ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»
Международный Институт Экономики и Финансов

Кондрашов Артем Александрович

ИСПОЛЬЗОВАНИЕ BVAR ДЛЯ ПРОГНОЗИРОВАНИЯ РОССИЙСКИХ МАКРОЭКОНОМИЧЕСКИХ ИНДИКАТОРОВ

USE OF BVAR MODELS FOR FORECASTING RUSSIAN MACROECONOMIC INDICATORS

Выпускная квалификационная работа -
БАКАЛАВРСКАЯ РАБОТА
по направлению подготовки 38.03.01 «Экономика»
образовательная программа **«Программа двух дипломов по экономике НИУ ВШЭ и Лондонского университета»**

Рецензент:	Научный руководитель:
Кандидат физико-математических наук	Магистр экономики
Виктор Александрович Лапшин	Борис Борисович Демешев

Москва 2018

Abstract

Contemporary economy is a complex subject, and accurate forecasting is important for conducting macroeconomic policies, as the effects of the latter are often lagged. Lately, many researchers, including employees of central banks find that Bayesian VAR models are generally more accurate than conventional approaches. Moreover, BVARs are capable to account for more data, solving the problem of overparametrisation.

In this paper I adopt Bayesian Compressed Vector Autoregression (BCVAR) model (Koop, Korobilis, Petenuzzo (2018)) with conjugate Normal-Inverse Wishart prior modified with «sum-of-coefficients» and «initial observation» (Sims, Zha, (1998)) and discuss it capabilities in terms of forecasting Russian macroeconomic indicators. The data consists of 30 monthly time series running from January 1999 to December 2015. The forecasts are compared against RWWN and VAR models.

The obtained results suggest that the BCVAR model generally beats the benchmark model, with some variables showing extremely good forecast quality.

Contents

Introduction

Economic forecasting is an important element in macroeconomic policies conduction. Inaccurate predictions, no matter whether they overestimate or underestimate, may incur additional costs. So, it is crucial to develop robust models with high forecasting accuracy.

One of the most frequently used models nowadays in the VAR (Vector Autoregression) model, which became popular after being introduced by Sims (1980). VAR is relatively easy to understand, enables modelling the dynamics between variables and is successfully used for forecasting and capturing macroeconomic shocks (Shevelev, 2017). However, for better handling the dynamics between variables, the VAR model needs a large number of lags in its specification. In addition, cental banks are interested in using large amount of variables in their analysis, which increases the number of coefficients to estimate. This may lead to the problem of overparameterization, when the amount of observations is not sufficient to estimate all desired parameters, and the researches have to switch to more parsimonious models. This is especially typical for Russian macroeconomic data: as it was mentioned in (Deryugina, Ponomarenko, 2015) a typical VAR model would often include no more than 5 indicators due to lack of sufficiently long times series. This drawback is addressed by BVAR (Bayesian VAR) models, which shrink the parameters by explicitly setting their prior distribution. More recent extensions of BVAR suggest adding random compression to explanatory variables (Koop, Korobilis, Pettenuzzo, 2018), facilitating estimation of models with over 100 variables.

In this paper, I adopt the BCVAR (Bayesian Compressed VAR) model to compare its forecast efficiency based on Russian macroeconomic data against more conventional models, notably the AR(1) model. The obtained results support the viability of BCVAR model in forecasting the data, especially on medium to long forecast horizons. Moreover, the author of this paper has created the *RcppBVAR* package for R, which can be used for both educational and scientific purposes.

The structure of the paper is the following: Chapter 1 covers the review on important literature. Chapter 2 describes the time series data chosen for the paper. Chapter 3 briefly introduces the basics of BVAR models with a deeper insight on BCVAR and forecasting techniques. Chapter 4 and 5 contain the results of estimations and forecast comparisons. In conclusion, I present the summary of the work.

Chapter 1. Literature review

A lot of research papers can be found on BVAR models and their success in handling over-parametrised data. Despite this fact, BVAR research, both theoretical and empirical, in Russia is quite scarce. In this chapter, I will make a review of relevant literature.

In their paper, Demeshev and Malakhovskaya (2016) [2] write a review on existing priors for Bayesian VAR models in a comprehensible way, mentioning their strong and weak points. The paper gives user-friendly guidelines for realisation of the priors as well as provides links to existing program codes for BVAR estimation including their own. Moreover, the authors describe the forecasting procedure. In their other work [3], the forecasting subject is given more insight, with BVAR model with Minnesota prior used to forecast CPI, Interbank rate, and IPI (Industrial Production Index). Despite the fact that Minnesota prior is not based on any economic model, the findings constitute that the BVAR model completely outperforms VAR, and outperforms Random Walk model in forecasting CPI and Interbank Rate, but not IPI. Another empirical work by Deryugina, Ponomarenko (2015), show findings that their Large BVAR model (BVAR with Minnesota, sum-of-coefficients and dummy-initial observation priors) is most accurate on real variables, especially on long forecast horizons, but cannot outperform vanilla Minnesota BVAR on price variables. It is noted by (Pestova, Mamonov, 2016) that the authors have used GINF (Generalized Impulse Response Functions), which do not account well for linkages between variables.

Considering the theoretical papers on BVAR, I would like to mention (Sims, Zha, (1998)) which describes prior modifications, notably «sums-of-coefficients» and «initial observation». In the article, the authors find that the abovementioned modifications give reasonable results, compared to using flat priors. There are also some technical papers, which provide guidelines for using BVAR. For example, (Blake, Mumtaz (2012)) provided a handbook covering Gibbs and Metropolis-Hastings algorithms for BVAR models, as well as Matlab code.

As for more empirical papers, paying special attention to forecasting, the following literature can be mentioned. In (Banbura, Giannone, Reichlin (2010)), the authors observe the performance of BVARs of different sizes, from 3-7 variables to 20 and 130. Their results suggest that BVARs can be effectively used for large data, providing better forecasting results than a conventional VAR model. Moreover, they found that BVAR models can be successfully used for monetary shocks effect analysis. Another article by (Carriero, Clark, Marcellino (2015)) finds that simulation forecasting method performs better than multi-step approaches, most evident for long horizons.

Both these papers adopt (Sims, Zha (1998)) SC and IO prior modifications mentioned above.

One of the most recent developments of BVAR model is the addition of random compression, described in (Koop, Korobilis, Pettenuzzo, 2018). The authors continued the idea of Bayesian Compressed Regression by Guhaniyogi, Dunson, 2014 by using a Random Projection Matrix to shrink the number of variables in BVAR model estimation. In addition, they compare forecasts from $BCVAR_C$ model (with covariance matrix compression) with BCVAR model without covariance matrix compression, Dynamic Factor model, Factor-Augmented VAR and BVAR with Minnesota prior. They used monthly US data from January 1960 to December 2014 with up to 129 variables of different types and 13 lags (216462 coefficients total). The data was divided into 3 clusters with 19, 46 and 129 variables and forecasted up to 12 steps ahead. For forecast accuracy measure, they used MSFE ratio[1], setting AR(1) model as a benchmark. The results show that BCVAR models beat the benchmark model and tend to forecast better than other approaches, especially on short horizons. As for specific indicators, BCVAR models tend to make good predictions for unemployment, prices and industrial production, but not very successful for 10-year T-bonds interest rates. In general, $BCVAR_C$ forecasts better than plain BCVAR, but there are some exceptions.

Overall, one may say that literature on Bayesian VARs is mainly focused on testing more sophisticated model specifications and priors in terms of forecasting quality. Moreover, many authors are also concerned by optimisation problems that arise while using data-heavy models. It is also important to mention the need for understandable and open-source software for reproducability of the results.

[1] See Chapter 3.5 for more details.

Chapter 2. Data overview

The dataset for estimation and forecasts includes 30 monthly time series (see Table 1) running from January 1999 to December 2015. The data was collected from various sources, including:

- FSSS (Federal State Statistical Services)
  Website: http://www.gks.ru/
- UAESD (United Archive of Economic and Sociological Data)
  Website: http://sophist.hse.ru/
- CBR (Central Bank of Russia)
  Website: http://cbr.ru/
- FINAM
  Website: https://www.finam.ru
- IFS (International Financial Statistics)
  Website: http://www.imf.org/en/Data
- OECD (Organisation for Economic Co-operation and Development)
  Website: https://data.oecd.org
- inflationinrussia.com (inflation data only)

FSSS, CBR and UAESD datasets were obtained from sophist.hse using R interface. OECD data was obtained from FRED (also using R commands).

All time series with seasonal fluctuations were decomposed using STL (Seasonal and Trend decomposition using LOESS). In estimations and forecasts, only the trend series are used. To counter the differences in scale within the data (see Table 2), standartisation is done. KPSS tests show strong evidence against trend stationarity[2][3]

[2]See Appendix A for the corresponding table
[3]For series graphs see Appendix B.

Table 1: Data

№	Name	Type	Basis (if any)	Source
1.	Real Monetary Income	Basis index	1992.12	FSSS
2.	Real Construction Volume Index	Basis index	1999.01	FSSS
3.	New Houses Commissioning	1000 of m^2		FSSS
4.	Real Wages Index	Basis index	1993.01	FSSS
5.	Agricultural Real Production Vol. Index	Basis index	1993.01	FSSS
6.	Unemployment Rate	Percent		FSSS
7.	Declared Need In Workers	1000 of people		FSSS
8.	Construction Works Price Index	Chain index		FSSS
9.	Transportation Tariffs Index	Chain index		FSSS
10.	Real Investments In Fixed Capital Index	Basis index	1994.01	UAESD
11.	Total Export	bln USD		FSSS
12.	Total Import	bln USD		FSSS
13.	Monetary Aggregate M0	bln RUR		CBR
14.	Monetary Aggregate M2	bln RUR		CBR
15.	International Reserves	bln USD		CBR
16.	Foreign Exchange Reserves	bln USD		CBR
17.	RTS index	points		FINAM
18.	Brent price index	USD		FINAM
19.	Inflation	Percent		website
20.	10-year Government Bond Yields	Percent		OECD
21.	90-day Interbank Rates	Percent		OECD
22.	CBR Immediate Rates: Less than 24 Hours	Percent		OECD
23.	Total Share Prices for All Shares	Basis index	2010	OECD
24.	Total Retail Trade	Basis index	2010	OECD
25.	Production of Total Industry	Basis index	2010	OECD
26.	Production in Total Manufacturing	Basis index	2010	OECD
27.	Real Effective Exchange Rate (b. on CPI)	Chain index		IFS
28.	Total Reserves excluding Gold	Spc. Draw. Rights		IFS
29.	Consumer Price Index	Basis index	2010	IFS
30.	Producer Price Index	Basis index	2010	IFS

Table 2: Descriptive statistics

№	Series	Mean	Median	StDev	Min	Max	Skewness	Kurtosis
1	Real Mon. Inc. idx	153.14	172.42	48.33	70.67	212.26	-0.37	-1.43
2	Constr. real vol. idx	327.57	380.18	97.44	152.84	446.06	-0.36	-1.54
3	New house comm.	4455.04	4814.77	1456.48	2374.22	7650.47	0.33	-0.91
4	Real wag. idx	162.46	177.00	63.11	58.64	251.81	-0.20	-1.41
5	Agr .real vol. idx	223.02	213.16	35.96	159.04	337.01	0.60	-0.33
6	Unemployment rate	7.55	7.44	1.93	5.17	13.13	1.03	0.81
7	Decl. need work	1147.06	1045.88	323.88	430.60	1900.43	0.42	-0.31
8	Constr. pi	101.10	100.94	0.77	99.69	103.30	0.94	0.69
9	Transp. pi	101.34	101.02	1.03	99.41	104.34	0.88	0.24
10	Real inv. FC idx	216.24	239.69	73.58	88.49	314.52	-0.23	-1.53
11	Total Exp.	25.24	25.28	13.31	6.27	44.82	0.03	-1.45
12	Total Imp.	15.19	15.48	8.85	3.10	28.60	0.11	-1.46
13	M0	3145.76	3047.67	2344.08	208.52	6855.68	0.27	-1.40
14	M2	12547.76	10777.83	10771.24	591.38	33310.14	0.51	-1.17
15	Int. reserves	284.55	361.85	201.45	10.77	534.11	-0.16	-1.71
16	FOREX reserves	264.95	315.55	187.43	6.60	520.65	-0.17	-1.70
17	RTSi	1010.61	1025.38	606.90	43.04	2135.10	0.01	-1.26
18	BRENT	63.79	64.53	32.57	10.90	113.06	0.13	-1.39
19	Inflation	1.02	0.90	0.61	0.27	4.67	3.00	12.57
20	BondYield 10yrs	15.66	8.37	19.36	6.62	101.26	3.04	8.46
21	IntrBank90	9.74	7.79	4.62	4.30	27.69	1.59	2.49
22	CBR imm. rate	15.83	11.84	12.54	5.17	61.24	2.08	4.02
23	Total share price	69.90	88.19	40.55	2.78	123.16	-0.30	-1.56
24	Total Retail trade	80.22	86.63	29.40	35.44	121.95	-0.14	-1.51
25	IPI	93.17	96.89	15.87	61.18	113.13	-0.48	-1.14
26	MPI	92.82	94.93	19.99	56.78	120.43	-0.24	-1.33
27	Real Eff. EXCH rate	83.27	87.38	18.86	44.63	109.28	-0.43	-1.03
28	Total rsv. w/o gold	1.7e+11	2.3e+11	1.2e+11	5e+09	3.3e+11	-0.22	-1.70
29	CPI	79.25	73.37	36.71	22.53	157.36	0.28	-1.03
30	PPI	80.44	78.98	41.59	15.48	159.96	0.16	-1.24

Chapter 3. Methodology

In my work I will be using a Bayesian Compressed VAR model with conjugate Normal-Inverse Wishart prior. In addition, I adopt modifications to the conjugate prior, notably sum-of-coefficients (SC) prior and initial obervation (IO) prior. The former gives an opportunity to account for possible unit roots in series. The latter reflects prior belief of stohastic trend in the series.

1. The BVAR model

A VAR(p) model in its general form is defined as follows:

$$y_t = \Phi_{const} + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma)$$

Where $y_t = [y_{1t}, y_{2t}, \ldots, y_{mt}]'$ is a vector of variables Φ_{const} is a vector of constants, Φ_s is an $(m \times m)$ coefficient matrix for lag s variables. As we define the Φ matrix in companion form: $\Phi = [\Phi_1, \ldots, \Phi_p, \Phi_{const}]'$ and x_t as $x_t = [y'_{t-1}, \ldots, y'_{t-p}, 1]'$, we can write the VAR model in reduced form:

$$y_t = \Phi' x_t + \varepsilon_t$$

Finally, as we define $Y = [y_1, y_2, \ldots, y_T]'$, $X = [x_1, x_2, \ldots, x_T]'$ and $E = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T]'$, we can rewrite VAR in its most compact form:

$$Y = X\Phi + E$$

This form is the most convenient for compuations using linear algebra.

In Bayesian econometrics, the estimation aims at finding posterior distribution of parameters $p(\overline{\Phi}, \overline{\Sigma}, \overline{\Omega}|Y)$ using maximum likelihood $p(Y|\Phi, \Sigma, \Omega)$ and prior distribution $p(\underline{\Phi}, \underline{\Sigma}, \underline{\Omega}|Y)$ together with the Bayes rule:

$$p(\overline{\Phi}, \overline{\Sigma}, \overline{\Omega}|Y) = \frac{p(\Phi, \Sigma, \Omega) p(Y|\Phi, \Sigma, \Omega)}{p(Y)}$$

Or, as $p(Y)$ is independent from Φ, Σ, Ω:

$$p(\overline{\Phi}, \overline{\Sigma}, \overline{\Omega}|Y) \propto p(\Phi, \Sigma, \Omega) p(Y|\Phi, \Sigma, \Omega)$$

2. Prior Distribution

Generally speaking, choice of prior distribution is a complicated task. Among the most popular ones (Demeshev B., Malakhovskaya O., 2016 [2]) is the Conjugate normal-inverse Wishart prior:

$$\begin{cases} \phi|\Sigma \sim \mathcal{N}(\underline{\phi}, \Sigma \otimes \underline{\Omega}) \\ \Sigma \sim \mathcal{IW}(\underline{S}, \underline{\nu}) \end{cases}$$

This prior has several advantages over other priors (e.g. Minnesota prior), as it does not require Gibbs sampling algorithm and enables Bayesian estimation of Σ. To facilitate the calculation of posterior hyperparameters and enhance the forecasting abilities of the model, I adopt the dummy observations method (see Sims, Zha, 1998).

The Dummy Observations approach is considered as a modification of priors. Its general idea is that one may use prior information to create artificial «dummy» observations and add them as extra rows to the data matrix. The approach is divided into two components: «sums-of-coefficients», initially proposed by (Doan, Litterman and Sims, 1984) and «initial observation» by (Sims, 1993).

Sums-of-coefficients component reflects the prior assumption that the mean of lagged values of a variable (e.g. $\overline{y}_i$) may be a good forecast for this variable. In terms of data, this prior component accounts for possible unit roots in variables. Another prior component, dummy initial observation, introduces only a single artificial observation such that all values of all variables are equal to respective averages. This prior component reflects the idea that the average of lagged values of a variable is a linear combination of others, thus introducing correlation between the variables. Implementation of these priors in the model is found to improve forecasts of economic time series, as they account for unit roots and cointegration.

The hyperparameters I use in the model are the following:

Table 3: Prior hyperparameters

Hyperparameter	Meaning	Prior value
$\underline{\nu}$	Degrees of freedom for Inverse-Wishart distribution	M + 2
$\underline{\Omega}$	Variance of coefficients' lags	10*eye[4](K)
$\underline{S}$	Covariance matrix	eye(M)
$\underline{\Phi}$	Coefficient matrix	$0_{K \times M}$

[4]Identity matrix

3. Random Compression

Compression methods have been successfully used in order to be able to account for a large number of variables. They are quite popular in domains like machine learning or image recognition. Moreover, these methods give the reseacher an opportunity to ignore some problematic phenomena in the data, e.g. multicollinearity. In this work, I will be following (Koop, Korobilis, Pettenuzzo, 2018) approach for data compression.

In their paper, KKP use random projection matrices, which reduce data matrix dimension. For example, let A be an $(M \times K)$ matrix, with M equations and K parameters ($K = kp + 1$). We can define a projection matrix, C $(m \times K)$, where $m \ll K$. Multiplying the two matrices, we obtain $\widetilde{A} = AC'$, an $(M \times m)$ matrix, which is smaller than A and thus, is easier to operate in computations. In other words, we want to decrease the number of parameters in one equation from K to m.

The compression matrix is drawn randomly, without referring to the data (so called «data oblivious» method), making the process of generation computationally inexpensive. A variety of schemes can be used, from generating a matrix from standard normal distribution to more complex ones, like in (Guhaniyogi, Dunson, 2014):

$$\text{C} = \begin{cases} P(C_{ij} = \frac{1}{\sqrt{\phi}}) = \phi^2 \\ P(C_{ij} = 0) = 2(1-\phi)\phi \\ P(C_{ij} = -\frac{1}{\sqrt{\phi}}) = (1-\phi)^2 \end{cases}$$

with m and ϕ as unknown parameters. According to KKP, the parameters should be drawn randomly from:

$$m \sim \mathcal{U}(1, 5log(K))$$
$$\phi \sim \mathcal{U}(0.1, 0.9)$$

It is advised that the columns of the matrix should have unit lengths. To achieve this, the matrix is normalised by dividing each column by the square root of column-wise sum of element squares.

Compression and estimation is done several times and the result of each estimation is viewed as having estimated a separate model. In order to assess the quality of compressions, each compressed model is assigned a weight based on the value of information criterion obtained from the model. As such, good compressions receive higher weights, while bad compressions' weights are lower. This process is called Bayesian Model Averaging (BMA) and can be intepreted as «importance sampling».

It is worth noting, that both coefficient and covariance matrices can be compressed. The latter is especially important in case of large dimension models, as covariance matrix may hold a large number of free parameters for esitmation. This paper, however, focuses only on compression of coefficient matrix.

4. Estimation of a BCVAR model

As compression is added to the matrix of coefficients, the model can be rewritten as:

$$Y^* = \Phi_C(CX^*) + E$$

In order to estimate the model, the following procedure is used. As it was stated in previous chapters, I adopt Dummy Observations method incorporate prior information. The artificial observations are obtained in the following way:

$$X^+ = \begin{pmatrix} (chol(\underline{\Omega}))^{-1} \\ 0_{M\times K} \end{pmatrix}, \quad \text{define } X_1 = (chol(\underline{\Omega}))^{-1}, \quad Y^+ = \begin{pmatrix} (X_1^+)^{-1}((X_1^+)'X_1^+\underline{\Phi}) \\ chol(S) \end{pmatrix}$$

And afterwards these dummy observartions are used to form data matrices:

$$X^* = \begin{pmatrix} X^+ \\ X \end{pmatrix}, \quad Y^* = \begin{pmatrix} Y^+ \\ Y \end{pmatrix}$$

The matrix X^* is afterwards compressed by a random compression matrix, thus giving X_c^*. Furthermore, it is used to obtain posterior hyperparameters analytically, by applyng the following formulae:

$$\begin{gathered} \overline{\nu} = \underline{\nu} + T \\ \overline{\Omega} = V diag(1/s)V' \\ \overline{\Phi}_C = \overline{\Omega}(X_C^* Y^*) \\ \hat{E}^* = Y^* - X_C^* \overline{\Phi}_C \\ \overline{S} = (\hat{E}^*)'\hat{E}^* \end{gathered}$$

Φ matrix is afterwards uncompressed to be used in model averaging:

$$\Phi_U = C'\Phi_C$$

Where V and s are an orthogonal matrix and diagonal vector from Singular-value decomposition of X^* respectively. This process is repeated several times, for each m from 1 to $5log(K)$ and R times for each m. This way, about $Z = 5 \cdot R \cdot log(K)$ models are estimated. For each model, we calculate Bayesian Information Criterion and store it in a vector.

After all models are estimated apply BMA to find the weighted values of hyperparameters. For BMA, I follow KKP and use a vector of obtained Bayesian Information Criterion values to create weights for each model.

$$\Psi = BIC - min(BIC)$$
$$Weights = \frac{e^{-0.5\Psi}}{\sum_{i=1}^{Z} e^{-0.5\Psi}}$$

To obtain the averaged $\overline{\Phi}$ one may either calculate a weigthed sum or use the model with maximum weight. In my analysis, I use the former method. The obtained posterior hyperparameters are used to draw a random sample of parameters. One can use the following scheme:

1. Draw a covariance matrix and a coefficient matrix from:

$$\Sigma^{(i)} \sim \mathcal{IW}(\overline{S}, \overline{\nu})$$
$$\Phi^{(i)} = \overline{\Phi} + chol(\overline{\Omega})V chol(\Sigma^{(i)})'$$

2. Increase i by 1 and repeat.

The obtained sample is further used for forecasting.

5. Forecasting

In my paper, I follow (Carriero *et al* (2013)) simulation approach for forecasting, as it was found to be more efficient than multiple-step approaches. The forecasting procedure is the following: for each element in the drawn sample of Φ we create h step ahead iterative forecasts, thus obtaining an $(N_{MCMC} \times m \cdot h)$ matrix of forecasts. Afterwards, we may use column means or medians as point forecasts.

$$\hat{y}_{t+h} = \frac{1}{N_{MCMC}} \sum_{i=1}^{N_{MCMC}} (\Phi_i)$$

The data is divided into training and test samples. Training sample is taken from January 1999 to December 2014 and test sample is, accordingly, from January 2015 to December 2015, 12 steps total.

6. Comparison of Forecasts

To compare the forecasts, I adopt the RMSFE method, or calculation of ratios between MSFE (Mean Square Forecast Error) from observed model with MSFE from a benchmark model:

$$MSFE_i = \frac{1}{M} \sum_{i=1}^{M} e^2$$

Where e^2 is the squared error of the forecast. As for the benchmark, I follow the approach by KKP, and use AR(1) model, as all the time series in the data are non-stationary.

$$\hat{y}_{t+h} = c + y_{t+h-1} + u_t$$

I shall call this model RWWN, as in (Demeshev, Malakhovskaya (2016)[3]).

Also I would like to mention indicators of special interest, notably: Unemployment, Inflation, Industrial Production Index, 10-year Government Bond Yields and CPI.

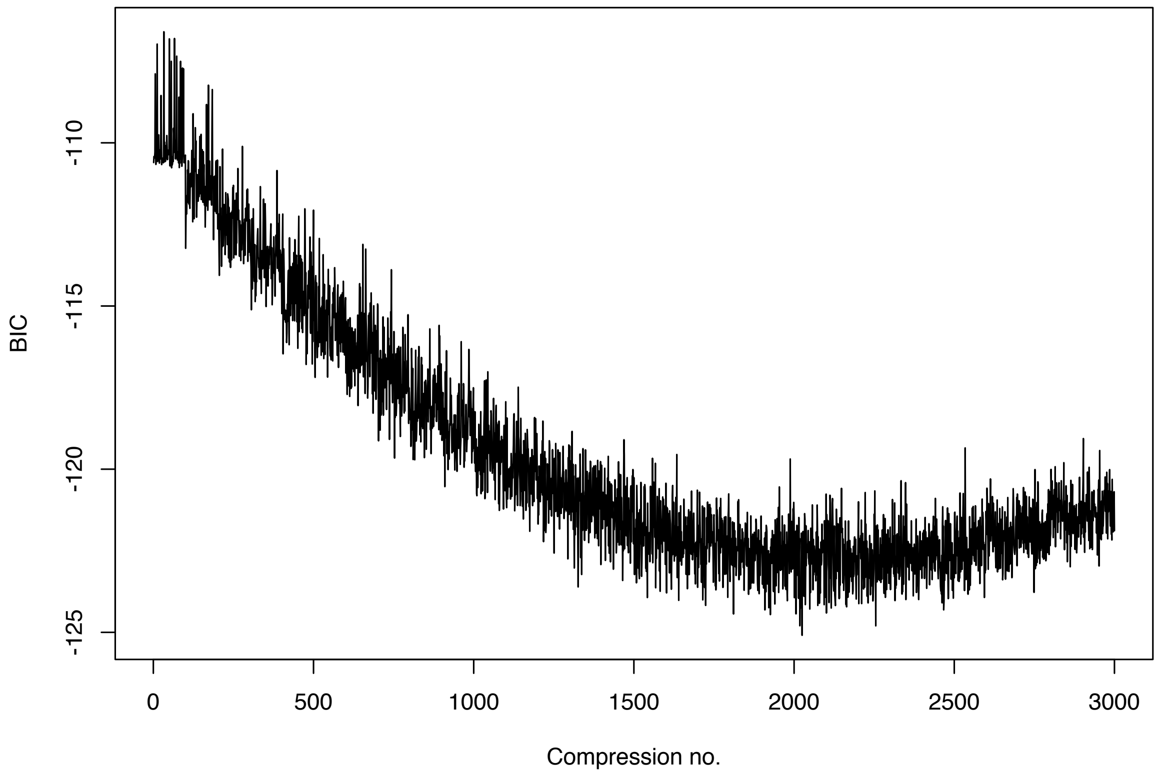
Chapter 4. Results of Estimations

In this chapter, I present the results of estimations and forecasts.

In my dataset, I have 30 variables. Following (Koop, Korobilis, Petenuzzo (2018)) and (Banbura *et al* (2010)), I choose a high number for lags ($p = 13$). In total, there are $K = (30 \times 13 + 1) = 391$ coefficients in one equation and $K \times M = 30 \times 391 = 11730$ coefficients in total. In addition, there are $30 \times (30 + 1)/2 = 465$ elements in covariance matrix, making 12165 parameters overall. As such, I compress the model from K to m parameters, with m ranging from 1 to $5 \times log(K) \approx 30$. For each m, $R = 100$ compressions is done.

The results of model estimation are the following[5]. As for compression quality we may see that best compressions are around 2000-2200 iterations, which correspond to shrinkage to 20-22 parameters.

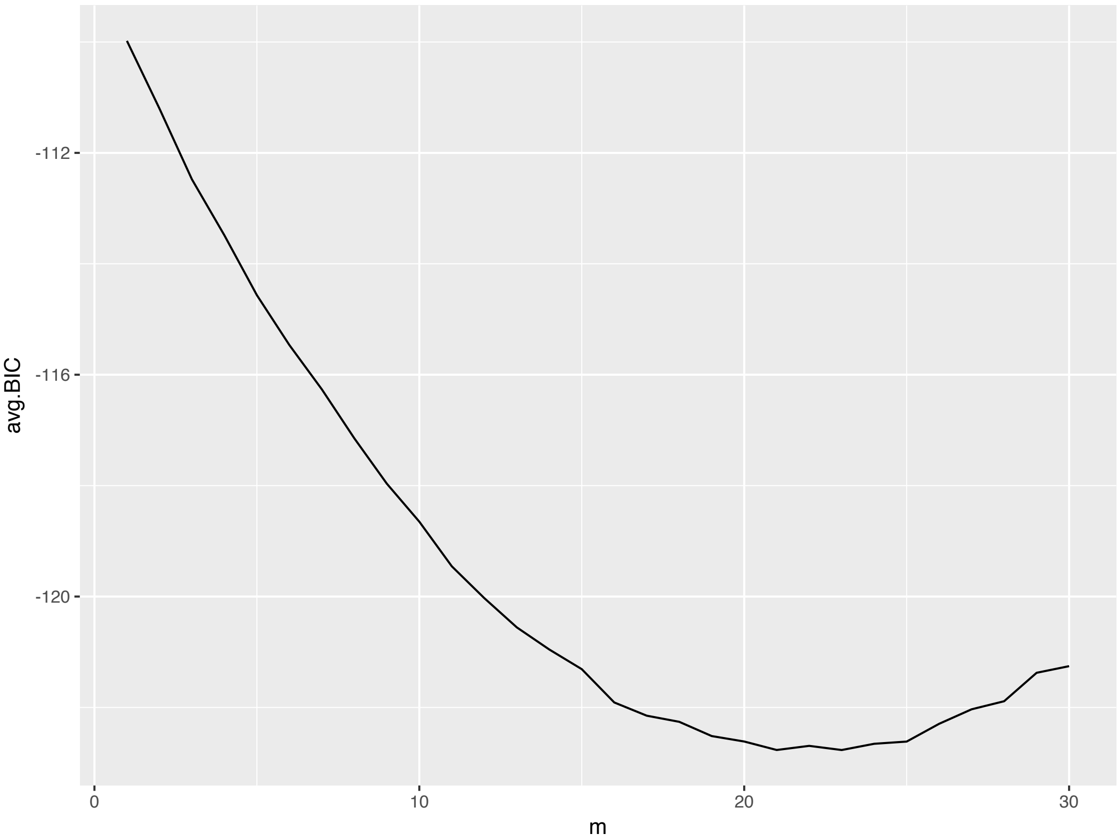
Figure 1: BIC values for Compression number



[5]For posterior parameter coefficients see Appendix A

Indeed, as we take average BIC for every m we may notice that best compressions are around $m = 21$:

Figure 2: BIC values for Compression number



Chapter 5. Forecast Comparison

To compare the forecasts I implement the Ratio of Mean Squared Forecast Error (RMSFE) method. This method is widely used in various literature concerning comparison of Bayesian model forecasts with other models.

$$RMSFE_i = \frac{MSFE(\hat{y}_{BCVAR,i})}{MSFE(\hat{y}_{bcmk,i})}$$

As benchmark I use AR(1) model (also called RWWN, as stated above). It is a commonly used benchmark model (Koop, Korobolis, Petenuzzo (2018)), (Demeshev, Malakhovskaya (2016)). I calculate RMSFE for several forecasting horizons: $h = 1, 4, 6, 12$ and for all available variables. For results, see Table 4.

The results suggest that BCVAR model generally beats the benchmark model. The model rarely fails on short term periods. Some indicators are generally forecasted better by the benchmark model, notably Construction Works Price Index, Inflation, Total Share Prices and Transportation Tariffs. The variables of special interest, notably Unemployment, Industrial Production Index, 10-year Government Bond Yields and CPI show great results on medium to long horizons.

Overall we can state that BCVAR model forecasts well all available macroeconomic indicator types: price variables, real variables, monetary variables, rates and indices.

Table 4: RMSFE values

	h = 1	h = 4	h = 6	h = 12
Agricultural Real Production Vol. Index	0.19174	0.45991	0.49588	0.20617
10-year Government Bond Yields	4.95990	0.50690	0.42307	0.27589
BRENT	0.57407	0.15174	0.08978	0.03086
CBR Immediate Rates: Less than 24 Hours	1.30949	0.23271	0.19896	2.14713
Construction Works Price Index	2.34850	2.16925	1.43584	0.22742
Real Construction Volume Index	0.08551	0.08514	0.08207	0.46781
Consumer Price Index	2.61490	1.03731	0.80907	0.69565
Declared Need In Workers	0.33118	0.22660	0.19059	0.09176
Foreign Exchange Reserves	0.55408	0.15275	0.09133	0.02970
Inflation	119.25434	12.30718	1.88049	0.48622
International Reserves	0.48173	0.12631	0.07505	0.02604
90-day Interbank Rates	0.33287	0.14252	0.08996	0.10705
IPI	1.88330	0.34630	0.23585	0.29838
M0	0.00000	0.00314	0.00203	0.01801
M2	1.87668	0.15245	0.05030	0.24970
MPI	0.37474	0.05043	0.02814	0.04829
New Houses Commissioning	1.44603	0.06168	0.02149	0.00808
PPI	1.90413	0.60131	0.53296	0.55223
Real Effective Exchange Rate	1.30351	0.22320	0.11425	0.02676
Real Monetary Income	0.00213	0.00815	0.00615	0.38831
Real Investments In FC Index	0.21919	0.08612	0.06838	0.31814
Real Wages Index	0.06480	0.01770	0.01447	0.00305
RTSi	0.03766	0.01093	0.02957	0.02963
Total Export	0.15590	0.04559	0.02657	0.01649
Total Import	0.50177	0.13093	0.07899	0.02510
Total Reserves excluding Gold	0.48939	0.14040	0.08089	0.02791
Total Retail Trade	0.00863	0.00779	0.00571	0.00478
Total Share Prices for All Shares	6.54336	1.69535	1.16986	0.46226
Transportation Tariffs Index	2.13960	1.39617	1.40312	1.90695
Unemployment Rate	0.02038	0.06264	0.03931	0.00795

Conclusion

In conclusion, I would like to restate the key moments.

In this paper, I have estimated the effectiveness of Bayesian Compressed VAR model in forecasting Russian macroeconomic indicators. As for specification, I used BCVAR model by (Koop, Korobilis, Petenuzzo, (2018)) with normal-inverse Wishart prior with Dummy Observations modification as in (Sims, Zha (1998)). The obtained results indicate, that BCVAR model can successfully beat the benchmark model in point forecasting. As such, I consider the goal of this paper completed.

There is still ways to improve this paper. Firstly, I only focused on one alternative model, although one may read the supporting literature and compare the values of RMSFE without the need to make extra estimations. Secondly, this paper only covers point forecasting, leaving density and interval forecating out of the scope. Finally, the Bayesian VAR models do not only allow precise forecasting, but also research of economic shock response. Basing on the results of this paper and relevant literature I can strongly recommend the readers to become more acquainted with Bayesian models in statistics and econometrics.

References

[1] Banbura A., Giannone D., Reichlin L. (2010) Large Bayesian Vector Auto Regressions, *Journal of Applied Econometrics*, v. 25, pp. 71-90.

[2] Blake A., Mumtaz H. (2012) Applied Bayesian econometrics for central bankers. Centre for Central Banking Studies, Bank of England.

[3] Carriero A., Clark T. E., Marcellino M. (2015) Bayesian VARs: Specification Choices and Forecast Accuracy, *Journal of Applied Econometrics*, v. 30, pp. 46-73.

[4] Demeshev B., Malakhovskaya O. (2016) BVAR mapping. *Applied Econometrics*, v. 43, pp. 118–141.

[5] Demeshev B., Malakhovskaya O. (2016) Macroeconomic Forecasting with a Litterman's BVAR Model. *HSE Economic Journal*, vol. 20, no 4, pp. 691–710.

[6] Deryugina E., Ponomarenko A. (2015) A large Bayesian vector autoregression model for Russia *Bank of Russia Working Paper Series* wps1, Bank of Russia. http://www.cbr.ru/Content/Document/File/16690/wps_1_e.pdf

[7] Doan T., Litterman R. B. and Sims C. A. (1984) Forecasting and Conditional Projection Using Realistic Prior Distributions, *Econometric Reviews*, Vol. 3, No. 1 pp. 1-100.

[8] Guhaniyogi R., Dunson D. B. (2014) Bayesian Compressed Regression. *Journal of the American Statistical Association*,
DOI: 10.1080/01621459.2014.969425

[9] Koop G., Korobilis D. and Pettenuzzo D. (2018) Bayesian Compressed Vector Autoregressions, Journal of Econometrics.

[10] Pestova A., Mamonov M. (2016) Estimating the Influence of Different Shocks on Macroeconomic Indicators and Developing Conditional Forecasts on the Basis of BVAR Model for the Russian Economy. *Ekonomicheskaya Politika*, vol. 11, no. 4, pp. 56-92

[11] Sims C. (1980) Macroeconomics and Reality, *Econometrica*, vol 48., no. 1., pp. 1–48.

[12] Sims C. (1993) A Nine-Variable Probabilistic Macroeconomic Forecasting Model, in J.H. Stock and M.W. Watson (eds.), *Business Cycles, Indicators and Forecasting*, University of Chicago Press for the NBER, 179-204.

[13] Sims C., Zha T. (1998) Bayesian Methods for Dynamic Multivariate Models, *International Economic Review*, vol. 39, issue 4, 949-68

[14] Shevelev A. (2017) Bayesian Approach to Evaluate the Impact of External Shocks on Russian Macroeconomic Indicators. *Mir Economiki i upravlenia*, vol. 17, no. 1, pp. 27-40

Additional Reading

[1] Conrad S., Ryan C. (2016) Armadillo: a template-based C++ library for linear algebra. *Journal of Open Source Software*, Vol. 1, pp. 26.

[2] Eddelbuettel D., François R. (2011) Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software*, Volume 40, Issue 8.

[3] Eddelbuettel D. (2013) Seamless R and C++ Integration with Rcpp. *Springer Use R!* DOI 10.1007/978-1-4614-6868-4

[4] Eddelbuettel D., Sanderson C. (2014) RcppArmadillo: accelerating R with high-performance C++ linear algebra. *Computational Statistics & Data Analysis*, 71, pp. 1054-1063.

[5] Olmsted Jonathan (2014) The Advanced Statistical Programming Camp (Lecture series) Retrieved from: https://q-aps.princeton.edu/book/advanced-statistical-programming-camp-2014

Appendix A

Table 5: KPSS.test results

	series	KPSS_stat	P-value
1	Real Mon. Inc. idx	5.08	0.01
2	Constr. real vol. idx	4.74	0.01
3	New house comm.	4.76	0.01
4	Real wag. idx	5.13	0.01
5	Agr .real vol. idx	4.56	0.01
6	Unemployment rate	3.84	0.01
7	Decl. need work	3.98	0.01
8	Constr. pi	2.81	0.01
9	Transp. pi	1.69	0.01
10	Real inv. FC idx	4.87	0.01
11	Total Exp.	4.49	0.01
12	Total Imp.	4.42	0.01
13	M0	5.14	0.01
14	M2	5.01	0.01
15	Int. reservess	4.55	0.01
16	FOREX reserves	4.39	0.01
17	RTSi	3.01	0.01
18	BRENT	3.93	0.01
19	Inflation	2.25	0.01
20	BondYield 10yrs	1.98	0.01
21	IntrBank90	1.07	0.01
22	CBR imm. rate	3.25	0.01
23	Total share price	4.25	0.01
24	Total Retail trade	5.12	0.01
25	IPI	4.74	0.01
26	MPI	4.80	0.01
27	Real Eff. EXCH rate	4.33	0.01
28	Total rsv. w/o gold	4.53	0.01
29	CPI	5.13	0.01
30	PPI	5.17	0.01

Table 6: Estimated coefficients for p = 1

	eq.Unempl_rate	eq.Inflation	eq.BondY10	eq.IPI	eq.CPI
Real_inc_idx, p = 1	-0.00	-0.00	-0.00	0.01	0.01
Constr_real_vol_idx, p = 1	-0.00	-0.00	-0.01	0.01	0.00
New_hous_comm, p = 1	-0.01	0.03	-0.01	0.02	0.01
Real_wag_idx, p = 1	-0.01	-0.00	-0.00	0.01	0.01
Agr_real_vol_idx, p = 1	-0.01	-0.03	-0.01	-0.00	0.02
Unempl_rate, p = 1	0.05	0.01	0.01	-0.01	0.00
DecNeed_work, p = 1	-0.03	0.04	0.01	0.00	-0.01
Con_pi, p = 1	-0.02	0.05	0.02	0.01	-0.01
Trans_pi, p = 1	-0.01	0.03	-0.00	-0.01	-0.00
Real_inv_FC_idx, p = 1	-0.01	-0.01	-0.00	0.01	0.01
T_EX, p = 1	-0.01	-0.01	0.00	0.00	-0.00
T_IM, p = 1	-0.01	-0.02	0.00	0.00	-0.00
M0, p = 1	-0.01	-0.01	-0.00	0.01	0.02
M2, p = 1	-0.01	-0.01	0.00	0.01	0.01
INT_res, p = 1	-0.01	-0.01	-0.00	0.00	-0.00
FEX_res, p = 1	-0.01	-0.01	-0.00	0.00	-0.00
RTSi, p = 1	-0.03	-0.01	0.01	0.02	-0.01
BRENT, p = 1	-0.02	-0.02	0.01	0.01	-0.00
Inflation, p = 1	0.01	0.13	0.00	0.00	-0.00
BondY10, p = 1	0.00	0.01	0.04	-0.00	-0.01
IntrBank90, p = 1	0.02	0.03	0.01	-0.02	-0.01
CBRimm, p = 1	0.01	0.02	0.03	-0.01	-0.01
T_shareP, p = 1	-0.02	-0.01	-0.00	0.03	0.01
T_Ret, p = 1	-0.01	0.00	-0.01	0.01	0.01
IPI, p = 1	-0.01	0.00	-0.01	0.02	0.01
MPI, p = 1	-0.02	-0.00	-0.00	0.03	0.01
R_effEX, p = 1	-0.01	-0.03	-0.00	0.01	0.00
T_res_woG, p = 1	-0.01	-0.01	-0.00	0.00	-0.00
CPI, p = 1	-0.00	-0.01	-0.01	0.01	0.02
PPI, p = 1	-0.01	-0.01	-0.01	0.01	0.03

Appendix B

Figure 3: Data graphs: series 1-10

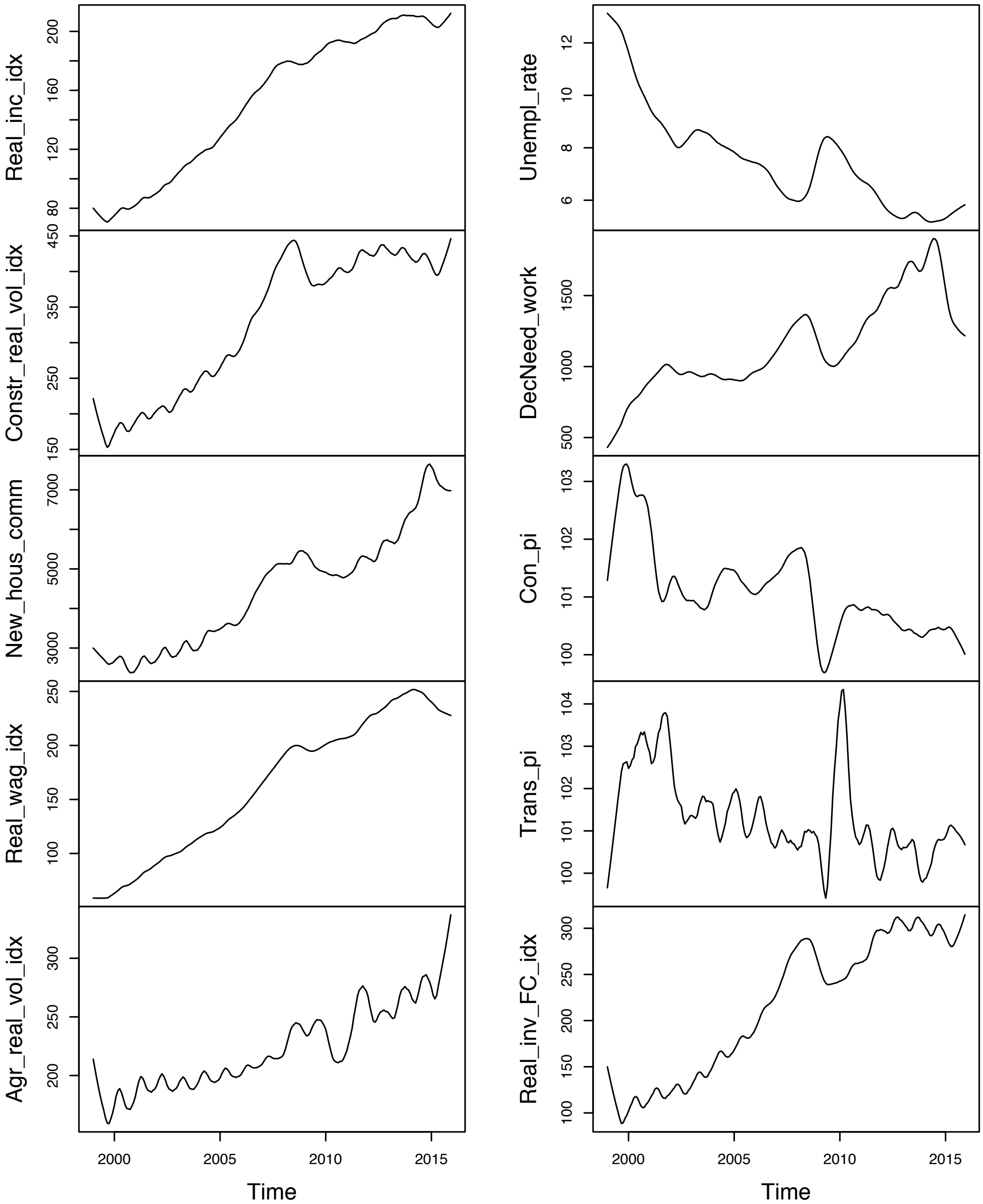


Figure 4: Data graphs: series 11-20

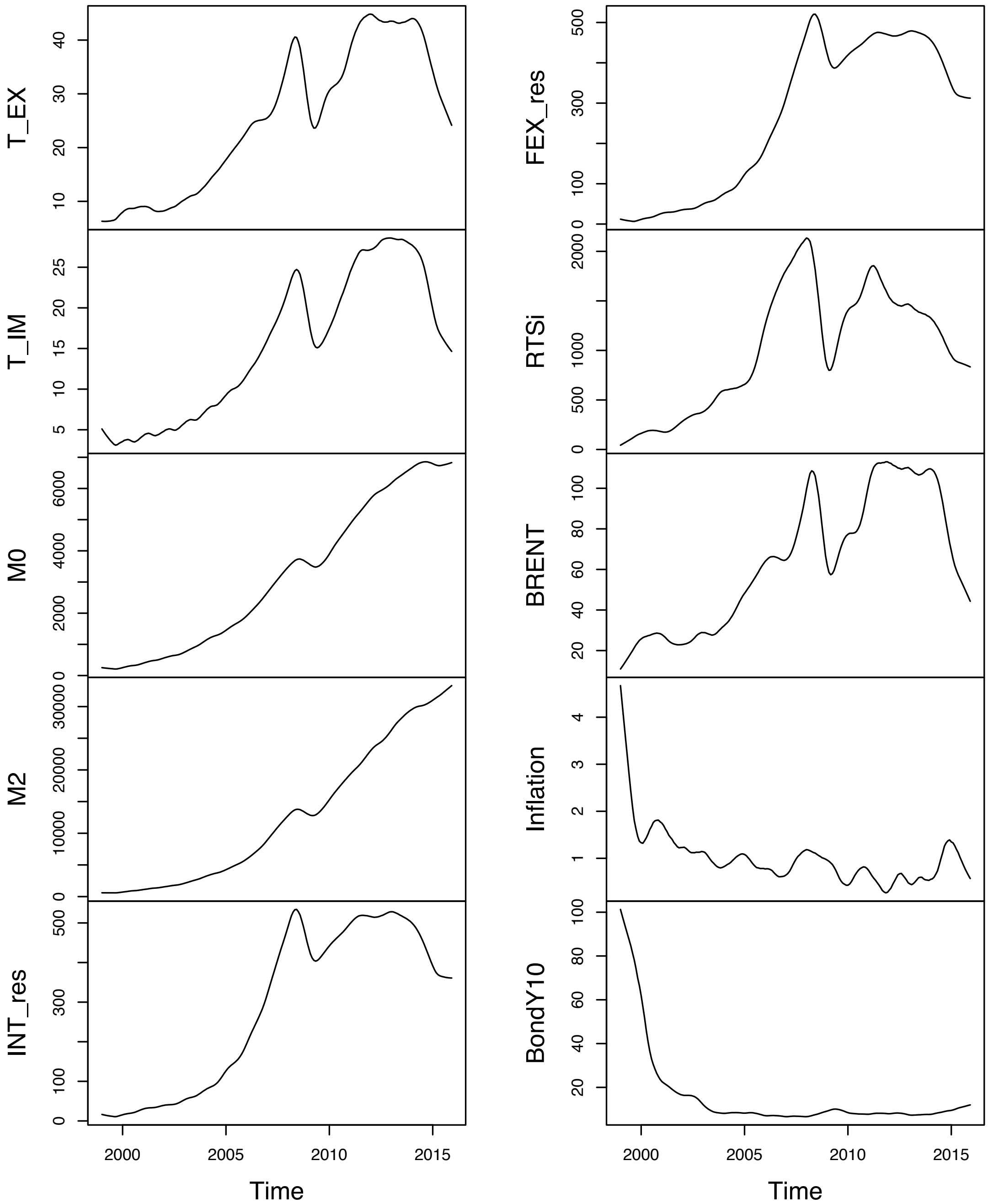


Figure 5: Data graphs: series 21-30

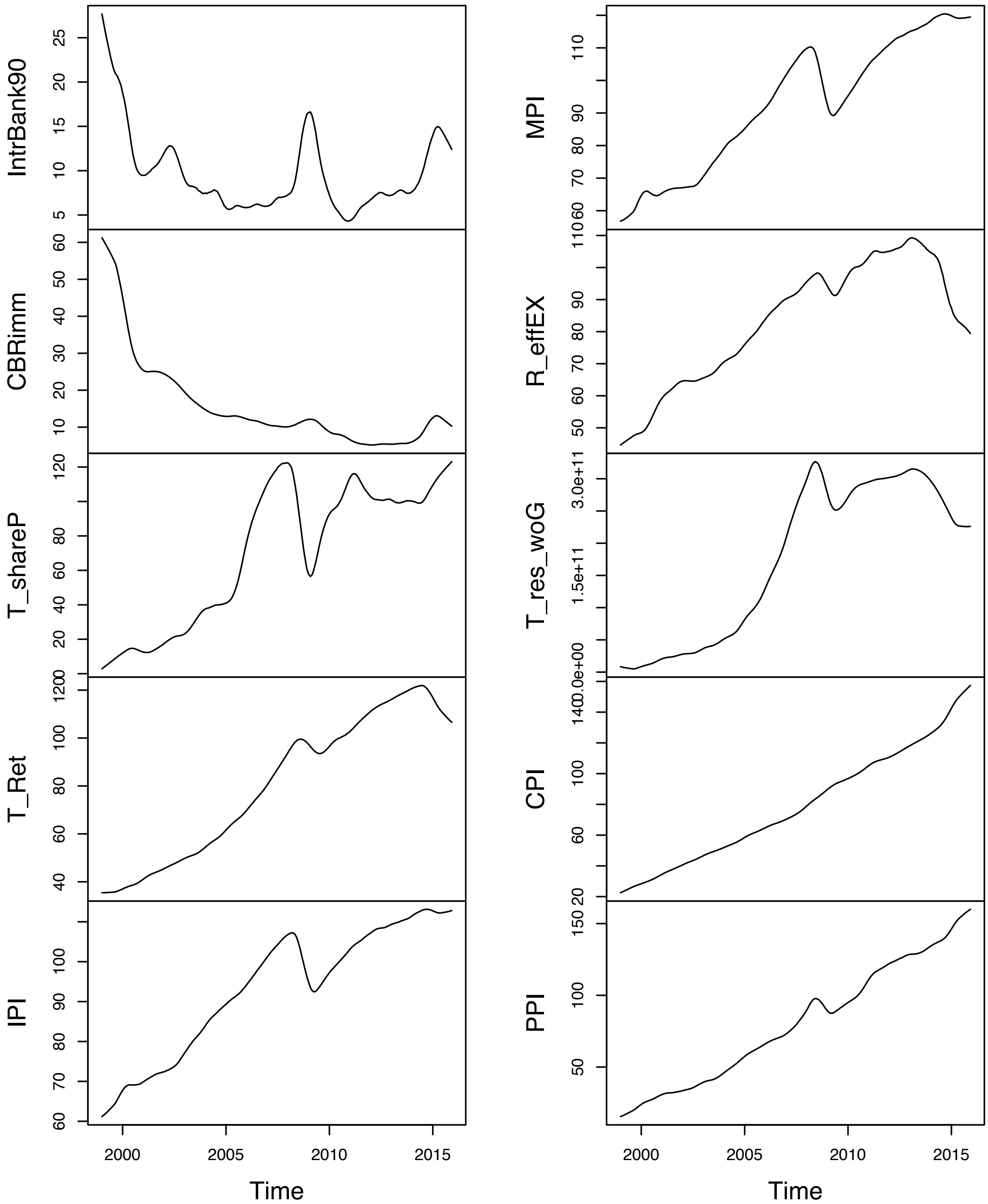


Figure 6: Forecasts of Unemployment rate

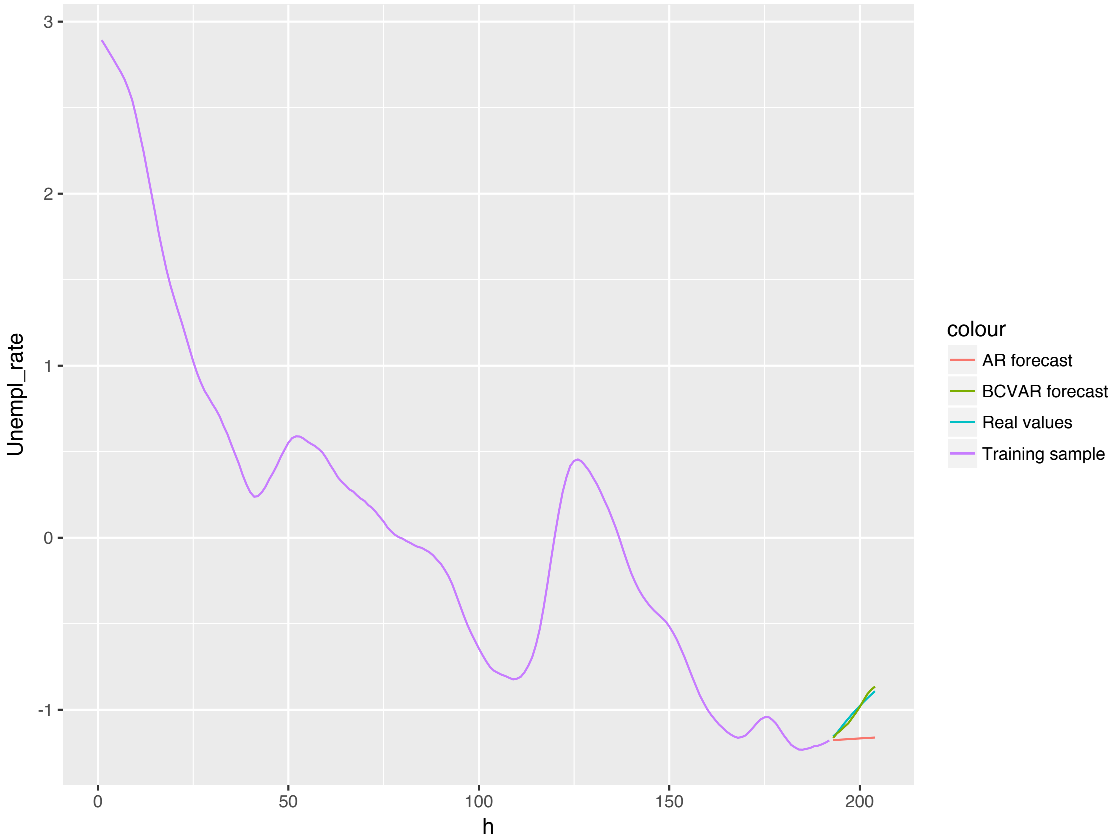


Figure 7: Forecasts of BRENT oil price

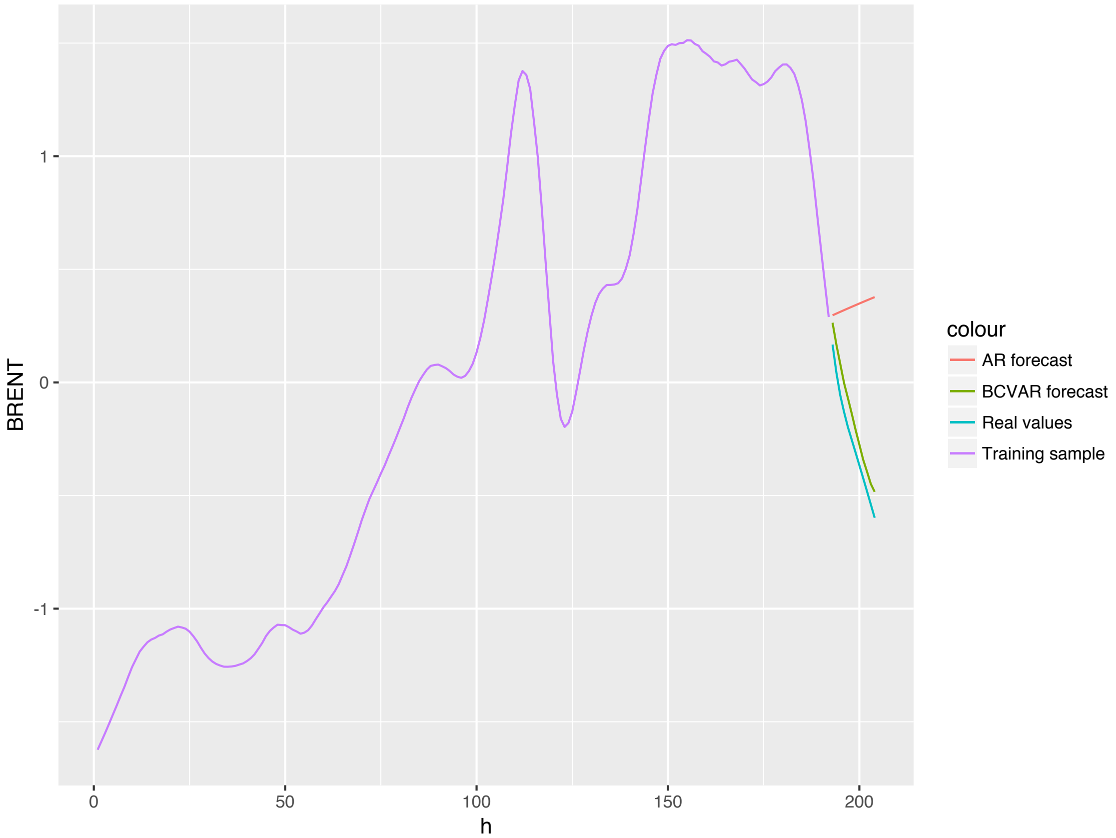


Figure 8: Forecasts of 10-year Bond Yields

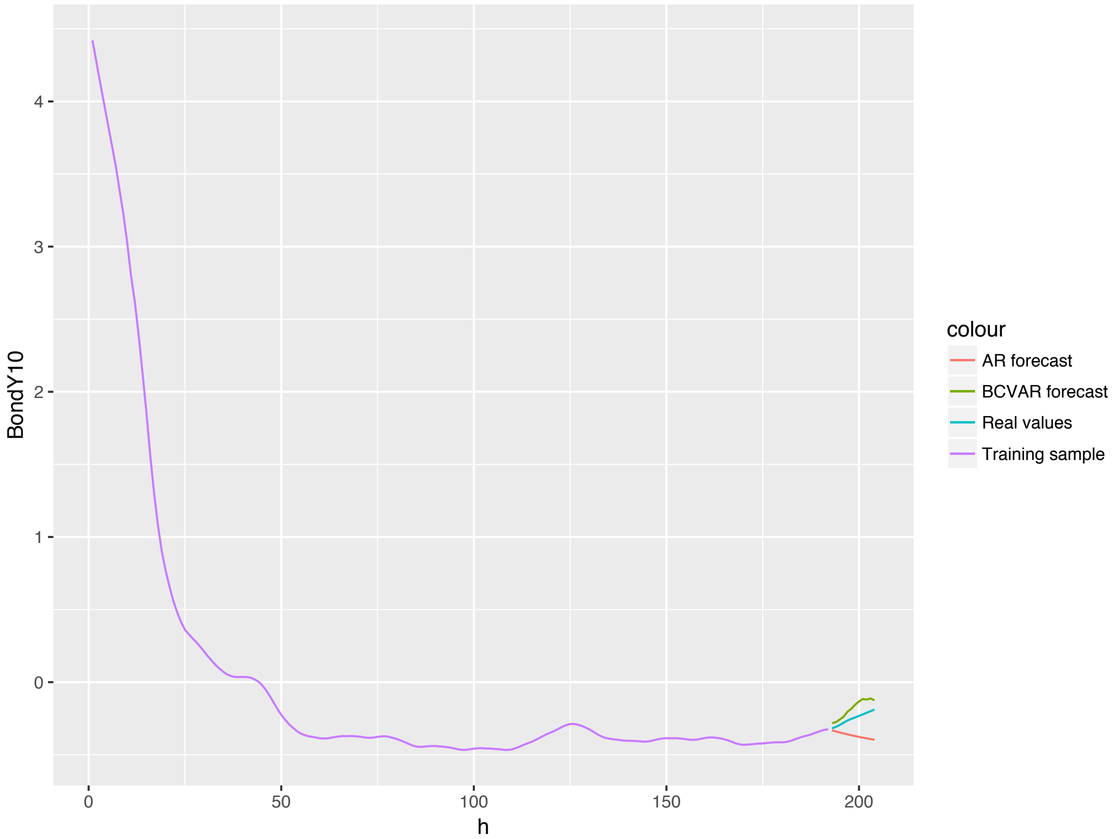


Figure 9: Forecasts of Inflation

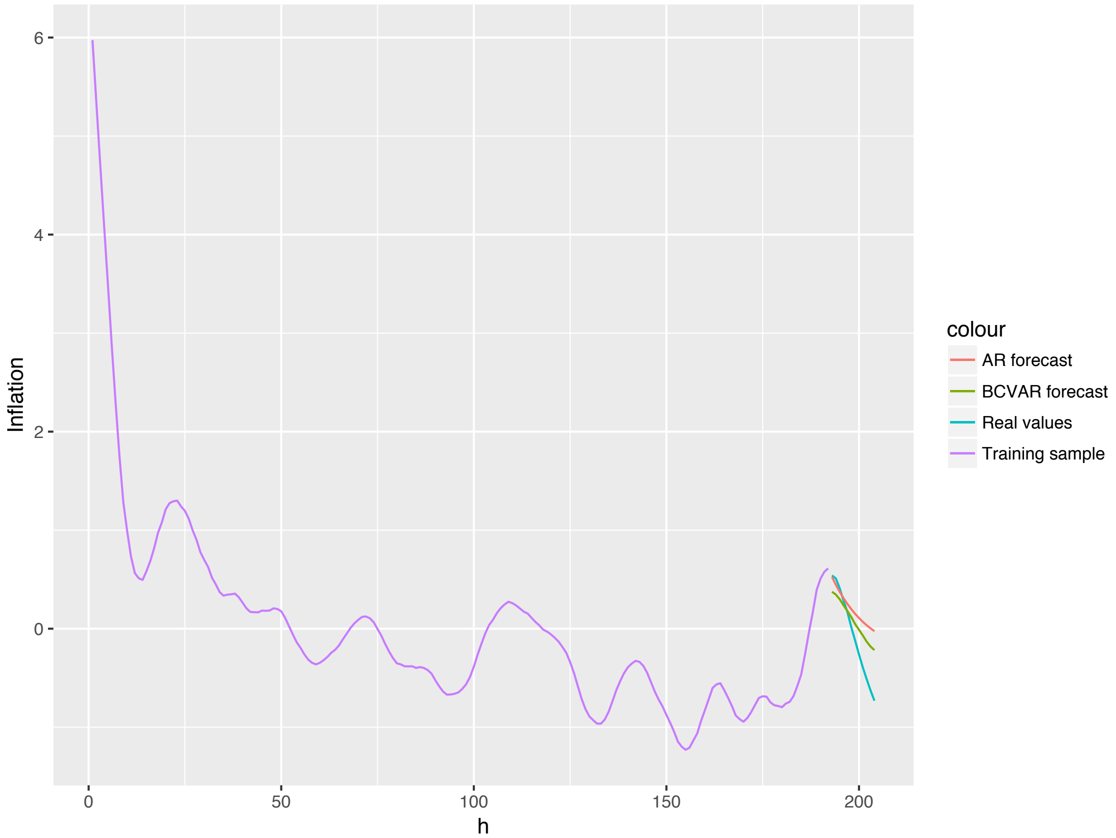


Figure 10: Forecasts of CPI

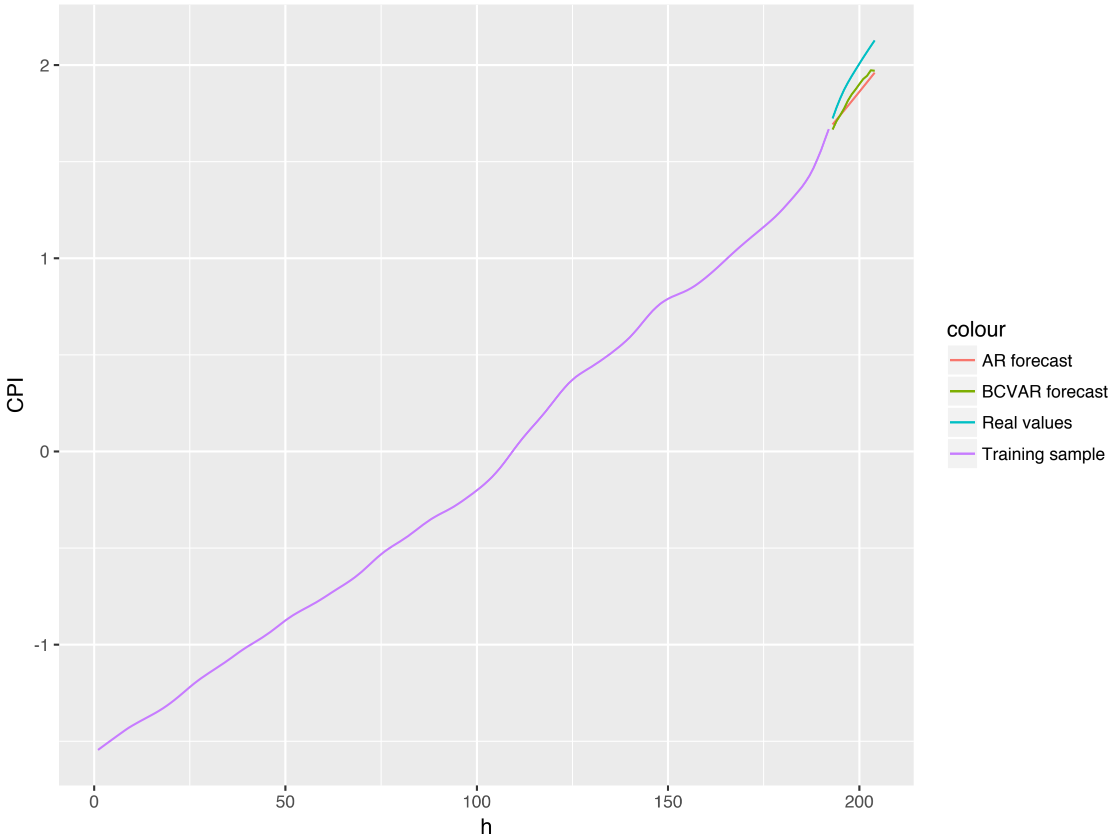


Figure 11: Forecasts of IPI