Graduate Institute of International and Development Studies
International Economics Department
Working Paper Series

Working Paper No. HEIDWP01-2020

Modeling Azerbaijan's Inflation and Output Using a Factor-Augmented Vector Autoregressive (FAVAR) Model

Vugar Rahimov
Central Bank of the Republic of Azerbaijan

Nijat Guliyev
Central Bank of the Republic of Azerbaijan

Vugar Ahmadov
Central Bank of the Republic of Azerbaijan

January 2020

Chemin Eugène-Rigot 2
P.O. Box 136
CH - 1211 Geneva 21
Switzerland

THE GRADUATE INSTITUTE | GENEVA

INSTITUT DE HAUTES ÉTUDES
INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL
AND DEVELOPMENT STUDIES

Modeling Azerbaijan's Inflation and Output Using a Factor-Augmented Vector Autoregressive (FAVAR) Model[1]

Vugar Rahimov[2]
Nijat Guliyev[3]
Vugar Ahmadov[4]

Abstract

In this study, we build and use a factor-augmented vector autoregressive (FAVAR) model to forecast inflation and output in Azerbaijan. The FAVAR model is particularly effective in data-rich environments, alleviating the curse of dimensionality of the standard VAR model and handling omitted variable bias. Using 77 variables for factor extraction and quarterly data for the period 2003 to 2018, we build several multivariate models, including a FAVAR model, and compare their performance with that of a benchmark univariate model. Our findings show that almost all of the multivariate models underperform in comparison with the univariate model. This result is in line with the literature, which finds that simple models are better forecasters of some macroeconomic variables, especially inflation. We acknowledge that the results might be affected by the relatively short length of the sample period and existence of irregularities in the data.

[1] The authors are greatly indebted to Professor Ugo Panizza for his guidance and useful feedback throughout the project, and are grateful to the BCC program, SECO, and Graduate Institute for their support. The authors would like to thank to Professor Cedric Tille and colleagues at the Graduate Institute for their helpful comments and suggestions related to the preliminary findings. The views expressed are those of the authors, and do not necessarily reflect the views of the Central Bank of the Republic of Azerbaijan.
[2] Vugar Rahimov – Central Bank of the Republic of Azerbaijan, e-mail: vugar_rahimov@cbar.az
[3] Nijat Guliyev - Central Bank of the Republic of Azerbaijan, e-mail: nijat_guliyev@cbar.az
[4] Vugar Ahmadov - Central Bank of the Republic of Azerbaijan, e-mail: vugar_ahmadov@cbar.az

Introduction

Forecasting is an important tool in central banks' decision-making processes. In this regard, forecasts for inflation and real output play a crucial role in these banks' policymaking processes, enabling efficient forward-looking monetary policies. However, finding the best model fit for a set of macroeconomic variables is a significant challenge. Many models have been proposed to produce robust forecasts. However, because countries are unique, no single model can be applied to all economies. The factor-augmented vector autoregressive model (FAVAR) is widely used in macroeconomic forecasting, and is considered particularly effective in data-rich environments. Factor models are useful because they alleviate the curse of dimensionality of the standard VAR model; handle omitted variable bias, providing some robustness in the presence of structural breaks; and require minimal conditions on the errors.[5] In FAVAR models, the information contained in a large number of variables is first expressed using a few (latent) factors, which are then used in conventional VAR models. In this study, we construct a FAVAR model to forecast inflation and output (non-oil GDP growth rate) for the Azerbaijani economy.

Although FAVAR models usually perform well in terms of forecasting, we find that simpler models are more successful in describing the dynamics of the macroeconomic indicators for Azerbaijan. In other words, we find that FAVAR models fail to outperform univariate and other multivariate models in terms of forecasting inflation, especially in the short term. The results of a Diebold–Mariano (DM) (Diebold & Mariano, 1995) test indicate whether the difference between the relative root mean squared errors (RMSEs) of the FAVAR and benchmark autoregressive (AR) models is significant. Although the FAVAR model underperforms relative to the simple VAR model, it outperforms the benchmark model in terms of forecasting output. However, the DM test results reveal that the differences between RMSEs are not statistically significant, implying that all models have similar forecasting ability. Lastly, graphs of the forecasts and forecast errors show that the poor performance of the FAVAR model is evident throughout the sample period.

To the best of our knowledge, this study is the first to use a FAVAR model to examine output and inflation in Azerbaijan. Huseynov *et al.* (2014) estimated inflation for Azerbaijan, using a FAVAR model to compare the forecasting ability of various univariate and multivariate models. However, they used the Bayesian method with Gibbs sampling, rather than using principal components (PCs). In addition, their estimation period runs until 2014, and their analysis is based on monthly models. Because inflation and output forecasts by the CBAR are performed on a quarterly basis, we compare the forecast ability of quarterly models. In addition, our sample period includes the most recent structural changes in the economy (i.e., after 2015).

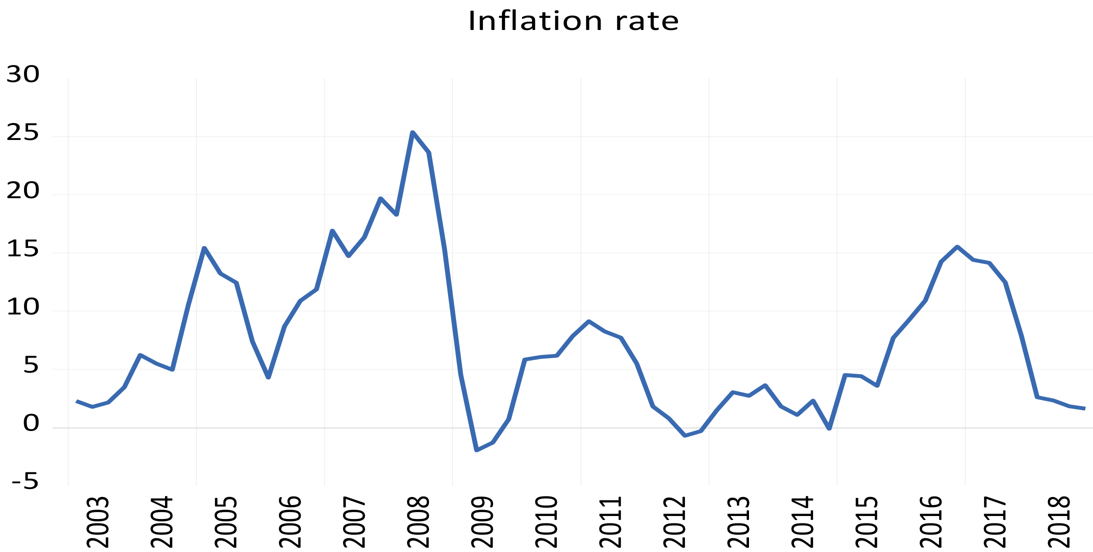
The rest of the paper is structured as follows. In section two, we provide a literature review. Sections three and four describe methodological and data issues, respectively. We present our results in section five, and section six concludes the paper.

[5] As of 2019, by applying a managed floating regime, the CBAR plans to move toward an inflation targeting regime in the near term. Therefore, one of the preconditions to transit to a new regime is to have better forecasting models available. This study is helpful in terms of reaching these goals.

Background information on the economy of Azerbaijan

Inflation in Azerbaijan over the last 18 years can be divided into three periods: high and upward trending (2000–2008), low and stable (2009–2015), and post-devaluation high-inflation (2016–2017) regimes (see **Figure 1**).

Figure 1. Annual inflation dynamics in Azerbaijan (2003–2018)



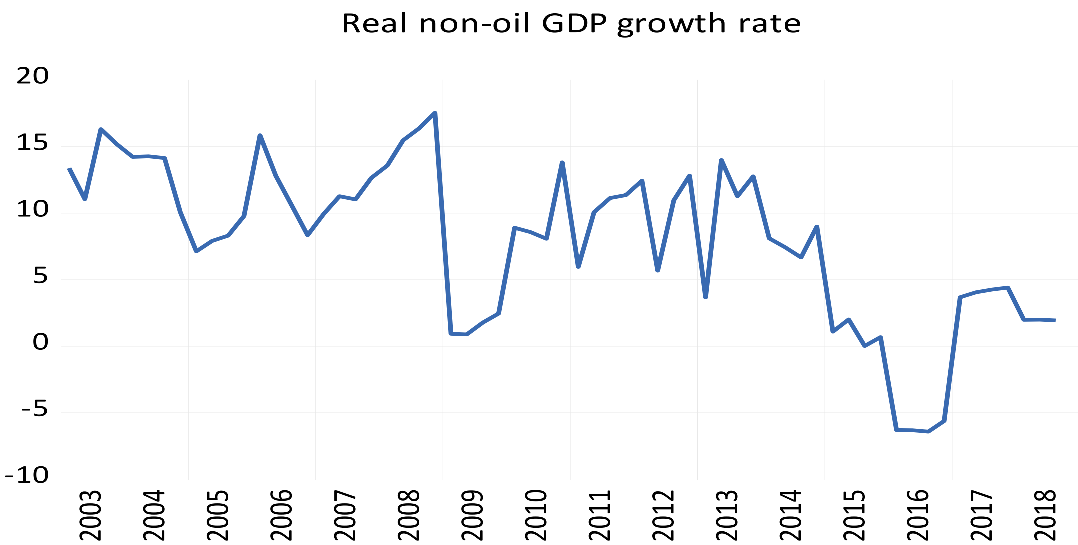
Source: State Statistics Committee of Azerbaijan

Increased oil revenues from the realization of oil pipeline projects created significant opportunities for Azerbaijan to implement social and infrastructure projects. As a result, government spending increased from AZN 0.8 billion in 2003 to AZN 11 billion in 2008. Positive demand shocks, accompanied by the persistent increase in government spending (consumption and investment) and sharp oil price hikes (from $56.75 per barrel in 2005 to $133.90 per barrel in August 2008), contributed to upward pressure on both actual inflation and inflation expectations.

After increasing until 2008, inflation became stable over the subsequent six years, despite the global financial crisis. However, it began increasing again in 2015 as a result of devaluation expectations and the Central Bank's (CBAR's) decision to move to a floating exchange rate regime after a sharp decrease in the oil price. Indeed, between 2008 and 2014, the CBAR's credible fixed exchange rate policy encouraged de-dollarization in the country. However, a reversal occurred when oil prices plummeted in 2014, forcing the CBAR to devalue the currency in 2015. The impact was immediate, and caused inflation to accelerate. In the aftermath of the exchange rate regime change, the annual average inflation rate climbed from 4 percent in 2015 to 12.4 percent in 2016, and to 12.9 percent in 2017. Then, in reaction to the stable bilateral exchange rate and the tight monetary and fiscal policies of 2016–2017, inflation decelerated, stabilizing at 2.3 percent in 2018.

Figure 2 plots the real non-oil GDP growth rate of Azerbaijan for the period 2003–2018. Here, we focus only on non-oil output, because economic policies can only affect the non-oil sector. The oil price and demand for oil production are formed externally to the country; thus, we treat this indicator as exogenous. The share of non-oil GDP to total GDP is around 60 percent. Like inflation, real non-oil output is affected significantly by budget expenditures. Azerbaijan boasted the fastest-growing economy in the world in 2006 and 2007. From 2009, the growth rate dropped to single digits, where it stabilized. After a long period of high and moderate GDP growth, the economic situation in the country began to deteriorate in 2016, with the growth rate turning negative, before recovering again in 2017.

Figure 2. Annual real non-oil output growth dynamics in Azerbaijan (2003–2018)



Source: State Statistics Committee of Azerbaijan

Literature review[6]

The FAVAR model gained in popularity after the seminal work of Bernanke *et al.* (2005), who examined the effect of monetary policy on different economic variables by augmenting a conventional VAR model with factors extracted from 120 variables. The authors estimate the FAVAR model in two ways: the one-step method, and the two-step method. In the two-step method, they first extract factors using the principal component method (other methods are possible), and then add these factors to a conventional VAR model. In the one-step method, they estimate the factor loadings and VAR jointly by applying Gibbs sampling to estimate a Bayesian VAR. The authors build alternative configurations of the FAVAR model by using one or more factors. They compare the performance of the FAVAR model against that of a benchmark AR and conventional VAR models. The results show that including even one factor significantly improves the performance of the model and

[6] See the following works for further information on factor models, including estimations and inferences using such models: Kilian & Lütkepohl (2016), Stock & Watson (2010), Bai & Ng (2002).

eliminates the "price puzzle." Because their main concern is analyzing the impulse-response functions (i.e., how one variable in the system affects the other variables), the authors face an identification challenge in order to recover the structural parameters. They address this challenge by dividing the observable variables into fast-moving and slow-moving groups, and placing the factor extracted from the slow-moving group first in the VAR ordering.

In constructing their model, Bernanke *et al.* (2005) rely heavily on the model and data of Stock & Watson (2002b). Stock & Watson (2002b) apply a factor model to forecast various macroeconomic variables, which they call estimated factor diffusion indexes (hence, the diffusion index (DI) model). They apply h-step-ahead forecasting by directly estimating y_{t+h} as a function of y_t:

$$y_{t+h}^{h} = a_h + b_h(L)F_t + j_h(L)y_t + e_{t+h}^{h}$$

Here, y_{t+h}^{h} is an h-step-ahead variable to be forecasted, F_t is factor vector, and L is a lag polynomial. They compare the DI model with different benchmark models, including the simple AR model, leading indicator model, and VAR model, at 6-, 12-, and 24-month forecasting horizons. The authors estimate three models. In the first, the explanatory variables are the DIs and the lags of both the DI and the target variables. In the second model, the explanatory variables are the DI and the lags of the target variable. The third model uses the DI only as the explanatory variable. The results show that the DI models outperform the benchmark models, and that the model based only on the DI performs best.

Using 105 macroeconomic variables, Lagana & Mountford (2005) use a FAVAR model to investigate UK monetary policy. Their findings show that the FAVAR model outperforms benchmark AR and VAR models in terms of forecasting interest rates. They further conclude that the FAVAR model can eliminate the "price puzzle" phenomenon.[7] Ajevskis & Davidsons (2008) apply Stock & Watson (2002a) and dynamic factor methodologies to forecast Latvia's GDP, and compare the results with those using benchmark AR models. They show that both factor models exhibit better forecasting ability, but that neither improvement is statistically significant. Eickmeier & Ziegler (2008) assess the relative forecast performance of large dynamic factor models for output and inflation using a meta-analytic approach, showing that these types of models exhibit better forecast performance than that of other models.

Employing a FAVAR approach, Reigl (2017) constructs a forecast model for Estonian headline and core inflation. As in Stock & Watson (2002b) and Bernanke *et al.* (2005), Reigl (2017) applies a two-step approach. First, he extracts factors from 388 quarterly economic and financial variables, which he then incorporates into the forecasting model. The results show that a small number of factors obtained from a large data set is more successful at forecasting both headline and core inflation than other models are.

Günay (2018) uses a factor model to forecast industrial production and core inflation for the economy of Turkey. The author estimates 339 model specifications by varying the data sets, number of factors,

[7] The puzzle refers to the fact that theory suggests that increasing the interest rate should lead to a decrease in prices. However, Sims (1992) found that higher interest rates lead to higher price levels.

and lags. The core model is similar to that of Stock & Watson (2002a). The findings show that, in general, the best models for industrial production and inflation differ. Thus, a model needs to be selected based on the target variables. Using a simple factor extraction model is generally preferred, because the additional computational difficulties associated with more complicated factor extraction models do not warrant the improvement in performance. Another finding of the research is that parsimonious models perform better. Thus, it is better to implement a few factors, and to choose a relatively small data set for factor extraction.

As noted above, Huseynov *et al.* (2014) were the first to investigate the forecasting ability of inflation models using a FAVAR model using data for Azerbaijan. For the period 2010–201, they found that univariate models have better predictive ability than multivariate models. Mammadov & Adigozalov (2014) identified leading and coincidence indicators of Azerbaijan business cycles for use in forecasting. However, they did not employ FAVAR models, and did not compare the performance of different models.

Methodology

Although the use of factor models in macroeconomic analyses has a long-standing history (Sargent & Sims, 1977, Engle & Watson, 1981, Quah & Sargent, 1993, Forni *et al.*, 2001), the use of FAVAR models is relatively new.

FAVAR models offer a proper specification that exploits valuable information provided by a large data set, without having to worry about degrees of freedom, overfitting, or increasing parameter uncertainty in the estimations. These models overcome the curse of dimensionality problem, while preserving the value-added information provided by relevant variables. In contrast, owing to issues with degrees of freedom, other models that use few variables often do not include necessary information (variables), thus yielding unreliable results. For instance, the "price puzzle" phenomenon, noted by Sims (1992), is the result of sparse data. Although this puzzle can be solved by including commodity prices in the model, FAVAR models solve such problems by including a rich data set and avoiding omitted variable bias.

Suppose Y_t is an $M \times 1$ vector of observable variables, and F_t is a $K \times 1$ vector of unobserved variables. Assume that the joint dynamics of (F_t, Y_t) is given by the following transition equation:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + v_t, \qquad (1)$$

where $\Phi(L)$ is a conformable lag polynomial of finite order d, which may contain a priori restrictions, as in the structural VAR literature. The error term v_t has mean zero and covariance matrix Σ.

Because F_t is unobservable, equation (1) cannot be estimated. However, this information can be extracted from a set of economic time series, denoted by the $N \times 1$ vector X_t. The number of informational time series N is large, and is assumed to be much greater than the numbers of factors and observed variables in the FAVAR model. In addition, X_t is assumed to be related to unobservable F_t and observable Y_t by the following equation:

$$X_t = \Lambda^f F_t + \Lambda^Y Y_t + e_t, \qquad (2)$$

where Λ^f is an $N \times K$ matrix of factor loadings, and Λ^Y is an $N \times M$ matrix. The $N \times 1$ vector of error terms e_t has mean zero, and is assumed to be normal and uncorrelated or to display a small amount of cross-correlation, depending on whether the estimation uses likelihood methods or principal components. In general, Y_t and F_t can be correlated. Equation (2) captures the idea that Y_t and F_t are both forces that drive the dynamics of X_t.

Thus, the forecast factors and variables in the FAVAR model can be written in the following form:

$$\begin{bmatrix} \hat{F}_{t+1} \\ \hat{Y}_{t+1} \end{bmatrix} = \hat{\Phi}(L) \begin{bmatrix} \hat{F}_t \\ Y_t \end{bmatrix}, \qquad (3)$$

where $\hat{F}_t$ is a factor estimate. The important part of the model is the estimation method. There are two major approaches to estimating the model. The first, usually called the two-step procedure, estimates the unobserved factors, and then uses these factors in conventional VAR models. The second method estimates the factors and the VAR jointly, and is known as the one-step estimation procedure. Both methods have advantages and disadvantages. Bernanke *et al.* (2005) estimate FAVAR models using both methods, but find little variation in the results. Thus, because the one-step procedure involves cumbersome technical procedures, without significantly improving the model performance, we apply the two-step procedure, where we use a PC analysis to estimate the factors.

The standard method used to evaluate the effectiveness of a model is to compare its performance with that of benchmark models, usually using the root mean square errors (RMSEs). Here, we compare the performance of the proposed FAVAR model with that of univariate AR and VAR models. The RMSE is given as follows:

$$RMSE = \sqrt{\frac{(Realization\ at\ time\ t+h) - (Forecast\ for\ time\ t+h\ at\ time\ t)^2}{Number\ of\ forecasts}} \qquad (4)$$

In the results, we express the RMSE of alternative models as a share of that of the AR model. Thus, a lower ratio indicates better model performance.

A Diebold–Mariano (Diebold & Mariano, 1995) test is performed to determine whether the differences between the RMSEs are significant. Define d_τ as follows:

$$d_\tau = g\left(e_\tau^{(2)}\right) - g\left(e_\tau^{(1)}\right), \qquad (5)$$

where

$$e_{\tau}^{(i)} = y_{\tau+h} - \hat{y}_{t+1|\tau}, \quad (6)$$

for models $i = 1, 2$, and g is a generic loss function ($g(e) = e^2$.
Then, the DM test is defined as follows:

$$DM = \sqrt{n}\frac{\bar{d}}{\sigma_d}, \quad (7)$$

where

$$\bar{d} = \frac{1}{n}\sum_{\tau=t}^{T+n-1} d_{\tau}, \quad (8)$$

and σ_d is the estimator of the variance of $\bar{d}$. The null hypothesis of the DM test is that the two models have the same RMSE.

Data and estimation

The full sample contains quarterly data for the period 2003Q1–2018Q4 on domestic and foreign variables. The sample period starts from 2003 because we do not have reliable earlier observations. The data are transformed before conducting the analysis. The first transformation seasonally adjusts the series using the Census X12 procedure. Most of the series have seasonal patterns. We also apply a log transformation and difference the series, which have a unit root. After performing these transformations, the data are stationary. Table 1 presents the descriptive statistics and Table 2 presents the unit root test results for inflation and real non-oil GDP, as well as for other observed variables used in the VAR models. The national data are obtained from the database of the National Statistics Office and the Central Bank of Azerbaijan Republic. The foreign variables are taken from international agencies' databases. We add a dummy variable to capture two major structural breaks: 2008–2009, and 2015–2016. The series used for the PC extraction are presented in Appendix I.

Azerbaijan's small open economy cannot affect the economies of other countries. Therefore, we use external forecasts for the out-of-sample analysis in our FAVAR, as well as in other multivariate models. For inflation forecasting, we use yearly forecasts for trade partners' inflation, taken from the IMF's Archive of World Economic Outlook Databases, that we interpolate to quarterly series. Similarly, yearly forecasts of the oil price are collected from the EIA's annual reports.

Table 1. Summary statistics*

Variables	Obs	Mean	Median	Std	Max	Min
Inflation	62	1.8	1.6	2.1	7.3	-3.5
RGDPN	62	1.8	1.7	2.6	9.2	-9.2
TP Inflation	62	1.6	1.6	0.5	3.2	0.7
M2	62	5.6	6.3	8.9	29.1	-26.1
NEER	62	-0.1	0.0	5.6	11.6	-22.3
Monetary base	62	5.1	5.2	8.5	24.4	-21.0

Oil price	62	1.2	3.7	14.8	25.2	-66.8
Budget capital expenditure	62	8.0	8.0	27.0	70.5	-64.6

**All variables are expressed in terms of growth*

Table 2: Unit root tests results (ADF)

Inflation						
Variables	**Level**			**First differences**		
	Intercept	**Intercept and trend**	**Status**	**Intercept**	**Intercept and trend**	**Status**
CPI	-1.489	-1.718	Nonstationary	-9.960***	-9.875***	Stationary
RGDPN	-3.105*	-0.263	Nonstationary	-8.229***	-9.366***	Stationary
M2	-2.170	-1.727	Nonstationary	-2.784*	-4.844***	Stationary
NEER	-2.324	-2.205	Nonstationary	-3.146**	-3.195*	Stationary
TP CPI	-1.819	-1.248	Nonstationary	-5.761***	-6.099***	Stationary
RGDPN						
Variables	**Level**			**First differences**		
	Intercept	**Intercept and trend**	**Status**	**Intercept**	**Intercept and trend**	**Status**
RGDPN	-3.105*	-0.263	Nonstationary	-8.229***	-9.366***	Stationary
CPI	-1.489	-1.718	Nonstationary	-9.960***	-9.875***	Stationary
Monetary base	-2.352	-1.063	Nonstationary	-3.703***	-4.331***	Stationary
Oil price	-2.902*	-2.611	Nonstationary	-5.779***	-5.911***	Stationary
Budget capital expenditure	-1.678	-1.419	Nonstationary	-7.477***	-7.585***	Stationary

*** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the percent level

To check the sensitivity of the forecasting performance to different sample periods, many authors split evaluation and forecasting periods into two parts, and conduct forecasting exercises for both. Although this is an effective strategy when comparing model performance across periods, it is not appropriate in our case, owing to the relatively short period of data availability and the relatively complicated nature of the Azerbaijan economy during the period 2015 to 2018. Instead, we provide graphs of the forecasted series and the actual series in Appendix III, and a graph of the forecast errors in Appendix IV.[8]

The estimation period always starts in 2003Q1. The shortest estimation period is from 2003Q1 to 2010Q4. Thus, the first forecasted period starts in 2011Q1 for one-period-ahead forecasting, 2011Q2 for two-period-ahead forecasting, 2011Q4 for four-period-ahead forecasting, and so on. We repeatedly increase the evaluation sample and decrease the forecasted period by one step, stopping when the forecasted period is the last quarter of 2018. Thus, the longest estimation period for one-step-ahead forecasting is 2003Q1 to 2018Q3, that for two-step-ahead forecasting is 2003Q1 to 2018Q2, and so on; the longest estimation period for eight-step-ahead forecasting is 2003Q1 to 2016Q4. PCs are always extracted from the data used in the estimation.

Choosing the correct number of factors in a factor model estimation can be challenging. Although formal criteria for such decisions are available (Bai & Ng, 2002), we refrain from using them, mainly because of their relatively poor ability to explain the variation in our data. Instead, we select the number of factors so that they together explain about 50 percent of the variation in the data. This approach leads us to use three or four factors, depending on the model specification. However, we also conduct an analysis using only one factor for all models, and find that the model performance usually deteriorates.

Two types of forecasting are popular: direct, and multistep (dynamic) forecasting. In direct forecasting, the model is constructed so that the dependent variable is a function of the h-step earliest values of the explanatory variables. In this case, there is no need to forecast values of the variables during intermediate periods while conducting the forecasting exercise. This approach is implemented by Stock & Watson (2002a). However, this method is not appropriate in FAVAR models, which we use here. Thus, we apply multistep (dynamic) forecasting in this study: we forecast one period ahead, and then use these values to forecast the next period.

Another issue that needs to be addressed when conducting a forecasting analysis is to choose the number of periods over which the forecast ability of the model will be compared. This issue becomes particularly important in a FAVAR analysis, because prior studies have shown that the FAVAR model performs better for long- and medium-term forecasts than it does for short-term forecasting. Thus, we compare the forecasting results of the FAVAR with that of benchmark models for one, two, four, six, and eight periods (corresponding to a quarter, a half year, a year, a year and a half, and two years, respectively).

The lag length in all VAR-type models is two. Formal criteria for lag selection (AIC, SC) usually suggest a lag length of between one and three (usually one, and rarely more than three). Thus, we use two lags[9] to ensure greater consistency between the VAR-type models. Our choice of two lags stems mainly from two reasons. First, the models employed by the CBAR for forecasting use two lags. Second, one lag may leave too much residual autocorrelation, and using three lags leads to a loss of degrees of freedom. For the ARIMA model, we use four lags because the data are quarterly.

---

[8] For brevity, we provide graphs for the most important models only. Other graphs are available upon request.

[9] We run the models with one lag as well; the results do not change significantly.

We use separate models to forecast inflation and non-oil GDP, for two reasons. First, existing research shows that the determinants of inflation and non-oil output in Azerbaijan differ (Rahimov *et al.*, 2016, Huseynov & Mammadov, 2016). Although the oil price and budget capital expenditure are key determinants of non-oil GDP, their effects on CPI are not straightforward. Similarly, trade partners' inflation and NEER are key determinants of the CPI, but are not significant in the GDP model. If we include all variables in a VAR model, this would result in a lead to FAVAR model with at least 10 variables. Second, the PCs used to explain inflation and non-oil GDP are extracted from different subgroups of variables; this point is elaborated on in the next section.

In general, we construct two types (specifications) of FAVAR models. In the first model, we use only the target variable and the PCs as endogenous variables. In the second model, we augment the VAR model used by the CBAR with the PCs. We refer to the first type as an SW-type model, because it is similar to that employed by Stock & Watson (2002a).[10] The second type is referred to simply as a FAVAR model.

A difficulty during forecasting is choosing forecasts for the foreign variables in our model, for example, the oil price and macroeconomic variables that characterize the economies of Azerbaijan's trade partners. In a VAR framework, the model automatically forecasts one period, and then uses these values for subsequent forecasting. Although this procedure is appropriate for domestic variables, it may not be accurate for foreign variables, because these are generally determined exogenously to our model. Thus, to be consistent, we use actual values of foreign variables during the estimation process, and use forecast values available from international organizations for these variables when forecasting domestic variables. In appendix II, we present the results for both types of models: those obtained using forecasted values provided by international organizations,[11] and those using values forecasted by our model.

Note that we do not apply the above-mentioned approach to the PC extraction; that is, the PCs are fully forecasted by the model. We acknowledge that this may be a subject of debate, because using both domestic and foreign variables for the PC extraction might mean that the PCs are not correctly forecasted endogenously. However, we refrain from correcting this issue, for two reasons. First, doing so would make the calculation extremely complicated, because it requires forecasting all domestic variables for the PC extraction, augmenting them with the forecasted values of foreign values, extracting PCs from this pool, and using these PCs to forecast the variables of interest. Second, of the variables used for PC extraction (32 variables for inflation, and 33 variables for non-oil GDP), few are foreign variables (seven and four, respectively). As a robustness check we exclude foreign variables from the respective groups, and extract PCs from the remaining domestic variables; the results are qualitatively similar.

In general, we use the following models for inflation and GDP (separately): the simple ARIMA model, existing VAR model (the benchmark model), SW model, FAVAR model, and BVAR model. Here, we estimate the existing models using the Bayesian estimation technique.[12] We estimate the models that contain foreign variables using two approaches, namely, with and without using forecasts

[10] Stock & Watson (2002a) use a single-equation model.
[11] The results obtained for this type of model are the main results of this study.
[12] We use the standard Bayesian estimation framework, which is the default in EViews.

provided by international organizations during the forecasting process. Appendix VI presents the VAR stability test graphs estimated from the full sample.

Results and discussion

We begin constructing the FAVAR model by using all X variables for the PC extraction, enabling us to capture general movements in the economy. We find that the performance of the PCs is very low, such that first five PCs explain only 35 percent of the total variation, and the first eight explain 50 percent. This result indicates that, in general, there are no strong correlations between the variables at the economy level, supporting the findings of previous works (Boivin & Ng, 2006 and Günay, 2018). Boivin & Ng (2006) find that adding data series to a model can reduce its forecasting ability when idiosyncratic shocks are cross-correlated. Furthermore, including additional variables in the factor extraction is not always beneficial. Instead, pre-selecting the variables may improve the model's performance. A potential explanation for this finding is that Azerbaijan is a small open developing economy, and the nature of its data changes regularly. Because the explanatory ability of the PCs is low, we would not expect the PCs to contribute much to the forecasting ability of the model. This is what we observe after running the FAVAR model with the full set of X variables. Other models outperform the FAVAR model in terms of forecasting inflation and output. Tables 3 and 5 provide the results for this framework. The DM test results (Table A1 and Table A3 in Appendix II) show that the differences are statistically significant for inflation, but nonsignificant for output.

After failing to identify strong factors from the overall data set that explain the movements in the economy, we select a subset of variables that are more closely linked to the target variables of the model. Although the set of all observable variables fails to characterize the economy using a few PCs, it may be possible for a smaller set to do so. If so, we may find a set of variables, the variation of which can be explained by a few PCs, that are correlated with the target variable. To identify these variables, we use a simple correlation between each of the X variables and the target variable. Then, we choose all variables that have a significant correlation coefficient, forming two new groups. The new group for inflation (CPI group) contains 32 variables, and that for GDP (GDP group) contains 33 variables. Henceforth, we use the CPI and GDP groups (not the whole data set) to extract PCs for the estimations of the respective FAVAR models.

The performance of the PCs extracted from the subgroups is significantly better than that based on all X variables. In both groups, the first three PCs explain more than 50 percent of the variation, and the first five PCs explain more than 70 percent. Next, we construct a FAVAR model using the PCs extracted from the new groups. Note that we choose the number of PCs such that they explain at least 50 percent of the total variation in the group. In addition, to provide a clearer view of the results, we construct models using one PC only. However, the results show that the models with three PCs outperform those with one PC; therefore we provide results for the former models only.

The benchmark model for inflation is the ARIMA model. However, we also estimate an existing VAR model based on the inflation forecasting model used by the CBAR, which is explained in Rahimov *et al.* (2016). The benchmark model for GDP is the ARIMA model.

The results for inflation are presented in Tables 3 and 4 in Appendix II. Table 3 presents the results based on 77 variables, and Table 4 shows those based on 32 variables. The number of variables used to extract the factors is the only difference between the two models; consequently, the results differ only in the factor models. In order to compare the RMSE values, we present the RMSEs of other models relative to that of the benchmark model (ARIMA). Thus, values less than one represent good forecasting performance, and values greater than one indicate poor performance.

Relative RMSE of the models for CPI forecasting.

AR: autoregressive model; VAR: vector autoregressive model; SW: Stock and Watson-type model; FAVAR: Factor-augmented VAR. MOD extension means that forecasts by international organizations have been used for foreign variables

Table 3: 77 variables

Relative RMSE					
	1p	**2p**	**4p**	**6p**	**8p**
AR	**1.00**	**1.00**	**1.00**	1.00	1.00
VAR	1.27	1.22	1.13	1.14	1.07
VAR - MOD	1.21	1.14	1.01	**0.99**	**0.93**
SW	1.33	1.22	1.24	1.26	1.20
FAVAR	1.35	1.38	1.27	1.44	1.37
FAVAR - MOD	1.35	1.36	1.25	1.18	1.09
BVAR	1.07	1.07	1.12	1.09	1.00
BVAR - MOD	1.05	1.04	1.08	1.05	0.99

Table 4: 33 variables

Relative RMSE					
	1p	**2p**	**4p**	**6p**	**8p**
AR	**1.00**	**1.00**	**1.00**	1.00	1.00
VAR	1.27	1.22	1.13	1.14	1.07
VAR - MOD	1.21	1.14	1.01	**0.99**	**0.93**
SW	1.30	1.32	1.26	1.27	1.12
FAVAR	1.54	1.33	1.27	1.48	1.61
FAVAR - MOD	1.44	1.22	1.36	1.38	1.45
BVAR	1.07	1.07	1.12	1.09	1.00
BVAR - MOD	1.05	1.04	1.08	1.05	0.99

The results show that the univariate inflation model outperforms the multivariate models up to one year. For longer horizons (six and eight periods), the VAR model outperforms both the benchmark and the other models. However, a DM test (Tables A1 and A2 in Appendix II) shows that the results are not statistically significant. The factor models perform poorly compared with both the univariate and other multivariate models. Both the SW and FAVAR models have relatively higher RMSEs, for all forecast horizons. Using the forecast values provided by international organizations for global variables improves the model performance slightly, in general.

The results for real non-oil output are presented in Tables 5 and 6. In this case, the FAVAR model (from the reduced data set) outperforms the univariate model in all horizons except one, but performs relatively poorly compared with the VAR model. The interesting finding is that most of the models, particularly those forecasting more than one period ahead, outperform the simple ARIMA model. However, the results of the DM tests (Tables A3 and A4 in Appendix II) show that these differences are not statistically significant, which again suggests that making models more complex does not necessarily improve their performance.

Relative RMSE of the models for RGDP forecasting.

Table 5: 77 variables

Relative RMSE					
	1p	**2p**	**4p**	**6p**	**8p**
AR	1.00	1.00	1.00	1.00	1.00
VAR	**0.95**	**0.84**	**0.78**	**0.85**	0.85
VAR - MOD	**0.95**	0.87	0.84	**0.85**	**0.78**
SW	1.18	0.92	0.87	0.98	1.03
FAVAR	1.17	0.94	0.82	0.92	0.88
FAVAR - MOD	1.15	0.94	0.85	1.01	0.88
BVAR	1.22	0.90	**0.78**	0.99	0.94
BVAR - MOD	1.22	0.91	0.92	1.01	0.97

Table 6: 33 variables

Relative RMSE					
	1p	**2p**	**4p**	**6p**	**8p**
AR	1.00	1.00	1.00	1.00	1.00
VAR	**0.94**	**0.84**	**0.78**	**0.85**	**0.84**
VAR - MOD	**0.95**	0.87	0.84	0.85	**0.78**
SW	1.12	0.83	0.81	0.88	0.90

FAVAR	1.10	0.85	0.81	0.94	0.78
FAVAR - MOD	1.11	0.88	1.09	1.19	1.14
BVAR	1.17	0.90	0.76	0.83	0.75
BVAR - MOD	1.17	0.90	0.79	0.83	0.76

Appendices III and IV compare the *forecast performance of the models for different periods*. Appendix III (a) presents a graph of forecasts, along with actual data; Appendix IV (a) presents graphs of the errors for selected inflation models; Appendix III (b) and Appendix IV (b) present similar graphs for non-oil GDP. To save space, we present the results for one, four, and eight period forecasts only. The main purpose of these graphs is to determine whether certain models (especially FAVAR-type models) perform better in particular periods. For example, it may be that complicated models perform better in turbulent periods. However, a close examination of the graphs reveals that the FAVAR models produce larger errors in almost all periods.

Complicated FAVAR models do not outperform simple models in terms of forecasting inflation or non-oil output. Our results are in line with the literature that shows that simple models exhibit better forecasting ability (Atkeson & Ohanian, 2001, Edge et al., 2010, Huseynov *et al.*, 2014, Duncan & Martínez-García, 2018). Given that Azerbaijan is a small resource-rich economy, with significant government ownership, this result is not surprising. However, such results may also stem from the short data sample period or poor data quality. One of the reasons why models, and especially complicated models, do not forecast inflation accurately in Azerbaijan is high inflation volatility, even after seasonally adjusting and differencing the variables. Similar volatility is evident in the case of the non-oil GDP variable as well.

Conclusion

We construct and use FAVAR models to forecast inflation and output for Azerbaijan. Conducting a PC analysis for the period 2003 to 2018, we first extract factors from a large set of variables, and then use these factors to construct the FAVAR model. We compare the forecasting performance of the FAVAR model with that of univariate and multivariate models for one and multiple-period horizons. The results show that for forecasting inflation, the univariate model outperforms the multivariate models, including the FAVAR model. For forecasting non-oil output, the FAVAR model and other multivariate models do not outperform the univariate model. Thus, our results support those in the literature, which finds that simpler models exhibit better forecasting ability.

We acknowledge that the results might be affected by the relatively short length of the sample period and the existence of irregularities in the data (structural breaks during this period limited the quality of the indicators used to extract the factors). Another potential reason for the better performance of the simple model is the presence of a strong inertial component in many macroeconomic series, including inflation.

Bibliography

Ajevskis, V., & Davidsons, G. (2008). Dynamic Factor Models in Forecasting Latvia's Gross Domestic Product. *Latvijas Banka*.

Atkeson, A., & Ohanian, L. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review, Vol. 25, No. 1, Winter 2001, pp. 2–11*.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*.

Bernanke, B., Boivin, J., & Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (VAR) approach. *The Quarterly Journal of Economics, Vol. 120, Issue 1, pp. 387–422*.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics, Vol. 132, No. 1, pp. 169–194*.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, Vol. 13, No. 3, p. 253*.

Duncan, R., & Martínez-García, E. (2018). New Perspectives on Forecasting Inflation in Emerging Market Economies: An Empirical Assessment. *Federal Reserve Bank of Dallas Globalization and Monetary Policy Institute, Working Paper No. 338*.

Edge, R.M., Gürkaynak, R.S., Reis, R., & Sims, C.A. (2010). How useful are estimated DSGE model forecasts for central bankers? *Brookings Papers on Economic Activity, pp.209–259*.

Eickmeier, S., & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting, Vol. 27, No. 3, pp. 237–265*.

Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association, Vol. 76, Issue 376, pp. 774–781*.

Forni, M., Hallin, M., & Lippi, M. (2003). Do financial variables help forecasting inflation and real activity in the EURO area? *Journal of Monetary Economics, Vol. 50, Issue 6, pp. 1243–1255*.

Günay, M. (2018). Forecasting industrial production and inflation in Turkey with factor models. *Central Bank Review, Vol. 18, pp*. 149–161.

Huseynov, S., & Mammadov, F. (2016). A small-scale forecasting and simulation model for Azerbaijan. *Central Bank of The Republic of Azerbaijan, Working paper series, N 08/2016*.

Huseynov, S., Ahmadov, V., & Adigozalov, S. (2014). Beating a random walk: "Hard times" for forecasting inflation in post-oil boom years? *MPRA Paper No 63515.*

Kilian, L., & Lütkepohl, H. (2016). Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey. *Cambridge University Press*.

Laganà, G., & Mountford, A. (2005). Measuring monetary policy in the UK: A factor-augmented vector autoregression model approach. *Manchester School, Vol. 73, No. S1, pp. 77–98*.

Mammadov, F., & Adigozalov S. (2014). Indicator Based Forecasting of Business Cycles in Azerbaijan. *Munich Personal RePEc Archive, MPRA Paper 64367*

Quah, D., & Sargent, T. (1993). A dynamic index model for large cross sections. *National Bureau of Economic Research*, *pp. 285–310*.

Rahimov, V., Adigozalov, S., & Mammadov, F. (2016). Determinants of inflation in Azerbaijan. *Central Bank of The Republic of Azerbaijan, Working paper series No 07/2016*.

Reigl, N. (2017). Forecasting the Estonian rate of inflation using factor models. *Baltic Journal of Economics, Vol. 17, No. 2, pp. 152–189*.

Sargent, T., & Sims, C. (1977). Business cycle modelling without pretending to have too much a priori economic theory. *Federal Reserve Bank of Minneapolis, Working Paper 55*.

Sims, C. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review, Vol. 36, Issue 5, pp. 975-1000.*

Stock, J. & Watson, M. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, Vol. 97, No. 460, pp. 1167–1179*.

Stock, J., & Watson, M. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business Economics and Statistics, Vol. 20, No. 2, pp. 147–162*.

Stock, J., & Watson, M. (2010). Dynamic Factor Models. *In: Clements MP, Henry DF Oxford Handbook of Economic Forecasting. Oxford: Oxford University Press ; 2010.*

APPENDIX I. List of all variables and summary statistics of the important variables

A. Real sector

1. Nominal GDP
2. Nominal non-oil GDP
3. Real GDP
4. Real non-oil GDP
5. Nominal income of population
6. Real income of population
7. Industrial production
8. Output gap
9. Non-oil output gap
10. Oil production
11. Agricultural, hunting, and forestry value-added
12. Fishery value-added
13. Mining and quarrying value-added
14. Manufacturing value-added
15. Construction value-added
16. Wholesale and retail trade, repair of cars, household and personal goods value-added
17. Hotels and restaurants value-added
18. Production of storage, transport, and communication value-added
19. Financial intermediation value-added
20. Real estate, renting and business activities value-added
21. Public administration, defense, and social security value-added
22. Education value-added
23. Health and social work value-added
24. Other community, social, and personal service activities value-added

A. Prices

25. Headline inflation
26. Food CPI inflation
27. Nonfood CPI inflation
28. Service CPI inflation
29. Aggregate PPI inflation
30. Mining and quarrying PPI inflation
31. Manufacturing PPI inflation
32. Agricultural products PPI inflation
33. GDP deflator

B. Foreign sector variables

34. Bilateral exchange rate vis-à-vis USD
35. Nominal effective exchange rate (NEER)
36. Non-oil NEER
37. Import-weighted NEER
38. Export-weighted NEER
39. Real effective exchange rate (REER)
40. Non-oil REER
41. Import-weighted REER
42. Export-weighted REER
43. Brent oil price
44. World food prices
45. Inflation in trading partners
46. Demand from trading partners
47. Aggregate export
48. Non-oil export
49. Aggregate Import
50. Non-oil import

C. Monetary variables

51. CB policy interest rate
52. Short-term interest rates
53. Long-term interest rates
54. Volume of loans to the economy
55. Base money
56. M0 money supply (cash outside banks)
57. M1 money supply (M0 + demand deposits in local currency)
58. M2 money supply (M1 + time deposits in local currency)
59. M3 money supply (M2 + deposits in foreign currency)
60. Velocity of money
61. Net domestic assets
62. Net foreign assets

D. Government finance

63. Budget expenditure
64. Current expenditures
65. Capital expenditures
66. Government debt
67. Personal income tax in local currency
68. Corporate income tax in local currency
69. VAT in local currency
70. Customs duty

E. Labor market indicators

71. Number of population
72. Labor force
73. Number of employed people
74. Number of civil servants
75. Number of employees with contracts
76. Unemployment level
77. Nominal wages

Appendix II. Diebold–Mariano tests

Abbreviations correspond to the following models:

AR: autoregressive model

VAR: vector autoregressive model

SW: Stock and Watson type model

FAVAR: Factor-augmented VAR

The MOD extension means that, in the forecast process, forecasts by international organizations have been used for foreign variables.

DM test results of the models for CPI forecasting.

Table A1: 77 variables

Diebold–Mariano test					
	1p	**2p**	**4p**	**6p**	**8p**
VAR	0.01**	0.03**	0.04**	0.08*	0.27
VAR - MOD	0.06*	0.17	0.47	0.43	0.10
SW	0.00***	0.02**	0.03**	0.00***	0.01**
FAVAR	0.02**	0.01**	0.06*	0.05**	0.08*
FAVAR - MOD	0.02**	0.04**	0.06*	0.06*	0.26
BVAR	0.16	0.11	0.06*	0.14	0.49
BVAR - MOD	0.23	0.26	0.10	0.24	0.40

Table A2: 32 variables

Diebold–Mariano test					
	1p	**2p**	**4p**	**6p**	**8p**
VAR	0.01**	0.03**	0.04**	0.08*	0.27
VAR - MOD	0.06**	0.17	0.47	0.43	0.10
SW	0.02**	0.07*	0.03**	0.05**	0.2
FAVAR	0.00***	0.00***	0.02**	0.07*	0.11
FAVAR - MOD	0.00***	0.02**	0.01**	0.1*	0.14
BVAR	0.16	0.12	0.06*	0.14	0.49
BVAR - MOD	0.23	0.26	0.10	0.24	0.40

DM test results of the models for RGDP forecasting.

Table A3: 77 variables

Diebold–Mariano test					
	1p	**2p**	**4p**	**6p**	**8p**
VAR	0.40	0.23	0.18	0.28	0.30
VAR - MOD	0.40	0.29	0.26	0.29	0.25
SW	0.08*	0.35	0.32	0.46	0.44
FAVAR	0.12	0.40	0.19	0.36	0.31
FAVAR - MOD	0.15	0.39	0.22	0.48	0.32
BVAR	0.18	0.34	0.18	0.49	0.43
BVAR - MOD	0.18	0.35	0.39	0.48	0.47

Table A4: 33 variables

Diebold–Mariano test					
	1p	**2p**	**4p**	**6p**	**8p**
VAR	0.39	0.23	0.18	0.28	0.30
VAR - MOD	0.40	0.29	0.25	0.29	0.25
SW	0.14	0.18	0.21	0.26	0.33
FAVAR	0.25	0.16	0.16	0.40	0.18
FAVAR - MOD	0.23	0.22	0.14	0.06*	0.30
BVAR	0.18	0.34	0.18	0.49	0.43
BVAR - MOD	0.18	0.35	0.39	0.48	0.47

Appendix III (a). Comparison of model forecasts with those of the AR model in different forecasting horizons (inflation).

Figure A1: One-period-ahead inflation forecasts

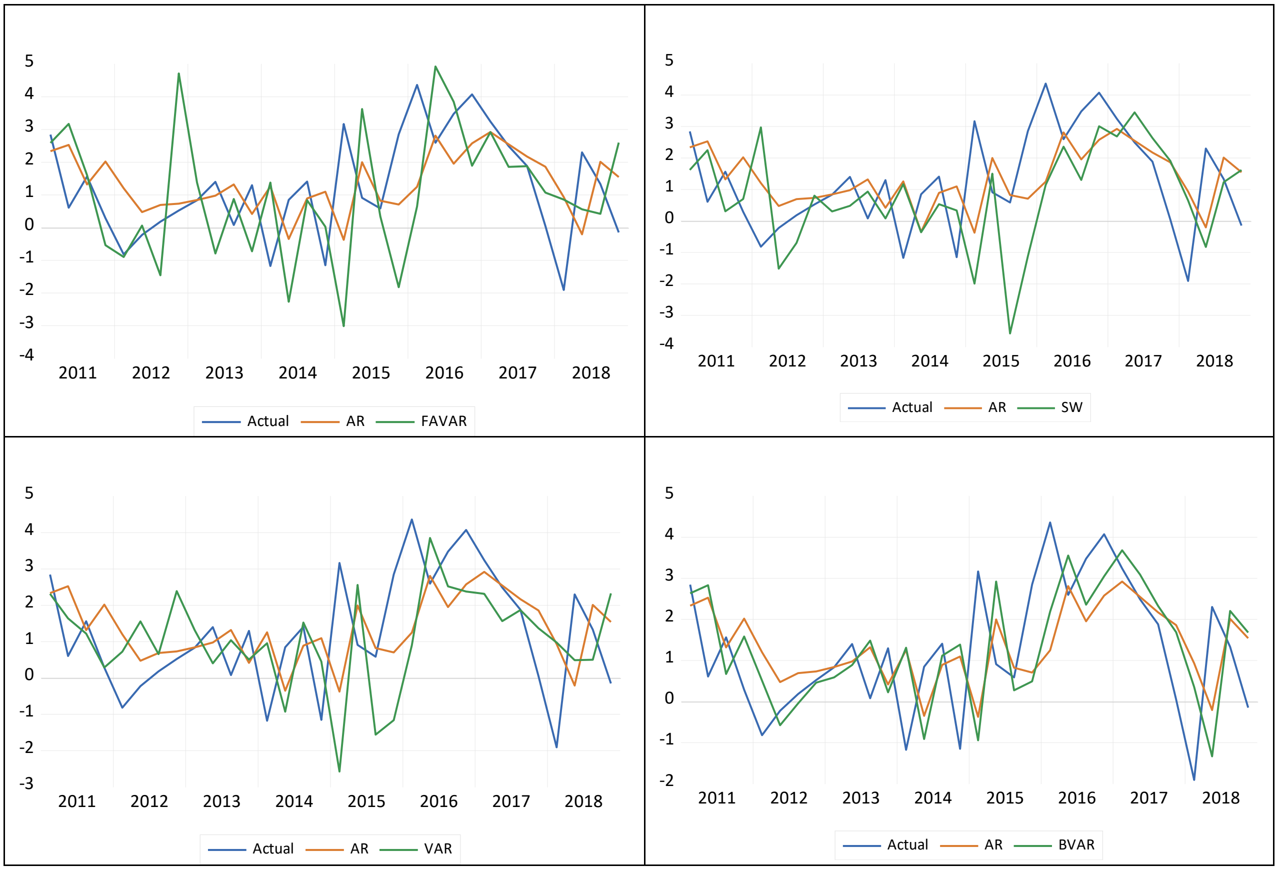
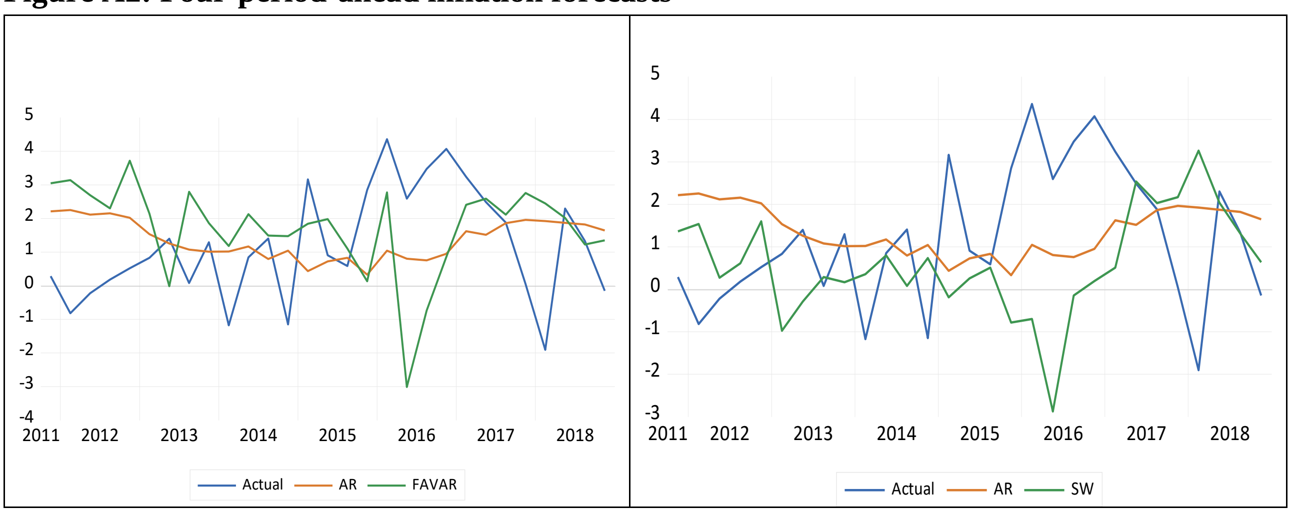


Figure A2: Four-period-ahead inflation forecasts



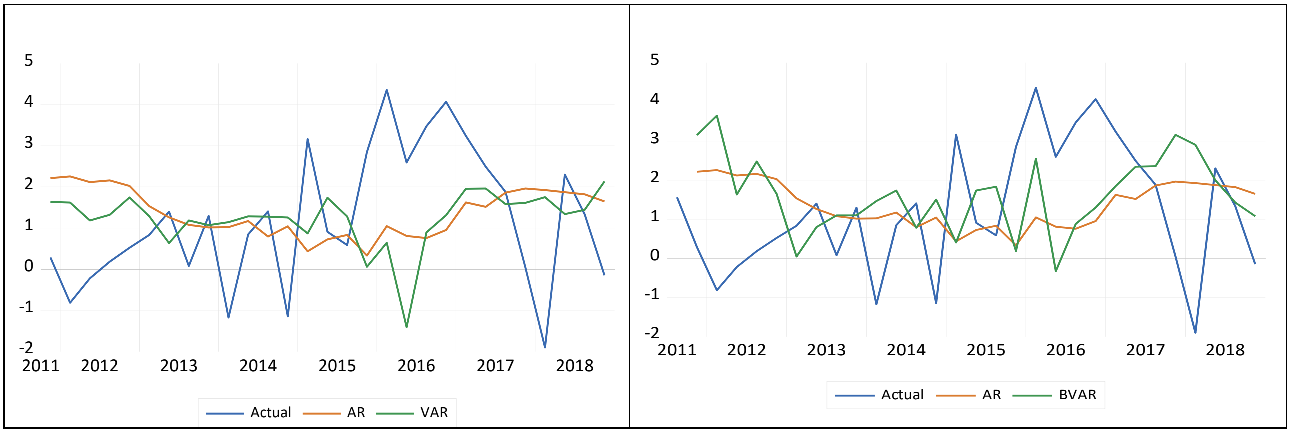
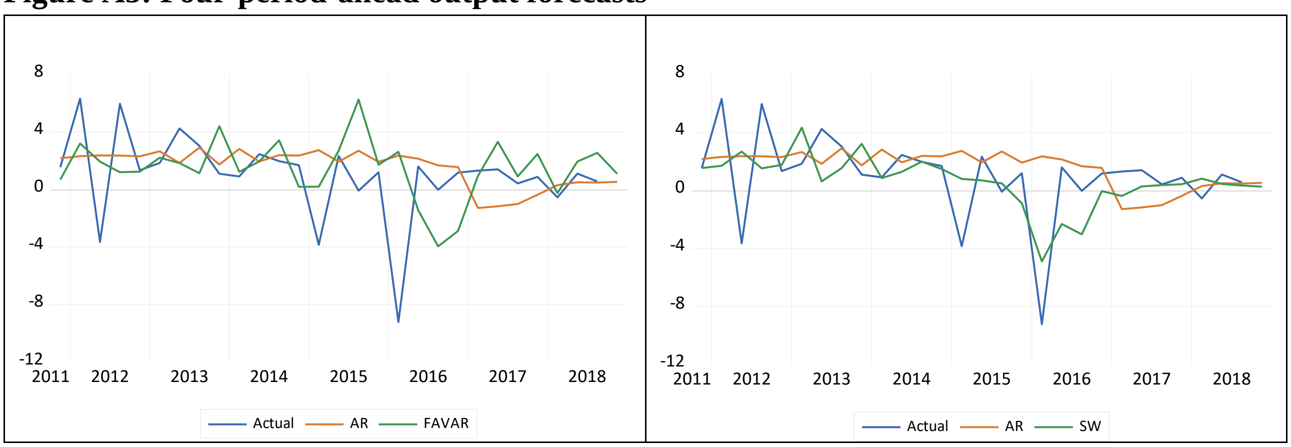


Figure A3: Eight-period-ahead inflation forecasts

Appendix III (b). **Comparison of model forecasts with those of the AR model in different forecasting horizons (output).**

Figure A4: One-period-ahead output forecasts

Figure A5: Four-period-ahead output forecasts



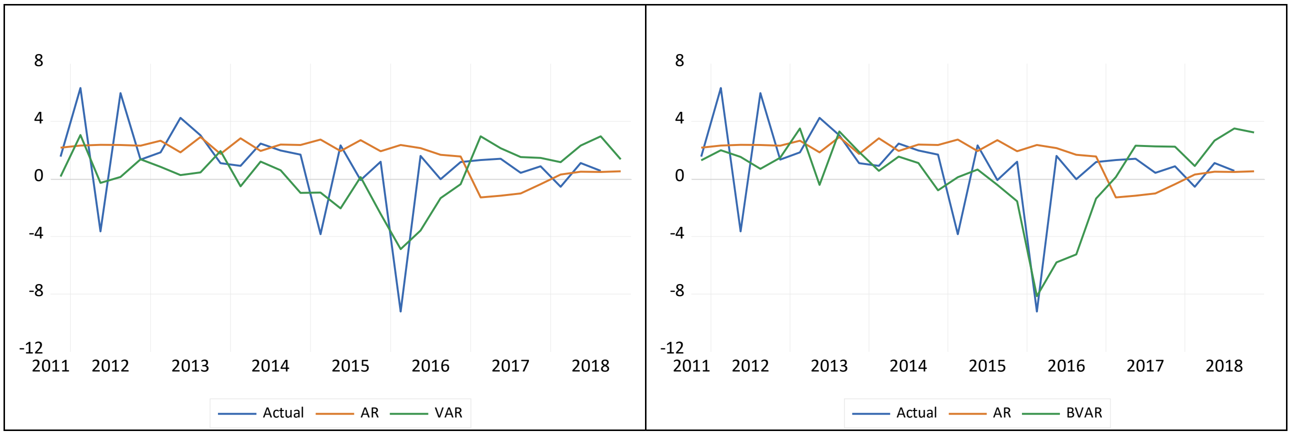


Figure A6: Eight-period-ahead output forecasts

Appendix IV (a). Forecast errors (inflation)

Figure A7: One-period-ahead inflation forecast errors

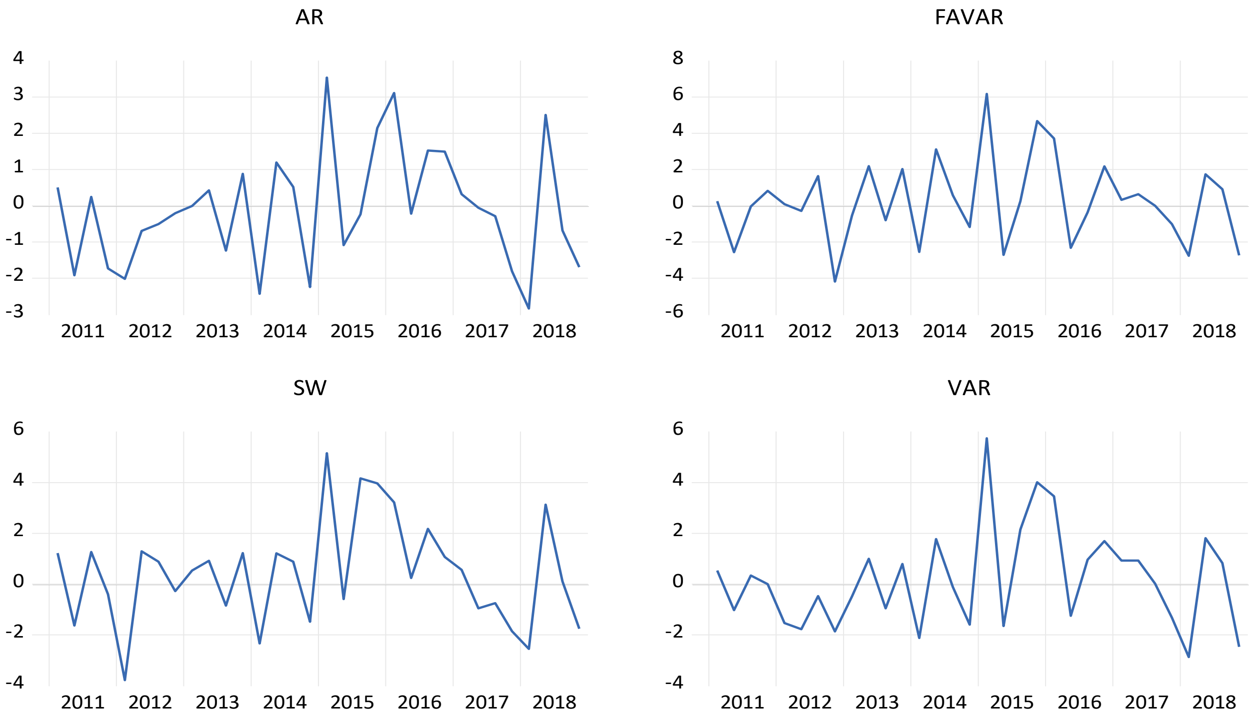


Figure A8: Two-period-ahead inflation forecast errors

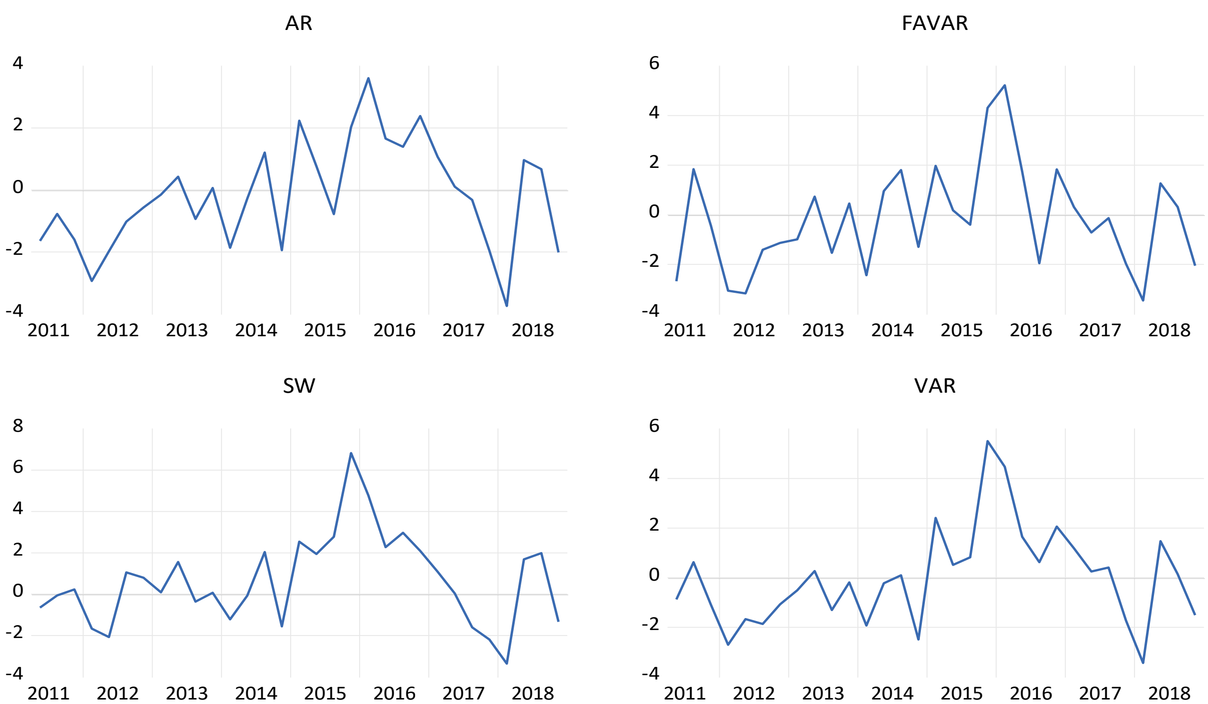


Figure A9: Four-period-ahead inflation forecast errors

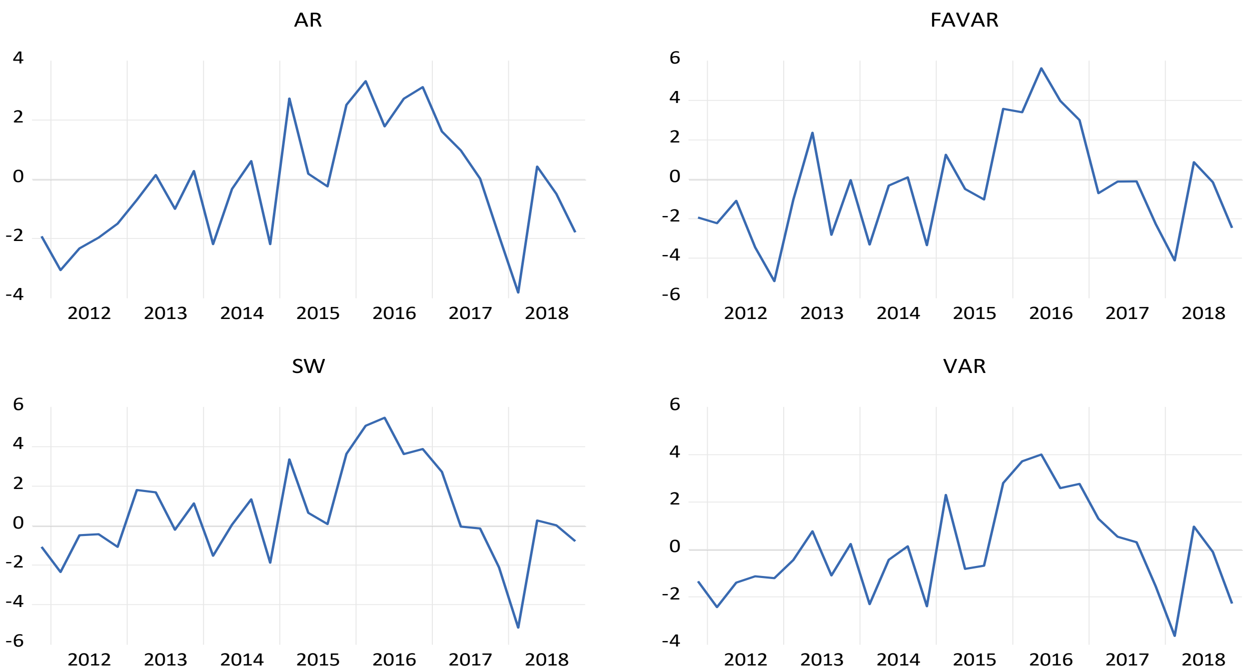


Figure A10: Six-period-ahead inflation forecast errors

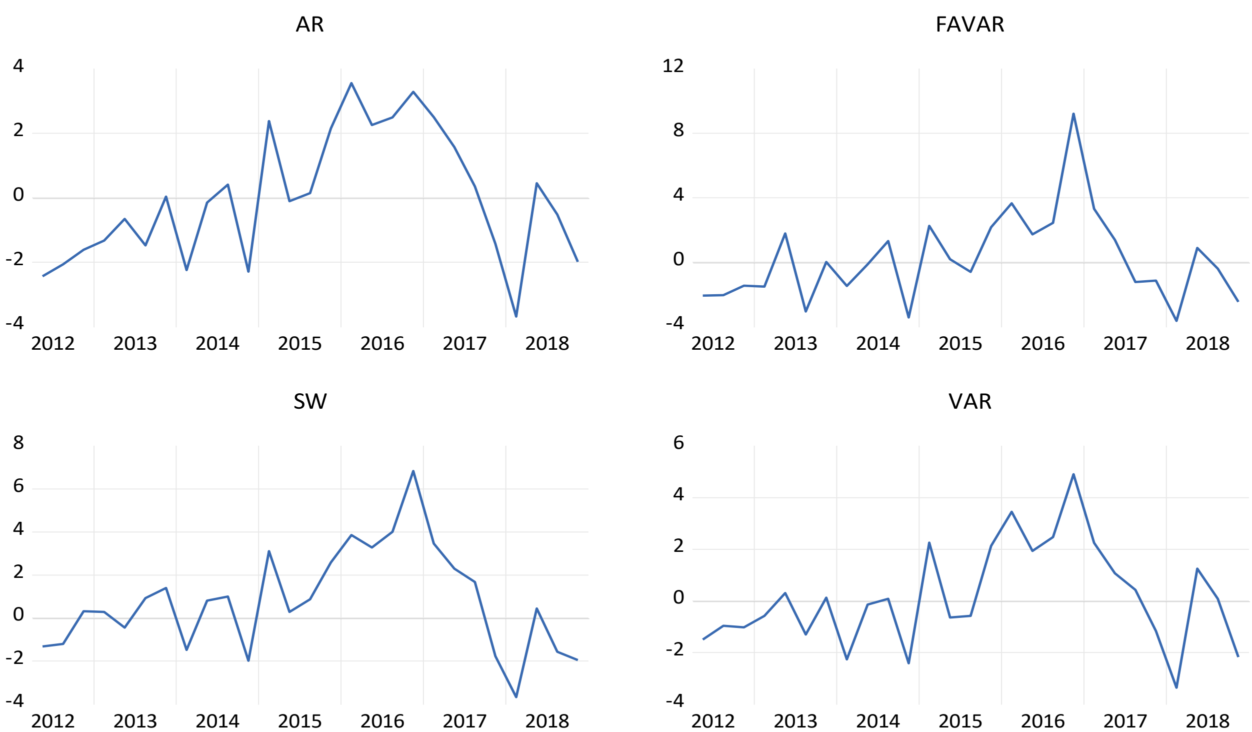
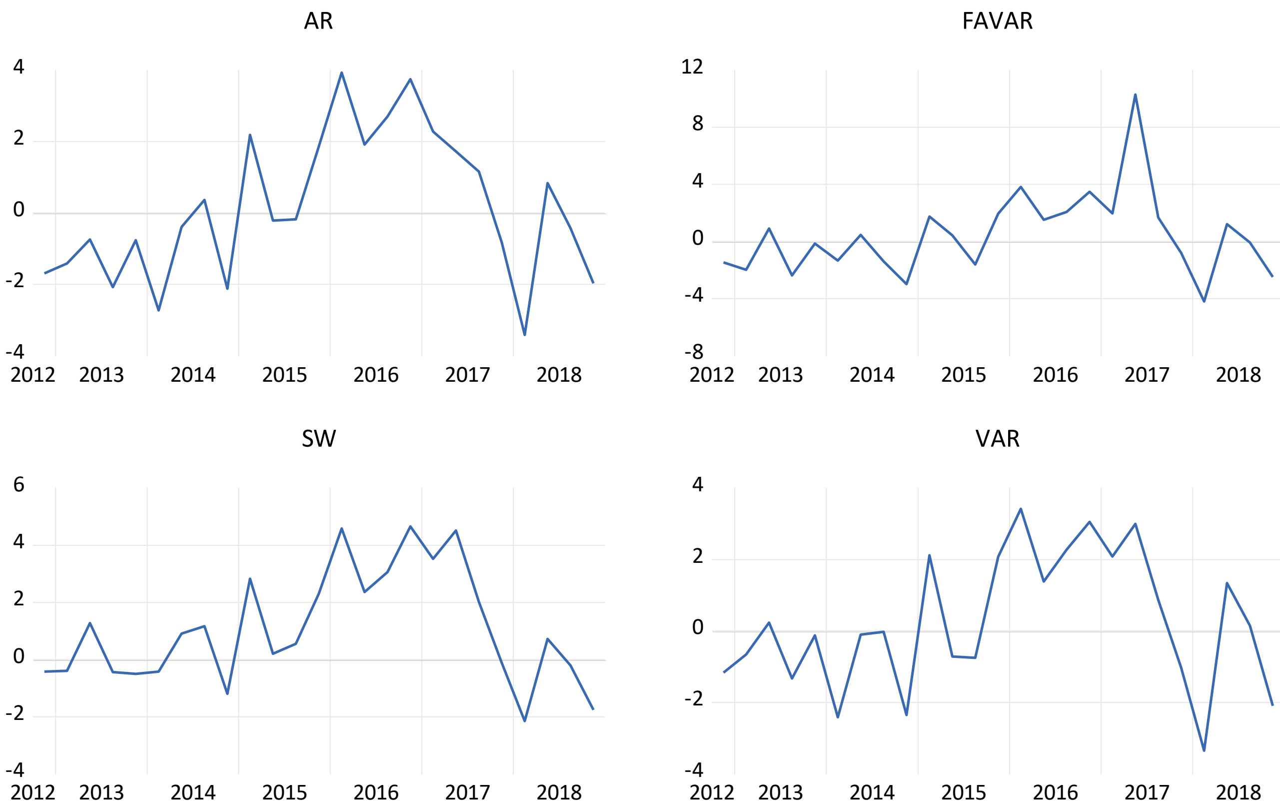


Figure A11: Eight-period-ahead inflation forecast errors



Appendix IV (b). Forecast errors (output)

Figure A12: One-period-ahead output forecast errors

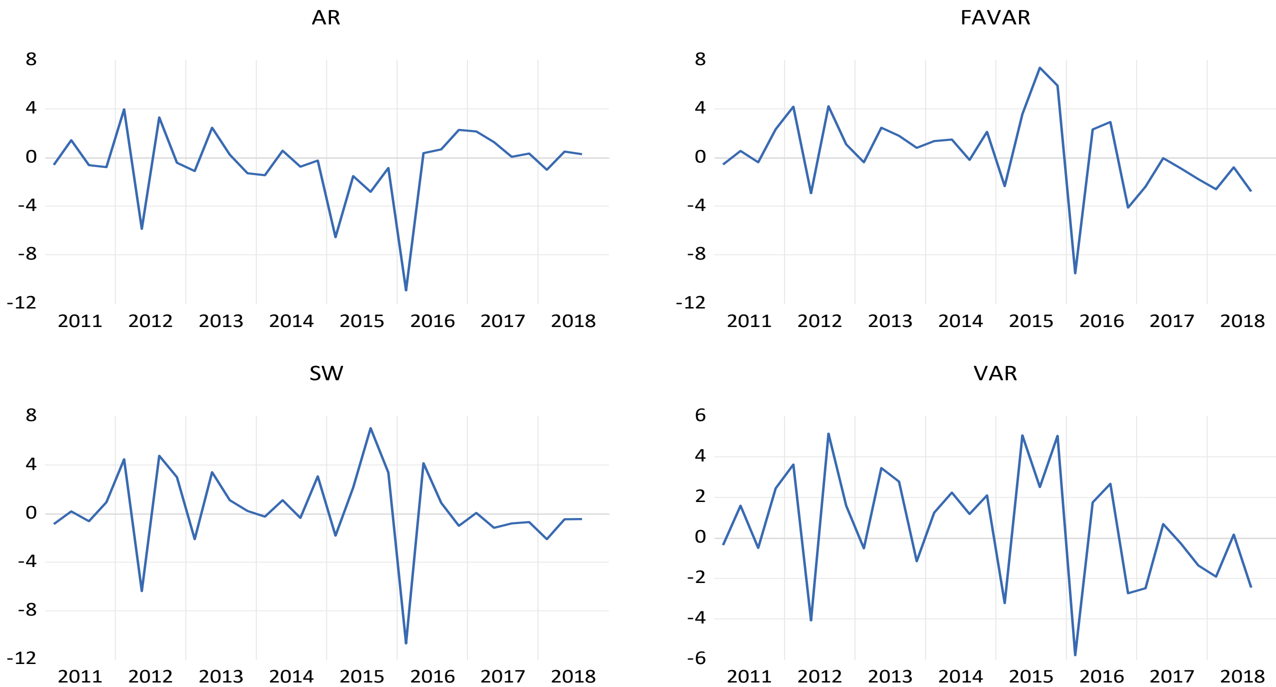


Figure A13: Two-period-ahead output forecast errors

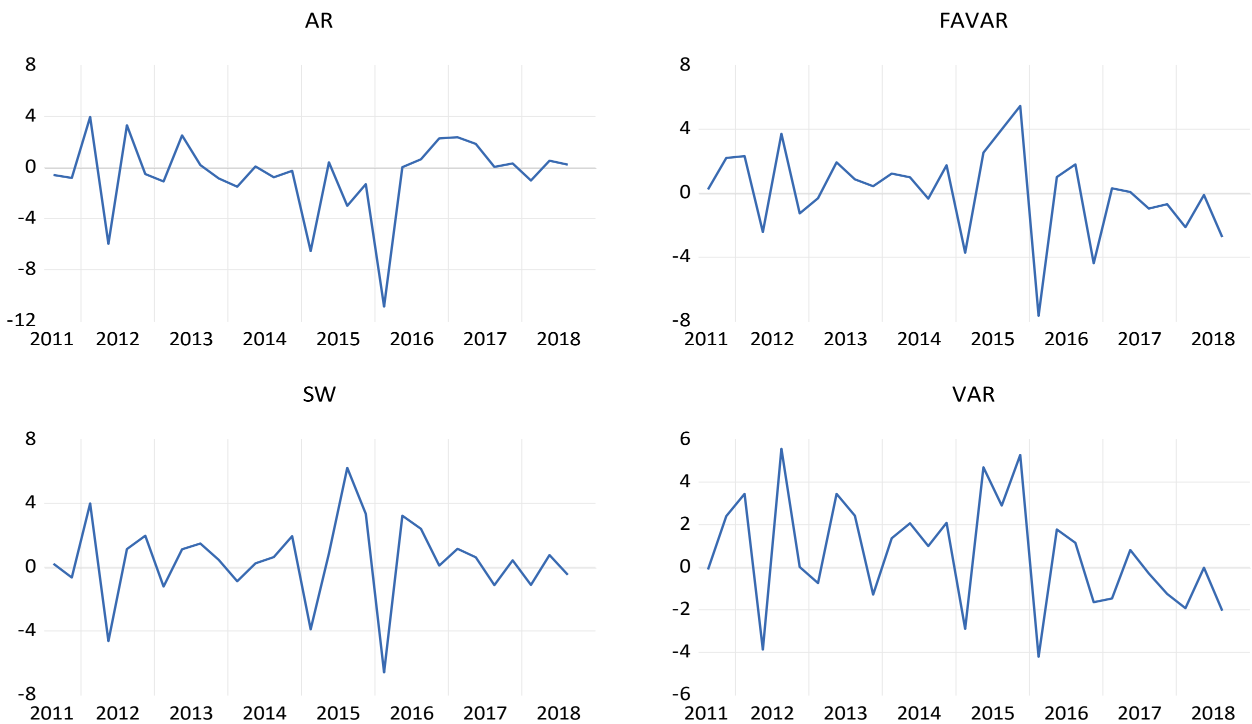


Figure A14: Four-period-ahead output forecast errors

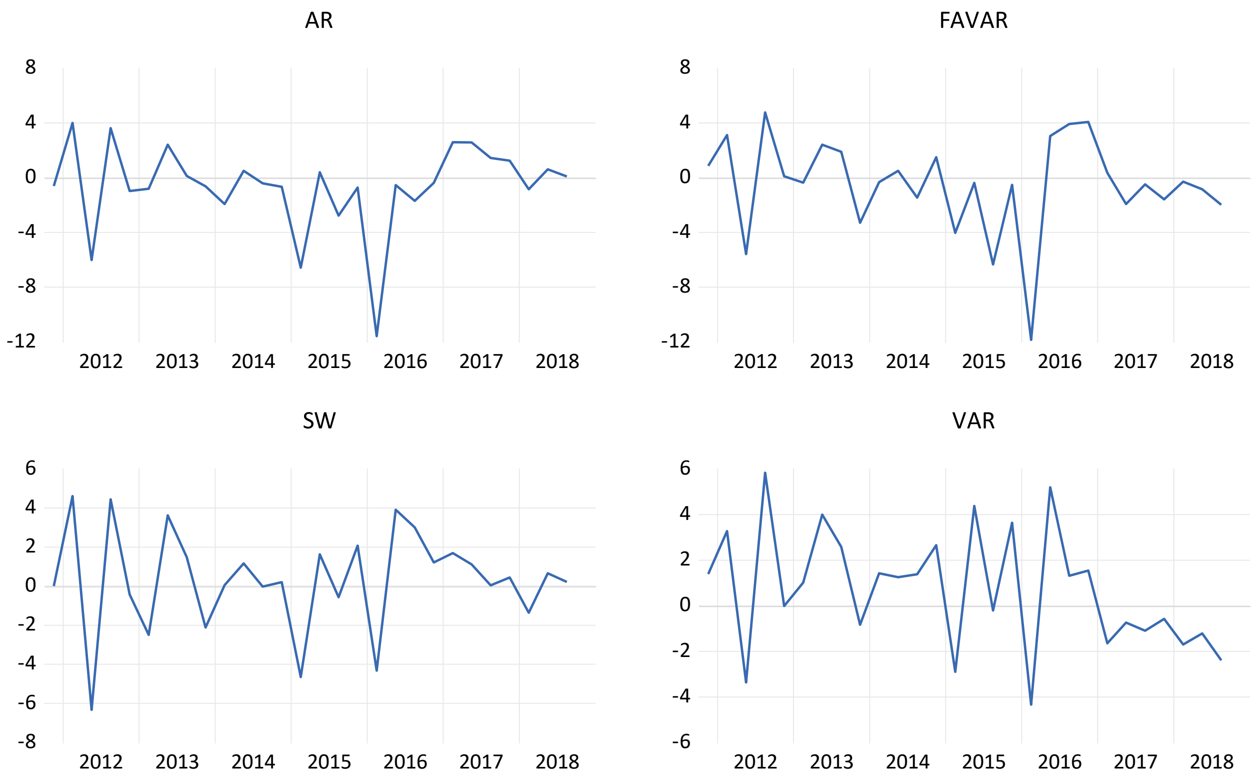


Figure A15: Six-period-ahead output forecast errors

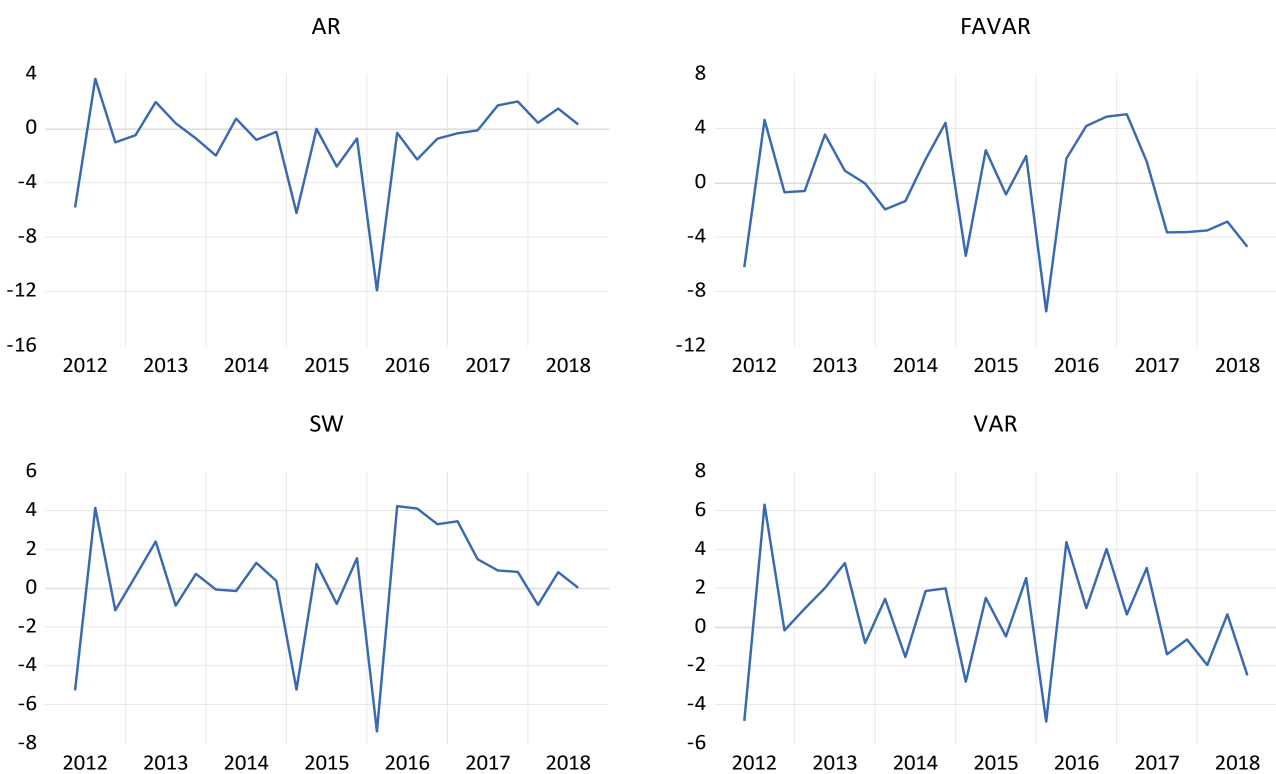
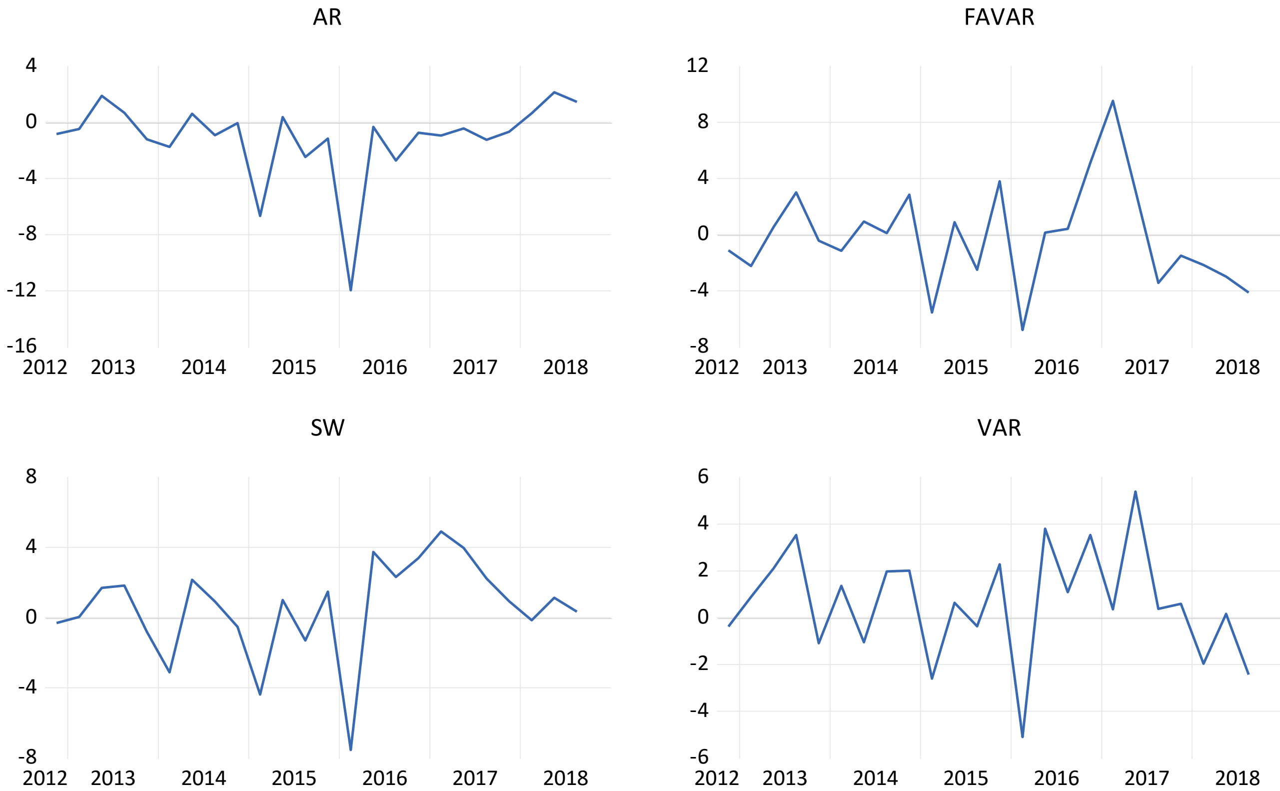


Figure A16: Eight-period-ahead output forecast errors



Appendix V. Impulse response functions.

Figure A17: Impulse response functions for inflation (VAR)

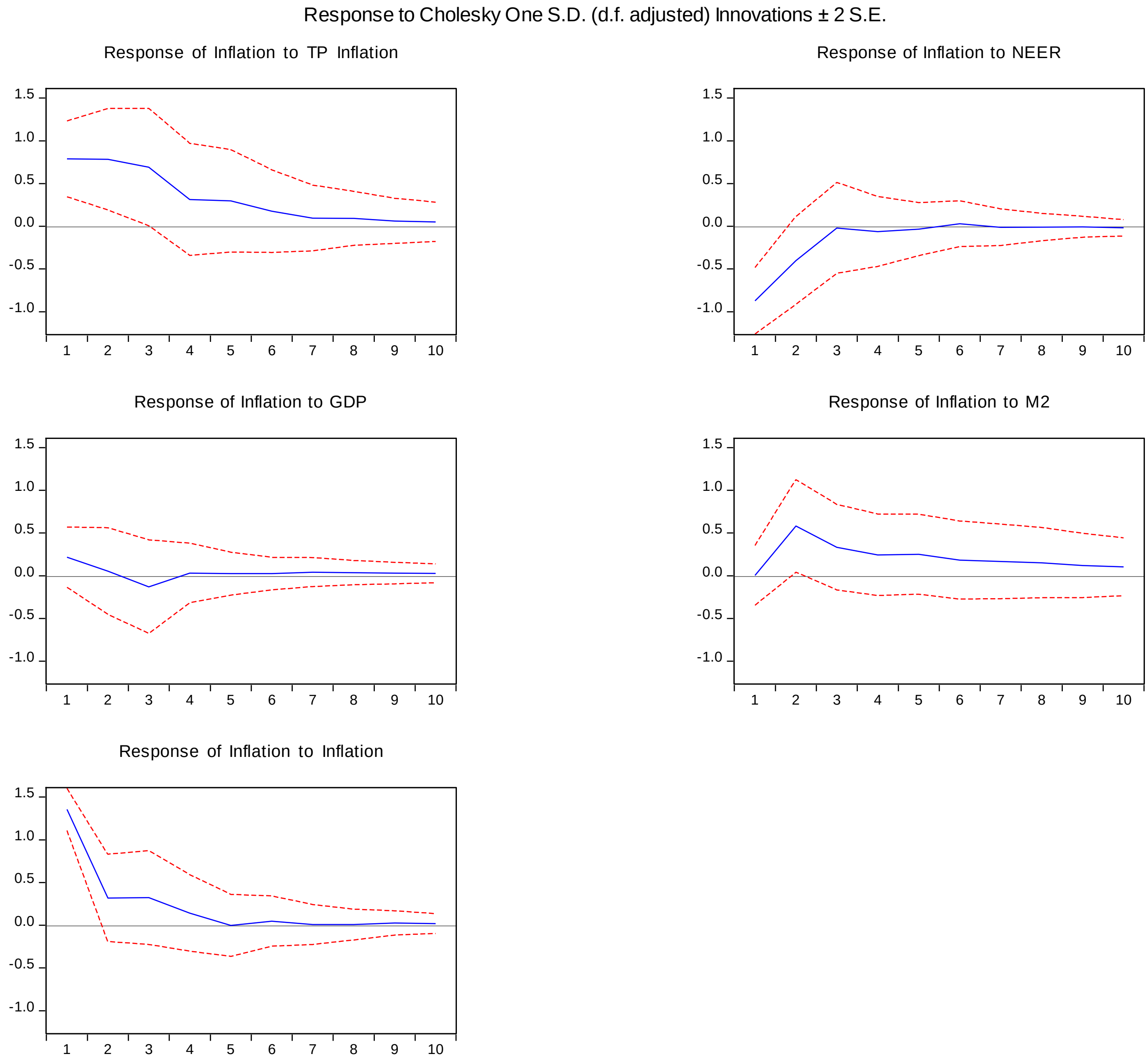


Figure A18: Impulse response functions for inflation (SW)

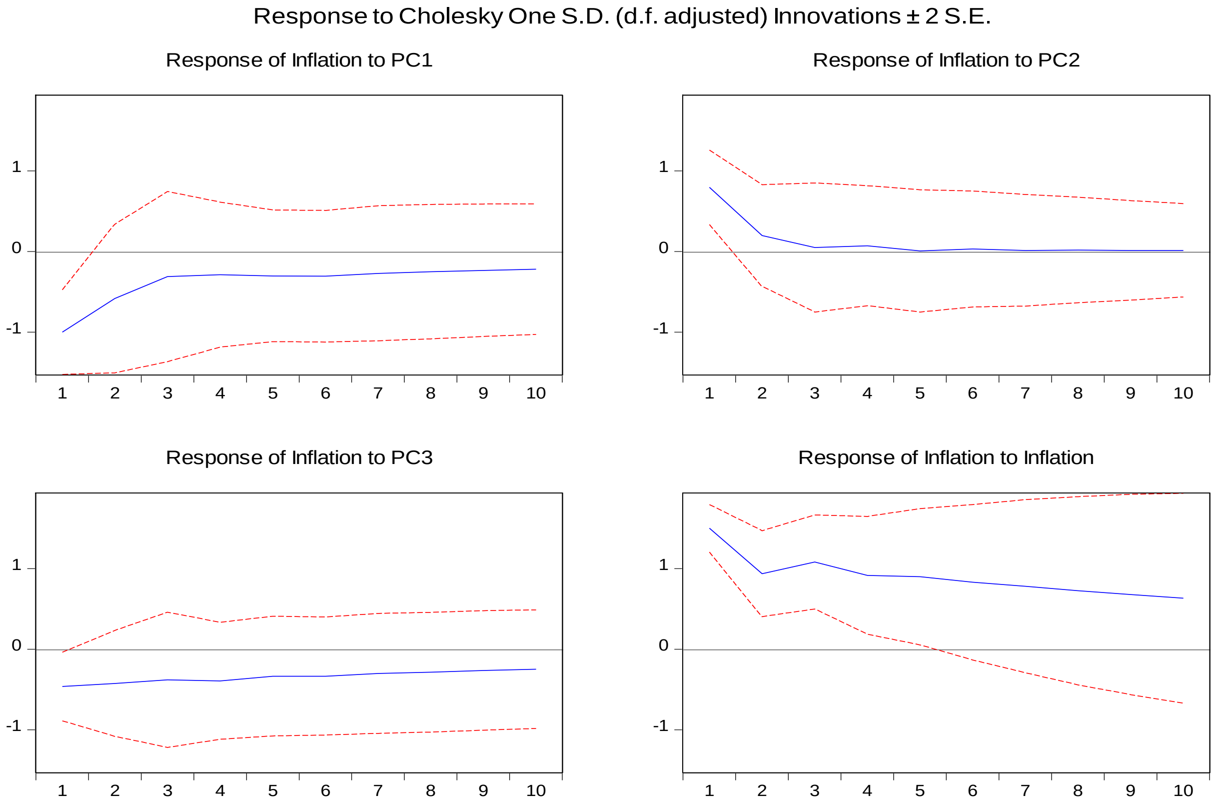


Figure A19: Impulse response functions for inflation (FAVAR)

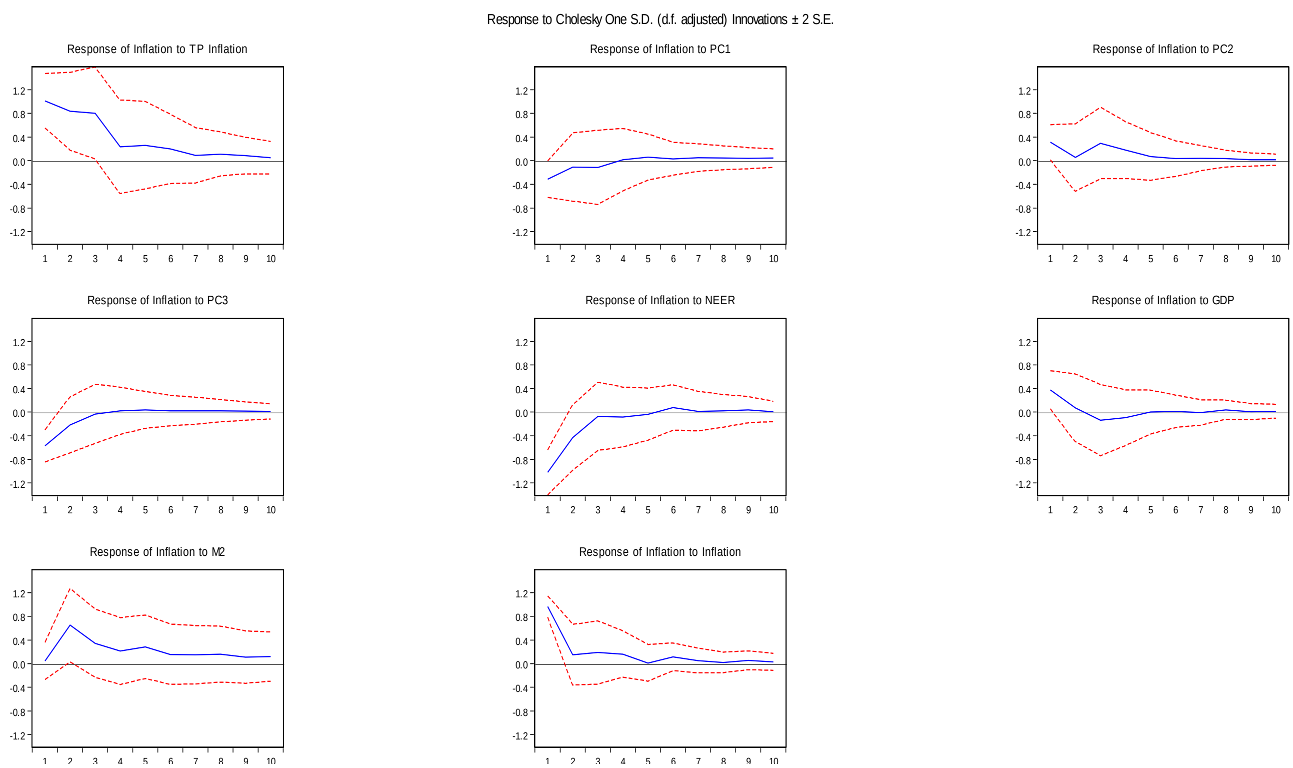


Figure A20: Impulse response functions for output (VAR)

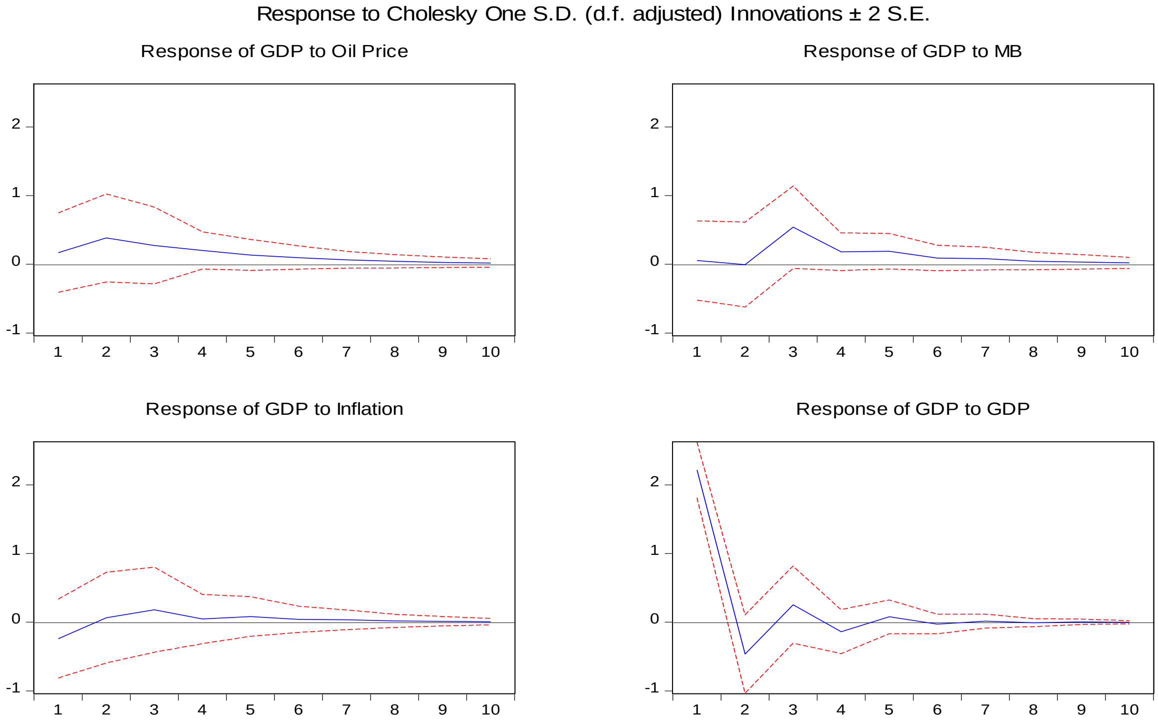


Figure A21: Impulse response functions for output (SW)

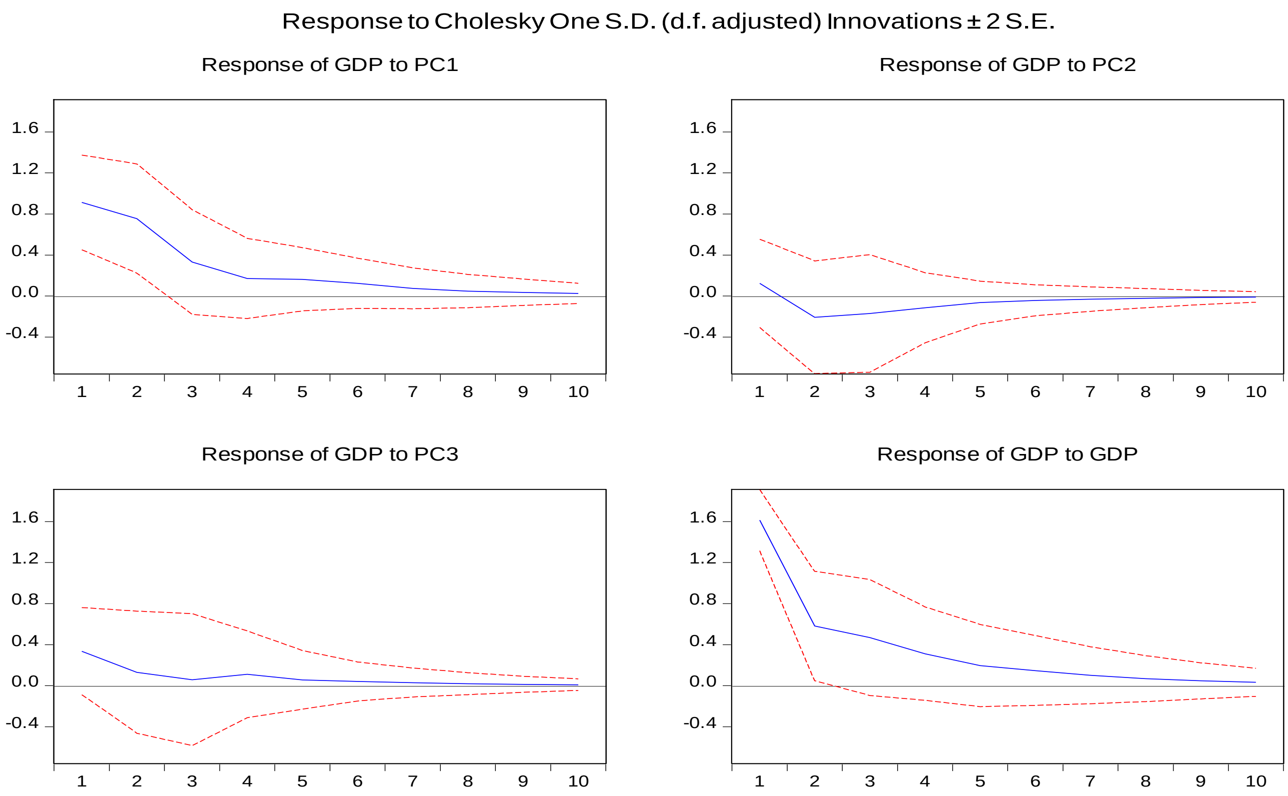
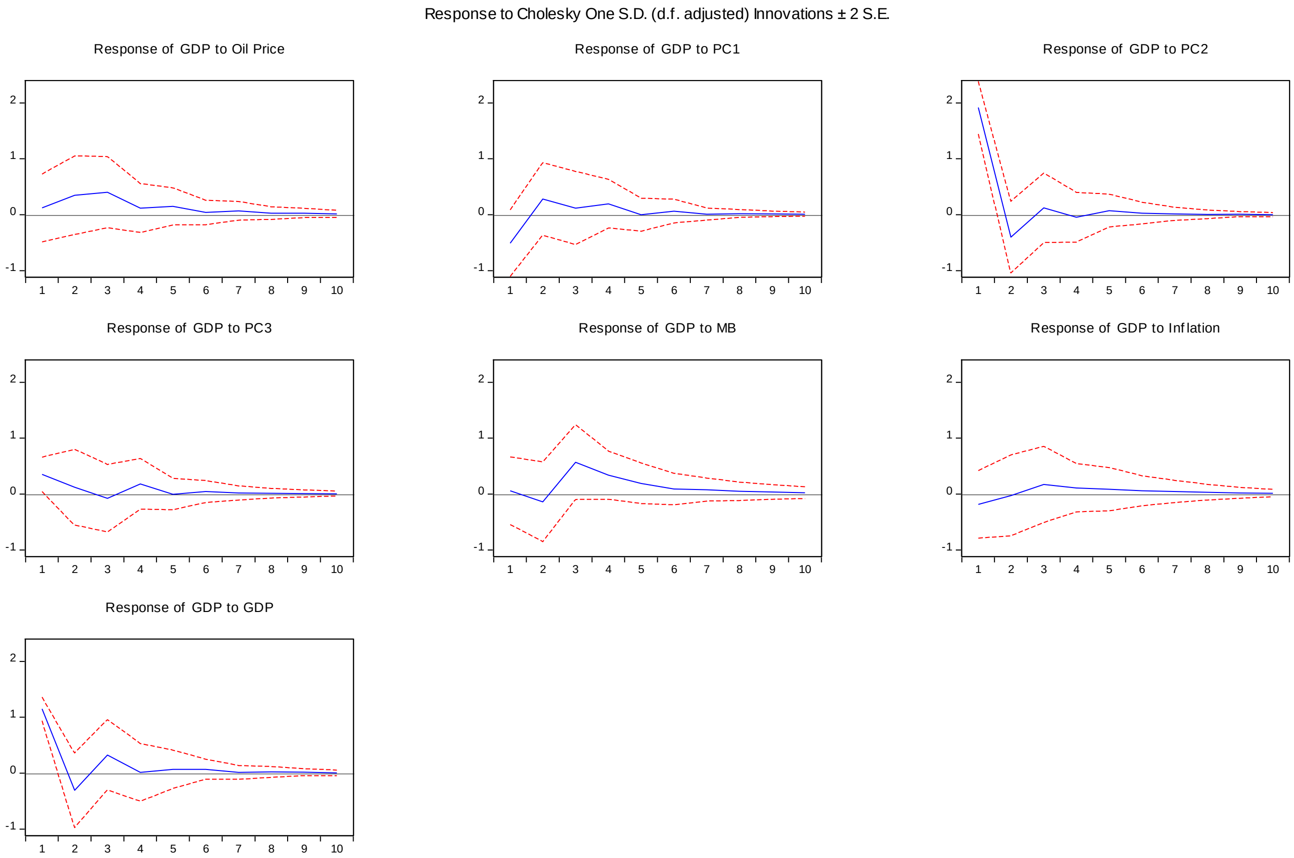
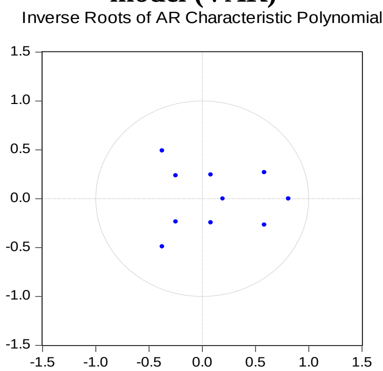


Figure A22: Impulse response functions for output (FAVAR)

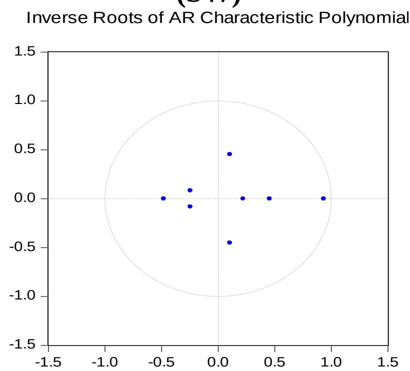


Appendix VI. Stability tests (full sample)

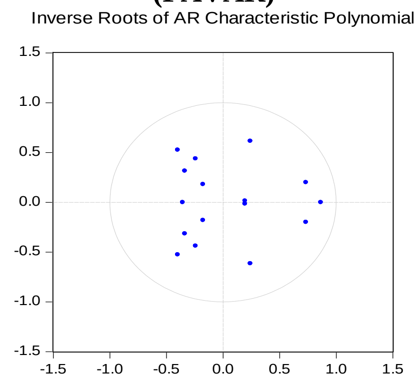
Stability test for inflation model (VAR)



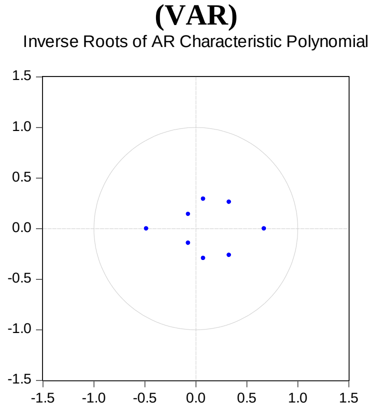
Stability test for inflation model (SW)



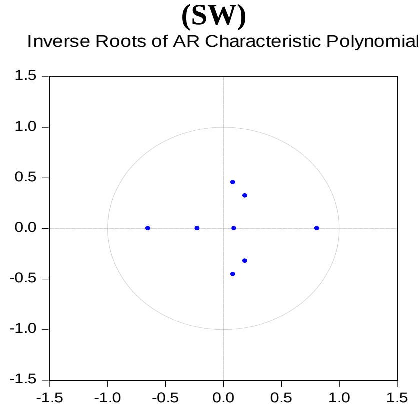
Stability test for inflation model (FAVAR)



Stability test for output model (VAR)



Stability test for output model (SW)



Stability test for output model (FAVAR)