National Bank of the Republic of Macedonia

Working Paper No. 6/2017

Forecasting Macedonian Inflation: Evaluation of different models for short-term forecasting[1]

Magdalena Petrovska, Gani Ramadani, Nikola Naumovski, Biljana Jovanovic[2]

December 2016

Abstract

The primary goal of this paper is to describe several models that are currently used at the National Bank of the Republic of Macedonia (NBRM) for short-term forecasting of inflation - Autoregressive integrated moving average models (aggregated and disaggregated approach), three equation structural model and a dynamic factor model. Additionally, we evaluate models' out-of-sample forecasting performance for the period 2012 q3 to 2016 q2 by using a number of forecast evaluation criteria such as the Root Mean Squared Error, the Mean Absolute Error, the Mean Absolute Percentage Error and the Theil's U Statistics. Additionally, we constructed several composite forecasts in order to test whether a combination forecast is superior to individual models' forecasts. Our results point to three important conclusions. First, the forecasting accuracy of the models is highest when they are used for forecasting one quarter ahead i.e. the errors increase as the forecasting horizon increases. Second, the disaggregated ARIMA model has the smallest forecasting errors. Third, majority of the forecast evaluation criteria suggest that composite forecasts are superior in comparison to the individual models.

JEL classifications: *C52, C53, E37*

Keywords: *Inflation, forecasting, forecast evaluation, composite forecast*

---

[1] The views expressed in this paper are those of the authors and do not necessarily represent the views of the National Bank of the Republic of Macedonia.

[2] Monetary Policy and Research Department, National Bank of the Republic of Macedonia. Corresponding author: RamadaniG@nbrm.mk.

Contents

Introduction

The primary goal of the National Bank of the Republic of Macedonia (NBRM) is maintaining price stability. Having in mind that monetary policy decisions affect real economy with a lag, precise forecast of the future inflation developments is of vital importance. The NBRM uses a set of different methods for forecasting inflation. The medium term inflation forecast is produced within the Macedonian Policy Analysis Model (MAKPAM), which is a consistent framework of the transmission mechanisms in the Macedonian economy. Besides the MAKPAM model, the NBRM uses suit of models starting from simple time series model to small structural models to produce short-term inflation forecast. In this paper, we discuss models for short term inflation forecasting used at the NBRM, present forecasts from different models and evaluate their forecasting performances.

Four models for inflation forecasting are discussed in the paper. The first one is small structural model consisting of three behavioral equations for the main Consumer Price Index (CPI) subgroups – food prices, energy prices and core inflation. The second one is based on the ARIMA modeling framework. A separate ARIMA model is estimated for all subcomponents of the CPI index and then, by using the individual subindicies' weights, a forecast of the overall CPI inflation is obtained. The third model is ARIMA model for the total CPI (aggregated approach) and the fourth model is a dynamic factor model, which uses large set of information to produce the inflation forecast. As an alternative to using one forecasting model, we present several composite forecasts by using different weighting schemes. The models' forecasting performances are evaluated using several criteria for comparing the models' forecasting performance – the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE) and the Theil's U Statistics (Theil).

The paper is organized as follows. Overview of methods for short term forecasting is provided in the next section. In section 3, we discuss the main characteristics of the three models used for short term inflation forecasting at the NBRM. The performance evaluation exercise, between model comparison and the combination forecast are given in section 4. Finally, section 5 concludes.

Overview of forecasting techniques

The literature on econometric forecasting is vast and numerous. One general classification of methods of forecasting is the one of Clements and Hendry (Clements & Hendry, 2004). They classify methods into seven groups: guessing, "rules of thumb" and "informal models"; expert judgment; extrapolation; leading indicators; surveys; time-series models and econometric systems. All these approaches have their advantages and weaknesses, and therefore, a reliable forecasting system will usually combine some/or all of them to produce a meaningful forecast. Given the topic of the research, this section will mainly deal with the core characteristics of time-series econometric methods and macroeconomic models. In brief, guessing and related methods rely solely on available information and predictability; leading indicators usually have good performance as long as there is strong connection between the indicator that is leading and the indicator that is led; surveys of consumers and businesses can be informative about future events, but rely on plans being realized. In summary, all these methods cannot be seen as a complete forecasting system and therefore, are rarely used as a sole forecasting technique. They are usually used for providing complementary information to the "core" forecast produced by time-series models or macroeconomic models.

Time-series models describe historical patterns of data and are often considered to be the workhorse of the forecasting industry. Work in this field has its traits back in the second half of the last century. One of the oldest and most influential time series forecasting methodology is the ***ARIMA (Auto-Regressive-Integrated-Moving-Average) model*** developed in the 1970s by Box and Jenkins (Box & Jenkins, 1976). ARIMA forecasting is relatively simple procedure based on the idea that any stationary stochastic process can be approximated well by an autoregressive moving-average (ARMA) process. Past observations contain information about the future developments and therefore, each time series might be expressed as a sum of past values – the autoregressive (AR) part and past error terms – the moving average (MA) part. The series should be stationary; if not stationarity is achieved by differencing the time series. Nowadays, many extensions exist of the simple ARIMA model such as the conditional heteroscedasticity models (ARCH, GARCH, etc.), threshold AR models (TAR, STAR, SETAR, etc.) and AR fractional integrated MA (ARFIMA) models. When it comes to forecasting inflation by using ARIMA method two approaches can be met in the literature - aggregated and disaggregated. Aggregated approach refers to using one ARIMA model for forecasting total CPI index, whereas disaggregated means using different ARIMA models for forecasting individual CPI subindicies and then aggregating these forecasts into one by using the individual subindicies' weights. The latter approach becomes especially popular nowadays. For example, Huwiler and Kaufmann (Huwiler & Kaufmann, 2013) used disaggregated ARIMA model for forecasting Swiss inflation.

Vector AR and ARMA (VAR; VARMA) models represent multivariate extensions of the univariate ARIMA models. Detailed overview of features of this class of models can be found in Lütkepohl (Lütkepohl, 1991) and Sims (Sims, 1980). In practice, VAR is simpler for estimation and therefore it is usually preferred over VARMA models. VAR models can be viewed as a generalization of the univariate AR models in a sense that includes not one, but a vector of time series. Moreover, all variables are treated as endogenous meaning that all of them are expressed as functions of the lagged values of the other variables in the system. VAR models can be used for forecasting and for economic analysis. Impulse response analysis or forecast error variance decompositions are typically used for disentangling the relations between variables in a VAR model. Besides higher flexibility, VAR models have one more additional strength i.e. unlike purely statistical ARIMA models, special extensions of VAR models, such as the structural VAR (SVAR), Bayesian VAR (BVAR) and Vector Error Correction Models (VECM) allow testing different economic theories. In the context of inflation, forecasting examples for the use of VAR model can be found in Lack (Lack, 2006).

Another class of time series models used for forecasting purposes are the ***Unobserved Components (UC) models***. Basically, these models decompose the time series into several unobserved components such as: a low-frequency stochastic trend; a periodic cycle; a seasonal component and an irregular component, normally considered to be a white noise. UC models are quite flexible framework in a sense that models can be univariate and multivariate; purely statistical or theoretical; stationary or augmented, to deal with non-stationary variables. The statistical representation of the UC models is the state-space form and the models are estimated by using the recursive Kalman filter algorithm. A detailed overview of the methodology of UC models can be found in Harvey (Harvey, 1989). A spectrum of UC models can be met in the empirical literature on inflation forecasting. The simplest one is purely statistical model that decomposes the CPI index into two unobserved components – trend and cycle (Clark, 1987). From the theoretical UC models, one of the most popular is the time varying Phillips curve model. Akdogan et al. (Akdogan, et al., 2012) estimated time varying Phillips curve model to forecast Turkish inflation in the short run. The specific of this model is that the parameters are treated as unobserved, time-varying components that evolve as random walk processes.

All the above-mentioned methods use limited number of time series for forecasting. The focus of the time series forecasting literature in the past decade has shifted to methods that exploit many predictors. As discussed by Stock and Watson (Stock & Watson, 2006) examples of such methods are ***forecast combination (FC), factor models (FM) and the Bayesian model averaging (BMA)***. ***FC method*** combines multiple individual model-based forecasts to produce a single, pooled forecast. Especially important issue in this method is the choice of the weights (Timmermann, 2006). Several options are available starting from equal weights to model based time varying parameter weights. The basic idea behind ***FM*** is that each economic variable have one common component that is driven by unobserved common factors and an idiosyncratic component, which is specific for each variable in the dataset. Once the common factors are estimated, they can be used to forecast variable of interest. ***BMA method*** can be viewed as a Bayesian approach to combination forecast. In this method, the weights are computed as formal posterior probabilities that the models are correct.

Time-series models, as emphasized previously, are mainly used for short-term forecasting of the macroeconomic variables. A reliable medium/long term economic forecast is usually conducted by using **macroeconomic model** that describes the linkages between different economic sectors. Macroeconomic models started to develop in the period around the Second World War. They were estimated using econometric techniques and the Keynesian theory. Tinbergen model and Cowles Commission Institute model were among the first large-scale macroeconomic models. The latest generation of macroeconomic models today are the Dynamic Stochastic General Equilibrium (DSGE) models, which, by structure, are New Keynesian models, with microeconomic foundation and a well-defined equilibrium. Recently, special attention has been devoted to the development and usage of Agent Based Models (ABM) for forecasting purposes and policy analysis. ABM models systematically model the individual behavior of all the agents in the system, and the system behavior emerges from all individual components.

Models for short term forecasting of Macedonian inflation

Small structural model

The Small structural model (SSM) is small, estimated model that defines the overall inflation as a sum of three sub-components – energy inflation ($cpi_t^{petroleum}$), food inflation (cpi_t^{food}) and core inflation (cpi_t^{core}), where the core inflation is defined as total inflation excluding the energy and the food component. The three subcomponents are modeled by separate behavioral equations.

$$\Delta cpi_t^{petroleum} = \mathrm{f}(\Delta cpi_{t-1}^{petroleum}, \Delta oil\_index_t \ , \Delta oil\_index_{t-1}) \tag{1}$$

$$\Delta cpi_t^{food} = \mathrm{f}(\Delta cpi_{t-2}^{food}, \Delta wm\_index_{t-6}, \Delta cpi_{t-3}^{petroleum}) \tag{2}$$

$$\Delta cpi_t^{core} = \mathrm{f}(\Delta cpi_{t-1}^{core}, \Delta cpi_{t-4}^{foreign}, \Delta cpi_{t-3}^{petroleum}) \tag{3}$$

where oil_index_t stands for world oil prices, wm_index is an index of wheat and maize prices and $cpi^{foreign}$ is foreign effective CPI index. Energy inflation is included as an additional explanatory variable in the equations describing food inflation and core inflation, with a suitable lag length, in order to capture the second round effects of developments in energy prices on food and core component. All variables are expressed in changes to achieve stationarity.

The total CPI inflation is obtained as a weighted sum of the sub-components, where the weights[3] are equal to the share of these sub-components in the total CPI index.

$$cpi_t^{total} = 0.37 * cpi_t^{food} + 0.04 * cpi_t^{petroleum} + 0.48 * cpi_t^{core} + 0.08 * cpi_t^{electricity} + 0.03 * cpi_t^{heating} \tag{4}$$

where $cpi_t^{electricity}$ and $cpi_t^{heating}$ are the domestic prices of electricity and heating.

In the forecasting horizon, we use external forecast for the explanatory variables. The prices of electricity and heating are mainly regulated, and the usual assumption is that they will remain unchanged, unless there is preannounced correction that will take place in the forecasting period.

[3] Consumer Price Index by COICOP classification and Retail Price Index – January 2017, News release of the State Statistical Office, No.4.1.17.10.

Dynamic Factor Model

Factor models are among the most popular methods for nowcasting and short-term forecasting. The basic idea behind this approach is that the variable of interest is a function of several unobservable factors. In other words, the covariance between large number of economic time series with their leads and lags can be represented by few unobserved factors. These factors, as shown by Stock and Watson (Stock & Watson, 2002), can be estimated by principle components.

By employing Principal Component Analysis, we estimate six principal components from a set of 73 economic and financial variables (Table 1) which are of importance for price dynamics in Macedonia. Data used is in monthly frequency, filtered by X12 ARIMA method and then standardized by using log differences. In the first step, the time series of the factors are estimated:

$$X_{it} = \lambda_i F_t + \varepsilon_{i,t} \tag{5}$$

With a VAR (1.1), the factors are forecasted over the forecast horizon:

$$f_t^Q = \sum_{k=1}^{p} A_k f_{t-k}^Q + B v_{i,t} \tag{6}$$

The forecasted CPI is obtained with the following specifications:

$$\Delta cpi_t^{total} = \mathrm{c} + a_1 * \Delta cpi_{t-4}^{total} + a_2 * \varepsilon_{t-12} + a_3 * f_1 + a_4 * f_2 + a_5 * f_4 + a_6 * f_5 + a_7 * f_6 \tag{7}$$

Total CPI is function of its own lags(Δcpi_{t-4}^{total}), moving average term (ε_{t-12}) and the estimated factors (f_1, f_2, f_4, f_5, f_6).

Table 1. Variables included in the DFM

List of variables

Prices and exchange rates	External sector	Monetary sector
CPI (total)	Exports, fob	Banks Foreign Assets
Food and non-alcoholic beverages	Imports, fob	Claims on central government
Alcoholic beverages, tobacco	Machinery and transport equipment – Exports	Claims on public nonfinancial corporations
Clothing and footwear	Machinery and transport equipment – Imports	Claims on private sector
Housing, Water, Electricity, Gas and other fuels	Other transportation equipment – Exports	Other Banks assets
Furnishings, household equipment and routine maintenance of the house	Other transportation equipment – Imports	Banks Foreign Liabilities
Health	Road vehicles – Exports	M4
Transport	Road vehicles – Imports	Currency in circulation
Communication	Oil – (Export)	Deposits
Recreation and culture	Oil – (Import)	Total short-term deposits
Education		Total long-term deposits
Restaurants and hotels		Total household deposits
Miscellaneous goods and services		Total enterprise deposits
Total CPI excl. Energy and Food (Core inflation)		NBRM international reserves
Total CPI excl. Energy		Required reserve
Producer price index (PPI)		
Brent crude oil		
HWWI Index (EUR) - Total		
HWWI index (USD) – Total excl. energy		
NEER		
REER(CPI)		
Real sector	**Labour market**	
Industrial production (total)	Total registered unemployment	
Energy	Newly registered unemployed	
Intermediate goods	Employed from the register	
Capital goods	Deleted from the register for reasons other than employment	
Durable consumer goods	Nominal net wage	
Non-durable consumer goods	Real net wage	
Mining and quarrying	Nominal net wage in public administration	
Manufacturing	Nominal net wage in industry	
Electricity, gas and water supply	Real net wage in public administration	
Tourist arrivals	Real net wage in industry	
Tourist nights	Nominal gross wage	
	Real gross wage	
	Nominal gross wage in public administration	
	Nominal gross wage in industry	
	Real gross wage in public administration	
	Real gross wage in industry	

Disaggregate CPI and total CPI ARIMA model

We estimate disaggregate models at the lowest level, for which price indices and weights are available from the State Statistical Office (SSO) of the Republic of Macedonia. Currently price indices are available at 4 digit COICOP CPI. At this level of disaggregation, the CPI basket comprises 87 subindicies. Estimating models at such a low level of disaggregation may improve forecast accuracy because of the heterogeneity across the CPI subindicies. The sources of heterogeneity are of economic and methodological nature. In either case, the price data exhibits statistical regularities which can be used in our forecasting models. For most of the CPI subindicies, we use data since January 2003. Some of the series are shorter and start from January 2010, when the SSO started to provide them in more detailed disaggregation.

Our forecasts for the vast majority of the CPI subindicies (around 90%) are based on ARIMA models. As stated in overview section above, the ARIMA methodology is based on the idea that any stationary stochastic process can be approximated well by an autoregressive moving-average (ARMA) process. To keep the exposition simple we omit the constant and assume that the process is integrated of order one. Thus, for a price index in logarithms, p_t, an ARIMA model of order (p, 1, q) can be written as:

$$\Delta p_t = \underbrace{\phi_1 \Delta p_{t-1} + \phi_2 \Delta p_{t-2} + \cdots + \phi_p \Delta p_{t-p}}_{AR\ part} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \underbrace{\theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}}_{MA\ part} \quad (8)$$

where p and q give the number of autoregressive and moving average terms, respectively, and Δ denotes the first difference. The error term ε_t is assumed to follow a white noise process with variance σ^2.

If the price changes display a seasonal pattern, we can extend the case above to a seasonal ARIMA model. For price data in monthly frequency, there is often a seasonality at lag 12 which motivates us to add a seasonal AR term with coefficient ρ to the specification. The seasonal ARIMA model can then be written as:

$$\Delta p_t - \rho \Delta p_{t-12} = \phi_1 (\Delta p_{t-1} - \rho \Delta p_{t-13)} + \phi_2 (\Delta p_{t-2} - \rho \Delta p_{t-14}) + \cdots + \phi_p (\Delta p_{t-p} - \rho \Delta p_{t-p-12)} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (9)$$

Note that if $\rho = 1$ the only difference to the non-seasonal case is that we remove a seasonal unit root by seasonally differencingΔp_t. If $\rho = 0$, we are back to the nonseasonal ARIMA model.

For each CPI subindex, we select a model in two steps. First, we analyze the statistical properties of each price series to determine the order of integration and decide whether it exhibits a seasonality. This analysis is performed on an irregular basis or only every few years. We assume that all CPI items are integrated of order one and use them in first log-differences (d = 1). This assumption is tested by means of two unit root tests: (Dickey & Fuller, 1979) and (Kwiatkowski, Phillips, Schmidt, & Shin, 1992). The tests support this assumption at the 5% level, for most of the 87 individual CPI items. Next, to identify seasonal patterns in the price changes, we examine the autocorrelation function (ACF). A seasonality at 12 months leads to a significant spike in the ACF at multiples of 12, and for those items, we allow for a seasonal AR term.

In the second step, we choose the lag order of the models on the basis of an automatic lag selection criterion. This procedure is automated and the lag order is selected every time the models are reestimated. For each CPI subindex, numerous models with different lag orders are estimated. The best model is then selected based on the Schwarz information criterion:

$$SIC = log\left(\frac{SSE}{T}\right) + k\frac{log(T)}{T} \tag{10}$$

where SSE denotes the sum of squared errors, T the number of observations and k the number of estimated parameters. The algorithm selects the model with the smallest SIC. As we know, more lags improve the fit of the model and therefore lead to lower SSE and SIC. However, choosing an over-large lag order or overfitting of the model can lead to inconsistency of the maximum likelihood estimator (Neusser, 2009, pp. 89-91) and reduce the out-of-sample forecast performance. For that reason, the Schwarz information criterion contains a penalty term which increases with the number of parameters k so that the criterion favors a more parsimonious specification.

For some of the CPI subindicies, we do not use the ARIMA forecast but replace it with ad-hoc assumptions. These ad-hoc assumptions include an extrapolation of the index by its most recent value, the average month-on-month growth rate or the average year-on-year growth rate. These assumptions are used in several cases. First, a visual inspection of the forecast may point to implausible and/or extreme price movements. Next, we replace the ARIMA forecast if we have information on special factors and events which affect only some of the items, mainly administered prices in the CPI basket. Some examples are electricity and central heating price increases that have been announced, or the tax increase on tobacco. If the resulting price effects can be quantified, we include these as add-factors in the forecast.

For prices of oil products, we use external assumptions. Prices of oil products, which represent regulated prices, are set every two weeks by the Energy Regulatory Commission taking into account past developments of crude oil spot price in USD and USD/MKD exchange rate. Therefore, we use exogenous forecast about crude oil prices and USD/EUR exchange rate[4] from external source to make an assumption regarding these prices in the forecasting period.

At the end, in order to obtain forecast for total inflation, we aggregate the forecasts from estimated ARIMA models for individual CPI subindicies and the forecasts for subindices that include ad-hoc and external assumptions, by using their corresponding weights. For comparison and forecast evaluation purposes, we also apply direct ARIMA method to the total CPI (equation 8). Hence, in the next section, beside the inflation forecast from structural model and DFM, we evaluate the inflation forecast generated from disaggregate ARIMA forecast of individual CPI items and from ARIMA forecast of total CPI.

[4] Since MKD is pegged to EUR, forecast developments of USD/EUR are applied to USD/MKD exchange rate.

Forecast evaluation

The forecast evaluation exercise is performed in two parts. First, we make "out-of-sample" forecast with the four models described in the previous section for 16 periods - from 2012 q3 to 2016 q2. In particular, we use historical data to estimate models for the period 1997 – 2012 q2 and then start the forecast. We forecast in each quarter from 2012 q3 to 2016 q2 for a period of one year (four quarters). In order to follow closely the forecasting routine when producing actual short term forecasting we re-estimate the models in each quarter within the forecasting horizon and we employ new assumptions for the exogenous variables. In the evaluation part, we check the forecasting performance of each model for a period up to four quarters - one quarter ahead, two quarters ahead, three quarters ahead and four quarters ahead. The second part of the evaluation exercise presents several composite forecasts. The intention is to compare the accuracy of the composite forecasts alongside the accuracy of the forecasts produced by the individual models.

We use several criteria for comparing the models' forecasting performance – the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), the Symmetric Mean Absolute Percentage Error (SMAPE) and the Theil's U Statistics (Theil). In essence, all these statistics provide a measure of the distance of the true from the forecasted values and each of them has its weaknesses and its strengths.

The RMSE and MAE are the most widely used forecast evaluation metrics.

$$RMSE = \sqrt{\sum_{i=1}^{N} \left(inf_{for} - inf_{act}\right)^2 / N} \tag{11}$$

$$MAE = \sum_{i=1}^{N} abs\left(inf_{for} - inf_{act}\right) / N \tag{12}$$

where N is the number of observations, inf_{act} is the actual (realized) inflation and inf_{for} is the forecasted inflation.

Both measures are relative and scale dependent i.e. should be used to compare forecasts of the same time series across different forecasting models. The smaller the RMSE and MAE, the better the forecasting performance of the model.

Unlike these two measures, MAPE is scale independent measure. On the other hand, its main issue is to be undefined when the denominator is null. In this case, it is recommended to use the SMAPE criterion.

$$MAPE = 100\sum_{i=1}^{N} abs(\frac{inf_{for} - inf_{act}}{inf_{act}})/N \tag{13}$$

$$SMAPE = 100\sum_{i=1}^{N} \frac{abs(inf_{for} - inf_{act)}}{(abs(inf_{for}) + abs(inf_{act}))/2}/N \tag{14}$$

Theil's U statistics or Theil's coefficient of inequality is another criterion that measures forecast accuracy. There are two Theil's coefficients labeled as Theil U1 and Theil U2 coefficient.

$$Theil\ U1 = \frac{\sqrt{\sum_{i=1}^{N}(inf_{for} - inf_{act})^2/N}}{\sqrt{\sum_{i=1}^{N} {inf_{for}}^2/N} + \sqrt{\sum_{i=1}^{N} {inf_{act}}^2/N}} \tag{15}$$

$$Theil\ U2 = \frac{\sqrt{\sum_{i=1}^{N}(inf_{for} - inf_{act})^2}}{\sqrt{\sum_{i=1}^{N} {inf_{act}}^2}} \tag{16}$$

Values closer to 0 for both, Theil's U1 and U2 criteria, indicate better forecasting performance of the evaluated models; if Theil's U1 and U2 are equal to zero than the forecast is perfect. Theil U1 is bounded between 0 and 1, whereas Theil U2 is not bounded.

Forecasts of the alternative models with different time span are shown in Figure 1. The calculated forecast evaluation criteria are presented in Table 2.

Figure 1. Comparison of the forecasts of inflation using different models and different forecasting horizon

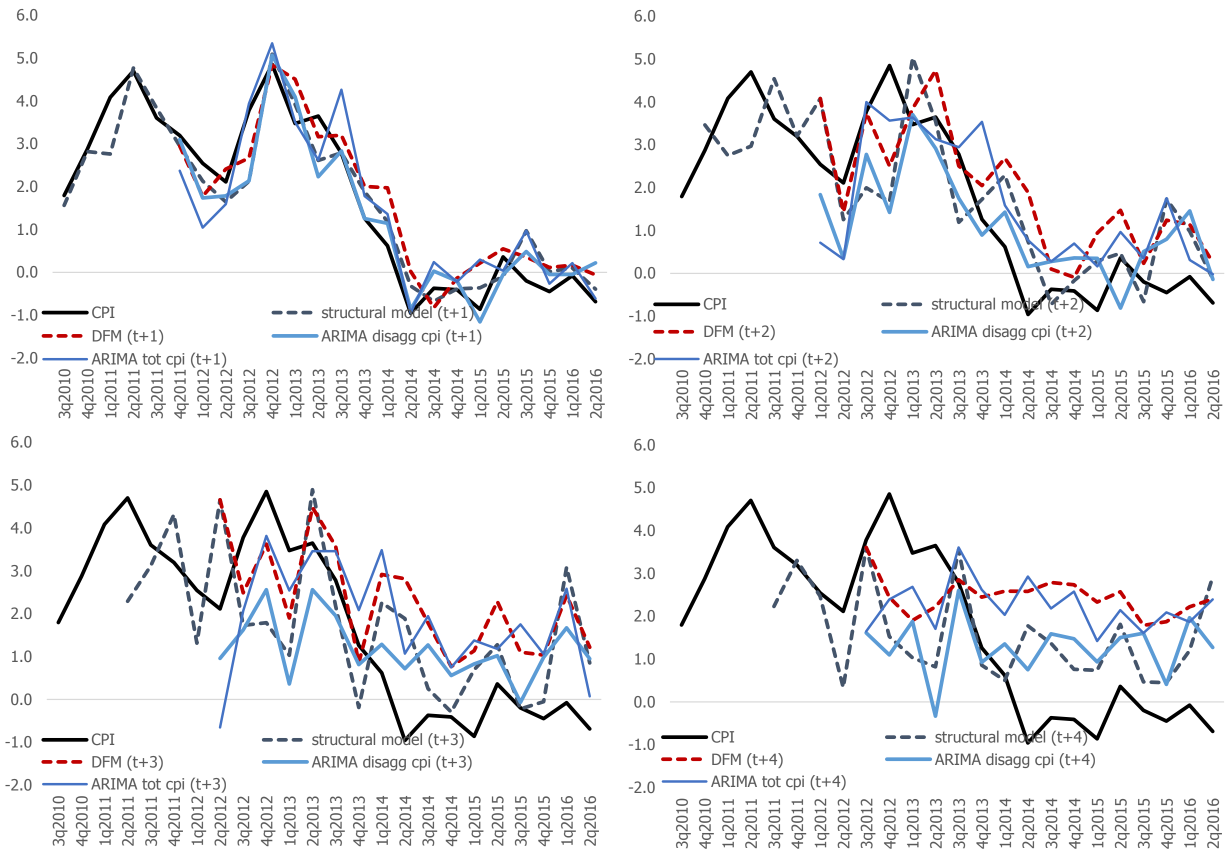


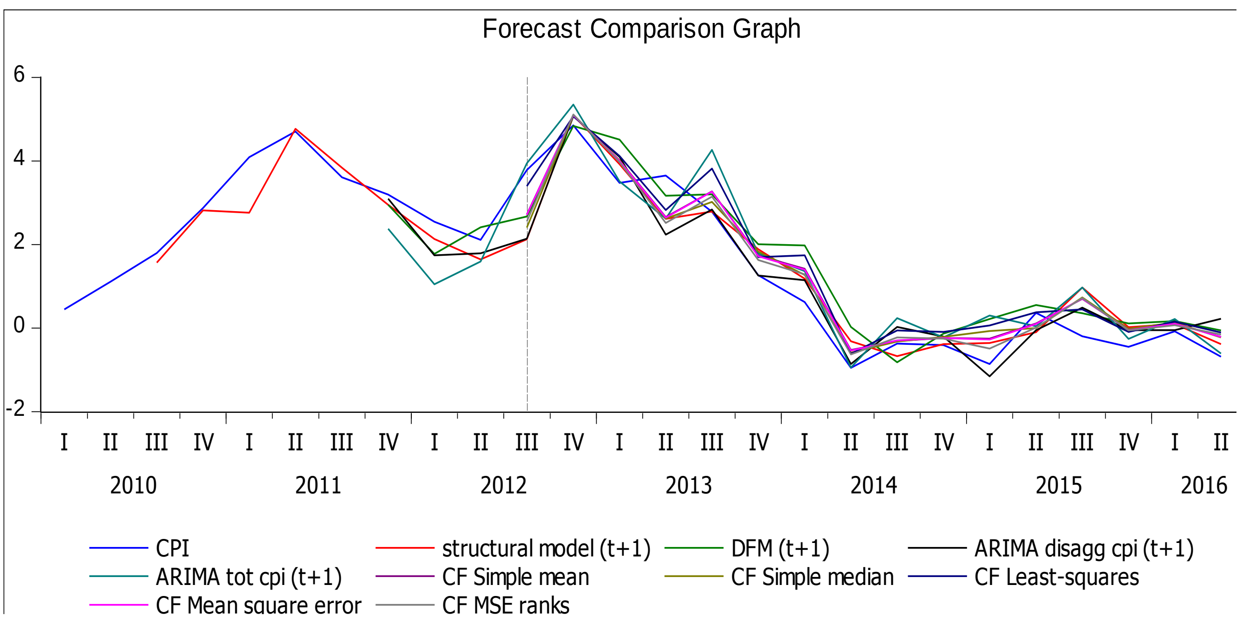
Table 2. Forecast evaluation criterions

	RMSE	MAE	MAPE	SMAPE	Theil U1	Theil U2
structural model (t+1)	0.68	0.54	95.19	81.76	0.16	0.67
DFM (t+1)	0.73	0.63	101.96	101.30	0.17	1.06
ARIMA disagg cpi (t+1)	**0.68**	**0.49**	**68.87**	80.60	0.16	1.37
ARIMA tot cpi (t+1)	0.69	0.53	106.14	**80.13**	**0.15**	**0.55**
structural model (t+2)	1.42	1.14	196.78	**102.49**	0.34	1.61
DFM (t+2)	1.37	1.11	236.30	120.83	0.31	1.89
ARIMA disagg cpi (t+2)	1.25	1.04	244.55	128.76	0.33	**1.28**
ARIMA tot cpi (t+2)	**1.12**	**0.91**	**166.57**	120.35	**0.25**	1.73
structural model (t+3)	1.78	1.48	365.55	**122.06**	0.44	2.51
DFM (t+3)	1.84	1.67	437.56	130.53	0.40	3.19
ARIMA disagg cpi (t+3)	**1.57**	**1.39**	**278.28**	130.19	0.44	2.66
ARIMA tot cpi (t+3)	1.67	1.48	438.83	127.25	**0.37**	**1.93**
structural model (t+4)	**1.91**	**1.57**	**286.87**	135.10	**0.48**	5.37
DFM (t+4)	2.33	2.11	523.08	**132.37**	0.50	5.05
ARIMA disagg cpi (t+4)	2.01	1.74	380.03	143.80	0.56	**3.14**
ARIMA tot cpi (t+4)	2.25	2.11	468.24	136.85	0.50	5.00

The graphical examination of the forecasts and the calculated statistics suggest two important conclusions. ***First***, the forecasting accuracy of all models is higher the shorter the forecasting horizon is; as the forecasting horizon increases the forecasting accuracy of the models declines. This is an expected result since statistical models and small structural models are usually superior when used for nowcasting and short-term forecasting of macroeconomic variables. On the other hand, meaningful medium term forecast should be based on a rich structural model that encompasses all linkages and transmission channels of the economy. The ***second*** conclusion refers to the performance of the individual models. In general, the vast of the evaluation criteria suggest that the disaggregated CPI ARIMA model has strongest forecasting performance and this is especially true when forecasting one and three quarters ahead. In addition, the total CPI ARIMA model and the structural model are superior in some cases, depending on the criterion and the forecasting horizon, whereas the performance of the DFM model is relatively weaker.

As an additional part of the evaluation exercise, we constructed different composite forecasts and compared their forecasting performance to the individual models. Timmerman (Timmerman, 2006) emphasizes at least three main reasons for why forecast combinations may produce better forecast on average than individual forecasting model. First, forecast combination can be motivated by a simple portfolio diversification (hedging) argument. Second, there may be unknown instabilities (structural breaks) that sometimes favor one model over another. By combining forecasts from different models, the decision maker may obtain forecasts that are more robust to these instabilities. Third, forecast combination may be desirable as individual forecasting models may be subject to misidentifications bias that are unknown to the model operators. In this case, combining forecasts may average out the biases, improving forecast accuracy. We constructed several composite forecasts by using different weights – mean, median, least square estimates, mean square error and MSE ranks. Having in mind that models' precision is highest in one-quarter ahead forecasting horizon, the forecast evaluation period is only one quarter ahead. Figure 2 shows the performance of the composite forecasts (CF)[5] against individual forecasts and the evaluation statistics are given in Table 3.

Figure 2. Comparison of one-quarter ahead forecasts of inflation using different models and composite forecasts



[5] The Composite forecasts are calculated as weighted averages of the individual forecasts by using the following weighting schemes - mean, median, least square estimates, mean square error and MSE ranks. The calculation was performed in Eviews 9.

Table 3. Forecast evaluation of individual models and composite forecasts

	RMSE	MAE	MAPE	SMAPE	Theil U1	Theil U2
structural model (t+1)	0.68	0.54	95.19	81.76	0.16	0.67
DFM (t+1)	0.73	0.63	101.96	101.30	0.17	1.06
ARIMA disagg cpi (t+1)	0.68	0.49	**68.87**	80.60	0.16	1.37
ARIMA tot cpi (t+1)	0.69	0.53	106.14	80.13	0.15	**0.55**
CF simple mean	0.58	0.50	83.83	**76.36**	0.13	0.82
CF simple median	0.63	0.53	89.10	86.71	0.15	0.82
CF least-squares	0.61	0.52	86.72	87.76	0.14	0.95
CF mean square error	**0.58**	0.50	83.71	76.39	**0.13**	0.82
CF MSE ranks	0.59	**0.49**	80.45	78.60	0.14	0.88

CF = composite forecast

In line with our expectations, the majority criteria suggest that the composite forecasts are superior in accuracy as compared to individual models, although the differences between composite forecasts and forecasts obtained from the disaggregated CPI ARIMA model in most of the cases are rather small. In addition, it should be acknowledged that more formal forecast combination exercises are performed by using more individual forecasts compared to the number of individual forecasts used in our research. For example, Kapetanious et al. (2008) used 16 competing models, Bjørnland et al. (2008) used 10 models, and Akdogan et al. (2012) used 14 models for constructing composite forecasts for the macroeconomic variables.

Concluding remarks

Forecasting is very important for policy makers. Effectiveness of policy choices and measures is largely dependent upon timely and accurate forecast of the macroeconomic variables because of the existence of considerable transmission lags.

This study provides a discussion on several models used for short-term forecasting of inflation at the NBRM and evaluation of their forecasting performance. Forecasting performance was evaluated by using standard evaluation criteria such as Root Mean Squared Error, the Mean Absolute Error, the Mean Absolute Percentage Error and the Theil's U Statistics (Theil). Additionally, we constructed several composite forecasts in order to test for the superiority of the combination method as opposed to individual models' forecast.

Our results point to three important conclusions. First, the forecasting accuracy of the models is inversely related to the forecasting horizon – the longer the forecasting horizon the higher the forecasting errors. Second, in general, the disaggregated CPI ARIMA model has smallest forecasting errors. Third, majority of the forecast evaluation criteria suggest that composite forecasts are superior in comparison to the individual models.

Future work in this field should be focused on further development and improvement of currently used models. In addition, efforts should be made in the area of new model development. In this sense, forecasting literature emphasizes the Bayesian VAR, Bayesian Model averaging and mixed-frequency BVAR as models with very good forecasting performances. Enlarging the suite of models used for inflation forecasting, will, in turn be beneficial for the forecasting accuracy of the composite forecast indicators, as well.

References

Akdogan, K., Baser, S., Chadwick, M. G., Ertug, D., Hulagu, T., Kosem, S., . . . Tekali, N. (2012). Short-Term Inflation Forecasting Models for Turkey and a Forecast Combination Analysis. *National Bank of Turkey Working Paper No:12/09*.

Benealal, N., Hoyo, J. L., Landau, B., Roma, M., & Skudelny, F. (2004). To Aggregate or Not to Aggregate? Euro Area Inflation Forecasting. *European Central Bank*(Working Paper Series 374).

Bjørnland, H., Jore, A., Smith, C., & Thorsrud, L. (2008). Improving and evaluating short term forecasts at the Norges Bank . *Norges Bank Staff Memo, No.4.*

Box, G. E., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco CA: Holden Day.

Chong, Y. Y., & Hendry, D. F. (1986). Econometric Evaluation of Linear Macro-Economic Models. *Review of Economic Studies, Oxford University Press, vol. 53(4)*, 671-690.

Clark, P. K. (1987). The cyclical component of U.S. economic activity. *Quarterly Journal of Economics*(102), 797-814.

Clements, M. P., & Hendry, D. F. (2004). *A Companion to Economic Forecasting.* Blackwell Publishing Ltd.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association 74(366)*, 427-431.

Figlewski, S. (1983). Optimal price forceasting using survey data. *Review of Economics and Statistics 65*, 813-836.

Figlewski, S., & Urich, T. (1983). Optimal aggregation of monay supply forecasts: Accuracy, profitability and market efficiency. *The Journal of Finance 28*, 695-710.

Harvey, A. C. (1989). *Forecasting, Structural Time Series and the Kalman Filter.* Cambridge: Cambridge University Press.

Huwiler, M., & Kaufmann, D. (2013). Combining disaggregate forecasts for inflation: The SNB's ARIMA model. *Swiss National Bank Economic Studies*(7).

Kapetanios, G., Labhard, V., & Price, S. (2008). Forecast combination and the Bank of England's suite of statistical forecasting models. *Economic Modelling 25*, 772–792.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics 54(1–3)*, 159–178.

Lack, C. (2006). Forecasting Swiss inflation using VAR models. *Swiss National Bank Economic Studies No.2*.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis.* Berlin: Springer-Verlag.

Neusser, K. (2009). *Time series analysis in economics. 2nd edition.* Wiesbaden: Teubner.

Sims, C. (1980). Macroeconomics and reality. *Econometrica*, 1-48.

Stock, J. H., & Watson, M. W. (2006). Forecasting with Many Predictors. In G. Elliott, C. W. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting, Volume 1.* Elsevier B.V.

Stock, J., & Watson, M. (2002). Macroeconomic Forecasting using diffusion indexes. *Journal of Business and Economic Statistics, 20*, 147-162.

Timmerman, A. (2006). Forecast combinations. *In G. Elliott, C. W. J. Granger, and A. Timmerman (Eds.), Handbook of Economic Forecasting, Volume 1*.

Timmermann, A. (2006). Forecast Combinations. *Handbook of Economic Forecasting*.