EUROSÜSTEEM

Working Paper Series
8/2016

# FORECASTING THE ESTONIAN RATE OF INFLATION USING FACTOR MODELS

NICOLAS REIGL

The Working Paper is available on the Eesti Pank web site at:
http://www.eestipank.ee/en/publications/series/working-papers

For information about subscription call: +372 668 0998; Fax: +372 668 0954
e-mail: publications@eestipank.ee

ISBN 978-9949-493-89-0 (hard copy)
ISBN 978-9949-493-90-6 (pdf)

Eesti Pank. Working Paper Series, ISSN 1406-7161; 8/2016 (hard copy)
Eesti Pank. Working Paper Series, ISSN 2504-5520; 8/2016 (pdf)

# Forecasting the Estonian Rate of Inflation using Factor Models

Nicolas Reigl *

**Abstract**

The paper presents forecasts of the headline and core inflation in Estonia with factor models in a recursive pseudo out-of-sample framework. The factors are constructed with a principal component analysis and are then incorporated into vector autoregressive forecasting models. The analyses show that certain factor-augmented vector autoregressive models improve upon a simple univariate autoregressive model but the forecasting gains are small and not systematic. Models with a small number of factors extracted from a large dataset are best suited for forecasting headline inflation. In contrast models with a larger number of factors extracted from a small dataset outperform the benchmark model in the forecast of Estonian headline and, especially, core inflation.

JEL classification: C32, C38, C53

Keywords: Factor models, factor-augmented vector autoregressive models, factor analysis, principal components, inflation forecasting, forecast evaluation, Estonia



*Bank of Estonia, E-mail: nicolas.reigl@eestipank.ee. The author would like to thank, without implicating, Dmitry Kulikov, Aleksei Netšunajev, Sulev Pert, Karsten Staehr, and Lenno Uusküla for useful comments to earlier versions. The paper has also benefited from comments received at presentations at Eesti Pank and at Tallinn University of Technology.

# Non-technical summary

Inflation dynamics have been an important topic for Estonian central bankers and policy makers. Forecasting the inflation rate with simple models which rely only on few variables has proven challenging, given the small and open structure of the Estonian economy. In recent years, more macroeconomic and financial time series have become available to researchers. One way to incorporate the increasing amounts of data is by using factor models. Factors summarise the information of large numbers of variables contained in an extensive dataset.

First, this paper examines whether and how factor models can be used to forecast the Estonian headline and core inflation rates. Second, I analyse how the number of factors in the forecast equation influences the forecast performance. And third, I investigate the impact of excluding presumably important variables from the large dataset on the factors and consequently on the forecasting results.

This paper uses a large dataset of 388 macroeconomic, microeconomic and financial time series spanning 2004 to 2014 to extract factors, which are then incorporated in a model to forecast the quarterly Estonian inflation rate from 2011 to 2014. To examine the effects of the size of the dataset on forecasting performance, I exclude domestic and foreign price indicators in the large dataset, creating a second smaller dataset of 246 variables. The extracted factors are later incorporated in what is called a factor-augmented vector autoregressive model, which also contains the inflation rate itself. The forecasts obtained from this factor model are compared to an autoregressive model, where the inflation rate is forecast using only its own history. The forecast errors are calculated by comparing both the factor and autoregressive model forecasts to the actual inflation rate.

The results show that factor model forecasts improve upon the autoregressive forecasts in many cases but the difference in forecast performance is rather small. In addition, the results show that including one factor in the model is sufficient for it to outperform the autoregressive benchmark model when the factor is extracted from the large dataset. Factor models fail to improve substantially upon the benchmark model when Estonian core inflation is forecast.

Removing domestic and foreign consumer price indicators from the dataset does not worsen the forecasting performance of the factor models. However, the results indicate that the first, second and third factors have to be included in the forecasting equation to obtain similar forecast results as in the benchmark case. Surprisingly, using the same three factor model shows forecasting errors that are up to 27 percentage points lower when the core inflation rate is forecast. However, robustness tests show that the distribution of the forecasting errors is less stable for models including the first three factors when those factors are extracted from the small dataset.

In conclusion, factor models can help to forecast the Estonian headline and core inflation rates. The forecast performance is dependent on the size of the dataset and the number of factors incorporated in the forecasting equation. For Estonia, the findings provide evidence in favour of using a fairly large dataset to extract the first factor, which should then be incorporated, together with the inflation rate, in a factor-augmented vector autoregressive forecasting equation.

# Contents

# 1. Introduction

Inflation and changes in inflation are key measures of macroeconomic performance. It follows that forecasting inflation is important in countries around the world, including Estonia. Volatile dynamics such as the pre-crisis rise in inflation, which has been largely attributed to the supply side shocks that hit the small open economy (Benkovskis et al., 2009), have challenged the forecasting skills of central bankers and policy makers.[1]

Forecasters relied earlier on models with only a few predictors, until increasing amounts of data became available at high levels of sectoral, regional and temporal disaggregation. Those macroeconomic, microeconomic and financial time series hold information that may be useful for economic forecasting and empirical analysis of monetary policy (Ibarra-Ramírez, 2010). Bernanke and Boivin (2003) point out, however, that researchers who use a small number of variables in their analysis can exploit only a limited amount of information. Small scale models have some advantages in their simplicity and tractability, but they are prone to omitted variable bias (Gavin and Kliesen, 2008).

Factor models in which the individual macroeconomic and financial time series are driven by a small number of factors can be used to address the shortfalls of small scale models. First, factor models summarise the information contained in a big dataset, which allows a richer information set to be incorporated in the analysis. Second, factor models are flexible in the way that they can simultaneously accommodate data released at different times, frequencies and areas. Finally, their methods for extracting driving factors are statistically rigorous, as they are agnostic about the structure of the economy (Bernanke and Boivin, 2003).

This paper investigates the properties of the factor model forecast of the Estonian headline and core inflation for the period from the second quarter of 2011 to the second quarter of 2014. Factors are constructed using a principal component analysis and are then incorporated into different parametrised forecasting models. To evaluate the relative performance of the forecasting methods, the forecasting errors of the factor-augmented models are compared to a univariate benchmark model to assess their predictive abilities.

This paper contributes to the growing literature on forecasting in a data rich environment in three ways.[2] It is the first systematic study to analyse the applicability of a factor-augmented vector autoregressive model for forecasting inflation in Estonia. Second, it examines the importance of the number of factors in the inflation forecasting model, when the factors are extracted from datasets where consumer price indicators are excluded. Third, the paper analyses the impact of small changes in the dataset on the forecast error distributions of different factor-augmented forecasting models.

[1]Detailed discussions of the dynamics of inflation in Estonia are also provided in Dabušinskas (2005), Dabušinskas and Kulikov (2007), Arratibel et al. (2009), and Errit and Uusküla (2014).

[2]In addition those discussed in the literature review, other methods have been used for summarising and extracting information from high-dimensional datasets. Forni et al. (2000, 2005) popularised generalised dynamic models where the factors are estimated in the frequency domain. Bai and Ng (2009) use boosting as a method of selecting the predictors in factor-augmented autoregressions. A factor-augmented VARMA model was introduced by Dufour and Stevanović (2013). Stock and Watson (2012) propose a general shrinkage model based on pretests such as Bayesian Model Averaging (BMA), empirical Bayes or bagging. Banerjee et al. (2014) present forecasts using a factor-augmented error correction model (FECM). Comparisons and reviews of various factor forecasting models can be found in Eickmeier and Ziegler (2008) and Kim and Swanson (2013).

Four findings can be highlighted. First, factor model forecasts can improve upon an autoregressive forecast but in most cases the forecasting gain is limited. Second, some models with one factor have smaller forecasting errors when the factors are extracted from a big benchmark dataset. Third, certain big factor models that contain three factors perform better than models with fewer factors when the factors are taken from a smaller dataset where the consumer price indicators have been excluded. This indicates that the dataset size and dataset composition matter for forecasting performance. Fourth, the forecasting performance is less contingent upon small arbitrary changes in the dataset composition when the factors are extracted from a large dataset than is the case with small arbitrary changes in a small dataset.

This paper is organised as follows. Section 2 reviews briefly the existing literature. Section 3 discusses the econometric framework. Section 4 presents the data used in the econometric model. Section 5 presents the empirical results. The final section concludes. The main tables (Table 1 to Table 4) are displayed in the main text. Appendix A contains the factor analysis result tables and graphs and Appendix B contains the robustness test results. The Online Appendix C displays the data used in the benchmark model.

# 2. Literature review

Forecasting using factor models has received a considerable amount of attention in recent years. Various studies have provided compelling evidence in support of the factor model forecast methodology. However, the literature is less conclusive in answering questions of how many factors to use in the model, the size of the dataset and the forecasting horizon.

Stock and Watson (2002) review the forecast performance of factors, which they call diffusion indexes. The authors extract those factors from large datasets and estimate the consistency of models with time variation. They show that their diffusion index models, or factor models, offer substantial improvements over univariate autoregressive models, leading indicator and vector autoregressive (VAR) models in an out-of-sample forecast of the Federal Reserve Board's Index of Industrial Production.

Lin and Tsay (2005) compare the forecasts of simple factor models with those produced by advanced large predictor models like partial least squares, Bayesian model averaging, and combination forecasts models. Their findings indicate that partial least squares outperform other models in short-horizon forecasts using a dataset of 141 predictors. The factor model provides good forecast accuracy when the number of common components is between three and five.

Gosselin and Tkacz (2001) compare the forecasting performance of four different factor models with that of univariate models. They conclude that the factor models are as accurate as more advanced models in forecasting the Canadian inflation rate. They include 344 Canadian variables together with 110 US macroeconomic and financial variables. Small factor models that contain one, two or three factors yield the best forecast accuracy. The researchers provide evidence that gains in forecast efficiency can be obtained for a small open economy by combining foreign macroeconomic and domestic time series.

Angelini et al. (2001) extract up to four factors from large cross-sectional datasets comprising 278 variables for 11 EMU countries. They conclude that factor models have relatively good forecasting performance in four and eight quarter-ahead forecasts. Their findings indicate that small models with two or three factors match the best or better alternative forecast models in an out-of-sample forecasting framework, especially if those factors are related to nominal developments.

Bruneau et al. (2007) investigate the forecasting performance of dynamic factors, which are extracted from 200 macroeconomic variables for France. Their results indicate that the dynamic factor model has good forecasting properties, especially when forecasting the core inflation rate. Factors extracted from datasets with blocks of homogeneous variables, particularly variables related to labour markets, improve their forecasts considerably. They also provide small-horizon factor-augmented VAR forecasts, finding that the FAVAR forecasts outperform the standard dynamic linear regression forecasting equation models at times of rising core inflation.

Schumacher and Dreger (2004) study the performance of large-scale factor models for economic activity in Germany. They extract the factors from a dataset of 121 time series and calculate the prediction errors in out-of-sample forecasts, and they find that factor models outperform simpler univariate benchmark models. However, their forecasting gains prove to be limited and not systematic.

Artis et al. (2005) construct a dynamic factor model from a UK dataset consisting of 81 variables. They consider forecasting models with between four and twelve factors and up to three lags. Their results are in line with those of previous studies for the US, such as Stock and Watson (2002), who find that factor-based forecasts outperform standard benchmark models for price developments at both short and longer horizons.

The literature on factor model forecasts is less extensive for countries in Central and Eastern Europe (CEE), particularly for inflation forecasting. Ajevskis and Davidsons (2008) compare the forecasting performance of a diffusion index model with a generalised dynamic factor model for Latvia's gross domestic product (GDP). They use 126 quarterly time series to extract up to twelve factors. Both models outperform simpler models but the differences are not statistically significant. For short horizons a model with four factors and two lags provides the best forecasting performance but, models with more factors and zero lags lead to better forecasting results for longer horizons.

Stakenas (2012) focuses on Lithuanian GDP forecasting and uses simple and advanced principal component analysis to extract factors from a dataset of 52 monthly variables. He finds that that factor models outperform naive univariate benchmark models. The most suitable models for the Lithuanian case encompass two factors irrespective of whether the factors are extracted by a generalised or static principal component method. In addition, the forecasts produced by a state-space model give similar results to those from forecasting using the principal component method.

For Estonia, Schulz (2007) derives common factors with a small-scale state-space model and with a large-scale diffusion index model and subsequently forecasts real economic growth. The factor models show a better forecasting performance for most forecasting periods than univariate and multivariate benchmark models do. Schulz (2007) emphasises that even though many data series are available for the Baltic states, those series are not very long and this makes it difficult to compare the results with those from mature Western countries.

# 3. Empirical model

The forecasting model uses a two-step approach. First the factors are extracted and then they are incorporated in a forecasting model. This paper closely follows the static principal component approach of Stock and Watson (2002) for the factor extraction. The forecasting equation is based on the approach proposed by Bernanke et al. (2005).

## 3.1. Econometric framework

For the formal setup, assume $X_t$ to be an $N \times 1$ vector of time series with $t = 1, \ldots, T$. It is assumed that both $N$ and $T$ are large. Those time series are driven by a few ($q < N$) unobserved common factors. In the general formulation of a dynamic factor model, each element of the vector $X_{it} = [x_{1t} \ldots x_{it} \ldots x_{Nt}]'$, for $i = 1, 2, \ldots, N$ can be represented as:

$$X_{it} = \lambda_i(L) f_t + e_{it}, \tag{1}$$

where $f_t$ is the $q \times 1$ vector of common factors, $\lambda_i(L)$ is an $N \times q$ lag polynomial in non-negative powers of $L$ and $e_{it}$ is the idiosyncratic error term.

The lag polynomial adds dynamics to the factor loadings $\lambda_i$, which are the weights that form a linear combination of the original variable when multiplied with the latent component. The idiosyncratic disturbance $e_t = (e_{1t}, \ldots, e_{Nt})'$ is assumed to have limited cross-sectional and temporal dependence. In the dynamic form it is assumed that the idiosyncratic disturbance is uncorrelated with the factor innovation at all leads and lags, so $\mathbb{E} e_t \eta'_{t-k} = 0$ for all $k$. Moreover, it is assumed that $\mathbb{E} e_{it} e_{js} = 0$ for all $s$ if $i \neq j$, meaning the idiosyncratic errors are mutually uncorrelated at any leads and lags (Stock and Watson, 2011).

Equation 1 has an alternative formulation in finite lag form:

$$X_t = \Lambda F_t + e_t \tag{2}$$

where $F_t = (f'_t, \ldots, f'_{t-p})'$ is an $r \times 1$ vector, where $r = (p+1) \times q$ factors drive the variables. The $i$th row of $\Lambda$ is $(\Lambda_{i0}, \ldots, \Lambda_{iq})$, thus $\lambda(L) f_t$ would be the lag polynomial of the common component of the $i$th series. It can be observed that the high-dimensional time series variable vector, $X_t$ is driven by a vector of latent factors, $F_t$ and a vector of mean-zero idiosyncratic disturbances, $e_t$.

The static representation of the dynamic factor model yields the advantage that the factors can be estimated using principal components. It should be noted that since $X_t$ can contain lagged values, $F_t$ can be understood as containing arbitrary lags of factors. When the number of predictors $N$ and the number of observations $T$ grow large, the factors can consistently be estimated by the principal components of the $T \times T$ covariance matrix of $X_t$.[3] Stock and Watson (2002) show that consistency is even preserved in an approximate factor model with factor loadings and idiosyncratic errors that are serially

[3]The objective function for the estimation of the factors $F_t$ is given by:

$$V(F, \Lambda) = \min_{\Lambda, F} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \Lambda_i F_t)^2$$

and (weakly) cross-sectionally correlated. The intuition behind this property is that only the linear combination of factors will remain after the weighted averages of the idiosyncratic disturbances have converged to zero because of the law of large numbers (Stock and Watson, 2011).

The forecasting equation is based on the approach proposed by Bernanke et al. (2005), who extract the factors in a similar manner to Stock and Watson (2002) and then proceed by estimating a factor-augmented VAR. Though the variable of interest is the inflation rate, more economic variables could be incorporated in the VAR model. Let $Y_t$ denote an $M \times 1$ vector of observable macroeconomic variables. Along with the vector of observable time series, additional economic information is contained in a $k \times 1$ vector of unobserved factors, $F_t$. Given a vector $Y_t$ of important macroeconomic variables and a vector $F_t$ of unobserved driving factors, it is reasonable to assume joint dynamics for $(F_t, Y_t)$.

The joint dynamics are given by:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + \epsilon_t, \tag{3}$$

where $\Phi(L)$ is a conformable lag polynomial of finite order $d$ in the lag operator $L$ and $\epsilon_h$ is an error term with a mean of zero and a covariance matrix $Q$.

If the terms of $\Phi(L)$ that relate $Y_t$ to $F_{t-1}$ are all non-zero, the equation 3 is referred to as a factor-augmented vector autoregression, or FAVAR; otherwise this system reduces to a standard VAR in $Y_t$. Since it is assumed that $M + k \ll N$, the FAVAR model can handle more information than standard small-scale VAR models, as the informational content of the large $N$ size dataset is summarised in a small set of $k$ factors.

The $h$-step ahead forecast for $\begin{bmatrix} F_t \\ Y_t \end{bmatrix}$ is obtained recursively.

The point estimate obtained is compared to the actual observed value, forming the forecast error $\epsilon^h_{t+h}$ to calculate the root-mean-square errors (RMSE) (Hamilton, 1994).

## 3.2. Number of factors and lag structure

Factor forecast applications differ not only in the factor estimation method but also in the number of factors used. The basic factor approach suffers from an important shortcoming as the factors that are extracted are ordered on the basis of how they express the common movement in the whole dataset, but this does not take account of the specific variables being forecast. Nor is the forecast horizon considered, which could be of significance when targeted predictors co-move with the variable to be forecast more in certain periods than in others. Periods of stronger co-movement can be expected to yield better forecast performance (Eickmeier and Ziegler, 2008). Dias et al. (2010) point out that including only the first few factors in the forecasting equation might exclude other factors that have a high correlation with the target variable or the forecast horizon.

One important determinant for the predictive power of the factors and the number of them to be included in the forecasting equation is the size and composition of the dataset.

where $F = [F_1, \ldots, F_t, \ldots F_T]'$ and $\Lambda_i$ is the $i$- th row of $\Lambda$. $F$ and $\Lambda$ are subject to the constraint $\frac{F'F}{T} = I_q$ where $I_q$ is the $q \times q$ identity matrix. Hence, applying the principal components method mean the residual sum of squares is minimised subject to the normalisation that $\frac{F'F}{T} = I_q$.

Studies have shown the relevance of targeted predictors (Boivin and Ng, 2006; Bai and Ng, 2008). Oversampling problems are reported, somewhat in contradiction to the principle that large datasets are beneficial, when arbitrary variables, that are irrelevant for the time series to be forecast were added. Boivin and Ng (2006) point out that reducing the sample size can help sharpen the factor structure and as a result forecast efficiency improves when certain series show idiosyncratic error cross-correlations. That assumption is put to the test when the same number of factors is extracted from a benchmark sample set and a reduced size one.[4]

Where some studies base the number of factors on formal restrictions, others choose the number of factors heuristically. Following Bernanke et al. (2005), I use a heuristic approach and construct various FAVAR models with different numbers of factor and lag structures from different sized datasets, and use performance measures to assess their forecasting abilities. The reason for doing this is that the lag length and the number of statistically significant factors could be re-estimated in recursive out-of-sample forecasts for each period when the in-sample window is extended. However, the margin of in-between factor model comparison is reduced as it is more difficult to assess the impact of the number of factors and their lag structure on the forecasting performance.

## 3.3. Forecasting procedure and evaluation

Multistep ahead forecasts are made at one quarter to six quarter-ahead horizons, so $h = 1, \dots, 6$. I use a recursive pseudo out-of-sample forecasting method. The forecast performance is evaluated on the out-of-sample set. The in-sample set is used to initialise the methods of factor estimation, model estimation and lag order selection. The dataset starts at the first quarter of 2004 and ends in the second quarter of 2014. The choice of the starting date reflects the aim of incorporating a large number of balanced time series in the analysis. For every quarter, the forecast $h$-steps-ahead is obtained recursively.

From $y^h_{2011/2Q+h}$ to $y^h_{2014/2Q+h}$ the forecast mechanism reoccurs twelve times. The iterative forecasts at the end of the out-of-sample set produce forecast values that are not used for further analysis as the actual observed inflation and core inflation values were not available at the point of analysis. Therefore, fewer observations enter the forecast performance evaluation for larger $h$-steps-ahead forecasts.

To compare the forecast accuracy of the models, the root mean square errors (RMSE) are calculated for each model from the differences in the values for the quarter on quarter inflation rate.

So that the forecast results are comparable, the root mean square errors of all the forecast models are also computed relative to the root mean square error of the benchmark autoregressive forecasts. Therefore, the relative root mean square error of the benchmark AR is 1.00 or 100%.

[4]One way to separate targeted predictors from uninformative time series is proposed by Bai and Ng (2008). They suggest partitioning the panel of predictors into two subsets. The first subset should include all time series (targeted predictors) that are relevant for the specific variable to be forecast and the other subset should include all series that are non-informative. The partitioning is done with thresholds defined by the least absolute shrinkage and selection operator (LASSO) and the elastic net rules. While those shrinkage models are interesting from a technical perspective and most researchers in the field acknowledge the importance of targeted predictors, practitioners tend to rely on heuristics to determine which time series to include in their dataset.

I abstain from using the Diebold-Mariano test (Diebold and Mariano, 1995) to test formally the statistically significant difference between the models in predictive abilities.[5] Researchers tend to conduct forecasting exercises on different time periods but testing the model on different time periods proves difficult in the case of Estonia, as the length of the data sample for the factor estimation is limited. Instead I test for the impact of removing one observation by excluding the second quarter of 2014 from the calculation of the absolute RMSE for every forecast horizon. The new RMSE are calculated using data from the second quarter of 2011 to the first quarter of 2014. If the RMSE do not deviate by significant margins between the two time periods, the results obtained are considered to be robust for small changes.

In addition to testing for the impact of small changes to the forecasting period, I draw 2000 random samples from the benchmark dataset of 388 variables (see Section 4) and create datasets of 329 variables. The same principle is applied to the reduced dataset, with the number of variables in each random draw cut by 37, or about 15 per cent. Those 2000 different datasets are used to extract the factors and forecast the headline and core inflation rates in the way described earlier. In the next step the distributional properties of the 2000 consecutive individual model forecast errors are analysed. Specifically, I plot the frequency distribution of the FAVAR models and analyse their shape, centre, spread and position relative to the benchmark AR model.

## 3.4. Factor-augmented vector autoregressive forecast models

The FAVAR forecasts are constructed by choosing the number of factors to be included and the lag order. I estimate 13 FAVAR models, the results for seven of which are reported in detail.[6] All the 13 FAVAR forecasts share the same properties for the $M$ vector. The $M$ vector is a one-variable vector that contains either the headline inflation rate or the core inflation rate, depending on the forecasting exercise.

For the models with a fixed lag length, I start testing from small dimensional FAVAR models and then add more factors and lags. "FAVAR 1F.1 Lag" contains the first factor (1F.) and has a lag length of one (1 Lag). "FAVAR 12F. 1 Lag" is a three variable vector, containing the inflation rate plus the first two factors (12F.). The model "FAVAR 123F. 1 Lag" contains the third factor as well. Equal size $k$ - factor models where also tested for lag lengths of two and three.

The forecast results of the FAVAR models are compared to the results of the benchmark model. Following Stock and Watson (2002), a univariate autoregressive model of order $p$ is used as the benchmark. The benchmark AR is based on the headline inflation rate and the core inflation rate. The lag length of the estimated lag polynomial is iter-

[5] The Diebold-Mariano test suffers from two shortcomings when the forecasting approach of Bernanke et al. (2005) is followed. First, finite sample properties of the estimators on which the forecasts may depend are not preserved asymptotically. Second, the DM-test is prone to nested model bias (Giacomini and White, 2006). That presents a problem under the out-of-sample extending window forecasting procedure when the competing forecasts are obtained from autoregressive and factor-augmented vector autoregressive models.

[6] The six models not reported include a FAVAR forecast where the lag order is allowed to vary, a model including the first five factors and models including only the second factor at different lag lengths. The forecasting results for those models are available upon request.

atively estimated by BIC, and is allowed to vary between one and three ($1 \leq p \leq 3$).[7] Given that the ARMA model forecasts do not improve upon the AR model forecasts, they are not reported in the result section.[8]

The forecasting abilities of the FAVAR models are also tested against a fixed-lag VAR model. Stock and Watson (2002) use the monthly growth in real activity, the change in monthly inflation and the change in the 90-day US treasury bill rate in their study. In the absence of T-bill equivalents for Estonia, I construct a VAR with the quarterly change in total real GDP, the quarterly change in total unemployment and the quarterly change in the headline inflation rate. The lags are allowed to vary between one and three ($1 \leq p \leq 3$).

Stock and Watson (2010) also posit that since the financial crisis it has become increasingly difficult to improve systematically upon simple univariate forecasting models like the random walk model by Atkeson and Ohanian (2001). Therefore, a random walk model constitutes the last alternative benchmark model.

# 4. Data

The data section contains two parts. Subsection 4.1 briefly presents the variables and their treatment in the dataset and Subsection 4.2 reports the results of the factor analysis.

## 4.1. Variables

The series chosen for the panel used in the analysis are similar to the variables used by Stock and Watson (2002). First, credit aggregates such as credit to firms and households are included along with data for different credit maturities, such as long-term and short-term credit. Similarly, series such as deposits from companies and deposits from individuals have been included. State budget revenues and state budget expenditures series are used in addition.

Various interest rates such as the 6-month Euribor rate and short-term interest rates are included in the dataset as are money supply aggregates such as the M3 rate and key data on the balance of payments. Further statistics on trades in consumer and capital goods are used so as to account for Estonia's open economy structure. The series of the composite leading indicators (CLI) may help to predict the future economic climate and are also included.

Labour market dynamics can play a significant role in the development of wages and prices, and I include the unemployment and job vacancy rates among other statistics.

[7] To ensure that the AR model constitutes a competitive benchmark model, the RMSE of different lag length intervals were compared. Neither a fixed lag order of one, two or three lags nor intervals ranging from $1 \leq p \leq 2$ to intervals up to $1 \leq p \leq 12$ show lower forecasting errors for the benchmark model than the forecasts obtained from AR models where $p$ is allowed to vary between one and three.

[8] The results for the autoregressive moving average (ARMA) are identical to the results from the AR benchmark forecasts. Within the order constraints given, which are a maximum of three lags for any autoregressive component and a maximum of three lags for any moving average component, the Bayesian information criterion (BIC) determined unanimously that the given process does not include any moving average terms. Therefore, the lag structure is equal to the lag structure of the benchmark AR process, and the forecasting results are identical.

Next, I took in data on the output of total, intermediate and capital goods, and data on new orders such as new orders for manufacturing goods. Like the composite leading indicators, business survey statistics give information on economic expectations, so turnover and sales are included in the dataset as they can be seen as indicators of consumer sentiment.

Following the findings of Gosselin and Tkacz (2001), who conclude that the macroeconomic dynamics of trading partners are of importance for factor modelling of inflation and output in open economies, I also consider price aggregates and the composite leading indicator series of Estonia's biggest trading partners. The aggregate PPI index for the whole euro area enters the panel as a total, as well as the individual PPI indexes for Finland, Lithuania, Latvia, Germany and other main trading partners. In addition the indexes are split up into sub-categories such as producer prices for energy, and food and beverages.

Another major group of variables is the harmonised consumer price indexes (HICP) of Estonia and Estonia's trading partners. The foreign consumer price indexes can be interpreted as foreign inflation proxies. First, the HICP series from trading partners in the European Union are included in the dataset and second, the sub-indices such as the HICP energy series or HICP food and beverages series also enter the dataset. In total more than 140 different harmonised consumer prices indexes are included in the dataset.

Financial market dynamics should be also considered, so I include stock price data from the Helsinki stock exchange (OMXH) and the Russian RTS index. The effects of productivity changes are captured by the incorporation of data on the number of hours worked, average wages by employment and nominal and real unit labour costs. The last major items included in the dataset are various economic deflators.

Only a few variables on personal consumption are available for Estonia and the same applies to detailed payroll and housing sales statistics. There are no Estonian sovereign debt securities or Estonian inflation-protected securities. This is unfortunate as inflation-protected securities may be used to compute measures of inflation expectations (Shen and Corning, 2001).

The variables to be forecast are the Estonian headline inflation rate and the Estonian core inflation rate. Headline inflation is defined as the official measure of consumer price inflation in Estonia for goods and services. Core inflation is a sub-category of headline inflation that excludes energy, food, alcohol and tobacco items.

The first panel used in this paper consists of 388 domestic and foreign time series at 42 quarterly observations, ranging from the first quarter of 2004 until the second quarter of 2014. This panel is labelled the "benchmark dataset".

To test for panel size effects and targeted predictor effects, a second panel with 246 time series was created. Its basis is the benchmark dataset, with all domestic and foreign harmonised consumer price indexes excluded. This allows possible changes in the underlying variance and forecasting performance to be analysed. This panel was labelled the "reduced dataset", or for clarity, the "reduced-size dataset". A complete list of the variables used in the benchmark dataset is reported in the Online Appendix (see Table: C.1).

The untreated dataset contained monthly and quarterly time series, so the monthly series were transformed into quarterly series. The process of transformation involved averaging the monthly values to quarterly values, summing up the monthly values, or taking the end of the last month value as the quarterly value.

Missing observations were treated with a regularised iterative missing principal component analysis algorithm to avoid overfitting problems associated with using an expected-maximisation (EM) algorithm. In the next step the seasonal effects were removed from the set of variables. Time series that were already seasonally adjusted according to the issuing source were still put through this stage to remove any residual seasonality. The augmented Dickey-Fuller test was performed on all the seasonally adjusted time series. Non-stationary series were marked and then subjected to the stationarity inducing transformation. The transformations involved taking the log differences for series that included non-negative values. For series that included positive and negative values the first difference was taken. The exact treatment of every time series can be found in the Online Appendix C.1. In the last step, all the series were standardised to have sample mean zero and unit sample variance.

## 4.2. Factor analysis

I start the analysis of the factors with the benchmark dataset ($N = 388$). As described in section 3.4, a maximum of five factors is used in the vector autoregressive models. Principal components summarise the variance in a dataset. The first component explains approximately 21.94 per cent, the second one 16.68 per cent, the third 7.99 per cent, the fourth 5.65 per cent, and the last one 3.82 per cent of the total variance in the dataset. The cumulative share of the total variation of the macroeconomic variables explained by the first three factors is 46.61 per cent and that explained by the first five factors is 56.08 per cent.

For the reduced-size dataset, the variance explained by the first principal component is almost six percentage points more than the variance explained by the first principal component in the big dataset. The cumulative explained variance of the first three common components equals 46.24 per cent and for the first five components it is 56.44 per cent, which is approximately the same as in the big dataset.

In the next step, the latent common components are extracted. The dynamics over the span of the dataset of these factor indexes are captured in the time series plot of Figure A.1 in the appendix. To make the presentation clearer, only the first three factors are depicted. The initially unobserved factor dynamics are plotted together with the observed headline inflation rate. The visual analysis indicates that all three factors show either strong co-movements or converse movements with the inflation rate. Those movements seem either to coincide with or to lead the inflation rate, which should give them predictive abilities. For the smaller dataset ($N = 246$), co-movements of the factors and the inflation rate are visible but not as conspicuous.

The correlation between the observed variables and the unobserved common component can be analysed by extracting the variables that are most characteristic for each dimension obtained by principal component analysis. This mean that the statistically significant variables are identified and ranked by their correlation coefficient for the particular factors. The significance threshold at which a variable characterises the dimension is set at 0.05. Only variables with the 10 highest positive and negative correlation coefficients are extracted and analysed.

An example of the correlation between the observed variables and the unobserved common component is given in Table A.1 in the Appendix. The table reports the correlation of variables with the direction of the first factor. This factor loads heavily on the producer price indexes (PPI) of Estonia's trading partners, with the Finnish producer price index (PPI) excluding construction being the most important. PPI Industry Lithuania (ex construction) and PPI Intermediate goods of the European Union 15 are ranked as the fourth and fifth most important variables in terms of correlation. The turnover and sales of intermediate goods and the output of intermediate goods are also strongly positively correlated with the first factor.

# 5. Results

The factor analysis can provide useful insights into the factor dynamics. The next section tests the theory that more informative factors can be derived from larger datasets, creating more accurate predictions. The results section consists of three parts. In the first part the forecasting results for VAR models augmented with factors from the benchmark dataset are reported. Second, I repeat the forecasting exercise on the reduced-size dataset and analyse the results. The last part provides a short robustness check by analysing the sensitivity to small changes in the size of the datasets and the time period.

## 5.1. Benchmark dataset forecasting results

The results for the forecast errors from the benchmark dataset are reported in Table 1. First, the relative root-mean-square errors (RMSE) of the benchmark autoregressive (AR), the alternative forecasts and the FAVAR forecasts are reported. The RMSE of each of the forecasting models are shown relative to the RMSE of the benchmark AR model (so the autoregressive forecast has a relative RMSE of 1.00). The six columns show the relative forecast error for one to six quarter-ahead forecasts. To give an example, the forecast error of the simulated alternative random walk model (RW) is 114.7 per cent of the forecast error of the autoregressive forecast at the one quarter horizon. Obviously, low values of RMSE indicate smaller forecast errors. The results for the lowest relative RMSE, which indicates the highest predictive abilities of the factor models, are given in bold.

The last row in the table shows the absolute RMSE of the autoregressive benchmark for the given forecast horizon. The absolute RMSE of the benchmark AR model can be interpreted as the percentage deviation of the forecast point estimates from the actual observed values over the full forecast window.

First, in many cases the performances of the FAVAR forecasts are better than those of the benchmark forecasts, but the differences are generally quite small. For example, from the results of the one quarter-ahead forecasts in the first column, it can be observed that factor augmented vector autoregressive forecasts including only the first factor show an improvement in forecasting performance over the AR benchmark and the other alternative models.

In line with the results of Stock and Watson (2002), models with a low lag order tend to perform better for all horizons. In most cases the FAVAR models with two lags show the best forecasting performance, and they show a tendency to improve on the benchmark at short-horizons. The smallest forecast errors are usually obtained for forecasts two quarters ahead. Forecasting one year ahead, only the FAVAR model with the first factor

Table 1: Headline Inflation Out-of-Sample Forecasting Results 1-6 Quarters Horizon

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 1.308 | 1.503 | 1.560 | 1.527 | 1.539 | 1.624 |
| RW | 1.147 | 1.024 | 0.872 | 0.817 | 0.804 | 0.824 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.992 | 1.030 | 1.020 | 1.022 | 1.057 | 1.101 |
| FAVAR 1F. 2 Lags | 0.902 | **0.953** | **0.951** | **0.944** | 0.980 | **1.025** |
| FAVAR 1F. 3 Lags | **0.901** | 0.997 | 1.012 | 1.018 | 1.049 | 1.088 |
| FAVAR 12F. 1 Lag | 1.043 | 0.990 | 0.992 | 1.004 | 1.048 | 1.098 |
| FAVAR 12F. 2 Lags | 1.043 | 0.969 | 0.991 | 1.011 | **0.977** | 1.061 |
| FAVAR 12F. 3 Lags | 1.148 | 0.999 | 1.023 | 1.037 | 1.013 | 1.086 |
| FAVAR 123F. 1 Lag | 1.047 | 0.995 | 0.988 | 0.993 | 1.027 | 1.066 |
| | | | | | | |
| Benchmark AR, abs. RMSE | 0.474 | 1.004 | 1.666 | 2.256 | 2.698 | 3.000 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points.

and two lags offers an improvement of five percentage points over the benchmark. For the forecasts six quarters ahead no FAVAR model is able to outperform the benchmark AR model. In contrast, the random walk model seems to capture the inflation dynamics appropriately on longer horizons, outperforming the benchmark AR model by almost 20 percentage points.

Turning to the results for the core inflation in Table 2, it can be seen that the absolute RMSE of the benchmark model are smaller than those for the headline inflation forecasts. This is in line with the theoretical arguments; given that the core inflation rate is less volatile than the headline inflation rate, it should be easier to forecast and therefore should yield smaller forecasting errors.

Table 2: Core Inflation Out-of-Sample Forecasting Results Benchmark Dataset

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 0.926 | 1.042 | 1.414 | 1.709 | 2.294 | 2.371 |
| RW | 1.387 | 1.625 | 1.846 | 1.869 | 2.357 | 2.252 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.908 | **0.890** | **1.058** | 1.151 | 1.435 | 1.475 |
| FAVAR 1F. 2 Lags | 0.923 | 1.000 | 1.198 | 1.231 | 1.481 | 1.569 |
| FAVAR 1F. 3 Lags | 0.934 | 1.091 | 1.447 | 1.566 | 1.876 | 1.930 |
| FAVAR 12F. 1 Lag | 0.919 | 0.943 | 1.145 | 1.236 | 1.501 | 1.449 |
| FAVAR 12F. 2 Lags | 0.993 | 1.177 | 1.509 | 1.628 | 2.003 | 2.090 |
| FAVAR 12F. 3 Lags | 1.034 | 1.154 | 1.464 | 1.546 | 1.959 | 1.932 |
| FAVAR 123F. 1 Lag | **0.906** | 0.903 | 1.076 | **1.150** | **1.387** | **1.285** |
| | | | | | | |
| Benchmark AR, abs. RMSE | 0.339 | 0.516 | 0.635 | 0.807 | 0.797 | 1.003 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points.

The FAVAR models tend to have slightly higher predictive abilities for headline inflation than for core inflation and this is especially true at longer forecast horizons. The forecasts from the FAVAR models outperform the benchmark AR forecasts only on the one and two quarter forecast horizon for core inflation.

## 5.2. Forecasting results for the reduced dataset

The forecasting results for headline inflation using the reduced size dataset are shown in Table 3. It may be presumed that domestic and foreign consumer price indexes constitute important predictors of the Estonian headline and the core inflation rates. Removing those targeted predictors may change the factor structure and consequently the forecasting performance of the FAVAR model. Decreasing the size of the dataset and removing the targeted predictors may thus produce higher absolute RMSE, indicating lower predictive abilities. It is, however also possible that decreasing the sample size from 388 to 246 variables will lead to the removal of less important predictors that dilute the extracted factors, resulting in a set of factors which can be used to calculate FAVAR forecasts that have lower absolute RMSE than the FAVAR forecast with factors extracted from the benchmark dataset. As the underlying time series of the benchmark model have not changed, and so neither have their absolute RMSE, the absolute and relative forecasting errors of the FAVAR models can be directly compared between the two different sized datasets.

Table 3: Headline Inflation Out-of-Sample Forecasting Results 1-6 quarters horizon

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 1.308 | 1.503 | 1.560 | 1.527 | 1.539 | 1.624 |
| RW | 1.147 | 1.024 | 0.872 | 0.817 | 0.804 | 0.824 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 1.031 | 1.069 | 1.050 | 1.056 | 1.069 | 1.079 |
| FAVAR 1F. 2 Lags | 0.952 | 1.010 | 0.981 | 0.988 | 1.019 | 1.031 |
| FAVAR 1F. 3 Lags | **0.949** | 1.063 | 1.047 | 1.049 | 1.074 | 1.076 |
| FAVAR 12F. 1 Lag | 1.041 | 1.071 | 1.052 | 1.063 | 1.090 | 1.085 |
| FAVAR 12F. 2 Lags | 0.963 | 1.023 | 0.978 | 0.971 | 1.018 | 1.050 |
| FAVAR 12F. 3 Lags | 1.017 | 1.088 | 1.075 | 1.125 | 1.161 | 1.198 |
| FAVAR 123F. 1 Lag | 1.075 | **1.006** | **0.955** | **0.920** | **0.919** | **0.991** |
| | | | | | | |
| Benchmark AR, abs. RMSE | 0.474 | 1.004 | 1.666 | 2.256 | 2.698 | 3.000 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points.

Stock and Watson (2002) found that the performance of comparable models is usually better when factors from a full dataset are used than when those from a reduced size subset are used. However, the assumption that removing predictors from the dataset would lead to worse RMSE values cannot be confirmed by the results obtained for Estonia. First, the best performing headline FAVAR models from the benchmark dataset contain fewer factors than the best performing headline FAVAR models from the reduced dataset. Second, for forecasts within one to three quarters the FAVAR model with the first factor shows the lowest RMSE in the benchmark dataset, whereas the FAVAR model with the first three factors shows lower forecasting errors in the reduced size dataset for the four to six quarters horizon. Comparing these different models, I see the forecasting performance is quite similar and most models outperform the benchmark by small margins.

Table 4 shows the forecasting results for core inflation when the reduced dataset is used. I already know that the benchmark AR model is the main competitor to the FAVAR forecasts, as the random walk model does not seem to capture very well the less volatile dynamics of the core inflation rate.

Notable differences appear when the headline and the core inflation forecasts are compared within each of the two datasets and also between the benchmark and the reduced-size datasets. The best headline FAVAR forecasts show lower forecasting errors than the benchmark AR for the forecasts one, three, four, five and six quarters ahead even though the performance improvement is partly weak in economic terms. The best core FAVAR model forecasts improve upon the benchmark AR in all the one to four quarter-ahead forecasts. However, the forecast improvement gains are much higher, especially for the forecasts one, two and three quarters ahead. To give an example, for the forecasts one to four quarters ahead, the core FAVAR models improve on average upon the headline FAVAR forecasts by 11 percentage points.

Table 4: Core Inflation Out-of-Sample Forecasting Results 1-6 Quarters Horizon

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 0.926 | 1.042 | 1.414 | 1.709 | 2.294 | 2.371 |
| RW | 1.387 | 1.625 | 1.846 | 1.869 | 2.357 | 2.252 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.928 | 0.903 | 1.051 | 1.136 | 1.385 | 1.375 |
| FAVAR 1F. 2 Lags | 0.939 | 1.029 | 1.220 | 1.253 | 1.522 | 1.579 |
| FAVAR 1F. 3 Lags | 0.909 | 1.059 | 1.419 | 1.534 | 1.872 | 1.896 |
| FAVAR 12F. 1 Lag | 0.897 | 0.884 | 1.046 | 1.134 | 1.382 | 1.36 |
| FAVAR 12F. 2 Lags | 0.925 | 1.020 | 1.227 | 1.250 | 1.531 | 1.583 |
| FAVAR 12F. 3 Lags | 1.078 | 1.224 | 1.671 | 1.888 | 2.126 | 2.129 |
| FAVAR 123F. 1 Lag | **0.833** | **0.728** | **0.881** | **0.935** | **1.104** | **1.175** |
| | | | | | | |
| Benchmark AR, abs. RMSE | 0.339 | 0.516 | 0.635 | 0.807 | 0.797 | 1.003 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points.

Comparing the headline inflation forecasts of the benchmark dataset with the headline inflation forecasts of the reduced dataset, it can be seen that FAVAR models with one factor have the lowest forecasting errors in the benchmark dataset whereas FAVAR models with the first three factors have the lowest RMSE in the reduced dataset. The differences in forecasting errors between these two forecasting models are however small and not systematic.

The results for the core inflation forecasts are more conclusive. Not only can a tendency for multi factor models to have better forecasting abilities than models with only the first factor be observed, but the best performing core FAVAR forecasts are obtained when the factors are extracted from the reduced size dataset. In addition, those forecasts errors are the smallest of all the models for all forecasting horizons.

The results from Tables 1-4 indicate that the forecasting performance of the FAVAR models is directly related to the number of factors included in the model. There is a clear tendency for FAVAR models with the first three factors to have higher predictive abilities than models containing only the first factor when those factors are extracted from a reduced size dataset. The forecasting performance also depends on the number of factors and the inflation measure to be forecast. These dynamics are interesting and deserve some discussion.

One possible explanation why models including the first three factors have a similar forecasting performance to that of models with only the first factor, depending on the size of the dataset, is that the information content of the benchmark dataset is higher than the information content of the reduced size dataset. When the factors are extracted and included in a FAVAR model, the number of factors needed in the model to exhibit good predictive abilities reflects the additional information content of the dataset. A FAVAR model with only the first factor from the benchmark dataset seems to capture an appropriate amount of additional predictive information. In contrast, the first three factors have to be included in a FAVAR model to obtain similar predictive abilities when those factors are extracted from a reduced size dataset with presumably lower information content. This is in line with the arguments in Boivin and Ng (2006) where it was shown that using more data would lead to inferior results when forecasting power is provided by a factor which is dominant in small samples but dominated in larger samples.

The second question that arises is why the core inflation forecasts with factors extracted from the reduced-size dataset improves upon the benchmark dataset and headline inflation forecasts by significant margins. One possible reason underlying this observation may be derived from the factor analysis in Section 4.2. When the factors are extracted from the small dataset, their dynamics are less pronounced. It may be conjectured that the interdependencies of these three factors and the inflation rate in the vector autoregressive system are more accurate in capturing the less volatile dynamics of the core inflation rate.

Finally, the results for the random walk forecasts deserve attention. The only model which consistently outperforms the benchmark AR by economically meaningful margins is a unit-root-based forecast, which is arguably a surprise. Comparing the results of the random walk forecast for headline inflation and for core inflation, it is clearly observable that random walk forecasts have substantially better forecasting abilities than all other models when headline inflation is forecast three to six quarters ahead. No such pattern is visible for core inflation, and the random walk forecast tends to worsen with an increasing forecast horizon.

The results for headline inflation are in line with the findings in Atkeson and Ohanian (2001) who found that backward-looking Philips curve forecasts can-not improve upon naive random walk models. Even though it has been shown that those findings are sensitive to the sample period and the parametrisation of the Philips curve model, Stock and Watson (2007) admit that on average, it is difficult for multivariate models to beat simple univariate models. Stock and Watson (2007) argue therefore that the value added of more complex multivariate models compared to simple univariate models is limited.

## 5.3. Robustness analysis

Analysing the sensitivity of the forecasts to small changes in the time period reveals that the results obtained for both datasets tend to be quite robust; see Tables B.1-B.4 in Appendix B. When one quarter is removed from each period for which the forecast errors are calculated, the root-mean-square errors change only a little in most cases. The models with the best forecasting abilities in the full datasets also tend to have the highest predictive abilities in the datasets where one period was removed. Overall, the robust forecasting performance can be attributed to the models obtained from the benchmarked and the reduced size datasets. The results for the core inflation forecasts indicate even less sensitivity to changes in the time sample.

Next, I analyse whether small changes in the composition of the dataset affect the forecasting results. The frequency distribution for the headline inflation forecasts with factors extracted from the benchmark dataset can be found in Figure B.1. The vertical line represents the AR benchmark, while the frequency plots report the distribution of the forecasting errors of the FAVAR models. Model distributions that are to the left of the vertical line have lower forecasting errors than the benchmark AR model does.

The distribution of the forecast errors supports the hypothesis that FAVAR models with the first factor and two lags (FAVAR 1F. 2 Lags) outperform the benchmark model. Up to the forecast horizon five quarters ahead, the mass of the distribution is centred clearly to the left of benchmark AR model with a low spread. On the five quarter-ahead horizon, the mass is still centred to the left of the benchmark, although for some samples the RMSE are larger than for the benchmark. Forecasting six quarters ahead, the FAVAR models do not manage to outperform the benchmark in the majority of the sample cases.

The sampling distribution for core inflation (see Figure B.2) shows that five FAVAR models have lower forecast errors than the benchmark AR model has on the one and two quarter forecasting horizons. In particular, the RMSE of the samples from the FAVAR models with the first factor and one lag (FAVAR 1F. 1 Lag) and the first three factors and one lag (FAVAR 123F. 1 Lag) clearly outperform the benchmark AR. Analysing the graphs over longer forecasting horizons leads to the same conclusions as those in the analysis of Table 2, as all the FAVAR model forecasts fail to outperform the benchmark AR forecasts in most cases.

When the number of variables in the reduced size dataset is decreased by 15 per cent, the forecasting error frequency distribution for headline inflation (see Figure B.3) shows less stable behaviour. At the one quarter forecasting horizon, the distribution of the forecasting errors of the FAVAR model with the first factor and two lags (FAVAR 1F. 2 Lags) and the FAVAR models with the first factor and three lags (FAVAR 1F. 3 Lags) are clearly to the left of the benchmark AR model. At the two quarter forecasting horizon, the mean of the FAVAR model with the first factor and two lags (FAVAR 1F. 2 Lags) is centred slightly to the right of the benchmark. More interestingly, the FAVAR model with the first three factors and one lag (FAVAR 123F. 1 Lag) has a spread distribution with a mean slightly to the left of the benchmark AR. For the one factor two lags or one factor three lags FAVAR models, the distributions tend to be to the left of the benchmark value for the three and four quarter-ahead forecasts. The distributions of the first three factors model (FAVAR 123F. 1 Lag) are spread between a range of approximately 0.9 and 1.1, with a tendency to be centred slightly to the left of the benchmark AR.

For core inflation, the distributional properties (see Figure B.4) of the forecasting errors of the FAVAR models are similar to those of the headline inflation forecasts. On short forecasting horizons of one and two quarters, the FAVAR model including the first factor and one lag (FAVAR 1F. 1 Lag), the model with the first and the second factor and one lag (FAVAR 12F. 1 Lag) and the model with the first three factors (FAVAR 123F. 1 Lag) clearly outperform the benchmark AR, even when the reduced size dataset is shrunk in size by 15 per cent. At the three and four quarters forecasting horizon, only the model with the first three factors (FAVAR 123F. 1 Lag) tends to outperform the benchmark AR. However, the distribution is spread out and asymmetric. Similarly, for the five and six quarter forecast horizon, some samples from the same model (FAVAR 123F. 1 Lag) show lower forecasting errors than those of the benchmark, but most values lie at the right side of the spread-out distribution.

In summary, the results show that small arbitrary changes to the number of variables in the two datasets have only a small impact on the forecast performance of FAVAR models including only the first factor. For the reduced size dataset, however, a different dataset composition has substantial effects on the FAVAR model with the first three factors and one lag (FAVAR 123F. 1 Lag). At the three to four quarter forecasting horizon in particular, the slightly asymmetric spread of the distribution around the benchmark AR value of one makes it complicated to draw a conclusion as to whether the FAVAR 123F. 1 Lag model forecasts outperform the benchmark model or not. This indicates that arbitrary changes to the number of predictors have a stronger impact on the reduced-size dataset than on the larger benchmark dataset.

# 6. Final comments

This paper investigates the performance of factor-augmented vector autoregressive models when used to predict the Estonian headline and core inflation rates. The factors are extracted by a principal component method from a big benchmark dataset with 388 quarterly economic and financial time series, and a reduced size dataset consisting of 246 series. The FAVAR forecasts range from the second quarter of 2011 to the second quarter of 2014 and their forecasting errors are compared to naive benchmarks, such as an autoregressive forecast.

The analysis of the forecasts of Estonian headline and core inflation at various forecast horizons and using different sample sets yields interesting and arguably surprising results. The results indicate that only a small number of factors are needed in order to forecast the Estonian headline and core inflation rates. FAVAR models can improve upon naive forecasts but the forecasting gains for most FAVAR models are, with notable exceptions, small. The evidence points to a pattern where models with fewer factors exhibit good predictive performance at short horizons when the factors are extracted from the benchmark dataset.

Surprisingly, essentially similar forecasting results for the Estonian inflation rate, and even better ones in certain cases, emerge when the factors are extracted from a reduced-size dataset that excludes domestic and foreign consumer price indicators. These well performing forecasts can be obtained from FAVAR models with the first three factors and one lag. However, the robustness analysis for this model indicates that small changes in the composition of the reduced-size dataset might have a substantial impact on the first three factors and therefore also on the forecasting performance.

Even though the results point to notable differences between the headline and core inflation forecasts, a clear statement of whether FAVAR models are better suited to forecasting one or the other is difficult to derive. Headline inflation forecasts show a tendency to perform better at longer horizons, whereas core inflation forecasts have slightly better predictive abilities at short horizons when the factors are extracted from the benchmark dataset. However, for the FAVAR models with more factors when the factors are extracted from the reduced-size dataset, the core inflation results clearly outperform the headline inflation results in the first four quarters.

The findings provide evidence that simple factor model forecasts such as factor-augmented vector autoregressive models can improve upon naive forecasts under certain circumstances. The forecast performance depends greatly on the number of factors included in the model, the size of the dataset from which the factors are extracted, the time series to be forecast, and lastly, the forecasting horizons.

Forecasting of inflation still remains challenging and this also applies to Estonian inflation. Among the models examined, substantial forecasting gains can only be reaped from two distinct models. Even from the perspective of an experienced forecaster it is still difficult to assess a priory how many factors should be incorporated in the model in relation to the size of the dataset. For Estonia, the results indicate that using a FAVAR model with the first factor extracted from a large dataset provides good forecasting performance, even when the exact dataset size and composition are unknown.

# References


Ajevskis, V. and Davidsons, G. (2008). Dynamic factors models in forecasting Latvia's gross domestic product. *Working paper: Bank of Latvia*, 2.

Angelini, E., Henry, J., and Mestre, R. (2001). Diffusion index-based inflation forecasts for the euro area. *Working paper: European Central Bank*, 61.

Arratibel, O., Kamps, C., and Leiner-Killinger, N. (2009). Inflation forecasting in the new eu member states. *Working paper: European Central Bank*, 1015.

Artis, M. J., Banerjee, A., and Marcellino, M. (2005). Factor forecasts for the UK. *Journal of Forecasting*, 24(4):279–298.

Atkeson, A. and Ohanian, L. E. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis Quarterly Review*, 25(1):2–11.

Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2):304–317.

Bai, J. and Ng, S. (2009). Boosting diffusion indices. *Journal of Applied Econometrics*, 24(4):607–629.

Banerjee, A., Marcellino, M., and Masten, I. (2014). Forecasting with factor-augmented error correction models. *International Journal of Forecasting*, 30(3):589–612.

Benkovskis, K., Kulikov, D., Paula, D., and Ruud, L. (2009). Inflation in the baltic countries. *Bank of Estonia, Kroon & Economy*, 1(2).

Bernanke, B. S. and Boivin, J. (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics*, 50(3):525–546.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120(1):387–422.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):pp. 169–194.

Bruneau, C., De Bandt, O., Flageollet, A., and Michaux, E. (2007). Forecasting inflation using economic indicators: the case of france. *Journal of Forecasting*, 26(1):pp. 1–22.

Dabušinskas, A. (2005). Money and prices in Estonia. *Working paper: Bank of Estonia*, 7.

Dabušinskas, A. and Kulikov, D. (2007). New Keynesian Phillips curve for Estonia, Latvia and Lithuania. - *Working paper: Bank of Estonia*, 7.

Dias, F., Pinheiro, M., and Rua, A. (2010). Forecasting using targeted diffusion indexes. *Journal of Forecasting*, 29(3):341–352.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, (13):256–263.

Dufour, J.-M. and Stevanović, D. (2013). Factor-augmented VARMA models with macroeconomic applications. *Journal of Business & Economic Statistics*, 31(4):491–506.

Eickmeier, S. and Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27(3):237–265.

Errit, G. and Uusküla, L. (2014). Euro area monetary policy transmission in estonia. *Baltic Journal of Economics*, 14(1-2):55–77.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics*, 82(4):pp. 540–554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association*, 100(471):830–840.

Gavin, W. T. and Kliesen, K. L. (2008). Forecasting inflation and output: comparing data-rich models with simple rules. *Federal Reserve Bank of St. Louis Review*, 90:175–192.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6):pp. 1545–1578.

Gosselin, M.-A. and Tkacz, G. (2001). Evaluating factor models: an application to forecasting inflation in Canada. *Working paper: Bank of Canada*, 18.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Ibarra-Ramírez, R. (2010). Forecasting inflation in Mexico using factor models: Do disaggregated CPI data improve forecast accuracy? *Working paper: Bank of Mexico*, 1.

Kim, H. H. and Swanson, N. (2013). Large dataset mining using parsimonious factor and shrinkage methods.

Lin, J. and Tsay, R. S. (2005). Comparisons of forecasting methods with many predictors. *Working paper: National DongHwa University*.

Schulz, C. (2007). Forecasting economic growth for Estonia: application of common factor methodologies. *Working paper: Bank of Estonia*, 9.

Schumacher, C. and Dreger, C. (2004). Estimating Large-Scale Factor Models for Economic Activity in Germany: Do they Outperform Simpler Models?/Die Schätzung von großen Faktormodellen für die deutsche Volkswirtschaft: Übertreffen sie einfachere Modelle? *Jahrbücher für Nationalökonomie und Statistik*, 224(4):731–750.

Shen, P. and Corning, J. (2001). Can TIPS help identify long-term inflation expectations? *Federal Reserve Bank of Kansas City: Economic Review*, 86(4):61–87.

Stakenas, J. (2012). Generating short-term forecasts of the Lithuanian GDP using factor models. *Working paper: Bank of Lithuania*, 16(1):49–67.

Stock, J. and Watson, M. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business and Economic Statistics*, 30(4):481–493. cited By 17.

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. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and banking*, 39(s1):3–33.

Stock, J. H. and Watson, M. W. (2010). Modeling Inflation After the Crisis. Working Paper 16488, National Bureau of Economic Research.

Stock, J. H. and Watson, M. W. (2011). Dynamic factor models. *Oxford Handbook of Economic Forecasting*, 1:35–59.

# Appendix A.

Table A.1: Correlation of the variables with the first factor. Benchmark dataset (N= 388)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|
| ppi_ind_finla_sa_no | PPI tradep: PPI Industry Finland | 0.920 | 0.00000 |
| tv_intermgood_sa_no | Turnover and Sales: Intermediate Goods | 0.881 | 0.00000 |
| tv_total_sa_no | Turnover and Sales: Total categories | 0.880 | 0.00000 |
| ppi_ind_lithu_sa_no | PPI tradep: PPI Industry Lithuania | 0.876 | 0.00000 |
| ppi_ind_eu15_sa_no | PPI tradep: PPI- Intermediate goods EU 15 | 0.860 | 0.00000 |
| op_intermgood_sa_no | Output: Intermediate Goods | 0.859 | 0.00000 |
| hicp_nrg_wa_sa_no | HICP Estonia: Energy Estonia | 0.856 | 0.00000 |
| op_total_sa_no | Output: Total categories | 0.852 | 0.00000 |
| ppi_ind_eu28_sa_no | PPI tradep: PPI Industry EU 28 | 0.847 | 0.00000 |
| price_import_sa_no | Import ex prices: Export price index | 0.846 | 0.00000 |
| spread_eur_e_sa_no | Interest margins 6-month Euribor NFC | -0.592 | 0.00004 |
| spread_eur_h_sa_no | Interest margins 6-month Euribor House loans | -0.568 | 0.00009 |
| xunempl_sa_no | Total unemployment | -0.499 | 0.00077 |
| neer_br_sa_no | Import ex prices: BIS, Nom. Broad Effective Exch. | -0.437 | 0.00384 |
| ca_total_sa_no | BOP: Current Account Total | -0.417 | 0.00602 |
| hicp_i_nd_se_sa_no | HICP partners: NEIG non-dur only Sweden | -0.411 | 0.00679 |
| hicp_i_nd_de_sa_no | HICP partners: NEIG non-dur only Germany | -0.405 | 0.00785 |
| empl_u_tkh_av_sa_no | Labor: The average unemployment insurance benefit | -0.397 | 0.00917 |
| ca_goods_net_sa_no | BOP: Estonia, Current Account, Goods, Net | -0.354 | 0.02151 |
| hicp_i_nd_wa_sa_no | HICP partners: NEIG non-dur only Euro area | -0.338 | 0.02842 |

Notes: p.value = 0.05

Table A.2: Correlation of the variables with the second factor. Benchmark dataset (N= 388)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|
| hicp_cp00_lt_sa_no | HICP partners: All-items HICP Lithuania | 0.883 | 0.00000 |
| hicp_serv_lt_sa_no | HICP partners: Services (ex goods) Lithuania | 0.856 | 0.00000 |
| hicp_x_nrg_f_sa_no | HICP partners: ex energy EU | 0.850 | 0.00000 |
| hicp_cp01_lt_sa_no | HICP partners: Food and non-alcoholic beverages Lithuania | 0.839 | 0.00000 |

| | | | |
|---|---|---|---|
| hicp_cp011_lt_sa_no | HICP partners: Food Lithuania | 0.837 | 0.00000 |
| hicp_gd_lt_sa_no | HICP partners: Goods (ex services) Lithuania | 0.829 | 0.00000 |
| hicp_cp01_wu_sa_no | HICP partners: Food and non-alcoholic beverages EU | 0.821 | 0.00000 |
| hicp_food_lt_sa_no | HICP partners: Food including alcohol and tobacco Lithuania | 0.816 | 0.00000 |
| hicp_cp00_lv_sa_no | HICP partners: All-items HICP Latvia | 0.816 | 0.00000 |
| hicp_food_wu_sa_no | HICP partners: Food including alcohol and tobacco EU | 0.814 | 0.00000 |
| cli_finl_a_sa_no | CLI: Finland, CLI, amplitude adjusted | -0.806 | 0.00000 |
| cli_finl_t_sa_no | CLI: Finland, CLI, tr | -0.756 | 0.00000 |
| cli_ez_amp_sa_no | CLI: EuroZone, CLI, amplitude adjusted | -0.707 | 0.00000 |
| cs_conf_sa_no | Surveys: CS, Confidence indicator | -0.706 | 0.00000 |
| sent_sa_no | Surveys: Economic Sentiment, Economic sentiment indicator | -0.706 | 0.00000 |
| cli_oecd_t_sa_no | CLI: OECD, CLI, amplitude adjusted | -0.704 | 0.00000 |
| cli_oecd_a_sa_no | CLI: OECD, CLI, amplitude adjusted | -0.697 | 0.00000 |
| ex_omx_sto_pr_sa_no | OMXS30 Index, Price Return,EUR | -0.687 | 0.00000 |
| cli_ger_t_sa_no | CLI: Germany, CLI, tr | -0.667 | 0.00000 |
| cli_ger_a_sa_no | CLI: Germany, CLI, amplitude adjusted | -0.667 | 0.00000 |

Notes: p.value = 0.05

Table A.3: Correlation of the variables with the third factor. Benchmark dataset (N= 388)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|
| hicp_igd_uk_sa_no | HICP partners: Industrial goods UK | 0.696 | 0.00000 |
| hicp_i_uk_sa_no | HICP partners: NEIG UK | 0.685 | 0.00000 |
| hicp_gd_uk_sa_no | HICP partners: Goods (ex services) UK | 0.641 | 0.00000 |
| hicp_i_d_uk_sa_no | HICP partners: NEIG dur only UK | 0.611 | 0.00002 |
| hicp_cp00_se_sa_no | HICP partners: All-items HICP Sweden | 0.598 | 0.00003 |
| hicp_gd_se_sa_no | HICP partners: Goods (ex services) Sweden | 0.591 | 0.00004 |
| hicp_i_nd_uk_sa_no | HICP partners: NEIG non-dur only UK | 0.553 | 0.00015 |
| hicp_cp00_fi_sa_no | HICP partners: All-items HICP Finland | 0.544 | 0.00020 |
| hicp_cp00_uk_sa_no | HICP partners: All-items HICP UK | 0.538 | 0.00024 |
| hicp_igd_se_sa_no | HICP partners: Industrial goods Sweden | 0.531 | 0.00030 |
| hicp_i_d_lv_sa_no | HICP partners: NEIG dur only Latvia | -0.677 | 0.00000 |
| st_it_usd_nfc_sa_no | Short-term interest (1 ¡ Year) rates USD NFC | -0.646 | 0.00000 |
| cred_st_ind_sa_no | Credit: Individuals | -0.619 | 0.00001 |
| ret_rt_food_sa_no | Retail Sales: Food, beverages and tobacco in non-specialized stores | -0.607 | 0.00002 |
| cred_st_lt_10_sa_no | Credit: Long-term | -0.583 | 0.00005 |
| cred_st_cu_sa_no | Credit: Cooperations | -0.582 | 0.00005 |

| | | | |
|---|---|---|---|
| xkgd_sa_no | Deflator: GDP total | -0.577 | 0.00006 |
| fin_tg_soc_pe_sa_no | State budget tax revenues, soc.security, pension | -0.570 | 0.00008 |
| fin_tg_soc_sa_no | State budget tax revenues, soc.security | -0.566 | 0.00010 |
| xkpr_sa_no | Deflator: Private | -0.560 | 0.00011 |

Notes: p.value = 0.05

Table A.4: Correlation of the variables with the first factor. Reduced dataset (N= 246)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|
| tv_intermgood_sa_no | Turnover and Sales: Intermediate Goods | 0.912 | 0.00000 |
| op_intermgood_sa_no | Output: Output: Intermediate Goowds | 0.907 | 0.00000 |
| tv_total_sa_no | Turnover and Sales: Total Goods | 0.904 | 0.00000 |
| op_total_sa_no | Output: Total Goods | 0.898 | 0.00000 |
| ft_total_sa_no | Foreign Trade: Total Commodities | 0.862 | 0.00000 |
| ppi_ind_finla_sa_no | PPI tradep: PPI Industry Finland | 0.857 | 0.00000 |
| empl_u_reg_ne_sa_no | Labor: Unemployment, Total Registered | 0.846 | 0.00000 |
| nord_intermgo_sa_no | New Orders: Production of intermediate consumption goods | 0.822 | 0.00000 |
| cci_us_sa_no | CLI: US, Coincident Index, Total | 0.820 | 0.00000 |
| empl_u_tkh_ne_sa_no | Labor: The new unemployment insurance benefit recipients | 0.820 | 0.00000 |
| spread_eur_e_sa_no | Interest margins 6-month Euribor NFC | -0.613 | 0.00002 |
| spread_eur_h_sa_no | Interest margins 6-month Euribor House loans | -0.570 | 0.00008 |
| xunempl_sa_no | Total unemployment | -0.524 | 0.00037 |
| cs_u_n12_sa_no | Surveys: CS, Unemployment exactions over 12 months | -0.466 | 0.00187 |
| empl_u_tkh_av_sa_no | Labor: The average unemployment insurance benefit | -0.465 | 0.00193 |
| neer_br_sa_no | Import ex prices: BIS, Avg. Nom. Broad Effective Exch. Rate | -0.460 | 0.00215 |
| ca_total_sa_no | BOP: Current Account Total | -0.432 | 0.00428 |
| cred_blnc_prc_sa_no | Credit: % of loan portfolio (balance) | -0.416 | 0.00615 |
| cred_cntr_prc_sa_no | Credit: % of loan portfolio (cntrct val.) | -0.383 | 0.01220 |
| ca_goods_net_sa_no | BOP: Estonia, Current Account, Goods, Net, Total | -0.357 | 0.02021 |

Notes: p.value = 0.05

Table A.5: Correlation of the variables with the second factor. Reduced dataset (N= 246)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|

| | | | |
|---|---|---|---|
| fin_tg_soc_pe_sa_no | State budget tax revenues, soc.security, pension | 0.689 | 0.00000 |
| fin_tg_soc_sa_no | Finance: State budget tax revenues, soc.security | 0.683 | 0.00000 |
| fin_tg_soc_me_sa_no | State budget tax revenues, soc.security, health | 0.669 | 0.00000 |
| xcgd_sa_no | NULC by hours: GDP total | 0.648 | 0.00000 |
| xdge_sa_no | Estonian deflators: General government consumption expenditure | 0.647 | 0.00000 |
| ret_rt_food_sa_no | Retail Sales: Food, beverages and tobacco in non-specialised stores | 0.644 | 0.00000 |
| empl_wages_sa_no | Labor: Monthly wages | 0.616 | 0.00001 |
| ppi_total_sa_no | PPI: Producer Prices, Total | 0.594 | 0.00003 |
| ppi_food_sa_no | PPI: Producer Prices, Food and beverages, Index | 0.590 | 0.00004 |
| st_it_eur_nfc_sa_no | Interest rates: Short-term interest rates NFC | 0.575 | 0.00007 |
| cli_finl_a_sa_no | CLI: Finland, CLI, amplitude adjusted | -0.831 | 0.00000 |
| cli_finl_t_sa_no | CLI: Finland, CLI, tr | -0.706 | 0.00000 |
| sent_sa_no | Surveys: Economic Sentiment, Economic sentiment indicator | -0.700 | 0.00000 |
| cli_ger_a_sa_no | CLI: Germany, CLI, amplitude adjusted | -0.666 | 0.00000 |
| cli_ger_t_sa_no | CLI: Germany, CLI, tr | -0.645 | 0.00000 |
| cli_oecd_a_sa_no | CLI: OECD, CLI, amplitude adjusted | -0.644 | 0.00000 |
| ex_omx_sto_pr_sa_no | External indic: OMXS30 Index, Price Return, EUR | -0.622 | 0.00001 |
| cli_ez_amp_sa_no | CLI: EuroZone, CLI, amplitude adjusted | -0.609 | 0.00002 |
| cli_oecd_t_sa_no | CLI: OECD, CLI, tr | -0.572 | 0.00008 |
| cs_conf_sa_no | Surveys: CS, Confidence indicator | -0.555 | 0.00014 |

Notes: p.value = 0.05

Table A.6: Correlation of the variables with the third factor. Reduced dataset (N= 246)

| variable code | variable full name | correlation | p.value |
|---|---|---|---|
| imf_pfood_sa_no | IMF: Food Price Index, 2005 = 100 | 0.599 | 0.00003 |
| imf_pfandb_sa_no | IMF: Food and Beverage Price Index, 2005 | 0.595 | 0.00003 |
| oilfutures_cs_sa_no | ECB commod: IMF IFS (U.K. Brent) | 0.537 | 0.00025 |
| ex_brentoil_i_sa_no | External indic: World, ICE, Crude Oil Index, USD | 0.533 | 0.00028 |
| imf_pallfnf_sa_no | IMF: All Commodity Price Index, 2005 = 1 | 0.531 | 0.00030 |
| ppi_food_eu28_sa_no | PPI Manuf of food EU 28 | 0.529 | 0.00032 |
| imf_poilapsp_sa_no | IMF: Crude Oil (petroleum), Price index | 0.528 | 0.00033 |
| imf_pnrg_sa_no | IMF: Fuel (Energy) Index, 2005 = 100 | 0.525 | 0.00035 |

| | | | |
|---|---|---|---|
| ex_crudeoil_i_sa_no | External indic: World, Energy, Oil, Brent, ICE, Average, USD | 0.517 | 0.00045 |
| ppi_food_eu15_sa_no | PPI Manuf of food EU | 0.516 | 0.00047 |
| cs_fin_l12_sa_no | Surveys: CS, Financial situation of households over l 12 months | -0.635 | 0.00001 |
| cs_ec_l12_sa_no | Surveys: CS, General economic situation over l 12 months | -0.559 | 0.00012 |
| st_it_usd_nfc_sa_no | Short-term interest rates (up to one year) USD NFC | -0.468 | 0.00176 |
| st_it_usd_hl_sa_no | Short-term interest rates (up to one year) USD households | -0.461 | 0.00212 |
| cs_fin_n12_sa_no | Surveys: CS, Financial situation of households over n 12 months | -0.411 | 0.00693 |
| xwse_sa_no | Nbr of hours worked by wage earners: Services | -0.382 | 0.01245 |
| xese_sa_no | Nbr of employed: Services | -0.378 | 0.01357 |
| fa_total_sa_no | BOP: Estonia, Financial Account, Balance, Total | -0.372 | 0.01522 |
| xose_sa_no | Nbr of hours worked in total economy: Services | -0.370 | 0.01573 |
| cs_s_n12_sa_no | Surveys: CS, Savings over n 12 months | -0.368 | 0.01646 |

Notes: p.value = 0.05

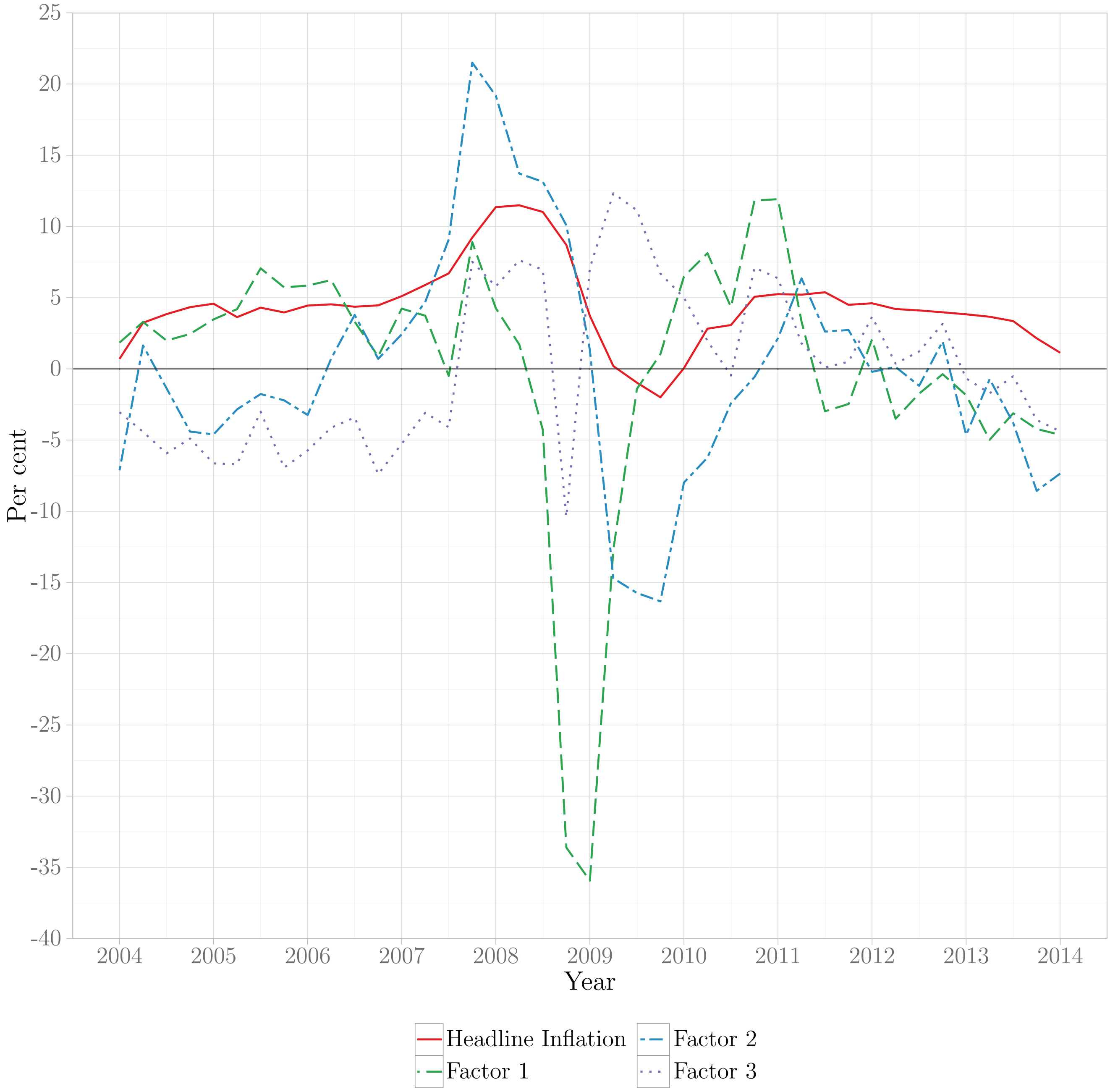


Figure A.1: Estimated common factors and Estonian headline inflation rate - Benchmark dataset

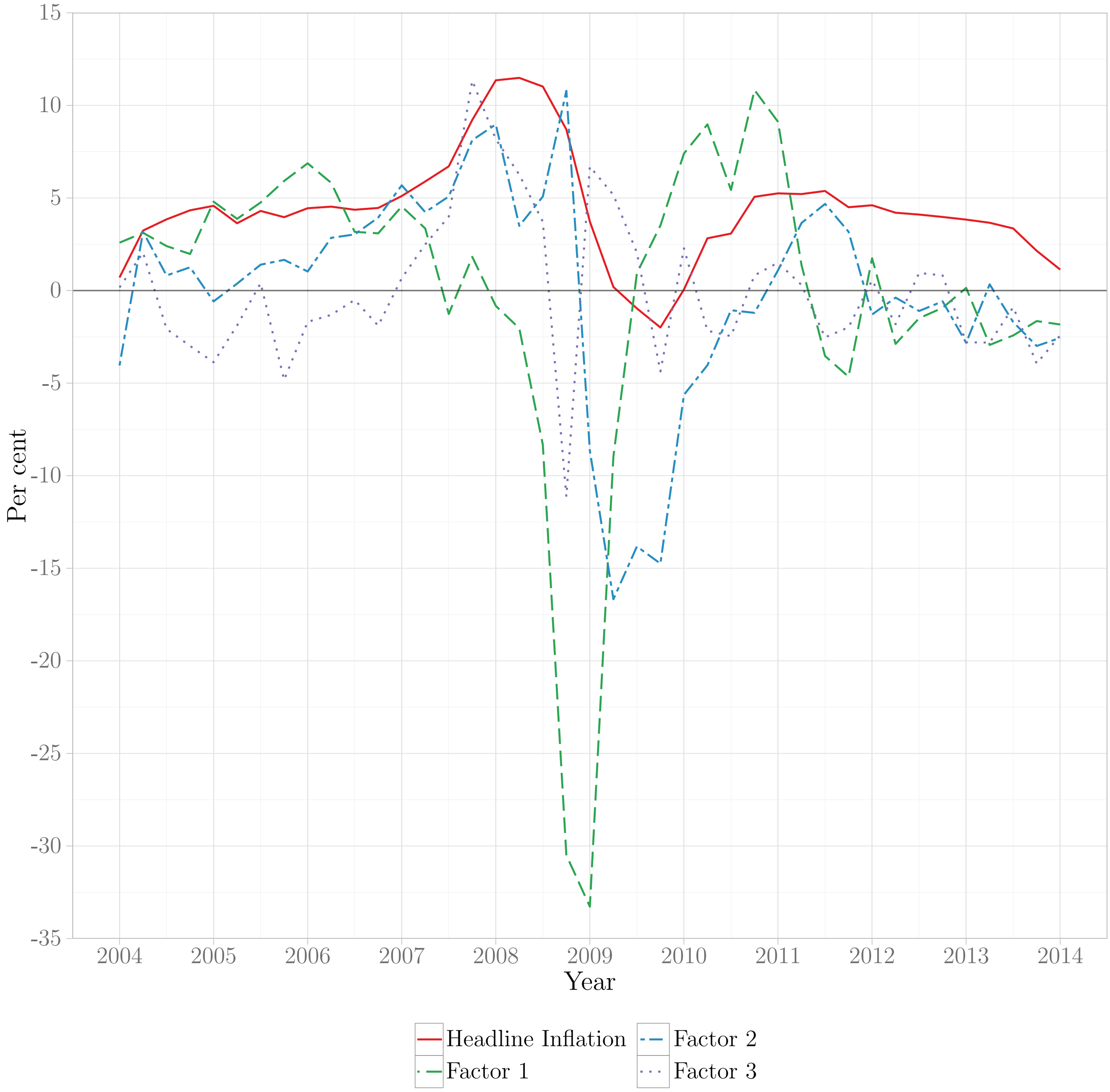


Figure A.2: Estimated common factors and Estonian core inflation rate - Reduced-size dataset

## Appendix B.

Table B.1: Headline Inflation Out-of-Sample Forecasting Results 1-6 quarters horizon obtained from the benchmark dataset

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 1.342 | 1.502 | 1.566 | 1.602 | 1.715 | 1.703 |
| RW | 1.174 | 1.024 | 0.958 | 0.968 | 1.042 | 1.044 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.984 | 0.991 | 1.000 | 1.059 | 1.114 | 1.082 |
| FAVAR 1F. 2 Lags | 0.939 | 0.953 | **0.930** | 0.986 | **1.019** | **1.011** |
| FAVAR 1F. 3 Lags | **0.936** | 1.006 | 1.003 | 1.065 | 1.082 | 1.099 |
| FAVAR 12F. 1 Lag | 1.051 | **0.951** | 0.987 | 1.050 | 1.128 | 1.079 |
| FAVAR 12F. 2 Lags | 1.076 | 0.971 | 1.000 | **0.973** | 1.091 | 1.046 |
| FAVAR 12F. 3 Lags | 1.183 | 1.008 | 1.017 | 1.002 | 1.109 | 1.046 |
| FAVAR 123F. 1 Lag | 1.056 | 0.952 | 0.980 | 1.029 | 1.094 | 1.051 |
| | | | | | | |
| absolute RMSE, AR model | 0.474 | 1.041 | 1.578 | 1.954 | 2.183 | 2.539 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points. Last observation has been removed from calculation of the RMSE.

Table B.2: Core Inflation Out-of-Sample Forecasting Results 1-6 quarters horizon obtained from the benchmark dataset

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 0.858 | 1.084 | 1.397 | 1.774 | 2.17 | 2.201 |
| RW | 1.438 | 1.678 | 1.819 | 2.117 | 2.324 | 2.176 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.889 | **0.923** | **1.021** | 1.145 | 1.275 | 1.322 |
| FAVAR 1F. 2 Lags | 0.938 | 1.033 | 1.156 | 1.291 | 1.404 | 1.485 |
| FAVAR 1F. 3 Lags | 0.931 | 1.124 | 1.405 | 1.643 | 1.788 | 1.849 |
| FAVAR 12F. 1 Lag | 0.898 | 0.979 | 1.102 | 1.219 | 1.289 | 1.362 |
| FAVAR 12F. 2 Lags | 1.006 | 1.223 | 1.459 | 1.71 | 1.94 | 2.069 |
| FAVAR 12F. 3 Lags | 1.001 | 1.195 | 1.363 | 1.627 | 1.844 | 1.932 |
| FAVAR 123F. 1 Lag | **0.871** | 0.94 | 1.027 | **1.112** | **1.134** | **1.214** |
| | | | | | | |
| absolute RMSE, AR model | 0.342 | 0.52 | 0.668 | 0.748 | 0.850 | 1.082 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points. Last observation has been removed from calculation of the RMSE.

Table B.3: Headline Inflation Out-of-Sample Forecasting Results 1-6 quarters horizon obtained from the reduced dataset

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 1.342 | 1.502 | 1.566 | 1.602 | 1.715 | 1.703 |
| RW | 1.174 | 1.024 | 0.958 | 0.968 | 1.042 | 1.044 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 1.018 | 1.033 | 1.042 | 1.076 | 1.087 | 1.059 |
| FAVAR 1F. 2 Lags | 0.991 | 1.008 | 0.991 | 1.042 | 1.025 | 1.036 |
| FAVAR 1F. 3 Lags | **0.985** | 1.067 | 1.051 | 1.115 | 1.076 | 1.09 |
| FAVAR 12F. 1 Lag | 1.025 | 1.033 | 1.045 | 1.093 | 1.093 | 1.071 |
| FAVAR 12F. 2 Lags | 1.003 | 1.014 | 0.976 | 1.040 | 1.037 | 1.077 |
| FAVAR 12F. 3 Lags | 1.061 | 1.097 | 1.131 | 1.196 | 1.207 | 1.29 |
| FAVAR 123F. 1 Lag | 1.056 | **0.951** | **0.908** | **0.928** | **1.023** | **0.916** |
| | | | | | | |
| absolute RMSE, AR model | 0.474 | 1.041 | 1.578 | 1.954 | 2.183 | 2.539 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points. Last observation has been removed from calculation of the RMSE.

Table B.4: Core Inflation Out-of-Sample Forecasting Results 1-6 quarters horizon obtained from the reduced dataset

| | 1 qt. | 2 qt. | 3 qt. | 4 qt. | 5 qt. | 6 qt. |
|---|---|---|---|---|---|---|
| Benchmark AR | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| VAR | 0.858 | 1.084 | 1.397 | 1.774 | 2.170 | 2.201 |
| RW | 1.438 | 1.678 | 1.819 | 2.117 | 2.324 | 2.176 |
| | | | | | | |
| FAVAR 1F. 1 Lag | 0.900 | 0.939 | 1.015 | 1.083 | 1.168 | 1.224 |
| FAVAR 1F. 2 Lags | 0.952 | 1.067 | 1.191 | 1.310 | 1.405 | 1.504 |
| FAVAR 1F. 3 Lags | 0.893 | 1.101 | 1.382 | 1.616 | 1.742 | 1.821 |
| FAVAR 12F. 1 Lag | 0.868 | 0.920 | 1.011 | 1.085 | 1.151 | 1.220 |
| FAVAR 12F. 2 Lags | 0.943 | 1.061 | 1.199 | 1.331 | 1.416 | 1.502 |
| FAVAR 12F. 3 Lags | 1.081 | 1.261 | 1.644 | 1.902 | 2.040 | 2.081 |
| FAVAR 123F. 1 Lag | **0.804** | **0.755** | **0.832** | **0.896** | **0.971** | **0.979** |
| | | | | | | |
| absolute RMSE, AR model | 0.342 | 0.520 | 0.668 | 0.748 | 0.850 | 1.082 |

Notes: Absolute root mean square errors (abs. RMSE) in percentage points. Last observation has been removed from calculation of the RMSE.

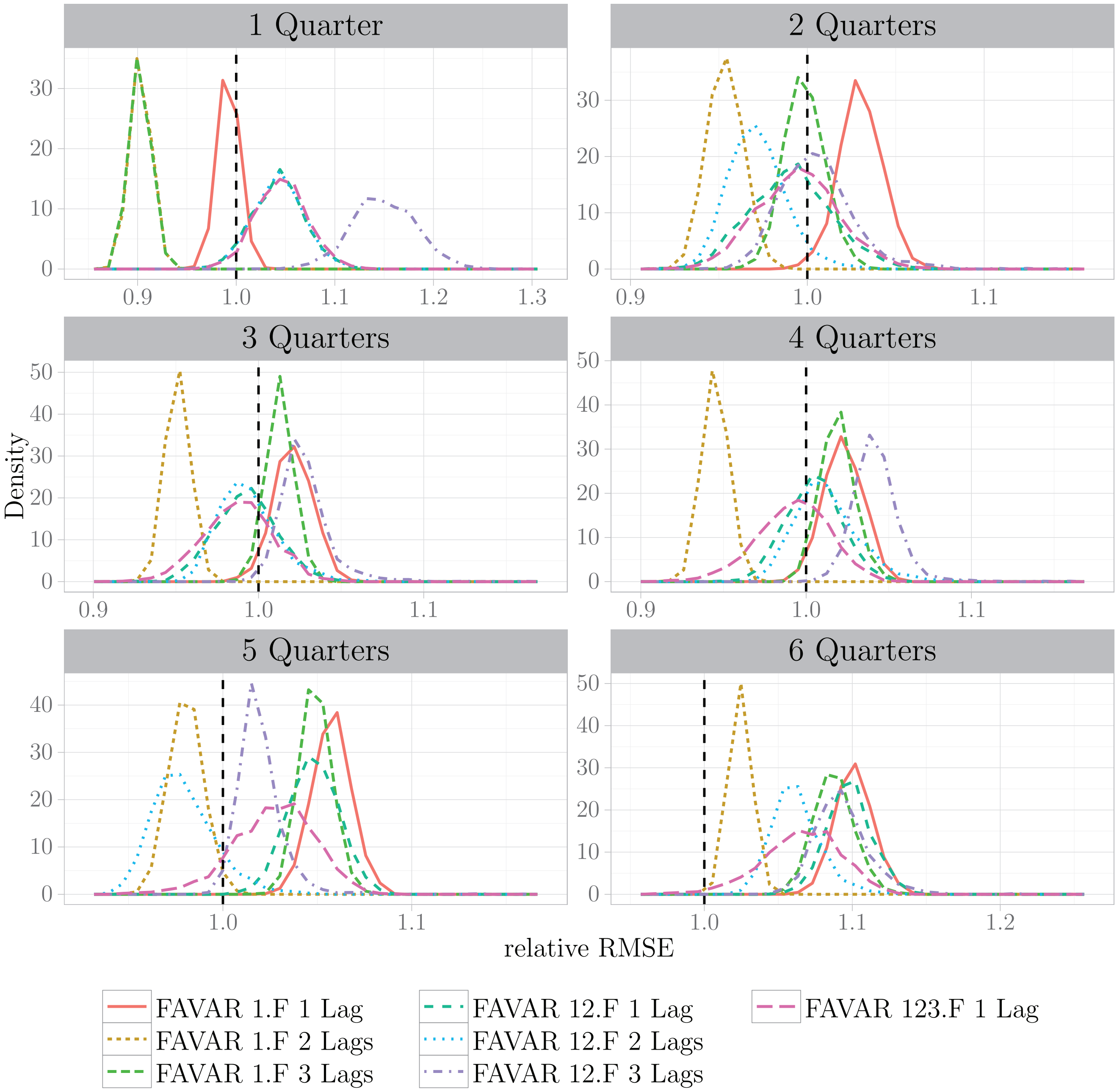


Figure B.1: Frequency distribution Estonian headline inflation - Benchmark dataset
Notes: Bins = 30. Density of points in bin, scaled to integrate to 1.

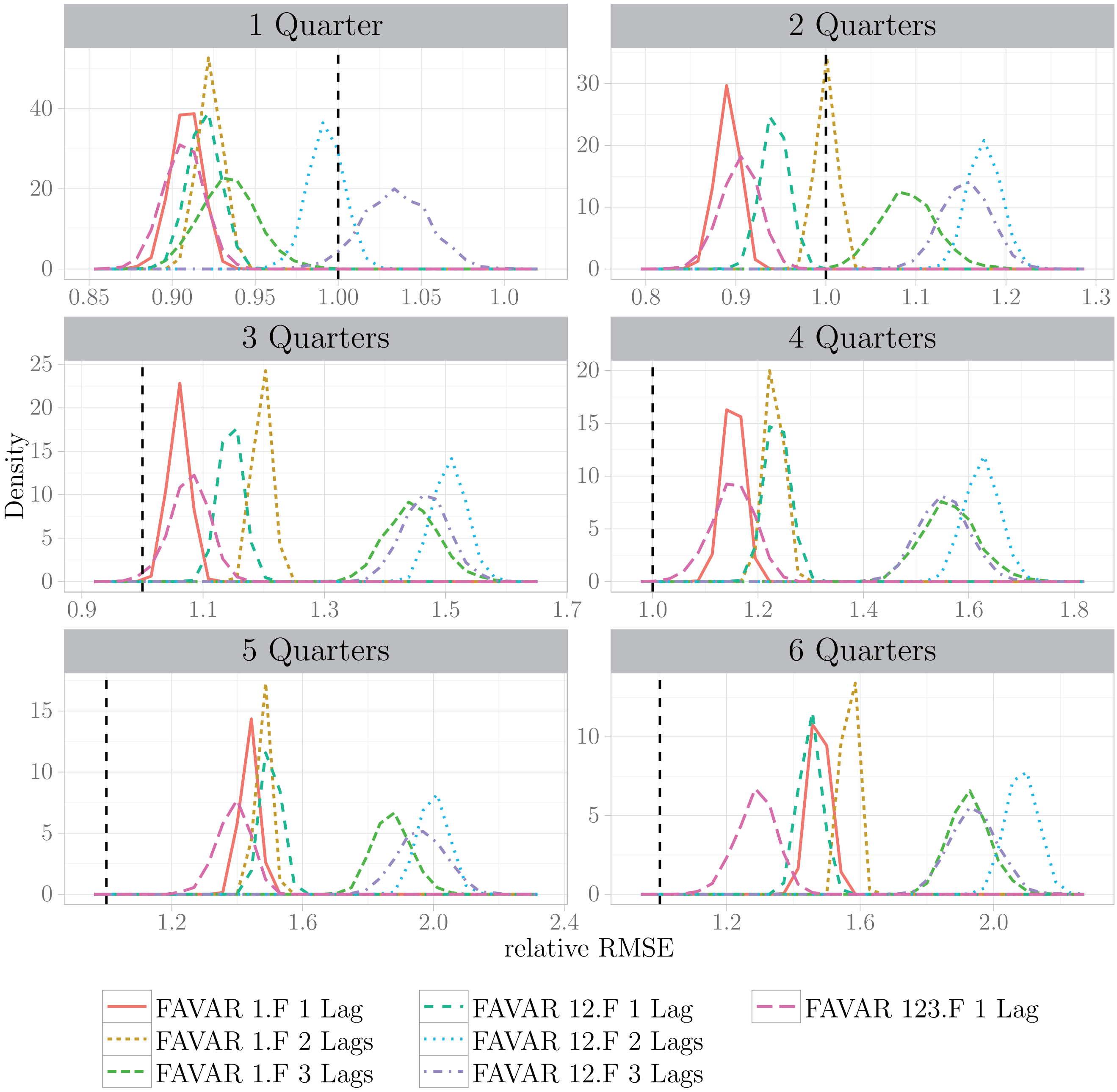


Figure B.2: Frequency distribution Estonian core inflation - Benchmark dataset
Notes: Bins = 30. Density of points in bin, scaled to integrate to 1.

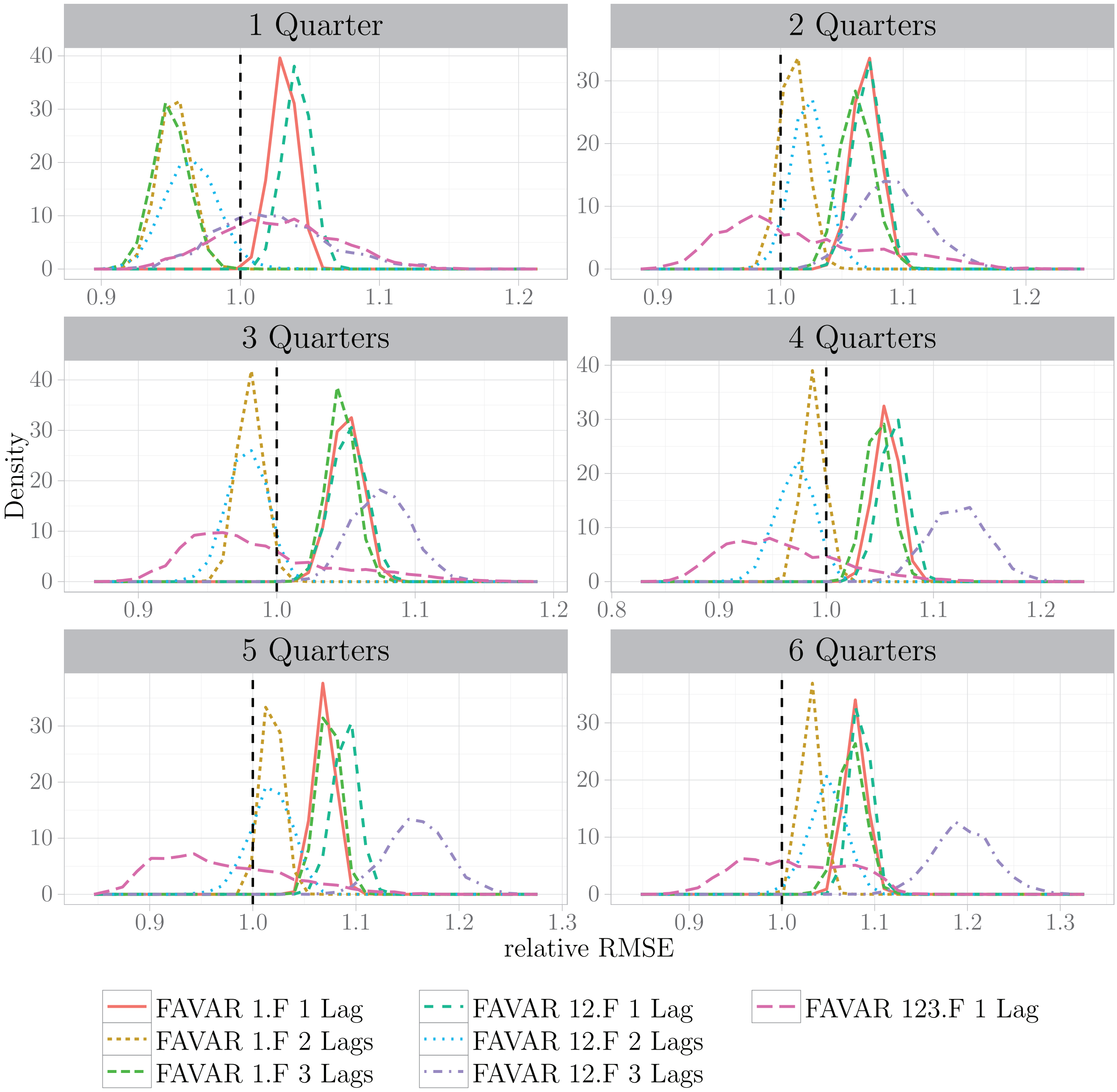


Figure B.3: Frequency distribution Estonian headline inflation - Reduced-size dataset
Notes: Bins = 30. Density of points in bin, scaled to integrate to 1.

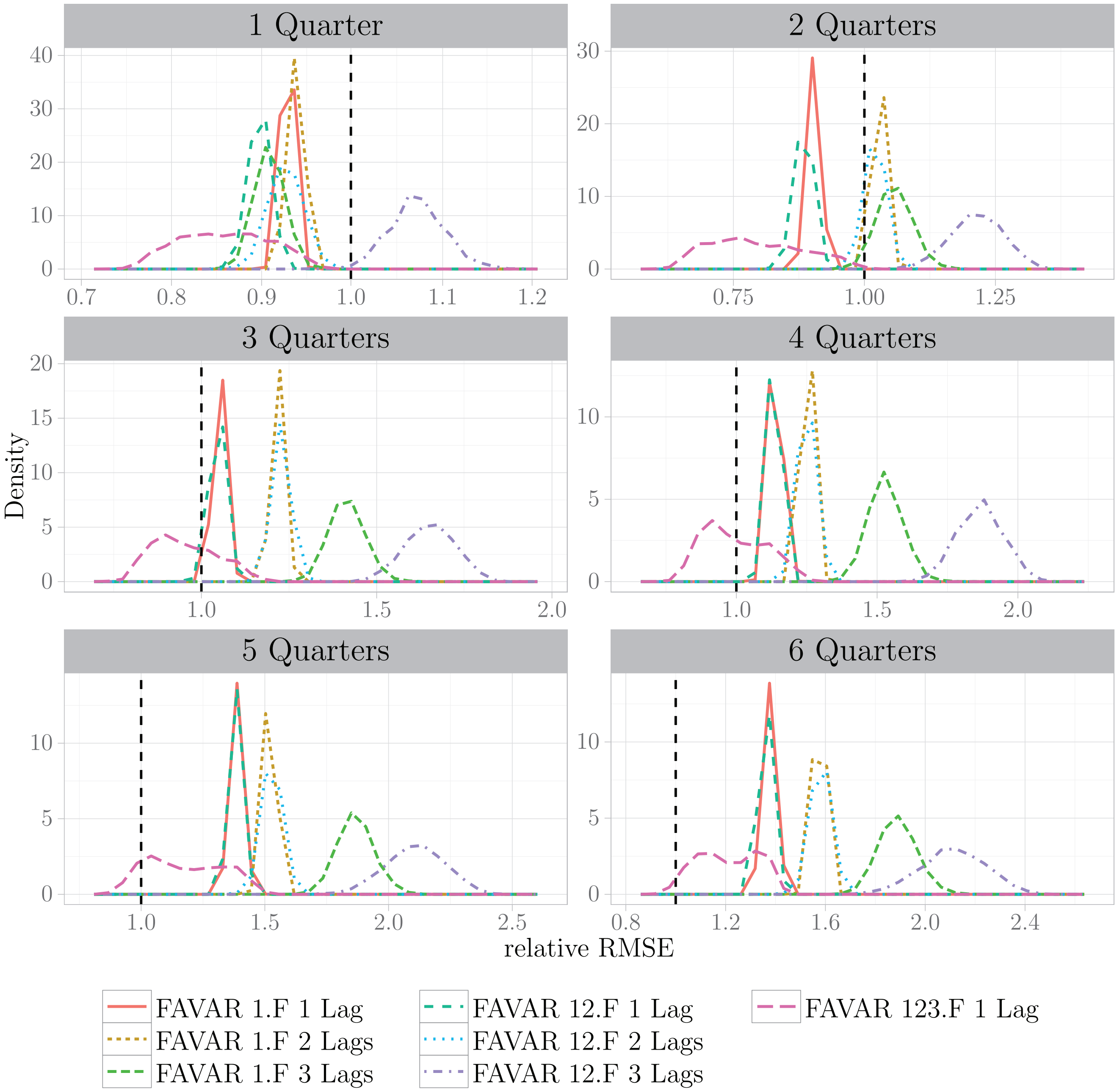


Figure B.4: Frequency distribution Estonian core inflation - Reduced-size dataset
Notes: Bins = 30. Density of points in bin, scaled to integrate to 1.

## Working Papers of Eesti Pank 2016

No 1
Merike Kukk. Debt repayment problems: what are the implications for consumption?

No 2
Alessandra Cepparulo, Juan Carlos Cuestas, Maurizio Intartaglia. Financial development, institutions and poverty alleviation: an empirical analysis

No 3
Urška Čede, Bogdan Chiriacescu, Péter Harasztosi, Tibor Lalinsky, Jaanika Meriküll. Export characteristics and output volatility: comparative firm-level evidence for CEE countries

No 4
Punnoose Jacob, Lenno Uusküla. Deep habits and exchange rate pass-through

No 5
Svetlana Makarova. ECB footprints on inflation forecast uncertainty

No 6
Simona Ferraro, Jaanika Meriküll, Karsten Staehr. Minimum wages and the wage distribution In Estonia

No 7
Lenno Uusküla. Monetary transmission mechanism with firm turnover