ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Gunay, Mahmut

Article

Forecasting industrial production and inflation in Turkey with factor models

Central Bank Review (CBR)

Provided in Cooperation with:
Central Bank of The Republic of Turkey, Ankara

Suggested Citation: Gunay, Mahmut (2018) : Forecasting industrial production and inflation in Turkey with factor models, Central Bank Review (CBR), ISSN 1303-0701, Elsevier, Amsterdam, Vol. 18, Iss. 4, pp. 149-161,
https://doi.org/10.1016/j.cbrev.2018.11.003

This Version is available at:
https://hdl.handle.net/10419/217326

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-nc-nd/4.0/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Central Bank Review 18 (2018) 149–161

Contents lists available at ScienceDirect

Central Bank Review

journal homepage: http://www.journals.elsevier.com/central-bank-review/

Forecasting industrial production and inflation in Turkey with factor models☆

Mahmut Günay*

Central Bank of the Republic of Turkey, Turkey

ARTICLE INFO

Article history:
Received 20 November 2018
Received in revised form
22 November 2018
Accepted 22 November 2018
Available online 4 December 2018

JEL classification:
E37
C53
C55

Keywords:
Forecasting
Factor models
Principal component

ABSTRACT

In this paper, industrial production growth and core inflation are forecasted using a large number of domestic and international indicators. Two methods are employed, factor models and forecast combination, to deal with the curse of dimensionality problem stemming from the availability of ever growing data sets. A comprehensive analysis is carried out to understand the sensitivity of the forecast performance of factor models to various modelling choices. In this respect, effects of factor extraction method, number of factors, data aggregation level and forecast equation type on the forecasting performance are analyzed. Moreover, the effect of using certain data blocks such as interest rates on the forecasting performance is evaluated as well. Out-of-sample forecasting exercise is conducted for two consecutive periods to assess the stability of the forecasting performance. Factor models perform better than the combination of bi-variate forecasts which indicates that pooling information improves over pooling individual forecasts.

© 2018 Central Bank of The Republic of Turkey. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Rapid expansion in the availability of data increases the challenge of using information efficiently. There are a lot of candidate indicators that can be used in the forecasting, and this number is increasing with the advances in the information technology. However, one can use only a limited number of variables in an OLS- or VAR-type forecasting model due to the degrees of freedom problem. Stock and Watson (2002a) state that some variable selection procedures may be used for determining a parsimonious forecasting model, but the performance rests on the few variables chosen. Hence, forecasters need techniques that enable them to use large amounts of data in the forecasting process.

Factor models have become popular in economics since the late 1990s after the seminal contributions such as Stock and Watson (2002a) and Forni et al. (2000, 2005). In this approach, information in a large data set is summarized with a few underlying factors and then these factors are used in the forecasting equation (Stock and Watson, 2002b). Those authors showed that one can use principal component type analysis to estimate the factors. This opened the door for handling large amount of data relatively easily. Researchers contributed by developing theory for consistent estimation of the factors for data sets with large N, number of indicators, and large T, number of time series observations, (such as Bai (2003)). Also, techniques for formally determining number of static and dynamic factors have been developed.

Use of factor models for forecasting Turkish macroeconomic variables is limited. Öğünç et al. (2013) conduct a comprehensive analysis for evaluating performance of various modelling techniques for forecasting inflation. In addition to a factor model, they consider univariate models, time varying Phillips curve models, decomposition based models, VAR and Bayesian VAR models. For the factor models, authors use a single equation model with factors and a VAR model augmented with the factors.

Their results show that factor models perform poorly. In the case of single equation model, factor model beats the benchmark random walk only for the two-quarter-ahead forecasts. FAVAR (Factor Augmented Vector Auto Regression) type models beat the random walk for one to three-quarter-ahead horizon. While FAVAR is relatively successful for two-quarter-ahead forecasts, for one and

☆ The views expressed here are of the author only and do not necessarily reflect the institutions that the author is affiliated with. This paper is based on the PhD dissertation of the author entitled "Forecasting economic and financial variables with factor models: the case of Turkey".

* TCMB İdare Merkezi, İstiklal Cad. No:10 16. Kat, 06050, Ulus, Ankara, Turkey.
E-mail address: mahmut.gunay@tcmb.gov.tr.

Peer review under responsibility of the Central Bank of the Republic of Turkey.

https://doi.org/10.1016/j.cbrev.2018.11.003

1303-0701/© 2018 Central Bank of The Republic of Turkey. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

three-quarter-ahead forecasts there are more successful models. Authors claim that world is not static and hence their dynamic factor model approach is expected to capture the workings of the economy more successfully. However, it is not clear *ex-ante* whether dynamic approach is preferable for the relatively short sample used by the authors. Soybilgen (2015) analyzes the performance of factor models for GDP, inflation and unemployment rate. For inflation, he states that small scale dynamic factor models outperform larger factor models. He also finds that rankings are not stable.

There are scant examples of the use of the factor models for real sector variables. In Akkoyun and Günay (2012), authors use a dynamic factor model for nowcasting Turkish GDP growth. Since GDP data for a quarter are published with a certain lag, nowcasting GDP growth is an essential ingredient of real time policy making. By using survey data, authors improve over the benchmark and obtain relatively successful nowcasts. Modugno et al. (2016) also use factor models for nowcasting Turkish GDP Growth. They find that financial variables can be as important as survey variables for accurate short term forecasts.

In this paper, we aim to contribute to the literature by analyzing the sensitivity of the forecasting performance of factor models for industrial production and core inflation for Turkish economy by analyzing different dimensions of factor models.[1]

Factor models enable one to incorporate as many series as he/she wants in the forecasting process. However factors are latent variables which are need to be estimated before using in the forecasting models. Therefore, depending on the factor estimation step, one may get different factors which in turn may affect forecast performance. From another angle, there is not a standard to follow for constructing the data set. Indeed, it is not clear whether there is a linear relation between forecasting performance and the number of series used for extracting factors. If this is so, one needs to target a large data set. On the contrary, if more data is not always better for factor forecasting, then one needs to work on the composition of the data set.

In addition to the composition and size of the data set, number of factors extracted from a given data set may affect outcomes as well. Using a few factors may be insufficient to summarize the information content of the data set while using a lot of factors may increase the parameter uncertainty in the estimation of forecast equation. These points imply that analyzing the effect of modelling decisions on the forecast performance of factor models may provide valuable information to the forecasters. In this respect, several authors analyze how certain modelling decisions affect the forecast performance of factor models by considering different dimensions of the modelling. In the paper, we take a relatively more comprehensive approach for understanding the effect of modelling choices.

Bookkeeping for all of the specifications that are used in this paper reveals the scope of the analysis. 84 alternatives (7 criteria for the number of factors x 2 factor extraction approach x 2 forecast equation type x 3 data sets) for the factor models are evaluated. In addition to the factor models, forecast combination for bivariate equations is considered as well. For the three data sets used in this paper, there are 3 alternatives for the pool of bivariate equations.[2] In order to understand the effect of certain data groups, these are removed from the data set and forecasts are obtained for the filtered data set. For these factor models that exclude data blocks one at a time, there are 3x84 = 252 alternatives. In total, 84+3+252 = 339 alternatives exist. Considering the simple benchmark, there are a total of 340 specifications.

In the literature, performance of forecast models are compared using a loss function, and in general a quadratic loss function is used. In the paper, models are compared using the Root Mean Squared Error (RMSE) as a metric. RMSE penalizes large errors more heavily. RMSEs for all of the 340 specifications described above are calculated for 3-, 6-, 9-, and 12-month-ahead horizons. Tabulations of relative RMSEs are done for two sub-periods. Namely, episode 1 is for January 2010–September 2011 and episode 2 is for October 2011–September 2013. These dates are selected based on the data availability and economic developments in the period. Our data set ends in September 2014. So, for 12-month-ahead forecasts, the last forecast that we compare with the realization is the forecast done in September 2013. We start the evaluation period after 2009, the year we observed significant effects of global financial crises on domestic variables. Splitting sample in September 2011 helps us to do the analysis with approximately same number of observations in two evaluation sample.

In summary our results point out that:

i. Best performing specifications change for the industrial production and core inflation forecasts. This suggests that careful examination of the sensitivity of the forecast performance for each target variable may be beneficial rather than following a one size fits all approach.
ii. Regarding factor extraction approach, findings show that using a simple factor extraction approach may be preferred in practical applications since compared to a more complicated method, difference in the forecast performance is in general marginal.
iii. Unlike factor extraction approach, number of factors used in the forecast equation affects the forecast performance considerably. While using a large number of factor summarizes a higher portion of the variance of the data set, this does not linearly translate to improved forecast performance. In general, using a few factors and hence working with a relatively more parsimonious forecast equation helps to obtain better forecast performance.
iv. Forecast equation type affects the conclusion about the effect of the number of factors and the effect of the size and the composition of the data set. Findings indicate that in the case of using a forecast equation with the lags of factors, using a few factors is preferable. It is also observed that for core inflation, using lags of the factors and the dependent variable seems to pay off while for industrial production picture is less clear.
v. Composition of the data set plays a crucial role on the forecast performance and more data is not always better. A small data set composed of aggregated variables produce the best performing specifications for a lot of cases. However, depending on the forecast horizon and the target variable, using disaggregated data by excluding certain blocks from the data set brings considerable improvement. So special care is needed for constructing the appropriate data set.

The paper is organized as follows. In the next section, we briefly summarize the methodology where we also discuss the papers in the literature analyzing the effect of these variables on the forecasting performance and we introduce data sets. In the third section, we present the results. Last section concludes.

[1] Gunay (2015) focuses on the static factor models for industrial production.

[2] Bivariate equations are estimated and used in the forecasting from each of the 22 series from the small data set, 63 series from the medium and 167 from the large data set. Then, average of the forecasts is calculated from the forecasts of these individual equations. This exercise enables one to compare the effect of pooling forecasts as opposed to factor models where one pools information. If bivariate equations perform relatively well, they will appear in tables A1 to A8 where we present results.

2. Methodology

2.1. Factor models

Let N be the number of cross-section units, T be the number of time series observations and x_{it} be the observed time series for variable i at time t. For $i = 1, \ldots, N$, $t = 1, \ldots, T$, a static factor model is defined as

$$x_{it} = \lambda_i' F_t + e_{it} \tag{1}$$

$$= C_{it} + e_{it}. \tag{2}$$

In the jargon of factor models, F_t are the (r x 1) static factors, e_{it} is named as the idiosyncratic error and λ_i is a vector of (r x 1). Elements of λ_i are named as the factor loadings. $C_{it} = \lambda_i' F_t$ is referred as the common component, i.e. for the variable x_i the part of the movement that is explained by the factors. As the intuition suggests, a few factors cannot explain all of the movement in a series. Part of the variation of the series that cannot be explained by the common component is shown with e_{it} and it is named as the idiosyncratic term. The challenge comes from the fact that variables on the right hand side of Equation (1) are not observed. Hence it is necessary to estimate factors and factor loadings.

In classical factor analysis, F_t and e_t are generally assumed to be serially and cross-sectionally uncorrelated. Bai and Ng (2008) observe that this assumption is fairly restrictive for economic data. This is due to the fact that economic time series data are in general serially correlated. Moreover, even if one can explain some part of the data with a common component and name the remaining portion as the "idiosyncratic errors", there may still be some cross-correlation in the errors. This may occur for instance when subcategories of a data block (such as durable good production and nondurable good production) are included simultaneously in the data set.

Due to this limitation of the classical factor models and before the seminal contributions by Stock and Watson (2002a) (SW hereafter) and Forni et al. (2005)'s (hereafter FHLR), the use of factor models in economic applications was fairly limited. Those authors show that principal component type analysis can be applied for estimating factors. These principal components in turn can be used as a regressor in a forecast equation.

In SW methodology, factors can be estimated as

$$\widehat{F}_t^{SW} = \widehat{S}' X_t \tag{3}$$

where $\widehat{S}_j$ corresponds to the r largest eigenvectors of the variance-covariance matrix of the data. So, factors are simply the eigenvectors times the data matrix.[3] FHLR method ends up solving a generalized eigenvalue problem in the following formula.

$$\widehat{\Gamma}_\chi(0)\widehat{Z}_j = \widehat{\mu}_j \widehat{\Gamma}_\xi(0) \tag{4}$$

where $\widehat{Z} = (\widehat{Z}_1, \ldots, \widehat{Z}_r)$ denotes the eigenvectors, $\widehat{\Gamma}_\chi(k) = \frac{1}{2H+1}\Sigma_{h=0}^{2H}\widehat{\Sigma}_\chi(\theta_h)e^{ik\theta_h}$ *for k=-M, ...,M* is the inverse discrete Fourier Transform that provides time-domain autocovariances of the common components, $\widehat{\Sigma}_\chi(\theta_h)$ is the spectral density of the common component, $\theta_h = \frac{2\pi h}{2H}$ for h = 0, ..., 2H and $\widehat{\Gamma}_\xi(k)$ are the autocovariances of the idiosyncratic terms (see Schumacher (2007) page 274 for more detail). Here, H refers to the number of frequency grids and M refers to the Bartlett lag window. So, H and M appear as additional parameters compared with SW approach where we only need to determine number of static factors. After obtaining the eigenvectors, factors can be calculated from the following formula.

$$\widehat{F}_t^{FHLR} = \widehat{Z}' X_t \tag{5}$$

There are a number of papers that compare performance of these two approaches in forecasting. Boivin and Ng (2005) show that while SW approach is relatively simpler to apply compared to FHLR, it gives competitive forecasts. Since it is easier to implement, in day to day use it may be preferred over dynamic approach. Schumacher (2007) compares alternative factor models for forecasting German GDP, namely SW, FHLR and a third approach proposed by Kapetanios and Marcellino (2009). He finds that effect of factor extraction method on the forecast performance depends on the modelling choices such as appropriately choosing auxiliary parameters, number of factors and forecast equation. D'Agostino and Giannone (2012) compare the forecasting performance of SW and FHLR approaches for the US economy for industrial production and inflation. What emerges from the paper is that one can see collinear forecasts from two factor approaches but there may be time-varying relative performance. Another important finding is that number of factors plays an important role on the relative forecast performance.

2.2. Determining the number of factors

Bai and Ng (2002) develop theory to determine the number of static factors in a formal and systematic way. They note that penalty for overfitting must be a function of both N and T in order to consistently estimate the number of factors (page 192). So, using classical form of the information criteria such as Akaike Information Criteria (AIC) or Bayesian Information Criteria (BIC) would not be appropriate for large panel of data.[4]

Bai and Ng (2002) aim to find the penalty function $g(N,T)$ such that the following froms of information criterion can consistently estimate the number of factors in the data set:

$$PC(k) = V\left(k, \widehat{F}^k\right) + kg(N,T) \tag{6}$$

$$IC(k) = \ln\left(V\left(k, \widehat{F}^k\right)\right) + kg(N,T) \tag{7}$$

Here, k is the number of factors used, V is the sum of squared residuals when one uses k factors, N is the number of variables (i.e. cross section dimension) and T is the number of time series observations (i.e. time dimension) and g is a function. Bai and Ng (2002) derive several forms from these general formulas.[5] Bai and Ng (2002) test their criteria with both simulated and actual data.

[3] Of course, we only get the space spanned by the factors.

[4] AIC and BIC are two measures that are used to compare models in model selection process. Broadly, they can be expressed as AIC = -2*ln(likelihood of the model)+2*k and BIC = -2*ln(likelihood of the model)+ln(N)*k. Here, k is the number of parameters estimated and N is the number of observations. In general, higher the number of variables used in a model the better the fit is. So, an increase in k will increase likelihood. However, higher k means a more complex model. Hence, depending on the marginal contribution of the added variable, AIC and BIC may increase or decrease. Models with lower score in the information criterion are preferred over models with larger scores.

[5] For example, from Equation (6) they define $PC_{p1}(k) = V(k, \widehat{F}^k) + k\widehat{\sigma}^2\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right)$ and from Equation (7) they define a criterion using log of the V: $IC_{p1}(k) = \ln(V(k, \widehat{F}^k)) + k\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right)$. Here $\widehat{\sigma}^2$ is a consistent estimate of $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}E(e_{it})^2$.

They find that PC_{p1}, PC_{p2}, IC_{p1} *and* IC_{p2} perform relatively well. It is worth emphasizing that they find that BIC_3 has very good properties in the presence of cross-section correlations (page 207 of Bai and Ng (2002)).[6] Thus, they conclude that this criterion can be helpful even though it does not satisfy all of the conditions of the theorem in the paper.

Bai and Ng (2007) work on determining number of dynamic factors. They start by considering a vector of observed stationary time series, F_t $(rx1)$ which follows the following VAR:

$$A(L)F_t = u_t \tag{8}$$

where A(L) are the lag polynomials of order p. If there exists an *rxq* matrix R with rank q such that

$$u_t = R\varepsilon_t \tag{9}$$

then Bai and Ng (2007) say that F_t is driven by a minimal number of q innovations. Here, ε_t is $(qx1)$ vector of mutually uncorrelated innovations (so variance-covariance matrix of the innovations is diagonal). From this logic, they come up with two criteria. So, compared to the SW approach, in FHLR approach number of dynamic factors is another input that we need to supply to the system.

Schumacher (2007) and D'Agostino and Giannone (2012) find that number of factors plays a non-negligible role on the forecast performance. In the case of Schumacher (2007), number of static factors is decided with IC1 and IC2. Comparing SW and FHLR approaches for two-quarter-ahead forecasts, FHLR approach performs slightly better than SW with IC2 information criterion while for IC1 reverse is true. In the case of D'Agostino and Giannone (2012), for forecasting industrial production, increasing the number of static factors from 1 to 3 decreases the relative RMSE by 41 percent (for p = 0). These observations show the importance of the effect of the number of factors used on the relative forecast performance. Barhoumi et al. (2013)'s paper entitled as "*testing the number of factors: an empirical assessment for a forecasting purpose*" analyze this issue thoroughly. They find that number of factors may have an important role on the forecasting performance. They conclude that rather than determining factors in an ad hoc manner, selecting the number of factors with an information criterion may produce better forecast results. In this paper, criteria developed by Bai and Ng (2002) for the number of static factors and Bai and Ng (2007) for the number of dynamic factors are used. These are frequently used in the forecasting applications and as suggested by Barhoumi et al. (2013) they are robust.

2.3. Forecast equation

When one wants to forecast for more than one-period-ahead, he/she needs to do multi-period ahead forecasting. For this task, there are two approaches: direct and iterated forecasting. In the case of iterated forecasting, one estimates a one step-ahead model and uses this model h times to get h-period-ahead forecasts. In the case of direct forecasting, one estimates a different model for each horizon h. In this paper, the direct approach for multi-step-ahead forecasting is used since it is a common method for many papers in forecasting such as Stock and Watson (2002a and 2002b) and Schumacher (2007), among others. Mechanics of this approach is shown following the presentation in Stock and Watson (2002b).

Stock and Watson (2002b) focus on the multi-step-ahead prediction. For industrial production they consider the following forecast equation (Page 149 of Stock and Watson (2002b)):

$$\widehat{y}^h_{T+h/T} = \widehat{\alpha}_h + \sum_{j=1}^{m} \widehat{\beta}'_{hj}\widehat{F}_{T-j+1} + \sum_{j=1}^{p} \widehat{\gamma}_{hj} y_{T-j} \tag{10}$$

where

$$y^h_{t+h} = \left(\frac{1200}{h}\right)\ln\left(\frac{IP_{t+h}}{IP_t}\right) \tag{11}$$

$$y_t = 1200\ln\left(\frac{IP_t}{IP_{t-1}}\right) \tag{12}$$

and $\widehat{F}_t$ is the vector of k estimated factors. Note that (1200/h) implies that annualized version of the h-period change is used. So, when h = 12, y^{12}_{t+12} is the annual growth rate of the industrial production. y_t is the annualized month-on-month growth rate of the industrial production.

Stock and Watson (2002b) use three versions of the above equation.

i. DI-AR Lag (Diffusion Index- Auto Regressive Lag): this version includes lags of the factors and lags of y with m and p estimated by the Bayesian Information Criterion. Authors use $1 \le m \le 3$ *and* $0 \le p \le 6$.
ii. DI-AR (Diffusion Index- Auto Regressive): This form of the forecasting equation uses the contemporaneous values of the factors (so m = 1) while picks the lag length of the y. For this case, authors use $0 \le p \le 6$.
iii. DI (Diffusion Index): This type of the forecast equation uses only the contemporaneous values of the factors so that m = 1 and p = 0.

In addition to the lag lengths of the factors and the y, number of factors that will be used in the forecasting equation should be decided as well. Stock and Watson (2002b) do not use a criterion in the spirit of Bai and Ng (2002). In this paper, following Stock and Watson (2002b) we use DI and DI-AR Lag type forecast equations, however unlike Stock and Watson (2002b) we choose the number of factors following Bai and Ng (2002).

2.4. Data set

A critical issue that a forecaster needs to address before setting up any forecasting model is to decide the structure of the data set that will be used. Data set structure can be even more challenging in the case of factor models since one can use as many series as he/she can collect for extracting the factors. There is no consensus on the ideal number of series to be used or on the distribution of indicators from different blocks in the data set from which the factors are extracted. For example, Rünstler et al. (2009) forecast GDP growth using large data sets for several European economies. Number of series used for different countries in Rünstler et al. (2009) ranges from 76 to 393. Moreover, distribution of the series in different blocks changes considerably. For instance, they do not use any price variable for euro area but use 42 price series for Belgium. Boivin and Ng (2006) note that adding more data may not always be useful for forecasting. They find that factors extracted from 40 pre-selected variables may yield better forecasting performance than using 147 series for factor extraction. Hence, composition of the data set may be crucial for forecasting performance.

In the empirical section, data from following groups are used: industrial production, foreign trade, consumer and business

[6] This criterion is defined as $BIC_3(k) = V(k, \widehat{F}^k) + k\widehat{\sigma}^2\left(\frac{(N+T-k)\ln(NT)}{NT}\right)$.

confidence, interest rates, exchange rates, European Union industrial production and confidence indicators, commodity prices, stock exchange, and global risk perception indicators (Table 1). Due to the technical requirements of principal components analysis, we work with stationary form of the series. For the series that exhibit seasonality, seasonally adjusted series are used. In the pseudo out-of-sample forecasting exercise, data are standardized at each point before extracting factors.

There are blocks that are frequently used for the factor models like real sector variables, prices and surveys. However, one can use a particular indicator from these blocks at different aggregation levels. For example, one can collect data on industrial production as headline index. Alternatively, he/she can use MIGS (Main Industrial Groupings) where industrial production is presented as sum of intermediate goods, consumer goods, investment goods, and energy. In yet another classification, one can see a more detailed picture of industrial production, such as production of food, textile, and so on for about 20 different sectors. So, from industrial production block one can use the headline series, five series from MIGS or around twenty from NACE or all of them at the same time (Table 2). Hence, deciding whether to use aggregated or disaggregated data and determining the level of detail for the disaggregation is another key issue that a forecaster faces when constructing a data set.

Following the approach of Barhoumi et al. (2010) for the empirical exercise of this paper three data sets are constructed with different aggregation levels: small (22 series), medium (63 series), and large (167 series). As an example, in the small data set for the industrial production only headline series are used. In the medium data set, industrial production components from MIGS are used. Note that in this case headline index for industrial production is not used. In the large data set, more detailed disaggregated sectoral classification for industrial production is used.

There is another angle that can help to understand the data set structure on the forecasting performance. Namely, analyzing the effect of different data blocks on forecast performance. This is the approach used, for example, by Forni et al. (2003) and Schumacher (2010). Forni et al. (2003) construct six alternative data sets: A master data set and five limited data sets constructed by excluding (one at a time) financial block/money block/price block/industrial production block/survey block. They find that, for inflation, for both SW and FHLR approaches excluding financial variables cause deterioration in the forecast performance. For industrial production picture is less systematic. Depending on the horizon and the factor extraction method, excluding financial block may increase or decrease the forecast errors. Schumacher (2010) analyzes forecasting performance with and without international data. He finds that adding international variables does not reduce forecast errors

Table 2
Example of increasing detail: Case of industrial production.

Small Data Set	Medium Data Set	Large Data Set
Industrial Production	Intermediate	Mining
	Capital	Food
	Non-durable	Beverage
	Durable	Tobacco
	Energy	Textile
		Apparel
		Leather
		Wood
		Paper
		Media
		Refined petroleum
		Chemical
		Pharmaceutical
		Rubber
		Other Mineral
		Basic Metal
		Fabricated Metal
		Electronic and Optical
		Electrical Equipment
		Machinery and Equipment
		Motor Vehicles
		Other Transport
		Furniture
		Other manufacturing
		Repair of machinery-equipment
		Electricity, gas and steam

Notes: An example of increasing detail level of the data set is shown in the table. In the small data set headline series is used, in the medium data set MIGS classification (headline index is not used in this data set) is used and in the large data set a more disaggregated sectoral detail is used.

Table 1
Indicators used in the small data set.

Indicator	Data Source
1. Industrial Production	TURKSTAT
2. Export Quantity Index	TURKSTAT, Author's Calculation
3. Import Quantity Index	TURKSTAT, Author's Calculation
4. Business Tendency Survey- Assessment of General Situation	CBRT
5. Capacity Utilization	CBRT
6. CNBC-e Consumer Confidence Index	CNBC-e
7. Inflation	TURKSTAT
8. Euro/Dollar Parity	CBRT
9. Dollar Exchange Rate	CBRT
10. TL Deposit Interest Rate	CBRT
11. Dollar Deposit Interest Rate	CBRT
12. TL Commercial Credit Interest Rate	CBRT
13. Euro Commercial Credit Interest Rate	CBRT
14. TL Consumer Credit Interest Rate	CBRT
15. Benchmark Interest Rate	CBRT
16. EU-Industrial Production	EUROSTAT
17. EU-Consumer Confidence	EUROSTAT
18. EU-Business Confidence	EUROSTAT
19. Commodity Price Index	Index mundi
20. VIX	Yahoo Finance
21. SP 500	Yahoo Finance
22. Borsa Istanbul-30	Istanbul Stock Exchange

Notes: Table shows the indicators that are used in the small data set. In the medium and large data sets, more disaggregated versions of these series are used.

of factor models. But, if a pre-selection is applied to the dataset though LARS-EN method, then there is improvement.

In Table 3, indicators that are excluded to construct three more data sets in addition to the small master data set are shown. For example for analyzing the effect of indicators from European Union on the forecast performance, forecasts from two data sets are computed: first one uses all the indicators and second one excludes industrial production, consumer and business confidence for the European Union. Similarly, data sets are constructed by excluding commodity and financial variables and the final one by excluding interest rates. Then forecasting performance of the master data set and three limited data sets constructed by excluding certain blocks one at a time are compared. In Table 3, indicators for the small data set are shown. For the medium and large data sets, the disaggregated versions of these indicators are excluded from the respective master data sets.

2.5. Forecast evaluation

In this paper, evaluation criterion for the forecasting performance is the RMSE that one gets from a pseudo out-of-sample forecasting exercise. Formally RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{[(Realization\ at\ time\ t+h) - (Forecast\ for\ time\ t+h\ at\ time\ t)]^2}{Number\ of\ forecasts}} \tag{13}$$

RMSE is calculated for a given evaluation sample. But forecast performance may be time varying. So, it can be informative to calculate and evaluate RMSE for different models for different time periods. Indeed, Stock and Watson (2003) note that relative performance of the models may change in different samples. They divide their evaluation sample into two parts and compare the relative performance of a large number of selected indicators for forecasting output relative to a benchmark for each of these samples. They find that only 10 percent of the indicators beat the benchmark in both periods, while around 20 percent of the indicators beat the benchmark in only one of the evaluation periods (page 811). Altug and Uluceviz (2014) analyze forecasting performance of selected indicators for the Turkish industrial production. Their results show that the forecast performance relative to an AR model changes depending on the evaluation sample. They find that recently it gets harder to beat the AR model.

Table 3
Indicators excluded for the construction of data sets for analyzing the effect of data blocks.

Data set excluding European Union variables
1. EU-Industrial Production
2. EU-Consumer Confidence
3. EU-Business Confidence
Data set excluding commodities and financial variables
1. Commodity Price Index
2. VIX
3. SP 500
4. Borsa Istanbul-30
Data set excluding interest rates
1. TL Deposit Interest Rate
2. Dollar Deposit Interest Rate
3. TL Commercial Credit Interest Rate
4. Euro Commercial Credit Interest Rate
5. TL Consumer Credit Interest Rate
6. Benchmark Interest Rate

Notes: Table shows which indicators are excluded to construct the data sets to analyze the role of data blocks on the forecasting performance. This table shows the excluded series for the small data set. In the medium and large data sets, disaggregated versions of these series are excluded from the master data sets.

In this paper, out-of-sample forecasting evaluation is done for two sub-samples to see whether the forecast performance is stable or not. Models are estimated starting from February 2005. In the first evaluation sample, out-of-sample recursion starts in January 2010 and ends in September 2011. For the second evaluation sample, the recursion starts in October 2011 and ends in September 2013. Data are available up until September 2014, and the longest horizon that the paper is interested in is the twelve-month-ahead. So, September 2013 is the last point in the recursion that one can compare twelve-month-ahead forecast with a realization.

In the pseudo out-of-sample forecasting exercise, it is aimed to mimic the situation that one would face if he/she produced forecasts at that point in time. At each step, factors are obtained with data that would be available at that time, lag lengths in Equation (10) are calculated, appropriate equation for h-step-ahead forecasting is estimated, and forecasts are obtained. Two versions of Equation (10) are estimated. In the first version, lags of the explanatory variables are used, as per the DI-AR Lag specification in Stock and Watson (2002b, page 149). The second specification is the DI of Stock and Watson (2002b), where one uses only contemporaneous values of the factors. For DI-AR Lag, lag lengths are determined using the Bayesian Information Criteria. Then, forecasts are obtained using the model suggested by the information criteria.

2.6. Benchmark model

In the forecasting literature, it is customary to compare models with a simple benchmark. This benchmark can be an autoregressive (AR) model or a random walk. Intuition behind comparing with a benchmark is the idea that going over all the intricate details of a complicated forecasting method may not be worth it if it cannot even beat a simple benchmark. Choosing an AR model as the benchmark suggests using just the time series properties of a variable to construct the forecasts. In the literature a frequently observed finding is inability of the sophisticated models in beating simple benchmarks.

Benchmark model in this paper is the average of the past realizations at the relevant recursion. For example, for twelve-month-ahead forecasting, the average of the twelve month cumulative growth until September 2013 is taken as the forecast for twelve-month-ahead forecast for September 2014. AR models are considered as the benchmark as well but this simple model outperformed them in most of the cases so it is chosen as the benchmark. In the tables where alternative specifications are compared, relative RMSE of the factor models with respect to the simple benchmark are presented. A figure lower than 1 means that, on average, the model is superior, i.e. it makes lower forecast error than the simple benchmark.

3. Results

In this section, we present the results obtained from pseudo out-of-sample forecasting exercise. First of all, the case of industrial production is discussed and then forecasting performance of the models for core inflation are analyzed.

In a nutshell, findings indicate that for industrial production and core inflation the best and the worst performing specifications are different. Another observation is that relative performance of the models is time varying. Yet, it is seen that it is highly likely that using a large data set with a high number of factors with a forecast equation that uses lags of those factors will provide poor forecasts. A specification with a carefully minted small or medium size data

set with a few factor may perform relatively successfully.

3.1. Industrial production

Four tables are presented for the analysis of the forecast performance of different specifications. In the first two, top 5 specifications are shown (Table A. 1 and Table A.2) while in the third and fourth the worst 5 are presented (Table A.3 and Table A.4). This exercise enables one to see whether there is a pattern in the best and the worst specifications. Following general points are worth highlighting:

- Best forecasts are not obtained by the largest dataset. In several cases, forecasts from a small data set with a forecast equation using only contemporaneous factors decreased forecast errors relative to benchmark considerably. So, parsimonious models can produce competitive forecasts.
- SW is not systematically worse than FHLR. Comparing SW and FHLR approaches, in the best specifications SW appears more frequently in the first evaluation sample while FHLR appears more frequently in the second evaluation sample. Yet, RMSEs one gets using different approaches are close to each other.
- Considering the best performing specifications, for the second evaluation sample for all the forecast horizons considered, DI-AR-Lag type forecast equations are used. For the first evaluation sample, DI appears relatively more frequently. It is interesting that for the worst specifications DI-AR Lag appears more frequently as well. Hence, it can be said that there are other determinants of the forecasting performance that interacts with the forecast equation type.
- In the literature review, it is seen that IC1 and IC2 are used more frequently for deciding the number of factors. While Bai and Ng (2002) point out the promising performance of BIC3, its use in practice is rare. However, tables for the top 5 specifications show that in addition to IC1 and IC2, BIC3 appears frequently as well. In the worst specifications PC3 and IC3 dominate the table indicating that using a large number of factors may harm forecasting performance.
- In the best performing specifications, it is seen that excluding interest rates and financial variables improve forecasting performance. Interestingly, for the worst specifications these data sets appear as well. This observation again shows that effect of a specific modelling choice is not independent from other choices.
- Excluding EU variables from the data set increases the forecast errors in the short run for the second evaluation sample. This may be due to the fact that EU is the major exporting partner for Turkey and developments in the export markets affect the demand.
- In general as the forecast horizon increases, we would expect a worsening in the relative performance of the factor models. However, for the first evaluation sample, as the forecast horizon increases relative performance of the factor models improve. Economic dynamics in the first evaluation sample may be the key reason for this result. That was the period just after the 2008–2009 global financial crisis. Three-month-ahead growth rates exhibit two peaks in this evaluation sample while factor model forecasts were relatively more volatile. For twelve-month-ahead growth rates there was a mean reverting behavior. Factor models did a better job tracking this trend-like movement than the benchmark. Hence, we cannot generalize from these result about the relative forecast performance of factor models at different horizons.
- Modelling decisions affect forecasting performance of the factor models considerably. For example, in the best equations one gets forty percent improvement relative to the benchmark while for the worst specifications one may get four times higher RMSE relative to the benchmark.

3.2. Core inflation

Similar to the industrial production, for core inflation four tables are presented and key points are highlighted. In the first two, the top 5 specifications are shown (Table A.5 and Table A.6) while in the third and fourth, the worst 5 are presented (Table A.7 and Table A.8). This exercise enables one to see whether there is a pattern in the best and the worst specifications. Following points are noted from the inspection of the tables:

- For three to nine-month-ahead forecasts, DI-AR Lag appears more frequently in the top 5 equations while DI appears more frequently in the worst 5 equations. In the literature, there are similar findings and observations suggesting that for core inflation, which a relatively more persistent series, using lags may help at forecasting.
- Both in the best and worst five equations FHLR approach appears relatively more frequently. It should be noted though that using SW approach results in similar RMSEs as well. Hence, while FHLR seems to perform better, its advantage is marginal.
- Tables showing the top 5 specifications for the first evaluation sample indicate that in addition to IC1 and IC2, BIC3 appears frequently as well.
- In the top performing specifications it is seen that excluding European Union variables decrease RMSE, while in the worst 5 specifications it is observed that excluding financial variables or interest rates cause an increase in RMSEs.
- Specifications that exclude financial variables from the medium data set are among the best performing models for twelve-month-ahead for the second evaluation sample. For the large data set, excluding those variables are among the worst performing specifications. So, for forecasting with factor models it is important to take into account other modelling dimensions as well when analyzing the effect of a certain modelling decision.
- For the first evaluation sample, even the best specifications cannot beat the benchmark for twelve-month-ahead forecasts. Öğünç et al. (2013), who forecast inflation with various models, do the out-of-sample forecast evaluation from fourth quarter of 2009 to second quarter of 2011 which is broadly same with our first evaluation sample (January 2010–September 2011). For the single equation factor model, they find that while for the two-quarter-ahead forecasts factor model can beat benchmark, for three and four-quarter-ahead factor model's forecast errors (evaluated with the RMSE) are higher than the random walk around 20 percent. Since the first evaluation sample covers the period after the 2008–2009 global financial crisis, inflation dynamics over this period may be playing certain role on the relatively poor performance of the factor models over 9–12 twelve-month-ahead forecasts. Indeed, a closer look at the forecast errors of factor models and the benchmark shows that forecast errors of the factor models were larger at the beginning of 2010. As a matter of fact, should we restrict the analysis to the after the May 2010, factor models would be able to beat the benchmark.
- Modelling decisions affect forecasting performance of the factor models considerably. For example, for three month ahead forecasts of the first evaluation sample, in the best equations one can get up to thirty percent improvement relative to the benchmark while for the same horizon using a different specification results in deterioration of up to 20 percent relative to the same benchmark.

4. Conclusion

Forecasting how key macroeconomic indicators, such as real economic activity or inflation, are going to evolve over the medium-term is essential for monetary policy making. This is a complicated task, though. Technological advances make it easier to construct data sets with hundreds of domestic and international variables easily. At first sight, increasing data set size may be thought to reduce forecast errors as one can get information about a wide range of areas. However, standard techniques such as OLS or VAR cannot handle a large number of indicators due to degrees of freedom problem. Therefore, the trend towards collecting big data that generates enormous amount of information requires using appropriate techniques to digest the information content of these data sets. Factor model approach is the natural candidate to serve as tool to process large data sets. Basic rationale of factor models is to summarize information in a large data sets with some few underlying factors.

There are different dimensions for evaluating the relative forecasting performance of the models. This is due to the fact that factors are unobservable, number of factors to extract from a data set is unknown, there is no formal guide for constructing a data set and multi-step forecast equation can be set up with or without the lags of the factors. Some papers concentrate on part of these dimensions while keeping others fixed. For example, some authors take a data set as given and analyze the effect of the number of factors on forecasting performance, while others look at the effect of changing the size of the data set while keeping the criterion for selecting the number of factors as fixed. Moreover, many papers evaluate models in a given period. However, different choices may not be mutually independent.

This paper takes a broader approach and makes a comprehensive analysis of the sensitivity of forecasting performance of factor models to inputs used in the models. Empirical exercise analyzes whether using aggregate or disaggregate data, whether number of factors extracted from the data set, whether using lags of the factors in the forecast equation and whether factor extraction approach affect the forecast performance. Moreover, part of the analysis is devoted to the role of certain data blocks on forecasting performance to see whether it is desirable at all to use the largest possible data set. This systematic and comprehensive approach can provide useful insights for practical applications as forecasters become more familiar about how forecasting performance changes with different parameters. In the end, this effort may help to optimize model selection for forecasting with factor models.

Findings indicate that for industrial production and core inflation the best and the worst performing specifications are different. Also, relative performance is time varying. Yet, it is seen that using a large data set with a high number of factors with a forecast equation that uses lags of those factors is more likely to provide poor forecasts. A specification with a carefully minted small or medium size data set with a few factor may perform relatively successfully.

These results are in line with the findings of the papers that analyze the forecasting performance of the factor models for advanced economies. Factor extraction method does not seem to change the forecasting performance considerably, more data is not found to be always better for forecasting performance, and number of factors extracted can change the relative forecast performance of alternative specifications. So, while our study focused on Turkish industrial production and inflation, results of the study contribute to the literature on forecasting with factor models by reinforcing some of the common findings in the literature.

In conclusion, *the answer to the question whether factor models are successful at forecasting is that it depends.* One may get substantial improvement relative to a benchmark with a well-crafted factor model while one may get rather poor forecast performance with ill-structured factor models. There is no systematic pattern to prescribe a recipe for the inputs of factor models that produces relatively successful forecasts at all times. So, as data accumulates, monitoring of the performance of the alternative factor model specifications and maintenance of the models accordingly would be beneficial for the effective use of factor models.

Appendix

Table A1
Rankings of the Models for Industrial Production. (The Best Performing Five Specifications, First Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation	Data Set	Evaluation Sample: Jan. 2010–Sept. 2011
Three-Month-Ahead						
1	DI	FHLR	PC3	M = H = 16	Large/Excl. Fin.	0.857
2	DI	FHLR	IC3	M = H = 16	Large/Excl. Fin.	0.861
3	DI_AR_Lag	FHLR	IC1	M = H = 16	Large/Excl. Fin.	0.877
4	DI_AR_Lag	FHLR	IC2	M = H = 16	Large/Excl. Fin.	0.877
5	DI_AR_Lag	SW	BIC3	–	Large/Excl. Fin.	0.880
Six-Month-Ahead						
1	DI	FHLR	IC2	M = H = 16	Small/Excl. Int. Rates	0.810
2	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. Fin.	0.824
3	DI	SW	IC2	–	Small/Excl. Int. Rates	0.824
4	DI_AR_Lag	SW	BIC3	–	Large/Excl. Fin.	0.830
5	DI	FHLR	BIC3	M = H = 16	Large/Excl. Fin.	0.830
Nine-Month-Ahead						
1	DI_AR_Lag	SW	PC3	–	Medium/All	0.620
2	DI_AR_Lag	SW	IC3	–	Medium/All	0.620
3	DI	SW	PC1	–	Small/Excl. Int. Rates	0.634
4	DI	SW	PC2	–	Small/Excl. Int. Rates	0.634
5	DI	SW	PC3	–	Small/Excl. Int. Rates	0.634

Table A1 (*continued*)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation	Data Set	Evaluation Sample: Jan. 2010–Sept. 2011
Twelve-Month-Ahead						
1	DI	SW	PC1	–	Small/Excl. Int. Rates	0.574
2	DI	SW	PC2	–	Small/Excl. Int. Rates	0.574
3	DI	SW	PC3	–	Small/Excl. Int. Rates	0.574
4	DI	SW	IC1	–	Small/Excl. Int. Rates	0.574
5	DI	SW	IC3	–	Small/Excl. Int. Rates	0.574

Notes: Table shows the best five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Barlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A2
Rankings of the Models for Industrial Production. (The Best Performing Five Specifications, Second Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation	Data Set	Evaluation Sample: Oct. 2011–Sept. 2013
Three-Month-Ahead						
1	DI_AR_Lag	FHLR	IC1	M = H = 16	Large/Excl. Int. Rates	0.797
2	DI_AR_Lag	FHLR	PC2	M = H = 16	Large/Excl. Int. Rates	0.815
3	DI_AR_Lag	FHLR	IC1	M = H = 16	Medium/Excl. Int. Rates	0.827
4	DI_AR_Lag	FHLR	IC2	M = H = 16	Medium/Excl. Int. Rates	0.827
5	DI_AR_Lag	SW	IC1	–	Large/Excl. Int. Rates	0.834
Six-Month-Ahead						
1	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. Int. Rates	0.768
2	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/All	0.781
3	DI_AR_Lag	SW	BIC3	–	Large/Excl. Int. Rates	0.804
4	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. Fin.	0.807
5	DI_AR_Lag	SW	BIC3	–	Large/All	0.856
Nine-Month-Ahead						
1	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/All	0.862
2	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. Int. Rates	0.892
3	DI_AR_Lag	SW	BIC3	–	Large/All	0.907
4	DI_AR_Lag	FHLR	BIC3	M = H = 16	Medium/Excl. Int. Rates	0.912
5	DI_AR_Lag	SW	IC2	–	Large/Excl. Int. Rates	0.932
Twelve-Month-Ahead						
1	DI_AR_Lag	SW	IC2	–	Small/Excl. Int. Rates	0.695
2	DI_AR_Lag	FHLR	BIC3	M = H = 16	Medium/All	0.745
3	DI_AR_Lag	SW	BIC3	–	Medium/All	0.759
4	DI_AR_Lag	SW	IC1	–	Large/All	0.783
5	DI_AR_Lag	SW	IC2	–	Large/All	0.783

Notes: Table shows the best five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Bartlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A3
Rankings of the Models for Industrial Production. (The Worst Performing Five Specifications, First Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation	Data Set	Evaluation Sample: Jan. 2010–Sept.2011
Three-Month-Ahead						
336	DI_AR_Lag	FHLR	PC1	M = H = 16	Large/Excl. Int. Rates	1.17
337	DI_AR_Lag	FHLR	IC3	M = H = 16	Large/Excl. Int. Rates	1.17
338	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/All	1.19
339	DI_AR_Lag	SW	PC1	–	Large/Excl. Int. Rates	1.25
340	DI_AR_Lag	SW	IC3	–	Large/Excl. Int. Rates	1.26
Six-Month-Ahead						
336	DI_AR_Lag	SW	IC3	–	Large/All	1.25
337	DI_AR_Lag	SW	PC3	–	Large/Excl. Fin.	1.26
338	DI_AR_Lag	SW	PC3	–	Large/All	1.34
339	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/All	1.36
340	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/Excl. Int. Rates	1.36
Nine-Month-Ahead						
336	DI	SW	PC3	–	Large/Excl. Int. Rates	1.34
337	DI_AR_Lag	SW	PC2	–	Large/Excl. Int. Rates	1.37
338	DI_AR_Lag	SW	PC1	–	Large/All	1.37
339	DI	FHLR	PC3	M = H = 16	Large/Excl. Int. Rates	1.39
340	DI_AR_Lag	SW	IC3	–	Large/Excl. Int. Rates	1.39
Twelve-Month-Ahead						
336	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/Excl. Fin.	1.34
337	DI_AR_Lag	SW	PC3	–	Large/Excl. Fin.	1.35
338	DI_AR_Lag	SW	PC3	–	Medium/Excl. EU	1.42
339	DI_AR_Lag	SW	IC3	–	Medium/Excl. EU	1.42
340	DI_AR_Lag	SW	PC1	–	Large/All	1.47

Notes: Table shows the worst five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Barlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A4
Rankings of the Models for Industrial Production. (The Worst Performing Five Specifications, Second Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation for FHLR Approach	Data Set	Evaluation Sample: October 2011–September 2013
Three-Month-Ahead						
336	DI_AR_Lag	SW	BIC3	–	Small/Excl. EU	1.22
337	DI	SW	IC2	–	Large/Excl. EU	1.24
338	DI_AR_Lag	FHLR	BIC3	M = H = 16	Small/Excl. EU	1.24
339	DI	SW	PC3	–	Large/Excl. EU	1.24
340	DI	SW	IC3	–	Large/Excl. EU	1.25
Six-Month-Ahead						
336	DI_AR_Lag	SW	IC3	–	Large/Excl. Int. Rates	2.10
337	DI_AR_Lag	FHLR	IC3	M = H = 16	Large/Excl. Fin.	2.14
338	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/Excl. Int. Rates	2.17
339	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/Excl. Fin.	2.26
340	DI_AR_Lag	FHLR	IC3	M = H = 16	Large/Excl. Int. Rates	2.32
Nine-Month-Ahead						
336	DI_AR_Lag	FHLR	PC3	M = H = 16	Medium/All	3.03
337	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/All	3.03
338	DI_AR_Lag	FHLR	PC3	M = H = 16	Medium/Excl. Int. Rates	3.09
339	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/Excl. Int. Rates	3.09
340	DI_AR_Lag	SW	PC1	–	Large/Excl. Int. Rates	3.11
Twelve-Month-Ahead						
336	DI_AR_Lag	FHLR	PC1	M = H = 16	Large/Excl. Int. Rates	3.73
337	DI_AR_Lag	SW	PC1	–	Large/Excl. Int. Rates	3.80
338	DI_AR_Lag	FHLR	PC3	M = H = 16	Medium/All	3.80
339	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/All	3.80
340	DI_AR_Lag	FHLR	IC3	M = H = 16	Large/Excl. Int. Rates	3.81

Notes: Table shows the worst five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Barlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A5
Rankings of the Models for Core Inflation. (The Best Performing Five Specifications, First Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation for FHLR Approach	Data Set	Evaluation Sample: Jan. 2010–Sept. 2011
Three-Month-Ahead						
1	DI_AR_Lag	FHLR	BIC3	M = H = 16	Small/Excl. EU	0.716
2	DI_AR_Lag	SW	BIC3	–	Small/Excl. EU	0.717
3	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. EU	0.725
4	DI_AR_Lag	FHLR	BIC3	M = H = 16	Small/Excl. Int. Rates	0.738
5	DI_AR_Lag	SW	BIC3	–	Small/Excl. Fin.	0.739
Six-Month-Ahead						
1	DI_AR_Lag	SW	PC3	–	Large/Excl. EU	0.851
2	–	Bivariate	–	–	Small	0.854
3	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. EU	0.855
4	DI_AR_Lag	FHLR	BIC3	M = H = 16	Small/Excl. EU	0.857
5	–	Bivariate	–	–	Medium	0.860
Nine-Month-Ahead						
1	DI_AR_Lag	FHLR	BIC3	M = H = 16	Small/Excl. EU	0.937
2	DI_AR_Lag	SW	BIC3	–	Small/Excl. EU	0.941
3	DI_AR_Lag	FHLR	BIC3	M = H = 16	Large/Excl. EU	0.957
4	DI_AR_Lag	SW	BIC3	–	Large/Excl. EU	0.965
5	DI_AR_Lag	FHLR	IC2	M = H = 16	Large/Excl. EU	0.967
Twelve-Month-Ahead						
1	Benchmark	–	–	–	–	1.000
2	DI	FHLR	PC3	M = H = 16	Large/Excl. EU	1.022
3	DI	FHLR	IC1	M = H = 16	Medium/Excl. EU	1.028
4	DI	FHLR	IC2	M = H = 16	Medium/Excl. EU	1.028
5	DI	FHLR	BIC3	M = H = 16	Small/Excl. EU	1.039

Notes: Table shows the best five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Barlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A6
Rankings of the Models for Core Inflation. (The Best Performing Five Specifications, Second Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation for FHLR Approach	Data Set	Evaluation Sample: Oct. 2011–Sept.2013
Three-Month-Ahead						
1	DI_AR_Lag	FHLR	PC1	M = H = 16	Small/Excl. EU	0.749
2	DI_AR_Lag	FHLR	PC2	M = H = 16	Small/Excl. EU	0.749
3	DI_AR_Lag	FHLR	PC3	M = H = 16	Small/Excl. EU	0.749
4	DI_AR_Lag	FHLR	IC1	M = H = 16	Small/Excl. EU	0.749
5	DI_AR_Lag	FHLR	IC2	M = H = 16	Small/Excl. EU	0.749
Six-Month-Ahead						
1	DI_AR_Lag	SW	IC2	–	Small/All	0.731
2	DI_AR_Lag	FHLR	IC1	M = H = 16	Small/Excl. Fin.	0.735
3	DI_AR_Lag	FHLR	IC2	M = H = 16	Small/All	0.735
4	DI	SW	IC2	–	Small/Excl. Fin.	0.741
5	DI_AR_Lag	SW	IC2	–	Small/Excl. Fin.	0.741
Nine-Month-Ahead						
1	DI_AR_Lag	SW	IC2	–	Medium/Excl. Fin.	0.789
2	DI_AR_Lag	FHLR	IC2	M = H = 16	Small/All	0.794
3	DI_AR_Lag	FHLR	PC2	M = H = 16	Medium/Excl. Fin.	0.794
4	DI_AR_Lag	SW	IC2	–	Small/All	0.794
5	DI_AR_Lag	FHLR	IC1	M = H = 16	Small/Excl. Fin.	0.797
Twelve-Month-Ahead						
1	DI_AR_Lag	SW	IC2	–	Medium/Excl. Fin.	0.886
2	DI_AR_Lag	FHLR	PC2	M = H = 16	Medium/Excl. Fin.	0.888
3	DI	SW	IC2	–	Medium/Excl. Fin.	0.888
4	DI	FHLR	PC2	M = H = 16	Medium/Excl. Fin.	0.890
5	DI	SW	PC1	–	Small/Excl. Fin.	0.891

Notes: Table shows the best five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Barlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A7
Rankings of the Models for Core Inflation. (The Worst Five Performing Specifications, First Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation for FHLR Approach	Data Set	Evaluation Sample: Jan. 2010–Sept. 2011
Three-Month-Ahead						
336	DI	FHLR	PC3	M = H = 16	Large/Excl. Fin.	1.22
337	DI	FHLR	IC1	M = H = 16	Large/Excl. Fin.	1.22
338	DI	FHLR	IC2	M = H = 16	Large/Excl. Fin.	1.22
339	DI	FHLR	IC1	M = H = 16	Large/Excl. Int. Rates	1.22
340	DI	SW	IC1	–	Large/Excl. Fin.	1.22
Six-Month-Ahead						
336	DI	SW	IC1	–	Large/Excl. Fin.	1.25
337	DI	SW	IC2	–	Large/Excl. Fin.	1.25
338	DI	FHLR	PC3	M = H = 16	Large/Excl. Int. Rates	1.25
339	DI_AR_Lag	SW	PC3	–	Large/Excl. Fin.	1.25
340	DI	SW	IC1	–	Large/Excl. Int. Rates	1.25
Nine-Month-Ahead						
336	DI_AR_Lag	FHLR	IC1	M = H = 16	Large/Excl. Int. Rates	1.31
337	DI_AR_Lag	SW	PC3	–	Medium/Excl. Int. Rates	1.34
338	DI_AR_Lag	SW	IC3	–	Medium/Excl. Int. Rates	1.34
339	DI_AR_Lag	SW	PC3	–	Large/Excl. Fin.	1.38
340	DI_AR_Lag	FHLR	PC3	M = H = 16	Large/Excl. Fin.	1.39
Twelve-Month-Ahead						
336	DI_AR_Lag	FHLR	PC1	M = H = 16	Medium/All	1.86
337	DI_AR_Lag	FHLR	PC1	M = H = 16	Large/Excl. Fin.	1.88
338	DI_AR_Lag	FHLR	IC3	M = H = 16	Large/Excl. Fin.	1.88
339	DI_AR_Lag	FHLR	PC2	M = H = 16	Large/Excl. Fin.	1.89
340	DI_AR_Lag	FHLR	PC2	M = H = 16	Large/All	1.93

Notes: Table shows the worst five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Bartlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

Table A8
Rankings of the Models for Core Inflation. (The Worst Five Performing Specifications, Second Evaluation Sample)

Rank	Multistep Ahead Forecasting Method	Factor Extraction Method	Number of Static Factor Selection Method	M and H For Spectral Density Estimation for FHLR	Data Set	Evaluation Sample: Oct. 2011–Sept. 2013
Three-Month-Ahead						
336	DI_AR_Lag	SW	PC3	–	Medium/Excl. Int. Rates	1.32
337	DI	SW	PC3	–	Medium/All	1.32
338	DI	SW	IC3	–	Medium/All	1.32
339	DI	FHLR	PC2	M = H = 16	Large/Excl. Int. Rates	1.32
340	DI	FHLR	PC3	M = H = 16	Medium/All	1.33
Six-Month-Ahead						
336	DI	SW	PC1	–	Large/Excl. Int. Rates	1.36
337	DI_AR_Lag	FHLR	PC3	M = H = 16	Medium/All	1.37
338	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/All	1.37
339	DI	FHLR	PC3	M = H = 16	Medium/All	1.37
340	DI	FHLR	IC3	M = H = 16	Medium/All	1.37
Nine-Month-Ahead						
336	DI_AR_Lag	SW	PC3	–	Large/Excl. Int. Rates	1.25
337	DI	SW	IC3	–	Medium/Excl. Int. Rates	1.27
338	DI	SW	PC3	–	Medium/Excl. Int. Rates	1.27
339	DI	FHLR	IC3	M = H = 16	Medium/Excl. Int. Rates	1.28
340	DI	FHLR	PC3	M = H = 16	Medium/Excl. Int. Rates	1.29
Twelve-Month-Ahead						
336	DI	SW	PC3	–	Medium/Excl. Int. Rates	1.15
337	DI	FHLR	IC3	M = H = 16	Medium/Excl. Int. Rates	1.16
338	DI	FHLR	PC3	M = H = 16	Medium/Excl. Int. Rates	1.17
339	DI_AR_Lag	FHLR	PC3	M = H = 16	Medium/Excl. Int. Rates	1.22
340	DI_AR_Lag	FHLR	IC3	M = H = 16	Medium/Excl. Int. Rates	1.22

Notes: Table shows the worst five specifications out of 340 alternatives. DI_AR_Lag and DI show the forecast equation types. In the DI_AR_Lag in addition to the contemporaneous factors one uses lags of the factors and the dependent variable while for DI one uses only contemporaneous factors. SW and FHLR show factor extraction approaches of Stock and Watson (2002b) and Forni et al. (2005). PC1, PC2, PC3, IC1, IC2, IC3 and BIC3 show the information criteria for the number of static factors from Bai and Ng (2002) as discussed in Section 2.2. M = H = 16 shows the parameters for the spectral density estimation for FHLR approach where H refers to the number of frequency grids and M refers to the Bartlett lag window. Three master data sets are used, Small, Medium and Large. By excluding European Union variables, financial and commodity variables and interest rates from these sets, new data sets are constructed. Final column shows the Root Mean Squared Error (RMSE) relative to the simple benchmark where the average of the past realizations is used for forecasting.

References

Akkoyun, H.Ç., Günay, M., 2012. Nowcasting Turkish GDP Growth. Central Bank of the Republic of Turkey Working Paper, (12/33).

Altug, S., Uluceviz, E., 2014. Identifying leading indicators of real activity and inflation for Turkey, 1988-2010. OECD J.: J. Bus. Cycle Meas. Anal. 2014 (1), 1–37.

Bai, J., 2003. Inferential theory for factor models of large dimensions. Econometrica 71 (1), 135–171.

Bai, J., Ng, S., 2002. Determining the number of factors in approximate factor models. Econometrica 70 (1), 191–221.

Bai, J., Ng, S., 2007. Determining the number of primitive shocks in factor models. J. Bus. Econ. Stat. 52–60.

Bai, J., Ng, S., 2008. Large dimensional factor analysis. Found. Trends Econom. 3 (2), 89–163.

Barhoumi, K., Darné, O., Ferrara, L., 2010. Are disaggregate data useful for forecasting French GDP with dynamic factor models? J. Forecast. 29 (1–2), 132–144.

Barhoumi, K., Darné, O., Ferrara, L., 2013. Testing the number of factors: an empirical assessment for a forecasting purpose. Oxf. Bull. Econ. Stat. 75 (1), 64–79.

Boivin, J., Ng, S., 2005. Understanding and comparing factor based macroeconomic forecasts. Int. J. Cent. Bank. 1 (3), 117–152.

Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? J. Econom. 132 (1), 169–194.

D'Agostino, A., Giannone, D., 2012. Comparing alternative predictors based on large-panel factor models. Oxf. Bull. Econ. Stat. 74 (2), 306–326.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The generalized dynamic-factor model: identification and estimation. Rev. Econ. Stat. 82 (4), 540–554.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the euro area? J. Monetary Econ. 50 (6), 1243–1255.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2005. The Generalized dynamic factor model: one sided estimation and forecasting. J. Am. Stat. Assoc. 100, 830–840.

Günay, M., 2015. Forecasting Turkish industrial production growth with static factor models. Int. Econom. Rev. (IER) 7 (2), 64–78.

Kapetanios, G., Marcellino, M., 2009. A parametric estimation method for dynamic factor models of large dimensions. J. Time Anal. 30 (2), 208–238.

Modugno, M., Soybilgen, B., Yazgan, E., 2016. Nowcasting Turkish GDP and news decomposition. Int. J. Forecast. 32 (4), 1369–1384.

Öğünç, F., Akdoğan, K., Başer, S., Chadwick, M.G., Ertuğ, D., Hülagü, T., Kösem, S., Özmen, M.U., Tekatlı, N., 2013. Short-term inflation forecasting models for Turkey and a forecast combination analysis. Econ. Modell. 33, 312–325.

Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., Den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Ruth, K., Van Nieuwenhuyze, C., 2009. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. J. Forecast. 28 (7), 595–611.

Schumacher, C., 2007. Forecasting German GDP using alternative factor models based on large datasets. J. Forecast. 26 (4), 271–302.

Schumacher, C., 2010. Factor forecasting using international targeted predictors: the case of German GDP. Econ. Lett. 107 (2), 95–98.

Soybilgen, B., 2015. Three Essays on Forecasting (PhD dissertation). İstanbul Bilgi Üniversitesi.

Stock, J.H., Watson, M.W., 2002a. Forecasting using principal components from a large number of predictors. J. Am. Stat. Assoc. 97 (460), 1167–1179.

Stock, J.H., Watson, M.W., 2002b. Macroeconomic forecasting using diffusion indexes. J. Bus. Econ. Stat. 20 (2), 147–162.

Stock, J.H., Watson, M.W., 2003. Forecasting output and inflation: the role of asset prices. J. Econ. Lit. 41 (3), 788–829.