Forecasting Turkish Industrial Production Growth With Static Factor Models

Mahmut Günay®

Yıldırım Beyazıt University and The Central Bank of the Republic of Turkey

ABSTRACT

In this paper, we forecast industrial production growth for the Turkish economy using static factor models. We evaluate how the performance of the models change based on the number of factors we extract from our data as well as the level of aggregation for the series in the data set. We consider two evaluation samples for the out-of-sample forecasting exercise to assess the stability of the forecasting performance. We find that the effect of the data set size on the forecasting performance is not independent from the number of factors extracted from this data set. Rankings of the models change in different evaluation samples. We conclude that using a dynamic approach to evaluate models from different dimensions is important in the forecasting process.

Key words: *Forecasting, Factor Models, Principal Components*
JEL Classifications: E37, C32, C33

1. INTRODUCTION

A quote attributed to the Nobel laureate Niels Bohr states that *"prediction is very difficult, especially if it is about future"*. However, forecasts of the key macro variables are vital for real time policy making due to lags in the transmission mechanism. Since we are living in a stochastic world, in general, realizations will be different from predictions and time to time by a high margin. Hence, over an evaluation period, it would be unrealistic to expect zero forecast errors from a forecasting model. In this respect, efficiency of forecasts is as important as accuracy. Inefficiency can occur due to various reasons, such as not using an indicator that has adequate forecasting power in the prediction process, not using a modelling technique that is known at the time of forecasting, or not considering the appropriate parameters in the models. Hence, it is important for forecasters to check whether all information in the economy is utilized to the greatest extent possible and in an efficient way. In this paper, we approach the issue from two perspectives: how to utilize the wide range of available data and understand the effect of model specification on forecasting performance.

There are a lot of candidate indicators that can be used in the forecasting, and this number is increasing with the advances in information technology. Due to increasing connectedness within the global economy, considering international data in addition to domestic indicators in the forecasting of local variables may be necessary. 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:147) state that some variable selection procedures may be

---

® Mahmut Günay, Yıldırım Beyazıt University and The Central Bank of the Republic of Turkey, (email: mahmutgunay@gmail.com).
This work is part of the PhD thesis that is going on at the Department of Economics in the Yıldırım Beyazıt University under the supervision of Asst. Prof. Sıdıka Başçı. The views attributed in this study are those of the author and cannot necessarily be attributed to the Yıldırım Beyazıt University or the Central Bank of the Republic of Turkey.

used for determining the 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 model.

Factor models became popular in the last decade for dealing with large data. In factor models, 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, 2002a and 2002b). Factor models enable us to incorporate as many series as we want in the forecasting process, but, there may not be a linear relation between forecasting performance and the number of series we use for extracting factors. Also, the number of factors we extract from a given data set may affect the forecasting performance. Hence, analyzing the effect of modelling decisions in factor models on forecast performance may provide valuable information to the forecasters.

Factor model approach is a tool that enables us summarize information in a, possibly large, data set with few underlying factors. The basic rationale of factor models is presented in Equation 3.1. We decompose each series (X) into a part that is explained by the factors (in the jargon of factor models, common part) and to a part that is specific to the series (in the factor model jargon, idiosyncratic component).

$$X_{it} = \lambda_i' F_t = + e_{it} \tag{3.1}$$

where X is a stationary series; F, is a vector of factors and λ, lambda is factor loadings.

Equation 3.1 is a theoretical representation of a factor model but we have several issues to address when applying these models in practice. First of all, we do not observe factors. There are methods to extract factors, but they require some parameters as input. Another issue is that we need to construct a data set. With the advance of information technology, the cost of accessing information has decreased considerably, and we can gather large amounts of data relatively easily. Increasing the size of data set, however, may not always improve forecasting power. Thus, we need to analyze the effect of the data set structure on the performance of the models. Once we decide the data set, we need to choose how many factors to extract from this set.

In a meta-analysis where the results of papers on factor models are analyzed, Eickmeier and Ziegler (2008) find that the relative performance of the factor models depends on several things such as the variable that is forecast, the country of the study, and the size of the data set. Hence, although factor models let us use large amounts of data in the forecasting, there is no guarantee for obtaining a better forecast than when using simpler methods. In this respect, we think careful analysis of the sensitivity of the forecasting performance to the modeling choices is necessary. Boivin and Ng (2005) is an example for an effort in this direction. They analyze direct and indirect approach and different factor extraction methods. They show that factor model specification indeed affects the forecasting performance.

In this paper, we forecast industrial production growth for the Turkish economy using a large number of indicators from different blocks of data with factor models. Our data set covers indicators from production, foreign trade, financial variables, confidence indicators, interest rates, commodity prices, and international variables. We set up three different data sets from these categories to see how the forecasting performance changes with data set size. The number of factors from these data sets is obtained with the different criteria suggested in the literature.

We find that modelling choices such as the number of factors and size of the data set affect the forecasting performance of factor models. More importantly, the effect of the data set size on the forecasting performance is not independent from the number of factors used in the forecasting. Another finding is that the evaluation sample of the models may play a considerable role on the relative performance of the models. In this respect, the forecasting models to be used in the future must be selected with caution based on past performance. All in all, our results point out the importance of considering, continuously, all the dimensions of modelling for efficient forecasting. In the next sections we describe the data and summarize the methodology we use in the paper; we then present our results and conclude.

2. DATA

A critical issue that a forecaster needs to address before setting up the forecasting model is the composition of the data set. This choice is even more important in the case of factor models since we can use as many series as we can collect for extracting the factors. Yet, there is no consensus on the ideal number of series 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. The number of series used for different countries in Rünstler et al. (2009) ranges from 76 to 393. Moreover, the distribution of the data 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, the composition of the data set may have some effect on the forecasting performance.

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. For example, we have data on industrial production as headline index; in MIGS (Main Industrial Groupings) we see industrial production as the sum of intermediate goods, consumer goods, investment goods, and energy. In another classification, we see a more detailed picture of industrial production, such as production of food, textile, and so on for about 20 different sectors. A similar picture arises for soft data. We can use consumer confidence as the headline index, or we can also consider subcomponents, which are questions about the recent state of the economy as well as about expectations. Angelini et al. (2010) use series from different detail levels in the same data set. On the other hand, Barhouimi et al. (2010) use different data sets depending on the level of detail. Barhoumi et al. (2010) find that Stock and Watson (2002b)'s static approach with a small data set, which uses headline series rather than subcomponents, led to competitive results.

We follow the same approach as Barhoumi et al. (2010) and construct three data sets with different aggregation levels resulting in three data sets: small (22 series,), medium (63 series), and large (167 series). In the small data set, we use only headline growth in industrial production. In the medium data set we adopt the definition of industrial production as the sum of five categories defined by the MIGS classification. Hence, we use the growth rate of the each of these five items in the data set. In the large data set, we use a more detailed disaggregated sectoral classification for industrial production. Table 2.1 demonstrates the increasing level of detail described above.

We include data about industrial production, foreign trade, consumer and business confidence, interest rates, exchange rates, European Union industrial production and confidence indicators, commodity prices, stock exchange, and global risk perception indicators. Details about these sets are provided in the Appendix (Table A.4 to Table A.6). The series are transformed by taking logs, if appropriate, and first differenced to ensure stationarity. For the series that exhibit seasonality, we use seasonally adjusted series. In the pseudo out-of-sample forecasting exercise, we standardize data at each point before extracting factors.

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 mach-eq
		Electricity, gas and steam

Table 2.1 Example of Increasing Detail: Case of Industrial Production.
Notes: We show an example of increasing detail level of the data set. In the small data set we use headline, in the medium data set we use MIGS classification and in the large data set we use a more disaggregated sectoral detail.

3. METHODOLOGY

In this paper, we use factor models for forecasting Turkish industrial production (cumulative) growth for 3 and 12 month-ahead with three types of data sets; small, medium, and large. In this section, we provide details about our modelling strategies.

3.1. Factor Extraction

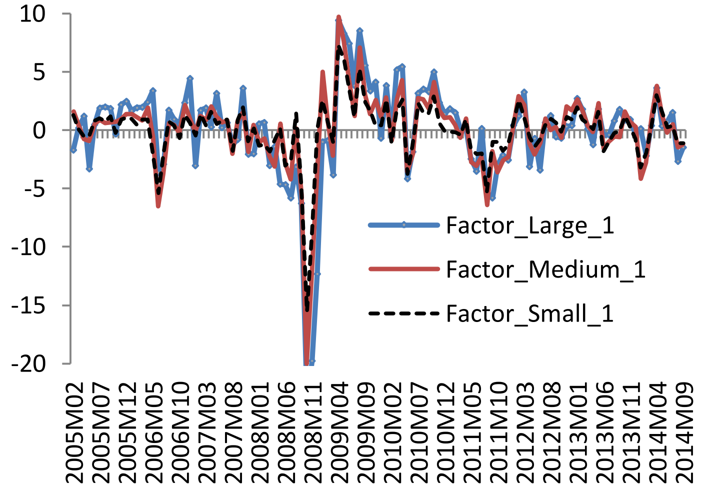
We extract factors with principal components as suggested by Stock and Watson (2002b). Factors can be obtained by using the formula in Equation 3.2. We obtain the eigenvectors of $X'X$, and using the eigenvectors corresponding to the largest r eigenvalues (we will discuss how we set r below in more detail) we can obtain factors.

$$\hat{F}_t = \frac{\hat{x}\hat{\Lambda}}{N} \qquad (3.2)$$

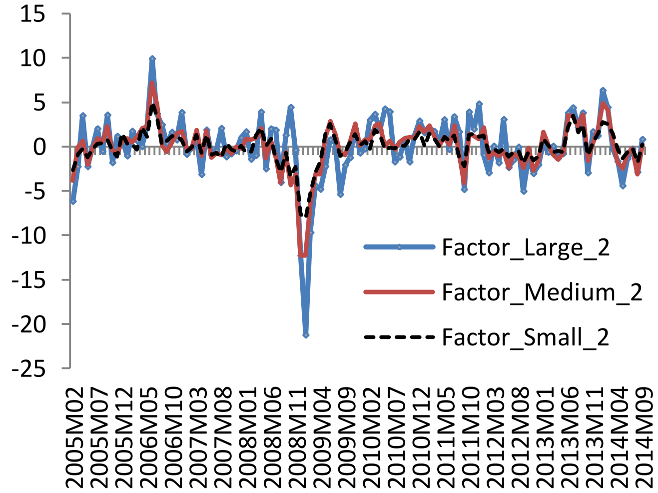
$\hat{\Lambda}$ = eigenvectors of $X'X$ corresponding tor largest eigenvalues where r is the number of factors.

Figure 3.1 Principal Components

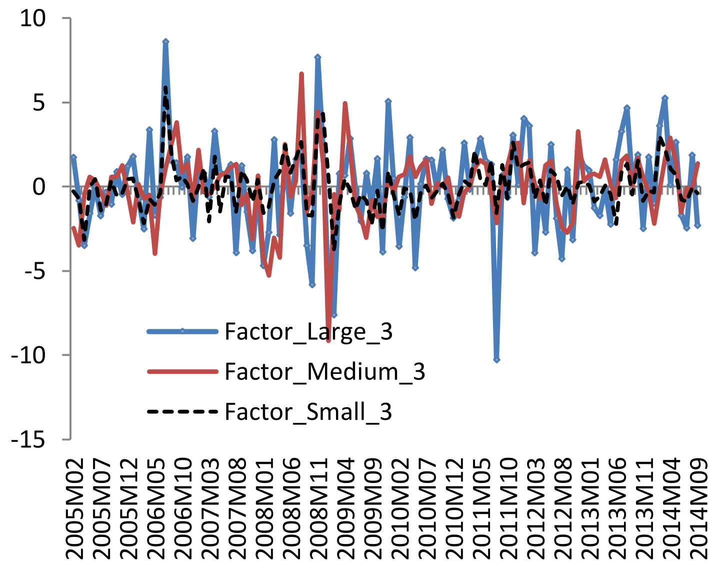
a. First Principal Components



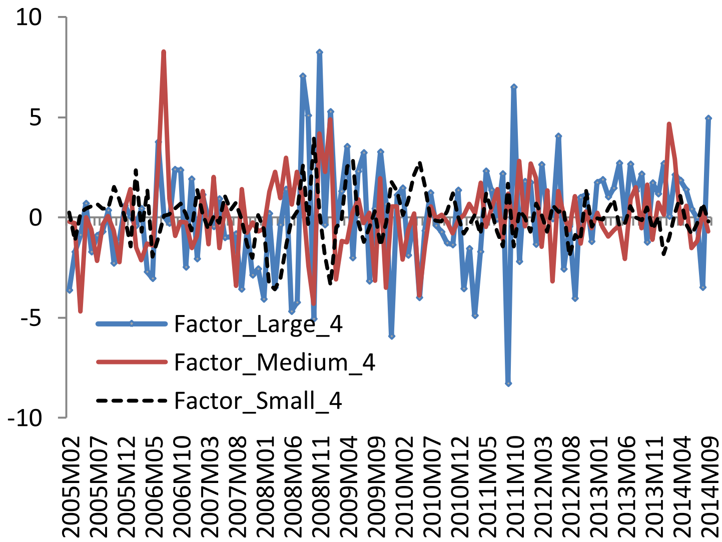
b. Second Principal Components



c. Third Principal Components



d. Fourth Principal Components

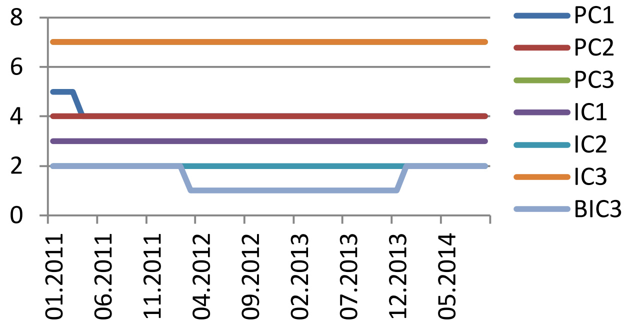


Notes: We show the principal components that we obtain from three different data sets namely, small, medium and large.

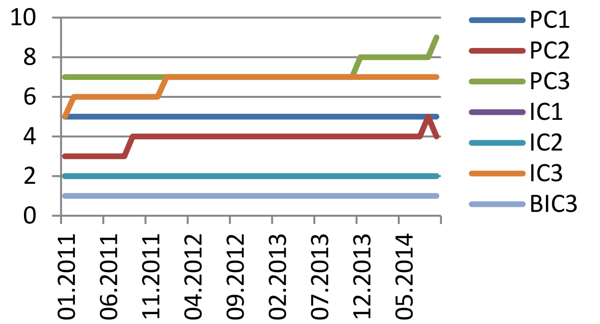
As we discussed in the data section, we work with three different data sets. In Figures 3.1.a to 3.1.d, we plot the first four principal components from these data sets. These principal components are the factors that we will use in the forecasting. Figure 3.1.a shows the first principal component from each of the three data sets. We see that they show similar patterns over the sample. Second, principal components are also fairly similar for medium and small data sets. When we go to the fourth principal components, the series become less similar. These observations suggest that, the number of factors we use in the forecasting may affect the conclusion about the effect of the size of the dataset. In particular, if we use only one factor, forecasts may be quite similar for three data sets. But if we use more than 2 factors, we may get different forecasts. In this respect, in the next section we discuss the choice of the number of factors.

Figure 3.2 Number of Factors Obtained from Different Information Criteria

a. Number of Factors Obtained from Different Information Criteria from Medium Data Set



b. Number of Factors Obtained from Different Information Criteria from Large Data Set



Notes: In the paper, we do recursive out–of-sample forecasting exercise. In the evaluation sample, at each month we get number of factors proposed by different criteria from Bai and Ng (2002). X axis shows the date where data set that we extract factors from ends. For small data set we set the maximum number of factors (required for PC1, PC2 and PC 3) as four, for medium data set as seven and for large data set as nine.

3.2. Number of Factors

In the Stock and Watson (2002b) approach to factor extraction, there is only one parameter that we need to set before we get the factors: number of factors. Bai and Ng (2002) note that if we know the true number of factors, we can use the Bayesian Information Criteria (BIC) to determine this number. When the factors are unknown and have to be estimated, however, the BIC will not always consistently estimate the true number of factors. Bai and Ng (2002) offered seven criteria to determine the number of factors. They find that PC1, PC2, IC1, and IC2 seem to perform better than PC3 and IC3 (for formulas see, Bai and Ng, 2002:201). In the presence of cross-section correlations, BIC3 has very good properties (see Bai and Ng, 2002:202 and 207). This criterion can be used despite not fulfilling all the conditions of Theorem 2 in their paper. Figure 3.2.a and Figure 3.2.b show the number of factors that we get by recursively expanding our medium and large data sets. As evident in the figures below, the seven criteria of the Bai and Ng (2002) give diverging results from each other in terms of the number of factors, and this number may change as we add more observations through time.

Some authors use one of these criteria and do not always check the role of using a certain information criterion on forecasting performance. For instance, Barhoumi et al. (2013) analyze the effect of the number of factors on forecasting performance. Although they are specifically interested in the effect of the number of factors on forecasting performance, they only employ IC1 among the Bai and Ng (2002) criteria. Gupta and Kabundi (2011) forecast South African variables with factor models. They find that PC1 and PC2 suggest seven factors, while IC1 and IC2 suggest five for their data set. They do not consider the BIC3 criterion for selecting the number of factors. They state that they use five factors. However, since the number of factors changes slightly over time and substantially depending on the choice of the criterion, it is still an empirical question to check whether using other criteria changes the forecast's performance. In this respect, we consider all of the seven criteria suggested by Bai and Ng (2002).

3.3. Forecast Equation

In this paper we present forecasting results for 3 and 12 month-ahead forecasts. Though other horizons for the analysis exist, we focus our attention for the sake of clarity in presentation. The 3 month-ahead forecast performance is expected to be informative about the short run performance of models, while the 12 month-ahead forecast is thought to be informative about the longer run. An important question emerges when we forecast more than one period ahead. Consider the month-on-month growth rate of industrial production as presented in Figure 3.3. We can define the 3 month-ahead forecast as the month-on-month growth from three months from now. For example, in a case where we have the January figures as the last data point, we can forecast what would be the monthly growth rate in April, which is three months from January. However, this is a highly volatile series, which would be very hard to forecast. Also, the monthly growth rate in April will depend on the monthly growth rate in March. Hence, the month-on-month growth rate from 3 or 12 month from now may not be very interesting from a policy maker's perspective. Rather, policy-makers may be interested in the over-all growth during these periods.

Figure 3.3 Month-on-Month Growth of Industrial Production (Annualized)

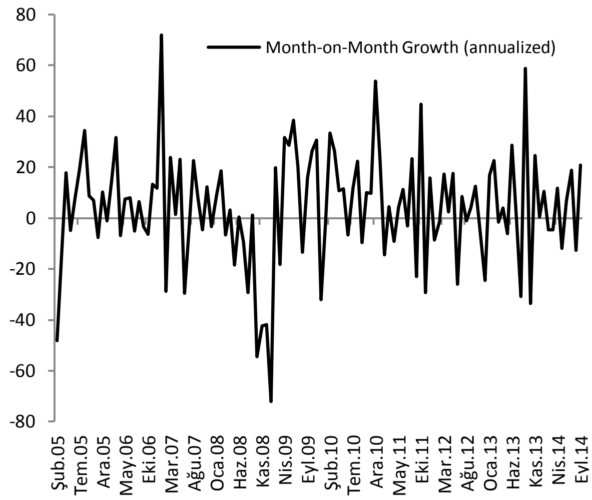
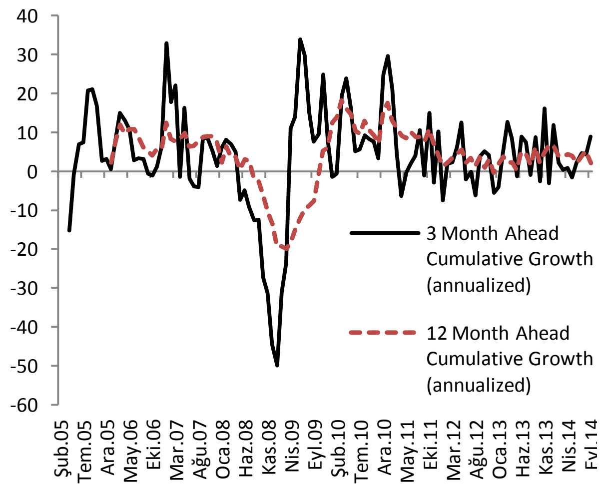


Figure 3.4 Three and Twelve Month Cumulative Growth of Industrial Production (Annualized)



In this respect, we follow Stock and Watson (2002a) and forecast the cumulative growth rate for 3 and 12 month-ahead periods. In this approach, for the case that we can access January data, we forecast the growth rate in April relative to the level in January; in other words, we work with the cumulative growth in the horizon of interest. The 3 and 12 month-ahead cumulative growth rates in Figure 3.4 show that, as expected, the 12 month-ahead cumulative growth rates are relatively more stable than the 3 month-ahead rates. We also observe that volatility of the three month growth increased after around mid-2011.

Equation 3.3 shows the forecasting model where we obtain coefficients with OLS. In this equation, the dependent variable is the cumulative growth rate from time t to time $t+h$ so that we are forecasting h-period ahead. We use month-on-month change of industrial production and the estimated factors as the independent variables. Using different letters in the notation of Equation 3.3 for the lag length (namely m and p) indicates that we can allow different number of lags for the lag of dependent variable and for the factors. We use a cap on the F which shows that we are working with estimated factors since we cannot observe actual factors.

$$\hat{Y}^{h}_{t+h/t} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}'_{hj} \hat{F}_{t-j+1} + \sum_{j=1}^{p} \hat{\gamma}'_{hj} Y_{t-j+1} \tag{3.3}$$

where Y is the variable that we want to forecast. In the direct forecasting approach, $\hat{\alpha}_h$ and $\hat{\beta}_{hj}$ change for each horizon. Subscript "h" in the dependent variable indicates that we define cumulative growth for each forecast horizon, h.

3.4. Forecast Evaluation

Our evaluation criterion for comparing the models is the Root Mean Squared Error (RMSE) that we get from a pseudo out-of-sample forecasting exercise. Stock and Watson (2003) note that the relative performance of the models may change in different samples. They divide their evaluation sample into two and compare the relative performance of selected indicators for forecasting output relative to a benchmark. 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. Altug and Uluceviz (2013) analyze the forecasting performance of selected indicators for 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.

We estimate models starting from February 2005 and do the evaluation for two samples to see whether the forecast performance is stable or not. In the first evaluation sample, the 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. We have data up until September 2014, and the longest horizon that we are interest in is 12 month-ahead. So, September 2013 is the last point in the recursion that we can compare our 12 month-ahead forecast with a realization. In other words, we will produce a forecast for the growth rate between September 2014 and September 2013. We will then compare this forecast with the realization.

At each step we get the factors, determine lag lengths in Equation 3.3, estimate the appropriate equation for h step-ahead forecasting, and derive the forecasts. We estimate two versions of Equation 3.3. In the first version, we use lags of the explanatory variables, as per the DI-AR Lag specification in Stock and Watson (2002b:149). The second specification is the DI of Stock and Watson (2002b), where we use only contemporaneous values of the factors. When we estimate the equations for DI-AR Lag, we determine the lag length using the Bayesian Information Criteria. After finding the appropriate model, using this model we get the forecasts.[1]

3.5. Benchmark Model

Our benchmark model is the average of the past realizations at the relevant recursion. For example, for 12 month-ahead forecasting, the average of the 12 month cumulative growth until September 2013 is taken as the forecast for 12 month-ahead forecast for September 2014. In the tables in the next section where we present results, we show the relative RMSE of the factor models compared to the simple benchmark. A figure greater than 1 means that we are making higher forecast errors than the simple benchmark. Although the benchmark is a very simple model, a frequently observed finding in the literature is the difficulty of beating the benchmark.

In summary, there are different dimensions for evaluating the relative forecasting performance of models. 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. We consider how changes in these dimensions affect the forecast performance at the same setup. This systematic search can provide useful insights in practice as forecasters become more familiar about how forecasting performance changes with different parameters, which may optimize model selection.

4. RESULTS

As we summarized in the methodology section, we observe the different dimensions that can affect forecasting performance. For the factor models, we change the number of factors, data set size, and evaluation period. Also, we specify the forecasting model in two different ways, namely the DI-AR lag and DI. Subsequently, we discuss results from the different dimensions; we have 42 models for each horizon, each with two evaluation episodes. We present the results by ranking the models

Beginning our discussion with the 3 month-ahead forecasts (Table 4.2), in a nutshell, we can say that the results are mixed and do not favor a single specification combination. Moreover, looking at the top and bottom of the list reveals that model specification plays a vital role on the conclusion about the performance of factor models. Factors extracted from a large data set with the PC2 rule, which gives around 4 factors, with a forecast equation set up according to the DI specification in Stock and Watson (2002a), give the least amount of forecast errors. On the other hand, if we extract factors from PC3, which tends to recommend using a large number of factors (in our case around 7), and set up the forecasting equation as DI-AR lag, we will do considerably worse than when using the simple benchmark. This observation is also valid for the second evaluation sample. In this regard, the sensitivity of the forecast performance of the factor models to the modelling choices should be taken into account when commenting about the relative performance of the models vis-à-vis a benchmark.

[1] Estimations are done in Eviews 7.1 with a code that we write for this project.

When we analyze the results with the aim of comparing DI and DI-AR Lag specifications, we see that in the first sample, the DI models dominate the top of the list, while in the second part we see that DI-AR lag specifications perform relatively better. In terms of data set size, a clear picture does not emerge. Yet, we can say that unlike Barhoumi et al. (2010) who find that small data size produces competitive forecasts, we see that using a more disaggregated data set for forecasting may be beneficial. Recalling the discussion about the possible benefits of using BIC3 in the selection of the number of factors, we see that extracting factors with BIC3 indeed produces relatively more successful forecasts.

Rank	Forecast Equation Specification	Number of Static Factor Selection Method	Data Set Size and Maximum Number of Factors	Evaluation Period: Jan 2010-Sept 2011	Forecast Equation Specification	Number of Static Factor Selection Method	Data Set Size	Evaluation Period: Oct. 2011-Sept. 2013
1	DI	PC2	Large/9	0.94	DI-AR Lag	IC2	Medium/7	0.86
2	DI	BIC3	Large/9	0.96	DI-AR Lag	IC1	Medium/7	0.89
3	DI-AR Lag	BIC3	Large/9	0.96	DI-AR Lag	PC1	Medium/7	0.89
4	DI	PC1	Medium/7	0.97	DI-AR Lag	PC2	Medium/7	0.89
5	DI	IC3	Large/9	0.97	DI-AR Lag	BIC3	Large/9	0.90
6	DI-AR Lag	IC1	Large/9	0.98	DI-AR Lag	PC1	Small/4	0.90
7	DI-AR Lag	IC2	Large/9	0.98	DI-AR Lag	PC2	Small/4	0.90
8	DI	IC1	Large/9	0.98	DI-AR Lag	PC3	Small/4	0.90
9	DI	IC2	Large/9	0.98	DI-AR Lag	IC1	Small/4	0.90
10	DI	PC1	Large/9	0.98	DI-AR Lag	IC3	Small/4	0.90
11	DI	PC1	Small/4	0.99	DI-AR Lag	IC1	Large/9	0.91
12	DI	PC2	Small/4	0.99	DI-AR Lag	IC2	Large/9	0.91
13	DI	PC3	Small/4	0.99	DI-AR Lag	IC2	Small/4	0.92
14	DI	IC1	Small/4	0.99	DI	IC2	Medium/7	0.95
15	DI	IC2	Small/4	0.99	DI-AR Lag	PC2	Large/9	0.95
16	DI	IC3	Small/4	0.99	DI-AR Lag	BIC3	Medium/7	0.97
17	DI	PC2	Medium/7	0.99	DI	IC1	Medium/7	0.97
18	DI	PC3	Large/9	0.99	DI	BIC3	Medium/7	0.97
19	DI-AR Lag	PC1	Medium/7	0.99	DI-AR Lag	PC1	Large/9	0.98
20	DI	PC3	Medium/7	1.00	DI	PC3	Medium/7	0.99
21	DI	IC3	Medium/7	1.00	DI	IC3	Medium/7	0.99
22	DI	IC2	Medium/7	1.00	DI	BIC3	Small/4	1.01
23	DI	BIC3	Medium/7	1.00	DI-AR Lag	PC3	Medium/7	1.01
24	DI	BIC3	Small/4	1.01	DI-AR Lag	IC3	Medium/7	1.01
25	DI	IC1	Medium/7	1.02	DI	PC2	Large/9	1.02
26	DI-AR Lag	IC1	Medium/7	1.03	DI	PC1	Large/9	1.02
27	DI-AR Lag	IC2	Medium/7	1.03	DI	PC1	Medium/7	1.03
28	DI-AR Lag	BIC3	Medium/7	1.03	DI	PC2	Medium/7	1.03
29	DI-AR Lag	PC2	Medium/7	1.04	DI	IC1	Large/9	1.03
30	DI-AR Lag	PC1	Small/4	1.06	DI	IC2	Large/9	1.03
31	DI-AR Lag	PC2	Small/4	1.06	DI	BIC3	Large/9	1.03
32	DI-AR Lag	PC3	Small/4	1.06	DI-AR Lag	BIC3	Small/4	1.03
33	DI-AR Lag	IC1	Small/4	1.06	DI	IC3	Large/9	1.03
34	DI-AR Lag	IC2	Small/4	1.06	DI	PC3	Large/9	1.04
35	DI-AR Lag	IC3	Small/4	1.06	DI	PC1	Small/4	1.07
36	DI-AR Lag	BIC3	Small/4	1.07	DI	PC2	Small/4	1.07
37	DI-AR Lag	PC3	Medium/7	1.09	DI	PC3	Small/4	1.07
38	DI-AR Lag	IC3	Medium/7	1.09	DI	IC1	Small/4	1.07
39	DI-AR Lag	PC2	Large/9	1.11	DI	IC3	Small/4	1.07
40	DI-AR Lag	IC3	Large/9	1.34	DI	IC2	Small/4	1.08
41	DI-AR Lag	PC1	Large/9	1.34	DI-AR Lag	IC3	Large/9	1.09
42	DI-AR Lag	PC3	Large/9	1.39	DI-AR Lag	PC3	Large/9	1.11

Table 4.2 Relative RMSE for 3 Month Ahead Forecasts.

	Forecast Equation Specification	Number of Static Factor Selection Method	Data Set Size and Maximum Number of Factors	Evaluation Period: Jan 2010-Sept 2011	Forecast Equation Specification	Number of Static Factor Selection Method	Data Set Size	Evaluation Period: Oct. 2011-Sept. 2013
1	DI-AR Lag	PC3	Medium/7	0.63	DI-AR Lag	BIC3	Medium/7	0.76
2	DI-AR Lag	IC3	Medium/7	0.63	DI-AR Lag	IC1	Large/9	0.78
3	DI	PC1	Small/4	0.70	DI-AR Lag	IC2	Large/9	0.78
4	DI	PC2	Small/4	0.70	DI-AR Lag	IC2	Medium/7	0.83
5	DI	PC3	Small/4	0.70	DI	BIC3	Medium/7	0.83
6	DI	IC1	Small/4	0.70	DI	IC1	Large/9	0.83
7	DI	IC2	Small/4	0.70	DI	IC2	Large/9	0.83
8	DI	IC3	Small/4	0.70	DI-AR Lag	BIC3	Large/9	0.84
9	DI-AR Lag	PC1	Small/4	0.74	DI-AR Lag	BIC3	Small/4	0.85
10	DI-AR Lag	PC2	Small/4	0.74	DI	IC2	Medium/7	0.85
11	DI-AR Lag	PC3	Small/4	0.74	DI	PC2	Large/9	0.89
12	DI-AR Lag	IC1	Small/4	0.74	DI	BIC3	Small/4	0.91
13	DI-AR Lag	IC2	Small/4	0.74	DI	BIC3	Large/9	0.97
14	DI-AR Lag	IC3	Small/4	0.74	DI-AR Lag	IC1	Medium/7	1.06
15	DI	PC3	Medium/7	0.78	DI	IC1	Medium/7	1.07
16	DI	IC3	Medium/7	0.78	DI	IC2	Small/4	1.08
17	DI-AR Lag	PC1	Medium/7	0.81	DI-AR Lag	IC2	Small/4	1.12
18	DI-AR Lag	PC2	Medium/7	0.82	DI-AR Lag	PC1	Medium/7	1.12
19	DI	PC1	Medium/7	0.84	DI-AR Lag	PC2	Medium/7	1.12
20	DI	PC2	Medium/7	0.84	DI	PC1	Medium/7	1.25
21	DI-AR Lag	PC2	Large/9	0.86	DI	PC2	Medium/7	1.25
22	DI	BIC3	Large/9	0.87	DI	PC1	Small/4	1.26
23	DI	IC1	Medium/7	0.89	DI	PC2	Small/4	1.26
24	DI-AR Lag	BIC3	Large/9	0.91	DI	PC3	Small/4	1.26
25	DI	PC2	Large/9	0.93	DI	IC1	Small/4	1.26
26	DI-AR Lag	IC1	Medium/7	0.94	DI	IC3	Small/4	1.26
27	DI	BIC3	Small/4	0.94	DI	PC1	Large/9	1.51
28	DI	IC2	Medium/7	0.95	DI	IC3	Large/9	1.80
29	DI	BIC3	Medium/7	0.95	DI	PC3	Large/9	1.81
30	DI-AR Lag	IC2	Medium/7	0.95	DI-AR Lag	PC1	Small/4	1.92
31	DI-AR Lag	BIC3	Medium/7	0.95	DI-AR Lag	PC2	Small/4	1.92
32	DI-AR Lag	BIC3	Small/4	0.96	DI-AR Lag	PC3	Small/4	1.92
33	DI	PC1	Large/9	0.96	DI-AR Lag	IC1	Small/4	1.92
34	DI	IC1	Large/9	0.97	DI-AR Lag	IC3	Small/4	1.92
35	DI	IC2	Large/9	0.97	DI	PC3	Medium/7	2.43
36	DI-AR Lag	IC1	Large/9	0.98	DI	IC3	Medium/7	2.43
37	DI-AR Lag	IC2	Large/9	0.98	DI-AR Lag	PC2	Large/9	3.13
38	DI	IC3	Large/9	1.03	DI-AR Lag	PC1	Large/9	3.33
39	DI-AR Lag	IC3	Large/9	1.21	DI-AR Lag	PC3	Large/9	3.34
40	DI-AR Lag	PC3	Large/9	1.22	DI-AR Lag	IC3	Large/9	3.36
41	DI	PC3	Large/9	1.28	DI-AR Lag	PC3	Medium/7	3.89
42	DI-AR Lag	PC1	Large/9	1.47	DI-AR Lag	IC3	Medium/7	3.89

Table 4.3 Relative RMSE for 12 Month Ahead Forecasts.

Next, we evaluate the 12 month-ahead forecasts, generally considered to be representative of the longer run. Our first observation about the 3 month-ahead forecasts, namely factor model specification may play a significant role on the verdict about relative performance, holds for 12 month-ahead horizon as well. As an example, in the second evaluation period, factor model beats the benchmark when the number of factors is selected with BIC3, medium data set is used to extract factors and DI-AR Lag type equation is used for obtaining the forecasts. Using same equation type with the number of factors decided by IC3, however, we could get substantially worse forecasts than those derived using the benchmark. These findings are related to the discussion about the efficiency of forecasts. We see that it is not enough to check the forecasting power of the available indicators; attention also needs to be paid to the model specification. For the 12 month-ahead forecasts, the medium data set tops the list in both samples. Regarding the choice about DI-AR Lag or DI, we obtain best forecasts with the DI-AR Lag type forecasting equation.

Finally, we concentrate on the information that we get from splitting the evaluation sample using the results for both forecast horizons. Many papers in the forecasting literature report out-of-sample forecasting performance for a single given sample. We evaluate our models in two different samples to see how stable our conclusions are. We see that the relative performance and best model specification changes considerably amongst two samples. For example, the best model in the first evaluation sample for the 3 month-ahead forecasts ranks 25th and performs worse than the benchmark. Similarly, the best model for the 12 month-ahead horizon for the second evaluation sample had a relatively poorer performance than the first evaluation sample.

Hence, if we had written this paper in 2012, or if we had concentrated only on the recent part of the sample, our conclusions would have been different. Of course, changing performance of the models in different samples may mean that our methodology is not stable and should not be used in practice. We do not interpret our finding in this way. Rather, we stress that due to the nature of the shocks and the changing source of the growth in the economy, such as private consumption or exports, different indicators may be more important for forecasting in different periods. Hence, rather than keeping a model fixed, we advise to adopt a more dynamic approach to evaluate the performance of the models at each period in the forecasting practice.

5. CONCLUSION

In this paper we analyze how the performance of factor models change with different specifications. We use industrial production growth trends for the Turkish economy as our subject, as well as various indicators from different areas of the domestic and international economy. We find that level of disaggregation in the data and the number of factors extracted from this set may play a significant role on the relative performance of the models. Moreover, different sample choices for the forecast evaluation may produce different rankings for different models.

APPENDIX

	Data (Abbreviations Used in the Table A.5 and Table A.6 are in Parentheses)	Source
1	Industrial Production (IP)	TURKSTAT
2	Export Quantity Index (QX)	TURKSTAT, Author's Calculation
3	Import Quantity Index (QM)	TURKSTAT, Author's Calculation
4	Istanbul Stock Exchange-30	Istanbul Stock Exchange
5	Business Tendency Survey (BTS)- Assessment of General Situation	CBRT
6	Capacity Utilization	CBRT
7	CNBC-e Consumer Confidence Index (CCI)	CNBC-e
8	Inflation (CPI)	TURKSTAT, Author's Calculation
9	Euro/Dollar Parity	CBRT
10	Dollar Exchange Rate	CBRT
11	TL Deposit Interest Rate	CBRT
12	Dollar Deposit Interest Rate	CBRT
13	TL Commercial Credit Interest Rate	CBRT
14	Euro Commercial Credit Interest Rate	CBRT
15	TL Consumer Credit Interest Rate	CBRT
16	Benchmark Interest Rate	CBRT
17	EU-Industrial Production (EU_IP)	EUROSTAT
18	EU Consumer Confidence (EU_CCI)	EUROSTAT
19	EU-Business Confidence (ESI_EU)	EUROSTAT
20	Commodity Price Index	INDEXMUNDI
21	VIX	YAHOO
22	SP 500	YAHOO

Table A.4 Small Data Set

1	IP_Intermediate	*32*	ESI_EU_Industry
2	IP_Durable	*33*	ESI_EU_Services
3	IP_Nondurable	*34*	ESI_EU_Construction
4	IP_Energy	*35*	ESI_EU_Retail
5	IP_Capital	*36*	ESI_EU_Building
6	QM_Investment	*37*	EU_CCI
7	QM_Intermediate	*38*	Euro
8	QM_Consumption	*39*	Yen
9	QX_Investment	*40*	Dollar
10	QX_Consumption	*41*	Interest Rate_deposit_One month_Euro
11	QX_Intermediate (excl. Gold)	*42*	Interest Rate_deposit_Euro
12	CNBCE CCI-Q1	*43*	Interest Rate_deposit_TL
13	CNBCE CCI-Q2	*44*	Interest Rate_deposit_Dollar
14	CNBCE CCI-Q3	*45*	Interest Rate_credit_cash_TL
15	CNBCE CCI-Q4	*46*	Interest Rate_credit_car_TL
16	CNBCE CCI-Q5	*47*	Interest Rate_credit_housing_TL
17	CPI-Clothing and Footwear	*48*	Interest Rate_credit_commerical_TL
18	CPI-Housing	*49*	Interest Rate_credit_commerical_Euro
19	CPI-Household equipment	*50*	Interest Rate_credit_commerical_Dollar
20	CPI-Health	*51*	Interest Rate_overnight
21	CPI-Transportation	*52*	Interest Rate_benchmark
22	CPI-Communications	*53*	Commodity Agricultural Raw Materials Price Index
23	CPI-Recreation	*54*	Commodity Beverage Price Index
24	CPI-Education	*55*	Commodity Fuel (energy) Index
25	Cpı-Hotels and restaurants	*56*	Commodity Food Price Index
26	CPI-Miscalleneous	*57*	Commodity Industrial Inputs Price Index
27	EU_IP_Intermediate	*58*	Commodity Non-Fuel Price Index
28	EU_IP_Energy	*59*	VIX
29	EU_IP_Capital	*60*	Istanbul Stock Exchange
30	EU_IP_Durable	*61*	BTS-Assesment of General Situation
31	EU_IP_Nondurable	*62*	Capacity Utilization
		63	SP500

Table A.5 Medium Data Set

1	IP_Mining	*56*	QX_Chemical	*111*	ESI_EU_Building
2	IP_Food	*57*	QX_Rubber and Plastic	*112*	EU_CCI_Q1
3	IP_Beverages	*58*	QX_Other Mineral	*113*	EU_CCI_Q2
4	IP_Tobacco	*59*	QX_Basic Metal	*114*	EU_CCI_Q3
5	IP_Textile	*60*	QX_Fabricated Metal	*115*	EU_CCI_Q4
6	IP_Apparel	*61*	QX_Machinery and Equipment	*116*	EU_CCI_Q5
7	IP_Leather	*62*	QX_Electrical Equipment	*117*	EU_CCI_Q6
8	IP_Wood	*63*	QX_Communication	*118*	EU_CCI_Q7
9	IP_Paper	*64*	QX_Motor Vehicles	*119*	EU_CCI_Q8
10	IP_Printing	*65*	QX_Furniture	*120*	EU_CCI_Q9
11	IP_Refined petroleum	*66*	CCF_Q1	*121*	EU_CCI_Q10
12	IP_Chemical	*67*	CCF_Q2	*122*	EU_CCI_Q11
13	IP_Pharmaceutical	*68*	CCF_Q3	*123*	EU_CCI_Q12
14	IP_Rubber and plastic	*69*	CCF_Q4	*124*	FX_Australian
15	IP_Other mineral	*70*	CCF_Q5	*125*	FX_Canadian
16	IP_Basic Metal	*71*	CPI-Clothing and Footwear	*126*	FX_Euro
17	IP_Fabricated Metal	*72*	CPI-Housing	*127*	FX_Japanese Yen
18	IP_Computer, Electronic	*73*	CPI-Household equipment	*128*	FX_Norwegian Krone
19	IP_Electrical Equipment	*74*	CPI-Health	*129*	FX_Dollar
20	IP_Machinery and Equipment	*75*	CPI-Transportation	*130*	Interest_deposit_1 month_Euro
21	IP_Motor Vehicles	*76*	CPI-Communications	*131*	Interest_deposit_3 month_Euro
22	IP_Other Transportation	*77*	CPI-Recreation	*132*	Interest_deposit_6 month_Euro
23	IP_Furniture	*78*	CPI-Education	*133*	Interest_deposit_12 month_Euro
24	IP_Other Production	*79*	Cpı-Hotels and restaurants	*134*	Interest_deposit_12 month+_Euro
25	IP_Installation of Machinery and Eq.	*80*	CPI-Miscalleneous	*135*	Interest_deposit_1 month_TL
26	IP_Electricity, Gas and Air Cond.	*81*	EU_IP_Mining	*136*	Interest_deposit_3 month_TL
27	QM_Agriculture	*82*	EU_IP_Food	*137*	Interest_deposit_6 month_TL
28	QM_Mining	*83*	EU_IP_Beverages	*138*	Interest_deposit_12 month_TL
29	QM_Food	*84*	EU_IP_Tobacco	*139*	Interest_deposit_12 month+_TL
30	QM_Tobacco	*85*	EU_IP_Textile	*140*	Interest_deposit_1 month_Dollar
31	QM_Textile	*86*	EU_IP_Apparel	*141*	Interest_deposit_3 month_Dollar
32	QM_Apparel	*87*	EU_IP_Leather	*142*	Interest_deposit_6 month_Dollar
33	QM_Leather	*88*	EU_IP_Wood	*143*	Interest_deposit_12 month_Dollar
34	QM_Wood	*89*	EU_IP_Paper	*144*	Interest_deposit_12 month+_Dollar
35	QM_Paper	*90*	EU_IP_Printing	*145*	Interest_credit_cash_TL
36	QM_Refined petroleum	*91*	EU_IP_Refined Petroleum	*146*	Interest_credit_car_TL
37	QM_Chemical	*92*	EU_IP_Chemical	*147*	Interest_credit_housing_TL
38	QM_Rubber and plastic	*93*	EU_IP_Pharmaceutical	*148*	Interest_credit_commercial_TL
39	QM_Other mineral	*94*	EU_IP_Rubber and Plastic	*149*	Interest_credit_commercial_Euro
40	QM_Basic Metal	*95*	EU_IP_Other mineral	*150*	Interest_credit_commercial_Dollar
41	QM_Fabricated Metal	*96*	EU_IP_Basic Metal	*151*	Interest_Overnight
42	QM_Machinery and Equipment	*97*	EU_IP_Fabricated Metal	*152*	Interest_Benchmark
43	QM_Office Equipment	*98*	EU_IP_Computer, optical	*153*	Commodity Agricultural Raw Materials Index
44	QM_Electrical Equipment	*99*	EU_IP_Electrical Equipment	*154*	Commodity Beverage Price Index,
45	QM_Communication Eqipment	*100*	EU_IP_Machinery and Equip.	*155*	Crude Oil (petroleum), Price index
46	QM_Motor vehicles	*101*	EU_IP_Motor Vehicles	*156*	Aluminum, 99.5% minimum purity
47	QX_Agriculture	*102*	EU_IP_Other Transport	*157*	Copper, grade A cathode,US Dollars per Metric Ton
48	QX_Mining	*103*	EU_IP_Furniture	*158*	Gold (UK), 99.5% fine, average of daily rates
49	QX_Food	*104*	EU_IP_Other Manufacturing	*159*	Lead, 99.97% pure,US Dollars per Metric Ton
50	QX_Tobacco	*105*	EU_IP_Installation of Machinery	*160*	Nickel, melting grade, US Dollars per Metric Ton
51	QX_Textile	*106*	EU_IP_Electricity, gas, air cond.	*161*	Silver (Handy & Harman), 99.9% grade refined
52	QX_Apparel	*107*	ESI_EU_Industry	*162*	Zinc, high grade 98% pure, US Dollars
53	QX_Wood	*108*	ESI_EU_Services	*163*	VIX
54	QX_Paper	*109*	ESI_EU_Construction	*164*	Istanbul Stock Exchange-30
55	QX_Refined Petroleum	*110*	ESI_EU_Retail	*165*	BTS-Assesment of General Situation
				166	Capacity Utilization
				167	SP500

Table A.6 Large Data Set

REFERENCES

Altug, S. and E. Ulucevíz (2013). Identifying Leading Indicators of Real Activity and Inflation for Turkey, 1988-2010: A Pseudo Out-of-Sample Forecasting Approach. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 2, 1-37.

Angelini, E., M. Banbura and G. Rünstler (2010). Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, 1, 1-22.

Bai, J. and S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica*, 70, 191-221.

Barhoumi, K., O. Darne and L. Ferrara (2010). Are Disaggregate Data Useful for Factor Analysis in Forecasting French GDP? *Journal of Forecasting*, 29, 132-144.

Barhoumi, K., O. Darne and L. Ferrara (2013). Testing the Number of Factors: An Empirical Assessment for a Forecasting Purpose. *Oxford Bulletin of Economics and Statistics*, 75, 64-79.

Boivin, J. and S. Ng (2005). Understanding and Comparing Factor Based Forecasts. *International Journal of Central Banking*, 1, 117-151.

Boivin, J. and S. Ng (2006). Are More Data Always Better for Factor Analysis? *Journal of Econometrics*, 132, 169-194.

Eickmeier, S. and C. Ziegler (2008). How Successful are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach. *Journal of Forecasting*, 27, 237-265.

Gupta, R. and A. Kabundi (2011). A Large Factor Model for Forecasting Macroeconomic Variables in South Africa. *International Journal of Forecasting*, 27, 1076-1088.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth and C. Van Nieuwenhuyze (2009). Short-term Forecasting of GDP Using Large Datasets: A Pseudo Real-time Forecast Evaluation Exercise. *Journal of Forecasting*, 28, 595-611.

Stock, J. and M. Watson (2002a). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics*, 20 (2), 147-162.

Stock, J. and M. Watson (2002b). Forecasting with Principal Components from a Large Number of Predictors. *Journal of American Statistical Association*, 97, 1167-1179.

Stock, J. and M. Watson (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, 16, 788-829.