Национальный исследовательский университет

"Высшая школа экономики"

Факультет экономических наук

Департамент Прикладной экономики

Выпускная квалификационная работа

"Использование модели FAVAR для прогнозирования российских макроэкономических рядов".

Гармидер Пётр, БЭК 165

Научный руководитель:

Борис Демешев,
Старший преподаватель Факультета Экономических Наук,
Департамента Прикладной Экономики

Москва

17 Апреля 2020

Contents

Abstract

There are several models used for time series forecasts that focus on incorporating a huge set of information. Conventional VAR fails to employ such an approach because of the "degrees of freedom" problem. The focus of this paper is to examine factor-augmented VAR's (FAVAR's) ability to forecast the main Russian macroeconomic indicators. Results will be displayed in a table of RMSEs for different horizons using time series cross-validation procedure for each one. I discovered that FAVAR, indeed, has potential in forecasting and create particularly accurate forecasts for long-horizons. FAVAR's cross-validation error is compared with benchmark models such as ETS, ARIMA, and VAR. Latent factors in model, surely, delivers additional information, that standard VAR fails to recover.

1 Introduction

1.1 Overview

Forecasting of macroeconomic time series is an important task for different economic institutions. Accurate forecasts for main macro indicators allow central banks to timely react to possible dangers for the economy. Proper actions conducted by the central bank may prevent a possible recession or the same way, wrong actions may result in initiating one.

Macro indicators forecasts are useful for commercial companies as well. There are many scenarios in which corporations are particularly interested in understanding possible future macro factor outcomes, such as: choosing the right price policy, signing long-term contracts with counterparties, making decisions on the possibility of entering a new market, etc. One cannot deny that a huge amount of financial instruments heavily dependent on macroeconomic situation within the country. Investors in attempts to value a particular financial asset always build a forecast of its determinants. In

this way, almost everything to some extent shows comovement with one or another macro factor. This demonstrates that forecasting macroeconomic time series is an extremely crucial part of quite a huge range of spheres.

There are a number of models used for time series prediction. Each model uses a different approach to dealing with time series. Some approaches do not require any assumptions about data and aimed to minimize proper loss function in a specific manner. The other includes restrictions on data, assuming, for example, the presence of the data-generating process with parameters that are to be estimated. State-of-the-art models, usually employ mixed approaches where forecast is based on the results of the two mentioned methods. Models of the first type show high accuracy of prediction, however, they are hard to interpret and are prone to outliers. Models of the second type, on the contrary, show moderate accuracy, though give information about moving forces and produce quite robust results. The last approach, tries to balance between advantages and disadvantages of the two previous approaches. Of course, there are plenty of other models with different views on data.

One can divide the parametric of models onto two subcategories: univariate and multivariate. To make a forecast for the considered time series more accurate, the last method tries to use additional data, while first assumes the data generating process of a particular series to be a function of past realizations of itself and only. This paper mostly deals with multivariate models vector autoregression (VAR) suggested by Sims (1980) and its variation factor-augmented VAR that aimed to eliminate drawbacks of classical version.

In this paper I will consider factor-augmented vector autoregression (FAVAR): method of estimation, its application for forecasting purposes, and accuracy evaluation of produced forecast. The goal is to understand,

whether the factor model outperform univariate approach. One also may be interested in comparing classical VAR with its augmented version. It will be shown, that VAR is a restricted model of FAVAR, therefore it is possible to test whether the difference between the two is statistically significant.

1.2   Literature review

Several studies discuss FAVAR model and employ its potential for particular purposes. Bernanke et al. (2005) were first to introduce a model to incorporate a huge set of information to assess structural shocks effect on monetary policy. They discuss Sims (1992) and his interpretation of the "price puzzle" when several papers showed that VAR models predict that contractionary monetary policy shock is followed by a rise in a price level, rather than decrease as traditional macroeconomic theory suggests. Authors believe that the explanation proposed by the inventor of VAR yet had weak points, therefore suggest their way to deal with "price puzzle". In the paper authors used 120 monthly time series pointing out that central banks while making decisions on interest rate act in a "data-rich environment", considering a larger set of time series than standard VAR approach allows to include. Therefore, econometrician, ignoring this fact, faces a problem of biased estimates. They suggested two ways of how FAVAR can be estimated and one of them will be presented as pseudo-code in this work. The main goal of model inventors was to obtain correct impulse responses for variables of interest rather than producing forecasts for them.

Nevertheless, there are some works focused on applying considered model for prediction task as well, for example, Mönch (2008) in his work uses FAVAR to forecast yield curve. The author mentions that state-of-the-art models in this sphere have a large number of parameters to estimate, which

means that users cannot simply include as many explanatory variables as one wishes, so there is a risk to lose important information. Therefore, the author used the approach of latent factor extraction described in Bernanke et al. (2005) and mixed it with traditional models for yield curve prediction.

Most of the papers agree, that for the time series prediction factor-augmented VAR shows better performance for long-term forecasts. In Figueiredo and Guillén (2013) FAVAR is used for time series forecasting of Brazilian consumer inflation (IPCA). Authors found that factor-based models outperform traditional AR models on long horizons. They also considered estimation methods pointing out that ML approach strictly outperforms two-step PCA method, that is the focus of study in this paper. Pang (2010) employed a factor model for predicting GDP growth rate, unemployment, and inflation rates of the Hong Kong economy. Author's results are in line with those obtained in Figueiredo and Guillén (2013) – FAVAR performs better than simple VAR and AR models especially for long horizons. The author used 76 monthly variables that are believed to reflect the economic situation in Hong Kong. Transforming original data, the author followed the methodology suggested in Stock and Watson (2005).

Berggren and Lodenius (2016) in their paper computed Diebold-Mariano test of equal predictive accuracy, comparing FAVAR with benchmark models. The authors found, that the forecast performance of FAVAR was not statistically better than traditional VAR or AR models. Nonetheless, they claim that further research should be done towards an optimal number of lags for the factors, since they found that FAVAR with only one lag for factors performed statistically better than with twelve.

Overall, the literature suggests that the potential of FAVAR model for time series prediction task still is not fully unlocked. There is still space for improvements and experiments, for example, employing FAVAR for non-

economic time series.

1.3   Methods

There are several approaches to measure the quality of time series prediction models. Each method, however, captures one side of the model. Traditional train test split is quite sensitive to split proportion and also to structural changes that occurred after the end of a training set. For this reason, it is not fully accurate to perform this method in their traditional form.

Nonetheless, it is possible to account for h-horizon prediction power performing time-series cross-validation and treat a forecast for each horizon separately. Still results of such procedure may be argued, since a robust estimate of a model's accuracy requires a large number of observations that is problematic for Russian data available. For this reason, I will measure models' performance using time series cross-validation with expanding window.

Forecasts are measured on RMSE (root of mean square error) metric as it is quite interpretable. The chosen measure will be evaluated for different horizons separately. The formula for h-th horizon as follows:

$$RMSE = \sqrt{\frac{1}{n}\Sigma^T_{t=T+h-J}\left[y_t - \hat{y}_{t|t-h}\right]^2}$$

where J – hold-out set size, $\widehat{y}_{t|t-h}$ – forecast for period t having available information up to the moment $t - h$.

There are several ways to estimate FAVAR model. Bernanke et al. (2005) in their work consider two, which are two-step estimation using principal components analysis to recover latent factors and one-step Bayesian likelihood approach estimation. Authors stated that they found no obvious evidence of one method being superior to another and whether the computational cost of the first one worths it. Since cross-validation requires multiple

reestimation, I simply choose the approach that is less time-consuming.

Particular properties of the two-step approach estimation make unclear distribution of estimated coefficients, therefore building right confident intervals becomes a separate task. Bernanke et al. (2005) use a bootstrap procedure for obtaining impulse response function intervals, the same way one may obtain predictive confident interval, however that is out of the scope of this paper.

2 Model

2.1 VAR

Lets Y_t be $M \times 1$ vector of variables under consideration. It is said that Y_t is a VAR(d) process if it has the following dynamic equation:

$$Y_t = A_0 + A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_d Y_{t-d} + \varepsilon_t, \tag{1}$$

where A_0 – $M \times 1$ vector; $\forall i \neq 0$ A_i – $M \times M$ matrix; ε_t – $M \times 1$ error vector with zero mean; components of Y_t are integrated of the same order. $\hat{A}_i$ can be obtained by equation-by-equation least squares. One can easily check that number of parameters to be estimated equals $M(Md+1) = O(M^2)$, $M \to \infty$. For example, an econometrician should estimate 57 coefficients if one wants to use VAR(6) model with 3 variables for forecasting purposes. This means, that M is strictly limited by the number of periods. In practice, it is uncommon to see more than 6 variables included in a VAR equation. This problem is particularly relevant while dealing with short time series, which is the case with Russian economic data. It is hard to believe that econometrician is able to select six variables without missing an important one. Consequently, researcher generally gets biased estimates which may result, for example, in

popular VAR issue that is called "price puzzle". In Bernanke et al. (2005) authors propose an explanation to this problem which completely relies on omitted equation assumption. Thus, ignoring one variable may result in incorrect economic dynamics modeling that may lead to weak forecasting performance. In order to address this issue, one can estimate FAVAR model, which allows incorporating a huge set of economic indicators without losing degrees of freedom[1].

2.2 FAVAR

This section will be based on II.A part in Bernanke et al. (2005). Let Y_t be $M \times 1$ vector of observable economic variables that are believed to reflect the dynamics of the economy. Standard VAR approach implies that Y_t contains price indicator, observable variable reflecting the economy's real output and policy instrument. Such a specification of Y_t is mainly used for structural analysis. As discussed earlier, one may argue that barely information captured by Y_t reflects the dynamics of the whole economy. Let F_t be $K \times 1$ vector of unobservable variables, where in some way K is a small number, that are assumed to summarise all the information that is not captured by Y_t. These factors are said to reflect movements of theoretical concept variable, such as: "economic activity", "investment environment", "price pressure" and etc.

Assume also that dynamics of vector (F_t', Y_t') follows the equation:

$$
\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + v_t, \tag{2}
$$

where $\Phi(L) = (\Phi_0 + \Phi_1 L + \Phi_2 L^2 + \ldots + \Phi_d L^d)$; d – finite number; Φ_i – $(K + M) \times (K + M)$ matrix; v_t – error term vector, such that: $\mathbb{E}(v_t) = 0$, $\mathbb{E}(v_t v_t') =$

---

[1]Number of observations - number of model's parameters

9

Q; Q is symmetric positive semi-definite $(K+M) \times (K+M)$ matrix; L – lag operator: $LY_t = Y_{t-1}$.

One may see, that if components of $\Phi(L)$ that relate Y_t to F_{t-1} are zeros, then Y_t is simply a VAR process. Otherwise, equation (2) is VAR in (F_t', Y_t'). The fact, that econometrician may obtain VAR in Y_t from (2) makes model (2) unrestricted to VAR in Y_t. This makes it possible to perform statistical tests, for instance, the likelihood ratio test, that allows to check whether an unrestricted model is statistically different from a restricted one. Model (2) is commonly called factor-augmented vector autoregression, or simply FAVAR. In the paper Bernanke et al. (2005) authors correctly note that if the true model is FAVAR, but instead VAR in Y_t is estimated, then econometrician generally gets biased estimates, that may lead to poor forecasting power of an estimated model.

Coefficients of $\Phi(L)$ cannot be estimated directly since vector F_t is unobservable. If one was provided with real values of F_t, he could apply conventional techniques of VAR estimation, equation-by-equation least squares, for example, in order to estimate (2).

It is obvious, that real dynamics of Y_t is affected by a huge number of time series, such as exchange rate, price indices, investment climate situation, import/export shocks, and many more. Let us introduce $N \times 1$ vector X_t – set of "informational" variables, where N is in some way a "large" number, specifically such that $K + M << N$. Unlike F_t, vector X_t is assumed to be observable. It is also believed that X_t is a linear function of observable Y_t and unobservable F_t:

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + e_t, \tag{3}$$

where Λ^f – $N \times K$ matrix; Λ^y – $N \times M$; e_t – error term vector, such that:

$\mathbb{E}(e_t) = 0$, $\mathbb{E}(e_{ti}e_{tj}) \approx 0$ $\forall i \neq j$.[2] Equation (3) reflects the idea that economy can be precisely described by just $K + M$ main indicators. All the other variables that we may observe are generated from those key ones. The main point in (3) is that F_t is still vector of unobservable variables and the idea is to exploit the fact that (F_t', X_t') is observed to get estimates of F_t vector. Having in some sense accurate $\hat{F}_t$, econometrician is able to estimate (2) in $(\hat{F}_t', Y_t')$. It is important to note, that X_t depends only on F_t and Y_t, not from lags of these vectors.

If N is a small number, it is possible to include X_t directly in (2) and estimate VAR in (X_t', Y_t'). However, barely it is the truth, that central bank, for example, monitor only three or four variables, setting the key rate. Though, if a researcher wants to incorporate an additional variable in a model, it is quite unlikely that he would have enough observations to get coefficients' estimates.

There are several possible FAVAR models may be constructed depending on what a researcher initially assume. For example, it is possible to include in Y_t variables such as CPI, GDP growth, and central bank key rate, which means that researcher is uncertain about the rest structure of the economy and believe in the presence of a finite number of latent factors that fill possible model misspecification. One may also argue, that in reality we don't observe truth GDP growth and CPI. Instead, we are provided with these series estimations by some federal institutions, that are still noisy measures for true variables. Such model specification is in line with the fact, that GDP growth and CPI series are subjects for several corrections, even in a year. However, for sure, we observe the truth central bank key rate.

---

[2]One can find formal requirements on e_t in Stock and Watson (2002)

2.3 Estimation

Authors in Bernanke et al. (2005) propose two ways of FAVAR estimation. The first is based on Gibbs sampling approach, which is not applicable if one desires to obtain a cross-validation model's error, since such an approach is enormously time-costly. Estimation of one model may take even an hour. The second approach is referred in Bernanke et al. (2005) as two-step principal component method, that is much more computationally efficient in time and easy in implementation. Moreover, the last approach does not require special assumptions on error-terms distributions, as well as allows some cross-correlation in it, which is more likely the truth dealing with real data.

Two-stage estimation approach includes:

1. Extraction of factors F_t from X_t, purifying it from the influence of Y_t.[3]

2. Estimation of model (2), using obtained $\hat{F}_t$ by standard VAR methods.

In order to perform stage 1 one should use the equation (3) and exploit the fact that (X_t', Y_t') is observable vector for each period t. Let us remind once again the equation (3):

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + e_t$$

We can rewrite (3) as:

$$X_t = \begin{bmatrix} \Lambda^f \Lambda^y \end{bmatrix} \begin{bmatrix} F_t \\ Y_t \end{bmatrix} + e_t \tag{4}$$

$$X_t = \Lambda \alpha_t + e_t, \tag{5}$$

where $\Lambda_{N \times (K+M)} \overset{\text{def}}{=} \begin{bmatrix} \Lambda^f \Lambda^y \end{bmatrix}$ and $\alpha_{t\,(K+M) \times 1} \overset{\text{def}}{=} \begin{bmatrix} F_t \\ Y_t \end{bmatrix}$.

---

[3]To get only additional information about economy dynamics, that is not captured by Y_t

The idea is simple. If we have estimated $\hat{\Lambda}$, $\hat{\alpha}_t$ and $\hat{e}_t$ is a small number for all t, then the estimates of parameters are "good". Assume X_t is normalized to have zero mean: $\sum_{t=1}^{T} X_t = \left(0\ldots0\right)'_{N\times1}$. Then, let us set the following minimization task:

$$\sum_{t=1}^{T}(X_t - \Lambda\alpha_t)'(X_t - \Lambda\alpha_t) \to \min_{\Lambda,\ \alpha_1,\alpha_2\ldots\alpha_T} \tag{6}$$

It is worth noting that (6) has in total $(K+M)N+(K+M)T$ parameters for estimation. It is also important to note, that there are infinite number of solutions for this optimization task. If we came up with a solution to (6) in form $(\tilde{\Lambda},\tilde{\alpha})$ then for any Z: $Z'Z = I$ follows $(\tilde{\Lambda}Z',Z\tilde{\alpha})$ is also solution of (6). Thus let us impose restriction on Λ: $\Lambda'\Lambda = I$, which basically means we are trying to obtain orthogonal and normalised unit-scale vectors. Problem (6) transforms to:

$$\begin{cases} \sum_{t=1}^{T}(X_t - \Lambda\alpha_t)'(X_t - \Lambda\alpha_t) \to \min_{\Lambda,\ \alpha_1,\alpha_2\ldots\alpha_T} \\ s.t.\ \Lambda'\Lambda = I \end{cases} \tag{7}$$

Assuming solution $\hat{\Lambda}$ being known, optimal $\hat{\alpha}_t = (\Lambda'\Lambda)^{-1}\Lambda'X_t$. Since α_t does not affect other periods error, then we can treat (6) as T independent minimization problems, such for fixed period t:

$$(X_t - \Lambda\alpha_t)'(X_t - \Lambda\alpha_t) \to \min_{\alpha_t} \tag{8}$$

With known Λ, problem (8) becomes nothing but a simple least-squares task with known analytical solution. Thus, (7) transforms to:

$$\begin{cases} \sum_{t=1}^{T}(X_t - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'X_t)'(X_t - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'X_t) \to \min_{\Lambda} \\ s.t.\ \Lambda'\Lambda = I \end{cases} \tag{9}$$

$$\begin{cases} \sum_{t=1}^{T}(X_t - \Lambda\Lambda'X_t)'(X_t - \Lambda\Lambda'X_t) \to \min_{\Lambda} \\ s.t.\ \Lambda'\Lambda = I \end{cases} \tag{10}$$

Let $L = \sum_{t=1}^{T}(X_t - \Lambda\Lambda'X_t)'(X_t - \Lambda\Lambda'X_t)$, then:

$$L = \sum_{t=1}^{T}\left[(I - \Lambda\Lambda')X_t\right]'\left[(I - \Lambda\Lambda')X_t\right] =$$

$$= \sum_{t=1}^{T} X_t'(I - \Lambda\Lambda')'(I - \Lambda\Lambda')X_t =$$

$$= \sum_{t=1}^{T} X_t'(I - \Lambda\Lambda')X_t = \sum_{t=1}^{T} X_t'X_t - \sum_{t=1}^{T} X_t'\Lambda\Lambda'X_t$$

Thus, problem (10) is equivalent to:

$$\begin{cases} \sum_{t=1}^{T} X_t'\Lambda\Lambda'X_t \to \max_{\Lambda} \\ s.t.\ \Lambda'\Lambda = I \end{cases} \tag{11}$$

Let $\tilde{L} = \sum_{t=1}^{T} X_t\Lambda\Lambda'X_t$, then:

$$\tilde{L} = tr(\sum_{t=1}^{T} X_t'\Lambda\Lambda'X_t) = \sum_{t=1}^{T} tr(X_t'\Lambda\Lambda'X_t)) =$$

$$\overset{*}{=} \sum_{t=1}^{T} tr(\Lambda\Lambda'X_tX_t') = tr(\sum_{t=1}^{T} \Lambda\Lambda'X_tX_t') = tr(\Lambda\Lambda' \sum_{t=1}^{T} X_tX_t'),$$

where (*) follows from the trace property: $tr(AB) = tr(BA)$, if A and B are matrices of suitable sizes.

Let us remind that if $X_{T \times N} = \left(X_1 X_2 \ldots X_T\right)'$ is matrix that has zero mean by columns, then $S_X = \frac{1}{T}\sum_{t=1}^{T} X_tX_t'$ – is empirical covariance matrix of X_t vector.

$$\tilde{L} = tr(\sum_{t=1}^{T} \Lambda\Lambda'X_tX_t') = tr(\Lambda\Lambda'TS_X) = Ttr(\Lambda'S_X\Lambda)$$

Thus (10) problem becomes equivalent to:

$$\begin{cases} tr(\Lambda'S_X\Lambda) \to \max_{\Lambda} \\ s.t.\ \Lambda'\Lambda = I \end{cases} \tag{12}$$

14

Still (12) has $N(K+M)$ parameters that should be estimated. Though, this optimization task has an exact analytical solution. One should compute the first $K+M$ eigenvectors of S_X corresponding to the largest eigenvalues. These vectors are columns of $\hat{\Lambda}$ that solves maximization task (12) The fact that S_X is a symmetric positive semi-definite matrix guarantees that it has real-valued eigenvector and eigenvalues. Columns of $\hat{\Lambda}$ are nothing but the first $K+M$ principal components of X. Knowing solution to (12) one can easily find $\hat{\alpha}_t$ from (8). Stock and Watson (2002) showed that under certain conditions on N, that should be "large", econometrician may use only first K principle components and obtain consistent estimator of space spanned by Y_t and F_t. It is important to note, that at this moment we did not rely on the fact that Y_t is observable. For now we got K linear combinations of F_t and Y_t. In Bernanke et al. (2005) authors then perform purification procedure, which consists in "removing" Y_t from the space spanned by the first K principal component of X_t.

For stage 2 authors divide variables into two groups: fast-moving and slow-moving. The first group assumed to contemporaneously react to unexpected shocks in Y_t, whereas the second one does not respond to unanticipated fluctuations in Y_t at a period t. Such variables' separation is necessary for obtaining purified factors F_t, that contains all the information apart from that in Y_t. Since impulse response functions are the main focus in Bernanke et al. (2005), an absence of immediate response in F_t to shocks in Y_t is one of the requirements in some SVAR identification schemes, such as recursive identification including Cholesky decomposition of residuals covariance matrix. However, there is a place for experiments if one tries to employ FAVAR methodology for forecasting purposes. In literature, there are several ways to remove dependence of obtained in stage 1 factors from Y_t. The one used in Bernanke et al. (2005) will be described below.

Let $\hat{C}(F_t, Y_t) \overset{\text{def}}{=} \hat{\Lambda}\hat{\alpha}_t \overset{(8)}{=} (\hat{\Lambda}'\hat{\Lambda})^{-1}\hat{\Lambda}'X_t \overset{(12)}{=} \hat{\Lambda}'X_t$. The authors estimate the following model:

$$\hat{C}(F_t, Y_t) = \beta_0 + \beta_{-0}\hat{C}^{slow}(F_t) + \gamma Y_t + v_t, \tag{13}$$

where $\hat{C}(F_t, Y_t)$ – are the first K principle components of X_t; $\hat{C}^{slow}(F_t)$ – the first K principle components of X_t^{slow}; b_0 – $K \times 1$ vector; b_{-0} – $K \times K$ matrix; γ – $K \times M$ vector ; v_t – $K \times 1$ error vector. Though, at this step one may include subset of Y_t variables. One can obtain $\hat{\beta}$ and $\hat{\gamma}$ using simple OLS method by rows.

$$\hat{F}_t = \hat{\beta}_{-0}\hat{C}^{slow}(F_t) - \hat{\gamma}Y_t$$

The estimates of F_t allows us to substitute F_t by $\hat{F}_t$ in (2), which result in:

$$\begin{bmatrix} \hat{F}_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} \hat{F}_{t-1} \\ Y_{t-1} \end{bmatrix} + v_t \tag{14}$$

Note once again that (14) in a standard VAR model in $(\hat{F}_t, Y_t)$, therefore $\Phi(L)$ can be estimated using traditional VAR techniques. Summarising all, we finally get:

Algorithm 1 FAVAR estimation

$C(F_t, Y_t) \leftarrow$ the first K components of X_t

$C^{slow}(F_t) \leftarrow$ the first K components of X_t^{slow}

OLS estimation results $\leftarrow \hat{C}(F_t, Y_t) = \hat{\beta}_0 + \hat{\beta}_{-0}\hat{C}^{slow}(F_t) + \hat{\gamma}Y_t$

$\hat{F}_t \leftarrow \hat{\beta}_{-0}\hat{C}^{slow}(F_t) - \hat{\gamma}Y_t$

$d \leftarrow$ VAR$(\hat{F}_t, Y_t)$ lag selection according to AIC

FAVAR estimation results $\leftarrow$ VAR(d) in $(\hat{F}_t', Y_t')$

The explained procedure was used in Bernanke et al. (2005) for producing impulse response functions to policy shock. Purification is necessary for producing factors that do not respond contemporaneously to a policy shock,

since authors used recursive identification with FFR (Federal Funds Rate) ordering last. For forecasting purposes, one may try to remove the impact of other variables (from X_t) or skip factor purification procedure at all.

2.4   Model specifications

Once an econometrician has estimated $\hat{\boldsymbol{\Phi}}(L)$ in (14), h-horizon forecast for $(\hat{F}_{t-1}, Y_{t-1})$ is the following:

$$\begin{bmatrix} \hat{F}_{t+h-1} \\ \hat{Y}_{t+h-1} \end{bmatrix} = \hat{\boldsymbol{\Phi}}(L)^h \begin{bmatrix} \hat{F}_{t-1} \\ Y_{t-1} \end{bmatrix}$$

For a particular variable forecast, a researcher takes the corresponding component of $\hat{Y}_{t+h-1}$ vector. One should also specify parameter K representing a number of latent factors. I find, that $K = 1$ is sufficient in most cases. However, $K = 2$ sometimes helps to obtain more accurate results in short-term prediction, Interesting that $K \geq 3$ generally increases the model's error for almost all horizons. There is a space for experiments still, one may receive more accurate results at some horizons selecting proper variables for Y_t.

Another approach, is to treat variable of interest z_t as "informational" (included in X_t), but I found no clear evidence that such a method is more preferable than assuming z_t being observed. These different FAVARs produce quite similar results, however, there is a still place for experiments in the choice of Y_t vector components.

3   Data

Data was collected from the Thomson Reuters database. Dataset includes 199 monthly variables, that represent the economic situation in Russia, start-

ing from 01/01/2000 and contains 240 observations maximum. Variables with missing values were dropped before the estimation procedure. To be in line with Bernanke et al. (2005) original time series were transformed (one or two times differencing) to assure stationarity according to KPSS test. In addition, for convenience original time series were standardized to have zero mean and unit standard deviation. Though, in order to obtain fair results, a seasonal component in considered time series was not removed, since benchmark models are capable of handling seasonality in data. One can still decompose original time series on seasonal and trend components, though FAVAR methodology cannot be applied to produce seasonal component forecasts, so it becomes a mix of FAVAR and another model for a seasonal component. Probably such an approach will yield more accurate forecasts, but the goal of this paper to check pure FAVAR model prediction power.

Depending on the variable being forecasted different number of observations were used: 238 in case of CPI and unemployment rate, 138 in case of GDP Y/Y % change forecasts.

4 Results

In order to understand FAVAR model's predictive power I decided to compare its results on cross-validation with popular univariate models, such as ETS, ARIMA, and simple random walk with drift; and restricted to FAVAR model – standard VAR ($K = 0$). I treat variable of interest as being observed, so include it directly in Y_t vector. For ARIMA and ETS optimal model selection I used auto-ARIMA and auto-ETS functions in the "forecast" R package, that select best model specification according to AIC. Optimal lag of VAR models is also selected using AIC.

	FAVAR	ARIMA	ETS	VAR	RWD
h=1	0.70*	0.86	0.76	0.83	0.93
h=2	0.95*	1.09	1.02	1.07	1.37
h=3	1.01*	1.07	1.06	1.12	1.54
h=4	0.96*	1.03	1.10	1.13	1.58
h=5	0.91*	1.03	1.05	1.13	1.49
h=6	0.95*	1.07	1.07	1.15	1.47
h=7	1.00*	1.05	1.19	1.17	1.65
h=8	1.06	0.96*	1.24	1.19	1.78
h=9	1.05	0.99*	1.26	1.18	1.75
h=10	0.94*	0.98	1.21	1.02	1.57
h=11	0.90*	0.99	1.23	0.97	1.35
h=12	0.90*	1.04	1.16	0.98	1.08

Table 1: RMSE $\times$ 100 for CPI forecasts

Table 1 shows cross-validation results for CPI forecasting. FAVAR specification includes two variables in $Y_t = (aRUCPI_t, RUCBIR = ECI_t)'$ and one latent factor $K = 1$. Interesting that expanding number of factors deteriorates the model's quality. One can see, that such FAVAR specification completely outperforms other models almost at each horizon.

	FAVAR	ARIMA	ETS	VAR	RWD
h=1	10.31	6.97*	9.90	7.13	11.39
h=2	10.25	6.95	9.75	6.74*	10.97
h=3	8.68	6.56	9.90	6.52*	10.06
h=4	8.04	6.60	9.86	6.36*	12.06
h=5	6.88	5.86	10.02	6.22*	12.00

h=6	6.10	5.94*	9.70	6.44	11.36
h=7	5.74*	6.00	9.27	6.56	12.64
h=8	5.71*	6.09	9.57	6.44	10.90
h=9	6.15*	6.15	9.66	6.17	10.40
h=10	5.90*	5.96	9.21	6.42	10.10
h=11	5.86*	5.89	9.67	6.52	9.32
h=12	6.43	5.70*	9.51	6.43	9.09

Table 2: RMSE $\times$ 100 for unemployment rate forecasts

Table 2 shows models' cross-validation errors in attempts to forecast unemployment rate in Russia. In this FAVAR specification $Y_t = (aRUCPI_t, RUUNR = ECI_t)'$ and $K = 1$. As with CPI, increasing of K leads to weak performance. As one can see, for unemployment rate FAVAR produces best forecasts only on long horizons. Yet one can experiment with proper Y_t selection to obtain lower error at short-term forecasts.

	FAVAR	ARIMA	ETS	VAR	RWD
h=1	19.88*	22.34	21.67	20.10	38.56
h=2	20.44*	23.22	22.07	20.79	25.65
h=3	21.92*	23.79	22.59	22.37	35.73
h=4	22.40	23.13	22.35	22.31*	31.02
h=5	22.96	23.49	22.94	22.90*	34.74
h=6	22.75	22.98	22.82	22.68*	30.00
h=7	23.06*	23.66	23.15	23.17	35.66
h=8	23.46*	24.45	23.53	23.51	34.69
h=9	24.52	25.21	24.29	24.25*	33.81
h=10	22.74	25.06	22.72*	22.75	33.99

h=11	22.98	24.59	22.74	22.70*	33.43
h=12	23.09*	24.29	23.15	23.17	25.37

Table 3: RMSE × 100 for GDP forecasts

In table 3 one can find cross-validation errors of compared models, that are used for GDP growth rate forecasts. It is important to mention again, that for GDP only 168 observations were used, because of missing values. That might be one reason for explaining the model's performance. Surprisingly, FAVAR produces best forecasts mostly for short-term, rather than for long-term horizons as was expected. FAVAR in $Y_t = (RUGDP = ECI_t, RUCBIR = ECI_t)'$ with $K = 2$ shows best performance on 6 out of 12 horizons, though difference is significant only for $h = \overline{1,3}$.

5   Conclusion

Results showed that FAVAR is quite competitive model in a forecasting task, even though it was initially developed as the solution for "price puzzle", that is more about impulse response functions, rather than predictions. For some series FAVAR can completely outperform traditional univariate models as well as multivariate VAR. Besides, FAVAR allows incorporating as many variables as one wants and there are reasons to believe that this will improve the model's performance. Nevertheless, it remains a question whether the cost of data collection outweighs potential forecast improvement for standard VAR.

Appendixes

A Data Description

The table below provides information about the used dataset. An asterisk *, after mnemonic of some series shows that such variables were assumed to be slow-moving. Each time series is provided with the transformation code applied to the original data: l – no transformations, Δ – first difference and Δ^2 – second difference.

Mnemonic	Description	
1. aRUPTTTT/C*	Passenger Transport Turnover, Price Index	l
2. aRUFTTCTT/C*	Freight Transport Turnover, Cargo transport, Volume Index	l
3. aRUCEXBA*	Exports of Goods, Balance of Payments Basis, Standardized, Current Prices	Δ
4. aRUCIMBA*	Imports of Goods, Balance of Payments Basis, Standardized, Current Prices	Δ
5. aRUCBOPA*	Visible Trade Balance, Balance of Payments Basis, Standardized, Current Prices	l
6. aRUCEXBPA*	Exports of Goods, Balance of Payments Basis, % month on month, Standardized, Chg P/P, Current Prices	l
7. aRUCIMBPA*	Imports of Goods, Balance of Payments Basis, % month on month, Standardized, Chg P/P, Current Prices	l
8. aRUCBOPPA*	Visible Trade Balance, Balance of Payments Basis, month on month, Standardized, Absolute change, Current Prices	l

9. aRUCEXBYA*	Exports of Goods, Balance of Payments Basis, % year on year, Standardized, Chg Y/Y, Current Prices	Δ
10. aRUCIMBYA*	Imports of Goods, Balance of Payments Basis, % year on year, Standardized, Chg Y/Y, Current Prices	Δ
11. aRUCBOPYA*	Visible Trade Balance, Balance of Payments Basis, year on year, Standardized, Absolute change, Current Prices	l
12. RUCBIR=ECI	Central bank key rate	Δ
13. aRUBCFBNKA	Bank Lending: Loans to banks (in frgn. cur.), Current Prices	Δ
14. aRUDOMDET	Domestic Debt, Current Prices	Δ^2
15. aRUBCDBNKA	Bank Lending: Loans to banks (in rubles), Current Prices	Δ
16. RUCPIY=ECI*	CPI, Chg Y/Y	Δ
17. aRUCPI*	CPI, Price Index	Δ
18. aRUCCPIYF*	CPI, % year on year, Standardized, Chg Y/Y, Price Index	Δ
19. aRUCCPIF/C*	CPI, Standardized, Price Index	Δ
20. aRUCPFYYF*	CPI, Food products, Chg Y/Y	Δ
21. aRUCPCORIF*	CPI, Core inflation, Chg Y/Y	Δ
22. aRUCCPIYE/A*	CPI, % year on year, Standardized, Chg Y/Y, Price Index, SA	Δ
23. aRUCCPIPF*	CPI, % month on month, Standardized, Chg P/P, Price Index	Δ

24. aRUCCORYE/A*	Core CPI, Standardized, Chg Y/Y, Price Index, SA	Δ
25. aRUCCPIPE/A*	CPI, % month on month, Standardized, Chg P/P, Price Index, SA	Δ
26. aRUCPNYYF*	CPI, Non-food products, Chg Y/Y	*l*
27. aRUCPIF*	CPI, Food and beverages, Price Index	*l*
28. aRUCPISERV*	CPI, Services, Price Index	Δ
29. aRUCPIXF*	CPI, Non-food goods, Price Index	*l*
30. aRUCPGOODF/C*	CPI, Goods, Price Index	*l*
31. aRUGOODSF*	CPI, Goods, Chg Y/Y	Δ
32. aRUCPFPFVF/C*	CPI, Food products without fruits and vegetables, Price Index	Δ
33. aRUCCPIE/CA*	CPI, Standardized, Price Index, SA	Δ
34. aRUCCORF/C*	Core CPI, Standardized, Price Index	Δ
35. aRUCCORE/CA*	Core CPI, Standardized, Price Index, SA	Δ
36. aRUCCORPF*	Core CPI, Standardized, Chg P/P, Price Index	Δ
37. aRUCCORPE/A*	Core CPI, Standardized, Chg P/P, Price Index, SA	Δ
38. aRUCCORYF*	Core CPI, Standardized, Chg Y/Y, Price Index	Δ
39. RUCPIY=ECIX*	CPI, Chg Y/Y	Δ
40. aRUCPFWFVF*	CPI, Food products without fruits .and vegetables, Chg Y/Y	Δ
41. aRUCPICORY/C*	CPI, Core CPI, Price Index	Δ
42. aRUCPNPSVF*	CPI, Paid services, Chg Y/Y	Δ
43. aRUCPICOR/C*	CPI, Core CPI, Price Index	Δ
44. aRUCLEAD	Composite leading indicators, Trend restored, SA	Δ

45. aRUEMPLMT*	Employment, Volume	Δ
46. aRUBISRXBR	BIS, Real Broad Effective Exchange Rate Index	Δ
47. aRUCXTWF/C	Trade Weighted nominal exchange rate, Standardized, Price Index	Δ
48. aRUCXTWPF	Trade Weighted nominal exchange rate, % month on month, Standardized, Chg P/P, Price Index	*l*
49. aRUCXTWYF	Trade Weighted nominal exchange rate, % year on year, Standardized, Chg Y/Y, Price Index	*l*
50. aRUCXTRF/C	Trade Weighted real exchange rate, Standardized, Price Index	Δ
51. aRUCXTRPF	Trade Weighted real exchange rate, % month on month, Standardized, Chg P/P, Price Index	*l*
52. aRUCXTRYF	Trade Weighted real exchange rate, % year on year, Standardized, Chg Y/Y, Price Index	*l*
53. aRUIRECE/C	Real effective exchange rate (REER) based on consumer price index, 2010=100, Not SA, Price Index	Δ
54. aRUINECE/C	Nominal effective exchange rate (NEER) based on consumer price index, 2010=100, Not SA, Price Index	Δ
55. aRUBISNXBR	BIS, Nominal Broad Effective Exchange Rate Index	Δ
56. aRUXRUSD	Russian roubles to US $	Δ
57. RUGDP=ECI*	GDP, Chg Y/Y	Δ
58. aRUTREV*	Revenue, Federal budget, Current Prices	Δ
59. aRUPFEXP*	Expenditure, Current Prices	Δ

60. aRUCBUDIC*	Revenue, Consolidated budget, incomes total, Current Prices	Δ
61. aRUEXPDT*	Expenditure, Current Prices	Δ
62. aRUGDEF*	Deficit/Surplus, Federal budget, Current Prices	l
63. aRUCGOVA*	Central government Deficit/Surplus, Standardized, Current Prices	l
64. aRUCGOVPA*	Central government Deficit/Surplus, month on month, Standardized, Absolute change, Current Prices	l
65. aRUCGOVYA*	Central government Deficit/Surplus, year on year, Standardized, Absolute change, Current Prices	l
66. aRSCGOVBLA*	Public Finances, Central Government, Budget, Balance, Deficit/Surplus, Current Prices	l
67. aRUDSCBUD*	Deficit/Surplus, Consolidated budget, Current Prices	l
68. aRUDCTDA	Dwellings Commenced, Total dwellings area, Volume	Δ
69. aRUDCTI	Dwellings Commenced, Total dwellings area, Price Index	Δ
70. aRUPPICONS*	Producer Prices, Investment products, Chg P/P	Δ
71. RUTRD=ECI*	Trade Balance, Total, Free On Board, Current Prices	l
72. aRUEXP*	Exports, FOB, Current Prices	Δ
73. aRUIMP*	Imports, FOB, Current Prices	Δ
74. aRUTBAL*	Trade Balance, FOB, Current Prices	l
75. aRUEXPNGC*	Exports, Natural gas, Current Prices	l

76. aRUEXPCRD*	Exports, Crude oil, Current Prices	*l*
77. aRUEXPDIESL*	Exports, Diesel fuel, Current Prices	*l*
78. aRUIMPPRO*	Exports, Petroleum products, Current Prices	*l*
79. aRUIMPMED*	Imports, Medicines, Current Prices	*l*
80. aRUIMPMNEQ*	Imports, Machinery and equipment, Current Prices	*l*
81. aRUIMPCLOTH*	Imports, Clothing, Current Prices	Δ
82. aRUEXPMOM*	Exports, Ferrous metals, Current Prices	*l*
83. aRUCEXPB/A*	Merchandise Exports, Standardized, Current Prices, SA	Δ
84. aRUCIMPB/A*	Merchandise Imports, Standardized, Current Prices, SA	Δ
85. aRUCEXPPB/A*	Merchandise Exports, % month on month, Standardized, Chg P/P, Current Prices, SA	*l*
86. aRUCEXPYB/A*	Merchandise Exports, % year on year, Standardized, Chg Y/Y, Current Prices, SA	Δ
87. aRUCIMPPB/A*	Merchandise Imports, % month on month, Standardized, Chg P/P, Current Prices, SA	*l*
88. aRUCIMPYB/A*	Merchandise Imports, % year on year, Standardized, Chg Y/Y, Current Prices, SA	Δ
89. aRUCVISB/A*	Visible Trade Balance, Standardized, Current Prices, SA	*l*
90. aRUCVISPB/A*	Visible Trade Balance, month on month, Standardized, Absolute change, Current Prices, SA	*l*
91. aRUCVISYB/A*	Visible Trade Balance, year on year, Standardized, Absolute change, Current Prices, SA	*l*
92. RUTRD=ECIX*	Trade Balance, Total, Free On Board	*l*

93. RUIP=ECI*	Production, IP Total , Chg Y/Y	l
94. aRUIP*	Production, Chg Y/Y	l
95. aRUIPOIL*	Production, Oil output, including gas condensate, Volume	Δ
96. aRUCINDG/A*	Industrial Production Index, Standardized, Volume Index, SA	Δ
97. aRUIPNAG*	Production, Natural gas output, Volume	l
98. aRUIPMANH/C*	Production, Manufacturing, Volume Index	l
99. aRUIPMANG*	Production, Manufacturing, Chg Y/Y	l
100. aRUINP/C*	Production, Industry, Volume Index	Δ
101. aRUCINDPG/A*	Industrial Production Index, Standardized, Chg P/P, Volume Index, SA	l
102. aRUCINDYG/CA*	Industrial Production Index, % year on year, Standardized, Chg Y/Y, Volume Index, SA	l
103. RUIP=ECIX*	Production, IP Total, Chg Y/Y	l
104. aRUPRATE	Policy Rates, Minimum Rate on 7 Day Repo	Δ
105. aRUINTRES	Reserves, Gross international, Current Prices	Δ
106. aRUFXRES	Reserves, Foreign currency reserves, Current Prices	Δ
107. aRUFCRES	Reserves, Foreign currency, Current Prices	Δ
108. aRUCRESA	Official international reserves, Standardized, Current Prices	Δ
109. aRURESGLD	Reserves, Gold, Current Prices	Δ
110. aRUCRESPA	Official international reserves, % month on month, Standardized, Chg P/P, Current Prices	Δ
111. aRUCRESYA	Official international reserves, % year on year, Standardized, Chg Y/Y, Current Prices	Δ

112. aRURTM1A1A	Official reserve assets, Foreign currency, Current Prices	Δ^2
113. aRURESSDR	Reserves, Special Drawing Rights, Current Prices	Δ
114. aRURTM1AA	Official reserve assets, Overall, Current Prices	Δ
115. aRURESIMF	Reserves, Reserve position in the IMF, Current Prices	Δ
116. aRUM2	Money supply M2 by national definition, Current Prices	Δ
117. aRUM0	Money supply M0, Current Prices	Δ
118. aRUNM1	Money Supply - M1, Current Prices	Δ
119. aRUMNBAS	Monetary base (narrow definition), Current Prices	Δ
120. aRUMBASMOC	Monetary base (broad definition), Current Prices	Δ
121. aRUMSMBBRA	Broad money liabilities, Current Prices	Δ
122. aRUCMS2B/A	Money Supply M2, Standardized, Current Prices, SA	Δ
123. aRUCMS2PB/A	Money Supply M2, % month on month, Standardized, Chg P/P, Current Prices, SA	l
124. aRUCMS2YB/A	Money Supply M2, % year on year, Standardized, Chg Y/Y, Current Prices, SA	Δ
125. aRUCMS1B/A	Money Supply M1, Standardized, Current Prices, SA	Δ
126. aRUCMS0B/A	Money Supply M0, Standardized, Current Prices, SA	Δ

127. aRUCMS1PB/A	Money Supply M1, Standardized, Chg P/P, Current Prices, SA	l
128. aRUCMS0PB/A	Money Supply M0, Standardized, Chg P/P, Current Prices, SA	l
129. aRUCMS1YB/A	Money Supply M1, Standardized, Chg Y/Y, Current Prices, SA	Δ
130. aRUCMS0YB/A	Money Supply M0, Standardized, Chg Y/Y, Current Prices, SA	Δ
131. RUPPI=ECI*	Producer Prices, Chg P/P	l
132. RUPPIY=ECI*	Producer Prices, Chg Y/Y	l
133. aRUPPI/C*	Producer Prices, Price Index	Δ
134. aRUPPITCP*	Producer Prices, Chg P/P	l
135. aRUPPIF*	Producer Prices, Chg Y/Y	l
136. aRUCPPIE/CA*	Producer Prices, PPI, Standardized, Price Index, SA	Δ
137. aRUPPIAR*	Producer Prices, Chg P/P	l
138. aRUCPPIF/C*	Producer Prices, PPI, Standardized, Price Index	Δ
139. aRUCPPIPE/A*	Producer Prices, PPI, % month on month, Standardized, Chg P/P, Price Index, SA	l
140. aRUCPPIPF*	Producer Prices, PPI, % month on month, Standardized, Chg P/P, Price Index	l
141. aRUCPPIYE/A*	Producer Prices, PPI, % year on year, Standardized, Chg Y/Y, Price Index, SA	Δ
142. aRUCPPIYF*	Producer Prices, PPI, % year on year, Standardized, Chg Y/Y, Price Index	Δ
143. RUPPI=ECIX*	Producer Prices, Chg P/P	l
144. RUPPIY=ECIX*	Producer Prices, Chg Y/Y	l

145. aRUPPIYAR*	Producer Prices, Chg Y/Y	*l*
146. aRUPPITC*	Producer Prices, Price Index	Δ
147. RURSLY=ECI*	Retail Sales YY, Price Index	Δ
148. aRURSLS*	Retail Trade Turnover, Current Prices	Δ
149. aRURSLSTO*	Retail Trade Turnover, Price Index	Δ
150. aRUCRETF/C*	Retail Sales, Standardized, Price Index	Δ
151. aRUCRETE/CA*	Retail Sales, Standardized, Price Index, SA	Δ
152. aRUCRETPF*	Retail Sales, Standardized, Chg P/P, Price Index	*l*
153. aRUCRETPE/A*	Retail Sales, Standardized, Chg P/P, Price Index, SA	Δ
154. aRUCRETYF*	Retail Sales, Standardized, Chg Y/Y, Price Index	Δ
155. aRUCRETYE/A*	Retail Sales, Standardized, Chg Y/Y, Price Index, SA	Δ
156. RURSLY=ECIX*	Retail Sales YY	Δ
157. aRURTTNFC*	Retail Trade Turnover, Non-food commodities, Current Prices	Δ
158. aRURSLS/C*	Retail Trade Turnover, Volume Index	*l*
159. aRURTTFC*	Retail Trade Turnover, Food commodities, Current Prices	Δ
160. aRUGBOND	Bid	Δ
161. aRUSHRPRCF	MICEX, Composite Index, Price Index	Δ
162. aRUWANMAVWG*	Wages, Nominal average monthly per worker, Current Prices	Δ
163. aRUWSAMTCPP*	Wages, Real average monthly, Volume Index	Δ
164. aRUCWAGF/C*	Wages, total, Standardized, Price Index	Δ

165. aRUCWAGPF*	Wages, total, % month on month, Standardized, Chg P/P, Price Index	*l*
166. aRUCWAGYF*	Wages, total, % year on year, Standardized, Chg Y/Y, Price Index	Δ
167. aRUAVGMANW*	Average monthly accrued nominal wages, Current Prices	Δ
168. RUUNR=ECI*	Unemployment, Rate	Δ
169. RURWGE=ECI*	Real Wages YY, Chg Y/Y	Δ
170. aRUUNR*	Unemployment, Rate	Δ
171. aRUUNTOTR*	Uneployment rate, ILO	Δ
172. aRUPOPEA*	Economically active population - Unemployed, Volume	Δ
173. aRUCUNPQ/A*	Unemployment rate, Standardized, SA	Δ
174. aRUUEMPILO*	Unemployment, Total persons (ILO), aged 15 and over, Volume	Δ
175. aRUUEMP*	Unemployment, Officially registered, Volume	Δ
176. aRUUNRAR*	Unemployment, Rate	Δ
177. aRUCUNPPQ/A*	Unemployment rate, month on month, Standardized, SA	*l*
178. aRUCUNPYQ/A*	Unemployment rate, year on year, Standardized, SA	*l*
179. aRUCUNPO*	Unemployment Level, Standardized, Volume	Δ
180. aRUCUNPP*	Unemployment Level, Standardized, Volume	Δ
181. aRUCUNPPO/A*	Unemployment Level, % month on month, Standardized, Chg P/P, Volume, SA	*l*
182. aRUCUNPPP*	Unemployment Level, % month on month, Standardized, Chg P/P, Volume	*l*

183. aRUCUNPYO*	Unemployment Level, % year on year, Standardized, Chg Y/Y, Volume	l
184. aRUCUNPYP*	Unemployment Level, % year on year, Standardized, Chg Y/Y, Volume	l
185. RURWGE=ECIX*	Real Wages YY, Chg Y/Y	Δ
186. RUUNR=ECIX*	Unemployment, Rate	Δ
187. aRUUNRRILO*	Unemployment, Rate, Officially registered (ILO)	Δ
188. .IRTS	RTS Index	Δ
189. .IMOEX	MOEX Russia Index	Δ
190. .RTSOG	RTS Oil & Gas Index	Δ
191. .RTSTL	RTS Telecom Index	Δ
192. BrentP	Brent Price nominal	Δ
193. BrentPP	Brent Price (%)	l
194. WTIP	WTI Price nominal	Δ
195. WTIPP	WTI Price (%)	l
196. RUEURP	Russian roubles to EU €, Close	Δ
197. RUEURMAX	Russian roubles to EsU €, Max	Δ
198. RUEURMIN	Russian roubles to EU €, Min	Δ
199. RUEURPP	Russian roubles to EU €, % change	l

References

Berggren, E. and Lodenius, E. (2016). Can favar improve swedish inflation forecasting?

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (favar) approach. The Quarterly journal of economics, 120(1):387–422.

Figueiredo, F. M. R. and Guillén, O. (2013). Forecasting brazilian consumer inflation with favar models using target variables. Technical report, mimeo.

Mönch, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented var approach. Journal of Econometrics, 146(1):26–43.

Pang, I. A. J. (2010). Forecasting hong kong economy using factor augmented vector autoregression.

Sims, C. A. (1980). Martingale-like behavior of prices.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. European economic review, 36(5):975–1000.

Stock, J. H. and Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics, 20(2):147–162.

Stock, J. H. and Watson, M. W. (2005). Implications of dynamic factor models for var analysis. Technical report, National Bureau of Economic Research.