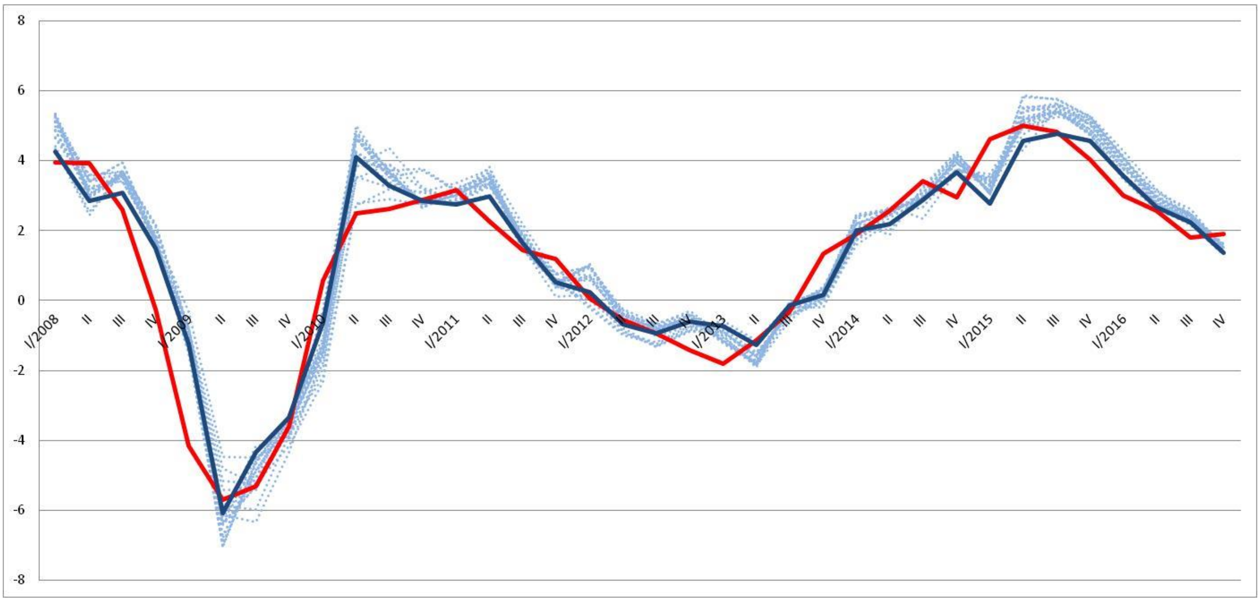

Figure 1: Forecast T+1 Dynamic model averaging

On Figure 1 we can see the forecasts on the horrizon one quarter ahead (T+1). The dark blue line represents the final Dynamic model averaging forecast. For the real growth of GDP is use red line. Mean square error on horizon T+1 decrease to 0.804, Root mean square error to 0.897, Mean absolute error to 0.654, THEIL 2 statistics to 0.657.

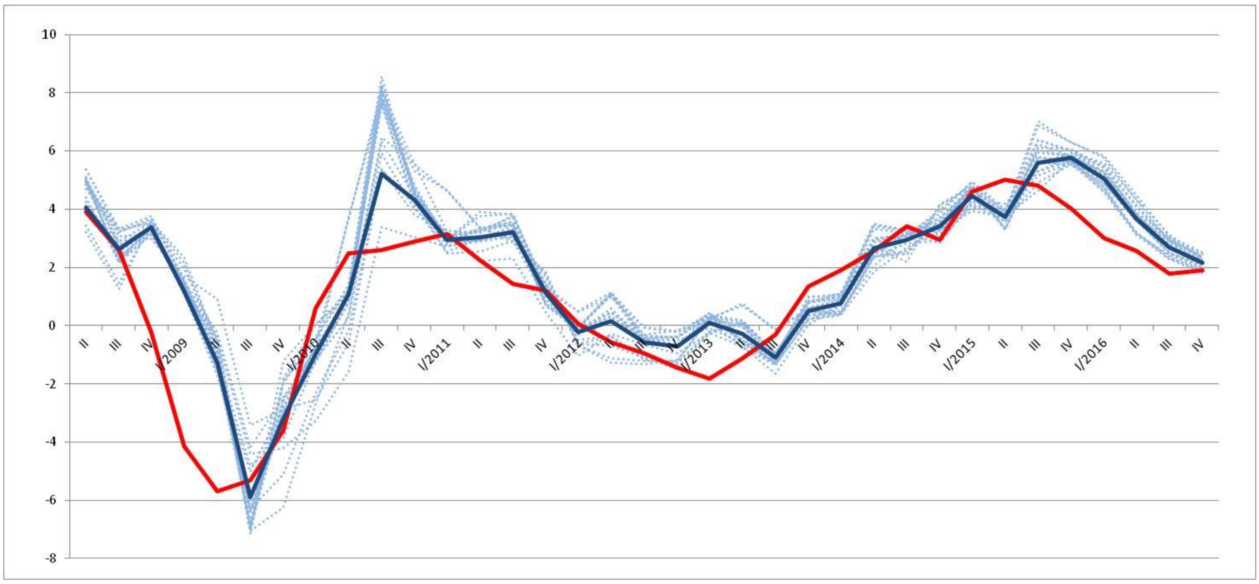

Figure 2: Forecast T+2 Dynamic model averaging

The forecast two quarter ahead forecast can be seen on the Figure 2. We can see that no version of Minnesota prior was able to forecast the beginning of the financial crisis in 2008 Q1. Mean square error decreases to 2.907, Root mean square error to 1.705, Mean absolute error to 0.656 and THEIL 2 statistics to 0.668.

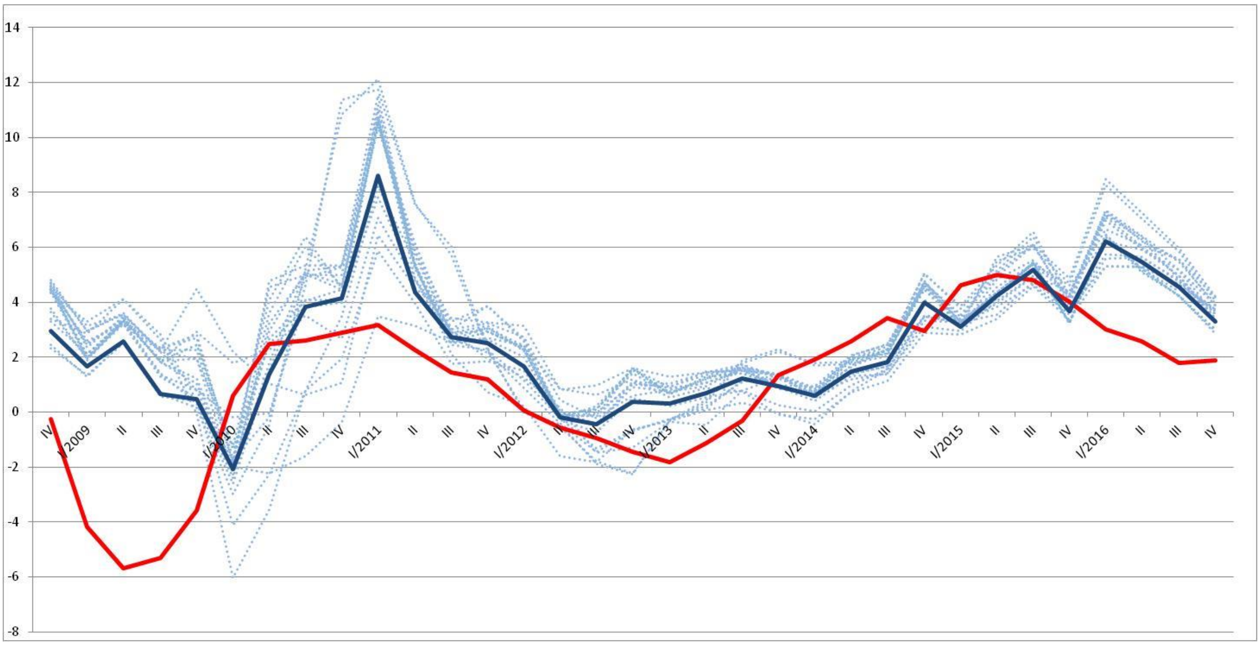

Figure 3: Forecast T+4 Dynamic model averaging

On the Figure 3 we can see the forecast 4 quarters ahead. The Mean square error was decreased to 8.136, root mean square error to 2.852, mean absolute error to 2.190 and THEIL 2 statistics to 0.684.

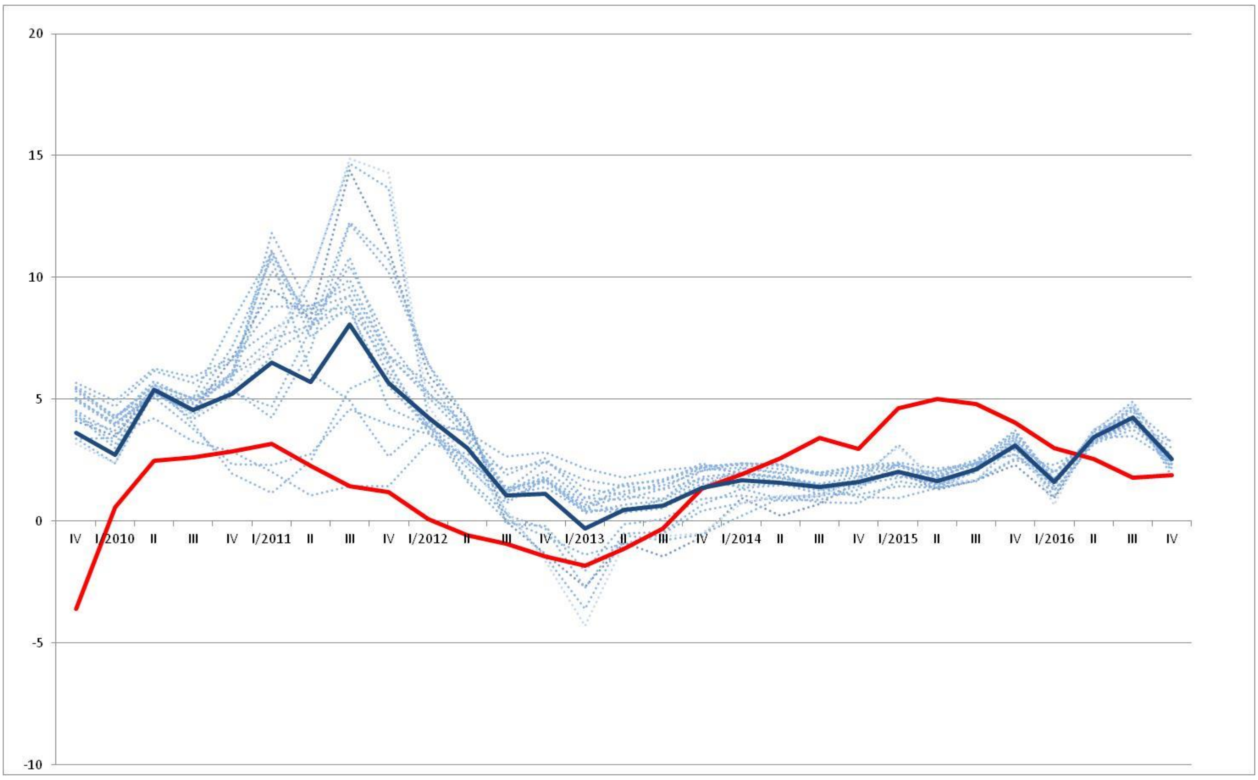

Figure 4: Forecast T+8 Dynamic model averaging

On the last Figure 4 we can see the forecast of Czech GDP on horizon 8 quarters (two years) ahead. Mean square error was decreased to 8.619, root mean square error to 2.935, mean absolute error to 2.427 and Theil 2 statistics to 0.631.

The final forecast on one-year horizon using Bayesian Dynamic averaging method is shown in Figure 5. We can see how the prediction looked like in every single quarter. The most problematic period for forecasting was when the economy was falling into the financial crisis between 2008 Q2 and 2009 Q2.

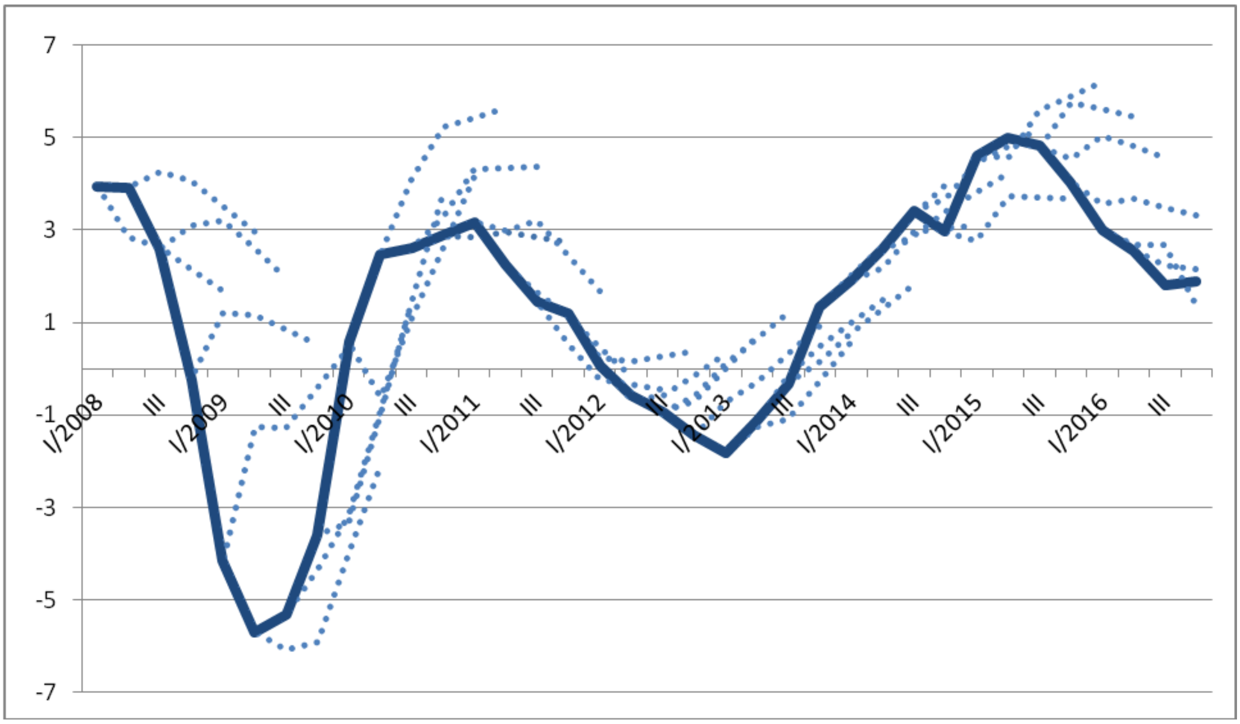

Figure 5: One-year horizon forecast using Bayesian Dynamic averaging

DISCUSING CONCLUSIONS

Fair forecasting of future develop of economic macroaggregates has become crucial for monetary and fiscal economic policy makers. Bayesian statistics and appropriate choice of prior helped to solved the problem of economic VAR models – over parametrization. In this paper we studied the problem of how to choose the best settings of informativeness of a prior distribution for VAR models. Every choice of the most appropriate model is always connected with a certain degree of uncertainty. We used Dynamic form of Bayesian model averaging as a tool to partly solve this problem. The weights of each models are over time changing depending on quality forecast in previous periods. We used KLIC criterion to evaluate quality of forecast and to implicitly penalize forecast intervals that are too wide. The use of dynamic weighing across the range of hyperparametres of each model has led to a significant increase of prediction quality characteristics in comparison to non-informative prior at all considered horizons. Dynamic model averaging caused decrease of Mean Square Error about 25.5% at T+1 horizon, about 29.8 % on T+2, 34.1% on T+4 and 55.2% on T+8 horizon. Theil statistic (THEIL2) was improved in average about 14.23 % at all horizons. Moreover, this approach reduces the number and importance of subjectivity choices in the setting f the prior. The following research will be focused on how weights of hyperparameter settings, tightness, variables or number of lags are changing during the economic cycle.

REFERENCES

Adam, T., & Plašil, M. (2014). The Impact of Financial Variables on Czech Macroeconomic Developments: An Empirical Investigation. *Czech National Bank Working papers 2014* .

Banbura, M., Giannone, D., & Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics* . https://doi.org/10.1002/jae.1137

Bikker, J. (1998). Inflation forecasting for aggregates of the EU-7 and EU-14 with Bayesian VAR models. *Journal of Forecasting* (2), pp. 147-165. https://doi.org/10.1002/(SICI)1099-131X(199803)17:2<147::AID-FOR674>3.0.CO;2-Q

Blake, A. P. a Mumtaz, H. 2012. Applied Bayesian econometrics for central bankers. *Technical Books.* 2012.

Canova, F. (2007). *Methods for applied macroeconomic research.* New Jersey: Princeton University Press.

ČNB. (n.d.). *Jak vzniká prognóza*. Retrieved 10 16, 2016, from Česká národní banka.

Félix, R. M., & Nunes, L. C. (2003). *Forecasting euro area aggregates with Bayesian VAR and VECM models.* Banco De Portugal: Economic Research Department.

Geweke, J., & Amisano, G. (2010). Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting* (2), pp. 216-230. https://doi.org/10.1016/j.ijforecast.2009.10.007

Giannone, D., Lenza, M., & Primiceri, E. G. (2015, May). Prior Selection for Vector Autoregressions. *Review of Economics and Statistics* , pp. 436-451. https://doi.org/10.1162/REST_a_00483

Gupta, R. (2006). Forecasting the South African Economy with VARs and VEC. *South African Journal of Economics* (4). https://doi.org/10.1111/j.1813-6982.2006.00090.x

Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical information and Modeling* , pp. 1-12.

Hoeting, J., Madigan, D., Raftery, A., & Volinsky, C. (1999). Bayesian model averaging: A tutorial. *Statistical Science* .

Koop, G., & Korobilis, D. (2010). *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics.* Now Publishers Inc.

Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review* (3), pp. 867-886. https://doi.org/10.1111/j.1468-2354.2012.00704.x

Koop, G., Poirier, D., & Tobias, J. (2007). *Bayesian Econometric Methods.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511802447

Kullback, S. a Leibler, R. 1951. On information and sufficiency. *The annals of mathematical statistics.* Vol. 22, 1951, 1, s.79-86.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics* (1), pp. 25-38.

Litterman, R. B. (1979). Techniques of forecasting using vector autoregressions. *Federal Reserve Bank of Minneapolis Working Paper. no. 115* .

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer Science & Business Media. https://doi.org/10.1007/978-3-540-27752-1

Mandel, M., & Tomšík, V. (2008). *Monetární ekonomie v malé otevřené ekonomice.* Praha: Management Press.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies* (3), pp. 821-852. https://doi.org/10.1111/j.1467-937X.2005.00353.x

Raftery, A. E. 1995. Bayesian model selection in social research. *Sociological methodology.* 1995, s. 111-163. https://doi.org/10.2307/271063

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association* (437). https://doi.org/10.1080/01621459.1997.10473615

Robert, C. P., & Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer Verlag. https://doi.org/10.1007/978-1-4757-3071-5

Spěváček, V., Žďárek, V., Zamrazilová, E., & Rojíček, M. (2016). *Makroekonomická analýza - teorie a praxe.* Praha: Grada Publishing a.s.

Sun, D., & Ni, S. (2004). Bayesian analysis of vector-autoregressive models with noninformative priors. *Journal of Statistical Planning and Inference* (2), pp. 291-309. https://doi.org/10.1016/S0378-3758(03)00116-2

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* Wiley.

