FORECASTING ECONOMIC AND FINANCIAL VARIABLES WITH FACTOR MODELS: THE CASE OF TURKEY

A THESIS SUBMITTED TO
THE INSTITUTE OF SOCIAL SCIENCES
OF
ANKARA YILDIRIM BEYAZIT UNIVERSITY

BY

MAHMUT GÜNAY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF ECONOMICS

FEBRUARY 2017

Approval of the Institute of Social Sciences

Yrd. Doç. Dr. Seyfullah Yıldırım

Manager of Institute

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy in the Department of Economics.

Prof. Dr. Erdal Tanas Karagöl
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Muhsin Kar
Supervisor

Examining Committee Members

Prof. Dr. Erdal Tanas Karagöl (YBU, Economics) ____________________________

Prof. Dr. Muhsin Kar (YBU, Economics) ____________________________

Doç. Dr. Mesut Murat Arslan (YBU, Economics) ________________________________

Doç. Dr. Şaban Nazlıoğlu (Pamukkale, Economics) ____________________________

Doç. Dr. Ahmet Şahbaz (Necmettin Erbakan, Economics) ______________________

I hereby declare that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work; otherwise I accept all legal responsibility.

Name, Last Name: Mahmut Günay
Signature:

ABSTRACT

FORECASTING ECONOMIC AND FINANCIAL VARIABLES WITH FACTOR MODELS: THE CASE OF TURKEY

Günay, Mahmut
Department of Economics
Supervisor: Prof. Dr. Muhsin Kar

February 2017, 161 pages

In this thesis, industrial production growth, core inflation and change in the stock market index are forecast using a large number of domestic and international indicators. Two methods are employed to deal with the curse of dimensionality problem stemming from the availability of ever growing data sets: factor models and forecast combination. Determining the best performing models requires a comprehensive analysis of the sensitivity of the forecast performance of factor models to various modelling choice. In this respect, effects of factor extraction method, number of factors, data aggregation level and forecast equation type on the forecasting performance are analyzed. Effect of using certain data blocks such as European Union variables and 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. Results show that best performing specifications change with the type of the variable that one wants to forecast, with forecast horizon and with the sample that is used to evaluate the out-of-sample forecasting performance. Factor models perform better than combination of the bi-variate forecasts.

Comparing models with alternative specifications shows that effect of modelling choices are not mutually independent. Hence, it is concluded that there is no “one size fits all approach” in forecasting with factor models. Thus, using a dynamic approach to continuously evaluate models from different dimensions is important in the forecasting process.

Key Words: Forecasting, Factor models, Principal components.

ÖZET

İKTİSADİ VE FİNANSAL DEĞİŞKENLER İÇİN FAKTÖR MODELLERİ İLE TAHMİNLER: TÜRKİYE ÖRNEĞİ

Günay, Mahmut
İktisat Bölümü
Danışman: Prof. Dr. Muhsin Kar

Şubat 2017, 161 Sayfa

Bu çalışmada, sanayi üretimi büyümesi, çekirdek enflasyon ve borsa endeksindeki değişim, yerel ve uluslararası alandan çok sayıda değişken kullanılarak tahmin edilmektedir. Bu amaç doğrultusunda, her geçen gün büyüyen veri setlerini işlemeye imkân veren iki yöntem kullanılmıştır; faktör modelleri ve bireysel tahminlerin birleştirilmesi. Faktör modelleriyle elde edilen tahminleri etkileyebilecek farklı boyutlar bulunmaktadır. Bu çerçevede, faktör elde etme yöntemi, faktör sayısı, verilerin hangi detayda kullanıldığı ve tahmin denkleminin faktör modellerinin tahmin performansına etkileri değerlendirilmektedir. Ayrıca, çeşitli veri kümelerinin, Avrupa Birliği’ne ilişkin göstergeler ve faiz oranları gibi, tahminler üzerindeki etkisi de incelenmiştir. Modellerin performansının zaman içindeki seyrinin istikrarını gözlemleyebilmek için tahminler iki ayrı dönem için değerlendirilmiştir. Bulgular, model tercihlerinin tahmin performansına etkilerinin birbirinden bağımsız olmadığına işaret etmektedir.

En iyi performans gösteren model seçimleri, tahmin edilmek istenilen değişkene ve tahminlerin değerlendirildiği döneme göre değişmektedir. Diğer yandan, faktör modelleri ile elde edilen tahminler, tek değişkenli modellerden elde edilen tahminlerin birleştirilmesine kıyasla daha az tahmin hatası yapmaktadır. Modellerin performanslarını etkileyen unsurların sürekli bir şekilde değerlendirilmesinin faydalı olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Öngörü, Faktör Modelleri, Temel Bileşenler

DEDICATION

To My Family

ACKNOWLEDGEMENTS

The author wishes to express his deepest gratitude to his supervisor Prof. Dr. Muhsin Kar and to Assist. Prof. Dr. Sıdıka Başçı who was his initial supervisor before her leave of absence.

The author would like to thank to Assoc. Prof. Mesut Murat Arslan and Prof. Dr. Ümit Özlale who served as the members of the Thesis Supervising Committee.

The author is grateful to his supervisors and to the members of the Committee for their guidance, advice, criticism, encouragements and insight throughout the research.

The author owes a debt of gratitude to his wife, Özlem Günay for her patience and emotional support as she took care of two sons while the author enjoys learning.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

FHLR: Forni, Hallin, Lippi and Reichlin (2005)'s Dynamic Principal Components Approach

RMSE: Root Mean Squared Error

SW: Stock and Watson (2002a)'s Principal Components Approach

CHAPTER I

1 INTRODUCTION

1.1 Scope of the Topic

Forecasting is a key ingredient of decision making process in many parts our lives. People forecast weather to plan their day. Commuters decide their routes based on their expectations about traffic intensity. Farmers pick the product that they will grow based on their predictions about the price of the alternative products. These forecasts are based on certain indicators that are thought to have forecast power. A blue sky may signal a low probability of rain, a weekend morning imply that roads will be open and high price on a product for this year may encourage a farmer to focus on that product for the next year. In these cases, part of the forecasting process rests on the judgement of the decision maker for picking up the indicator to use for forecasting based on his/her experience. In the meantime, advances in the technology help to produce forecasts in a more formal and systematic manner. Instead of looking at open skies to forecast whether it is going to rain or not, one can check the weather forecasts, which are based on data like information from satellites, to see the probability of the rain for each hour of the day. Instead of making wild guesses about the traffic conditions in certain routes, one can use navigation to see in which part of the city traffic is getting worse or running smoothly. A farmer may look at data to see whether there is a pattern in the weather conditions or whether there is a cycle in the prices of the products to make a better informed guess. In summary, forecasting is a key ingredient of decision making process and using advances in the technology may help to design a systematic forecasting process to reduce forecast errors.

Forecasting is also important for economic policy making and for designing investment strategies. Policy makers form expectations about growth and inflation, tax revenues and unemployment while investors try to predict the return of assets and probability of default. For these aims, incorporating tools brought by advances in the technology and basing forecasts on a more scientific and objective procedure may be beneficial as well. In a stochastic world with all sort of uncertainties, formal models enable one to set up an accountable forecasting process, test the forecasting power of potentially useful indicators or alternative models, identify sources of forecast errors and determine the possible ways to improve forecast performance based on these observations. As a matter of fact, there is a lively literature on the methods to improve forecasting performance for economic and financial variables.

There are different sort of challenges in the forecasting that the literature on forecasting tries to find solutions. To name just a few, there are instabilities in the economies. This can make it hard to predict the future with the estimated data generating process of the past. Another challenge is related to increasing interconnectedness of the world economies or increasing importance of the financial sector in modelling the real sector variables. So new transmission channels need to be incorporated to the forecast process. With the technological advances it gets easier to collect, store and process large amount of data. This brings another challenge to the forecasting process, namely incorporating available data to the forecasting process in a smart and efficient way.

1.2 Motivation of the Thesis

In this thesis, interest is in the performance of factor models that help one to deal with the challenge of forecasting in a data rich environment. In particular, forecasting performance of the factor models are evaluated for three variables from Turkey; industrial production as a real sector variable; core inflation as a variable from the price block; and a financial sector variable namely the stock market. For forecasting these variables, data from industrial production, export and import quantity indices, confidence indicators, exchange rates, interest rates, European Union variables, financial variables and commodity prices are utilized.

Factor model approach summarizes large data sets with few underlying factors and then use those factors for the desired aim from forecasting to impulse response analysis. In this approach, each series is thought to be composed of a common component related to a few factors and a part that is specific to the series. However, factors are latent variables so they have to be estimated. Estimation step involves several choices such as designing the data set, obtaining the factors and deciding the number of factors that will be used to summarize the data set. In forecasting applications, there are choices regarding the set-up of the forecast equation as well. Decisions in all of these steps may affect the forecasting performance. Moreover, effect of decisions may not be mutually independent. Literature on the analysis of factor models focus on some of these points by keeping other dimensions fixed. This thesis contributes to the literature by analyzing the effect of modelling decisions in the estimation of factors and forecasting steps in a more comprehensive way. For this aim, a pseudo out of sample forecasting analysis is conducted. Different factor models are constructed by changing model specifications and their forecast performances are evaluated. This practice enables one to see if he/she had used a given specification for forecasting in a period, how the models would perform. Upon determining the best performing models of the past, these models may be used to produce forecasts for future.

In the next section, it is discussed why the focus is on the factor models. In the section, contribution of the thesis is explained as well. Then, findings are summarized and in the last section topics in the chapters of the thesis are introduced.

Forecasts of key macro variables, such GDP and inflation, are vital ingredients of real time economic policy making. Considerable time and effort is devoted to producing forecasts, communicating them and assessing risks around those forecasts. Variables that will be forecast and forecast horizons change with the needs of the economic actor conducting forecasting. To fix ideas, consider why and how three institutions use forecasts: central banks, ministries of finance and investment banks. Effects of monetary policy decisions feed through the economy gradually. So, central banks produce forecasts for several variables such as inflation, GDP growth and unemployment rate in the monetary policy making process. In the budget making process, governments forecast economic growth for forecasting tax revenues. They may produce multi-year forecasts for growth and budget items. On the other side of the spectrum, investors need very short term forecasts. For instance daily forecasting of exchange rate movements may be used to design investment portfolio. These are just three examples showing the diversity of forecasting needs in the decision making process for policy makers and market players.

So, designing and optimizing forecasting procedures can be beneficial for a wide range actors.

In this thesis, forecasting and factor model approach are two main themes. In Figure 1.1, it is shown how factor models can be useful for forecasting. In this section following issues are discussed based on the logic in this diagram: why efficiently using information is as important as accuracy of forecasts and how increasing data availability increase the challenge of using information efficiently. After discussing these topics, contribution of the thesis is explained as well.

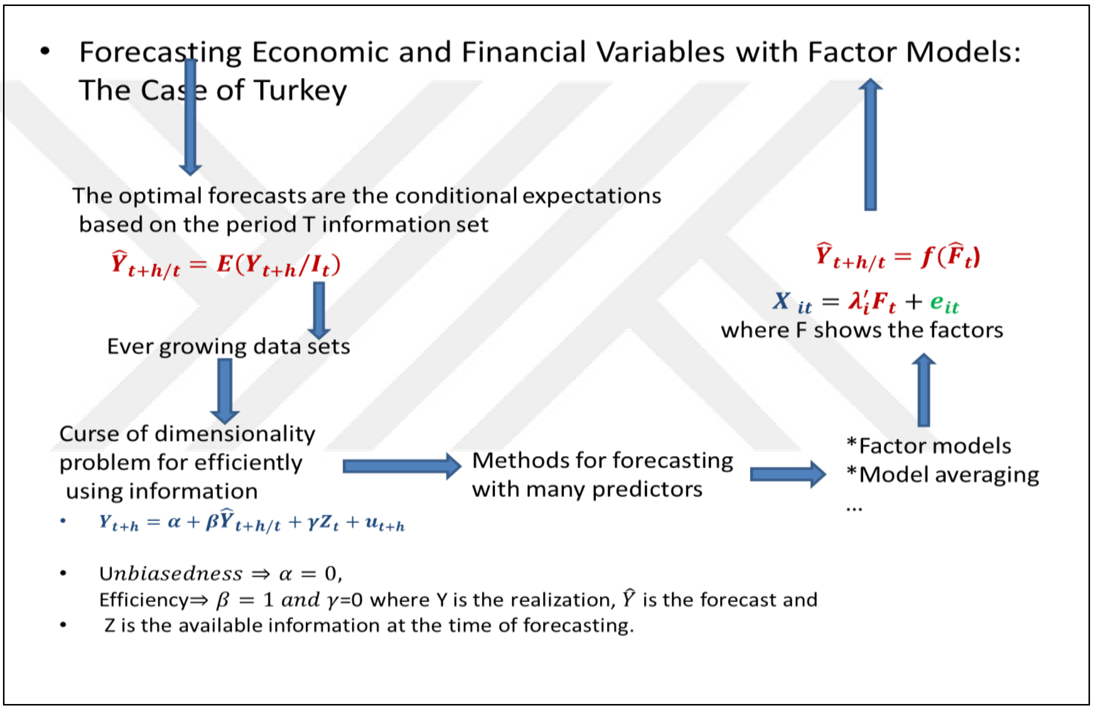


Figure 1.1. Forecasting and Factor Models

Source: Author's representation based on the literature

We are living in a stochastic world so forecasting comes with great challenges. There can be various events and shocks that affect the economic and financial variables which cannot be foreseen at the time of forecasting. So, in general realizations will be different from predictions and time to time by a high margin. Yet, forecasts, forecasters and forecast models are evaluated ex-post by their success and big forecast errors are criticized.

It would be unrealistic to expect zero forecast errors from a forecasting model/forecaster in a stochastic world. However, it is fair to expect that forecast errors should not be predictable with the information that would be available at the time of forecasting. This is due to the fact that forecastable errors imply inefficient use of resources. Inefficiency can occur as a result of 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. In this respect, efficiency of forecasts is as important as accuracy. Hence, it is important for forecasters to check whether all available and suitable information is utilized to the greatest extent possible and in an efficient way in the forecasting process.

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 information technology. Even if the aim is forecasting the local variables only, in addition to domestic indicators using international data may be necessary. However, one can use only a limited number of variables forecasting model estimated with OLS due to the degrees of freedom problem. Stock and Watson (2002a, page 147) 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.

1.3 Contribution of the Thesis

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 forecast equation (Stock and Watson, 2002a and 2002b). Factor models enable one to incorporate as many series as he/she wants in the forecasting process. But there may not be a linear relation between forecasting performance and the number of series used for extracting factors. Also, number of factors extracted from a given data set may affect forecasting performance. Using a few factors may be insufficient to summarize the information content of the dataset while using a lot of factors may increase the parameter uncertainty in the estimation of forecast equation. Moreover, combining forecasts from bivariate equations estimated with all of the available indicators is also an option. So, comparison of the forecast performance in the case of combining bivariate forecasts or combining information a la factor models is considered as well. Figure 1.2 illustrates these dimensions.

Analyzing the effect of modelling decisions in factor models on forecast performance may provide valuable information for the forecasters. There are papers that try to understand the effect of these dimensions on the forecasting performance. In this thesis, it is aimed to analyze the sensitivity of forecasting performance of factor models to the modelling choices for Turkish economic and financial data.

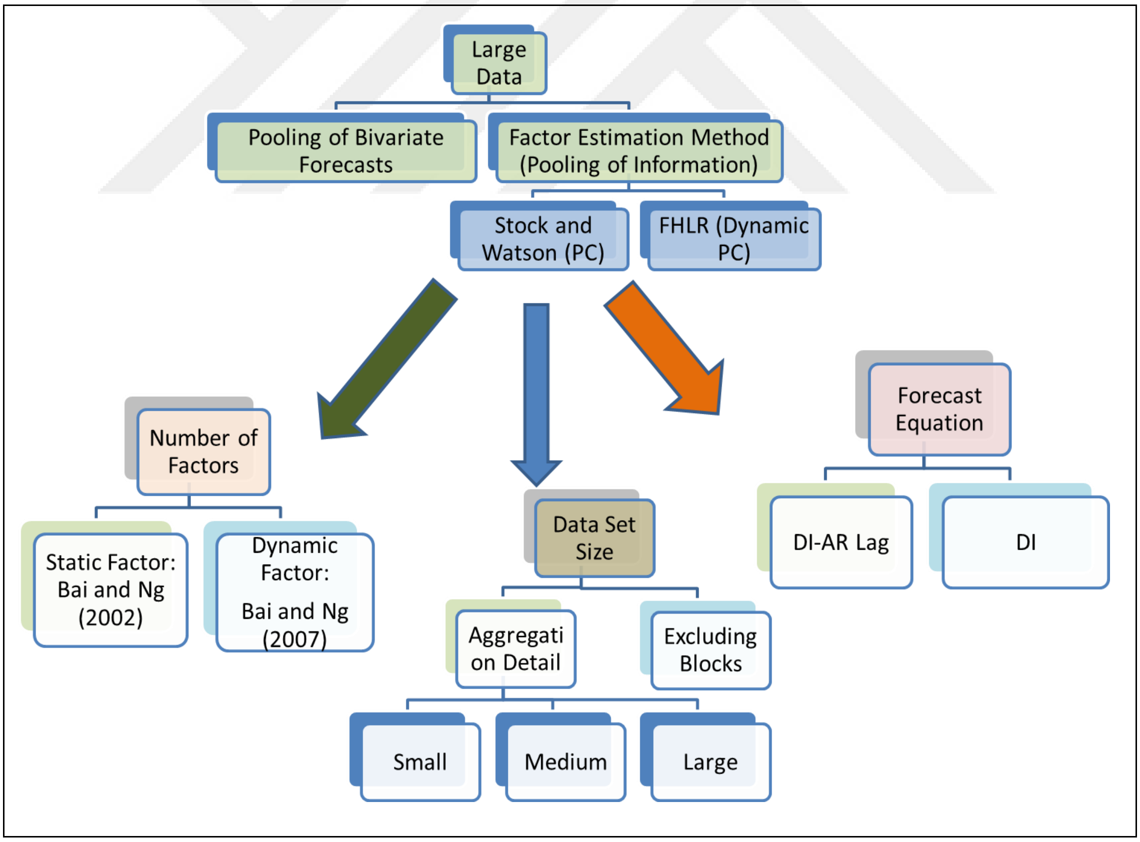


Figure 1.2. Model Choices in Forecasting with Factor Models

Source: Author's representation based on the literature

Contribution of this thesis is to analyze the sensitivity of the forecast performance of factor models to the modelling choices in a relatively more comprehensive manner. In particular, 336 factor model specifications are evaluated. This is achieved by estimating factors with two different approaches, deciding the number of factors with seven different information criteria, using variables in different aggregation level so that collecting three different data sets, analyzing the effect of three data blocks namely European Union variables, financial indicators and interest rates by excluding these series from the master data sets and finally using two different forecast equation. Apart from factor model specifications, forecasts from combining bivariate forecasts are considered as well. In addition to effect of model specifications, considering different type of target variables enables one to understand the forecast performance of the factor models for different sort of variables. Last but not least, computational system designed to see the most successful forecasting models for Turkish economy for the empirical sections of this thesis can be run regularly to produce forecasts from the best performing models of the recent past.

1.4 Summary of the Main Findings

In this section, main points emerging from the pseudo out of sample forecasting exercises are summarized. From a broad perspective, performance of the specifications changes over time and depending on the target variable. There are important specific points though that are worth highlighting.

i. Factor extraction approach does not play an important role on the forecast performance. Two alternative approaches are considered for estimating the factors. In one of the approaches, one only needs to choose number of static factors while in the other approach one needs to choose number of static and dynamics factors and also two more parameters need to be chosen for frequency domain analysis. This increase in the choice regarding auxiliary parameters may wipe out the benefit of using a more complicated approach to factor extraction. Findings show that using a simple factor extraction approach may be preferred in practical applications as difference in the forecast performance is in general marginal.

ii. 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. Indeed, using one or two factors from larger data sets results in competitive forecasts. Moreover, most of the worst performing specifications use a high number of factors. Hence, similar to the factor extraction approach, using a few factors and hence obtaining a relatively more parsimonious forecast equation helps to obtain better forecast performance.

iii. Forecast equation 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 and the dependent variable, using a few factors helps to improve forecast performance. 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 and stock market picture is less clear.

iv. Factor model approach allows one to use a large number of variables in the data sets. However, composition of the data set plays a crucial role on the forecast performance and more data is not always better. Using a high number of variables from a wide range of data blocks may affect forecast performance negatively. In particular,

   a. Excluding European Union variables harms the forecast performance for industrial production while for core inflation and stock market enlarging the data set with these variables increases the forecast errors. For industrial production, forecasts with factors estimated from a data set by excluding interest rates or financial variables make less error than factors from data sets that use series from these blocks as well. So, depending on the target variable composing data sets with indicators from different blocks may be beneficial.

   b. Using disaggregated data does not necessarily improve the forecast performance. 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.

In summary, our findings offer important guides about the sensitivity of factor models for forecasting purposes.

1.5 Structure of the Thesis

In Chapter 2, the literature on factor models is reviewed. In the first part of the chapter, a general overview of the use of factor models in different areas of the economics is presented. Starting from early examples of factor models, it is discussed how factor models evolved to the current use in forecasting. This part is expected to fix ideas about the handiness of factor models as a strong technique for dealing with big data for answering questions from a wide range of fields, be it monetary policy or international business cycle synchronization. In the sixth section of the chapter, focus is on applications of forecasting with factor models. In the seventh section, pooling of bivariate forecasts is discussed and in the last section literature on forecasting Turkish economy variables is reviewed. Papers about each required input in the process of obtaining forecasts with factor models are discussed. This strategy enables one to construct a road map for the methodology used in this thesis. Also, applied papers give hints about possible effect of each of the input on forecast performance.

In Chapter 3, methodology used in the thesis is introduced. Forecasting with factor models requires several inputs: extracting factors, deciding the number of static and dynamic factors and setting up a forecast equation. In the chapter, an analytical summary of the methods proposed for obtaining or deciding these inputs is presented.

In Chapter 4, forecast environment for factor models for Turkish economy is presented. In particular following topics are discussed: three data sets that are used to extract factors, factors that are obtained from these sets, number of factors suggested by different information criteria for these data sets, how the forecast equation is set up and how the forecast evaluation is made.

In the next three chapters, from 5 to 7, results for forecasting industrial production growth, core inflation and stock market growth, respectively, are discussed. These chapters are organized in five sections. After making an overview in the first section, in the second section effect of model specifications for factor models is analyzed. This is divided into four sub-sections. In these subsections role of factor extraction approach, number of factors, data set size and forecast equation on the forecasting performance are evaluated. In the next section, an alternative for dealing with large data sets namely pooling of bivariate forecasts is considered. In the fourth section, focus is on the effect of data blocks on forecast performance. In particular, the effect of European Union data, financial and commodity variables and interest rates on the forecasting results are studied. These three sections use graphs to present findings. In the last section of these three chapters, using tables the best and worst five specifications, out of 340, are presented for three, six, nine and twelve month ahead forecasts. Last chapter concludes our thesis.

CHAPTER II

2 LITERATURE ON THE APPLICATIONS OF FACTOR MODELS

2.1 Introduction

Information content of ever growing data sets, both in terms of time series and cross-section dimensions, cannot be extracted efficiently without using appropriate techniques. Factor model approach serves as a handy tool to utilize large amount of data relatively easily and analytically. In the next chapter, technical dimension of factor models are discussed in detail. In this chapter, literature on the development and use of the factor models is reviewed.

Factor models have been used in psychology since the beginning of the 20th century. First use of the factor analysis was based on the observation that students' grades for different subjects were correlated. It was postulated that a latent variable, intelligence, is the main deriver of those grades. Later in the century other disciplines, such as finance, started to use factor analysis. But its use in economics had been limited. This is due to the fact that some of the assumptions used to obtain factors are fairly restrictive for most of the economic time-series problems. However, in the last two decades there has been significant progress on the techniques for estimating factors in a theoretically consistent way for answering the problems that economists face. In this chapter, review of the papers will show the seminal contributions that make it possible to utilize factor models more widely.

In a nutshell, one can express factor model approach using the matrix representation given by the Equation 2.1. In this equation, X is the matrix of indicators used to extract factors, F shows the factors, Λ shows the loadinds. $F\Lambda'$ is the common component and e is the part that is not explained by the factors so it is the idiosyncratic component. Depending on the assumptions about the relation of factors and the data one may have static or dynamic factor model and depending on the assumption about the idiosyncratic component, one may have exact or approximate factor models.

$$X = F\Lambda' + \mathrm{e} \tag{2.1}$$

The challenge comes from the fact that variables on the right hand side of the Equation 2.1 are not observed. Hence it is necessary to estimate factors and factor loadings. As an example, Bai (2003) brings attention to the fundamental assumption of Arbitrage Pricing Theory where asset prices follow a factor structure. In this notation of Equation 2.1, X would be the matrix for return for asset i at time t and e would be the idiosyncratic returns. According to factor model approach, asset prices will be determined in part by the common factors affecting the economy such as GDP growth or weather conditions. Of course a common factor cannot explain all of the movement in each series. The part for each series that cannot be explained by those common factors are considered as specific to that asset, i.e. idiosyncratic.

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 and 2005). Those authors showed that one can use principal component type analysis to estimate the factors. This paved the way for handling large amount of data relatively easily. Researchers contributed by developing theory for consistent estimation of the factors for data 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.

In order to provide a more concrete idea about factor models, first of all applications in several areas are reviewed. Papers reviewed in these sections show the seminal contributions in the application of factor models and/or they are informative examples regarding the idea of factor model approach. Then, literature on the forecasting with factor models is discussed with reference to the applications of this thesis. Finally, some papers that deal with forecasting variables for Turkish economy are discussed. Each area is covered in brief by summarizing the research methodology of the selected papers in this area.

2.2 Asset Pricing Models

The first example of this section is Arbitrage Pricing Theory (APT). Roll and Ross (1980) is one of the first attempts to test this theory. Their model decompose the asset returns into two parts: the part that is explained by the common factors and part that is noise or an unsystematic risk component. They ask the natural question of what is a common factor in this case (page 1077). In this case, they expect these factors to be related to the fundamental economic aggregates such as GNP, interest rates or weather conditions. Ross (1976) and Roll and Ross (1980) are the seminal works explicitly taking into account the factor structure. However, strict factor models, which assume that idiosyncratic terms are independent, are used in these applications. As Connor and Korajczyk (1995) state, this may be a too severe restriction when the number of factors is substantially less than the number of variables. They note that strict factor model representation of Ross's diversification argument is sufficient but not necessary. In this respect, Chamberlain (1983) and Chamberlain and Rothschild (1983) are important contributions as they relax the assumptions of Ross (1976). In particular, the assumption that variance-covariance matrix of the errors is diagonal is relaxed which brings us the world of approximate factor models.

Chamberlain and Rothschild (1983, page 1282) have two aims. The first aim is to study a market with many assets that does not necessarily follow a factor structure. The second aim is to define an *approximate factor structure* (emphasis is original in the cited paper). As they highlight, this is a weaker concept than the standard strict factor structure. One of the key messages of the paper is that principal component analysis can be used to find the approximate factor structure. In this thesis, factors are obtained with principal component method and then they are used in the forecasting as in the seminal works of Stock and Watson (2002a and 2002b). Indeed, papers by Stock and Watson are extension of the Chamberlain and Rothschild (1983) by allowing serial correlation in the factor model structure. Hence, Ross (1976), Roll and Ross (1980) and Chamberlain and

Rothschild (1983) are essential building blocks of the methods of the approximate factor models that are used in this thesis.

2.3 Business Cycle Analysis

Forni and Reichlin (1998) use dynamic principal components to estimate the number of common shocks and then to get the factors. They use this framework to obtain the common factors from 4-digit industry level data from 1958 to 1986. Forni and Reichlin (1998, page 455) note that a static version of this framework has been proposed in the financial literature such as Chamberlain and Rothschild (1983). They use spectral analysis to be able to work in a dynamic set-up. Authors show that in the limit, variance of the average of idiosyncratic component defined as in the paper will go to zero. Forni and Reichlin (1998) note that this observation has two implications. First when the cross-section is large, sectoral averages can be used to determine the dimension of the common shocks. Second, common shocks can be estimated by q cross-sectional averages where q is the dimension of the shocks. In summary, this work extends the static form of Chamberlain and Rothschild (1983). One of the methods that is used in this thesis is the dynamic principal components approach of Forni et al. (2005). Forni and Reichlin (1998) is a pioneering work for the method developed and introduced in Forni et al. (2005).

Factor models are used in the international business cycle analysis as well. Eickmeier (2007) uses factors in a structural VAR to investigate the transmission of the US shocks to the German economy. Eickmeier (2007) finds that supply shock raises output and lowers prices and interest rates in both the US and the Germany while demand shock increases those three variables in both countries.

Final paper that attention is brought to in this sub-section is the Kose et al. (2008). They use a Bayesian dynamic factor model to estimate the common and country specific factors for the main macro variables such as GDP and consumption. They get factors for the G-7 aggregates and also for each country in this group. They find that variance explained by the G-7 factor for output and consumption increased substantially in the globalization period, from 1986 to 2003, relative to Bretton-Woods period of 1960 to

1972. So, they conclude that business cycle synchronization across the G-7 countries increased during the globalization period.

2.4 Monetary Policy Analysis

In the canonical VAR models, effect of monetary policy is analyzed using three variables: interest rate, inflation and the output growth. However using a limited information set may cause a number of problems such as omitted variable bias. Factor models can be very useful for monetary policy analysis since they enable incorporating more information in a VAR model and also by enabling using latent variables. Another problem from using a small-scale VAR model is that one can observe impulse responses only for the included variables. For example, for the canonical model, one can see the response of inflation or output growth to the monetary policy surprise. However, there are other variables that one may be interested in seeing the response to a monetary policy shock. Bernanke et al. (2005) use factors from a large data set and augment the VAR model with these factors. Hence they call their approach as Factor Augmented VAR (FAVAR). When they add one factor to the model, results become consistent with the theory.

Favero et al. (2005) is another paper that uses factor models to study monetary policy. These authors also note that omitting part of the information set that is used to make policy becomes a bigger problem when expectations of the variables are also in the policy reaction function. Authors use factors as instruments in the GMM estimations in addition to the lags of the output gap, inflation and commodity price index. In this case, standard error of the estimates decline which indicates that using factors as instruments reduce the uncertainty in the estimations.

Final paper that is reviewed is Baumeister et al. (2010). In this paper, authors use the FAVAR but extend the analysis by allowing time-varying coefficients and stochastic volatility of the shocks. Moreover, authors also focus on the disaggregated prices for understanding the “price puzzle”. They find that prize puzzle is observed for some sectors in the short run even if it is absent in the aggregate data.

2.5 Monitoring of the Economy

Factor models are used to produce indices for tracking the developments in the economy such as business cycle conditions index and financial conditions index. Some key examples of these indices are presented.

Chicago FED publishes an index for the state of the US economy (FED, 2001). The index is named as Chicago FED National Activity Indicator (CFNAI). Index is obtained as a weighted average of 85 monthly indicators. This index is inspired from the Stock and Watson (1999). In that paper, Stock and Watson (1999) show that first component from 85 series forecast inflation relatively successfully. CFNAI is the weighted average of these 85 series obtained by calculating the first principal components. The index is published monthly and it is normalized. So, if the index is equal to zero, this means the economy is close to its trend growth rate. After calculating the index for a sufficiently long period of time, one can find thresholds for better interpreting the index. For example, analysis of the CFNAI shows that in the last seven recessions, index was below -0.7.

Center for Economic Policy Research (CEPR) publishes an index for monitoring the developments in the state of the economy in the euro area. They name their index as "Eurocoin" (Altissiomo, (2010)). While CFNAI uses principal components analysis to get the common factors, Eurocoin is based on generalized principal components. Medium to long run growth is obtained by removing fluctuations shorter than one year. The methodology to estimate the medium to long run variance is from Forni et al. (2000).

Constructing indices is not only done for tracking the economic conditions. Hatzius et al. (2010) is an example of using factor models for producing a financial conditions index. They find that their methodology is better at forecasting compared to some other alternative financial condition indexes. Angelopoulou et al. (2014) is another example of financial conditions index using factor models. They use data such as interest rates, credit developments and bank surveys. It is found that financial conditions in the euro area differed from each other before and after the 2008 crisis.

2.6 Forecasting with Factor Models

Factor model approach serves as a handy tool to utilize large amount of data relatively easily and analytically. Hence, it is not surprising that they are especially popular for forecasting and business cycle analysis purposes at central banks and policy institutions where hundreds of variables are monitored. Indeed, authors of many papers that are going to be reviewed in this section are affiliated with central banks and policy institutions. First of all, papers that analyze factor models for forecasting key macroeconomic variables are presented. In the second part, papers that work on forecasting Turkish economy are discussed.

Before producing forecasts, it is necessary to define the forecast environment. In particular, one needs to provide following inputs (i to iv) to get a forecast (see Figure 1.2 for a systematic presentation):

i. Factors,
ii. Number of factors,
iii. The data set that factors are extracted from
iv. Type of the forecast equation.
v. Pooling of forecasts or pooling of information

Moreover ex-ante it is not clear whether using factors as a summary of the information in a large data set improves over pooling forecasts from bivariate equations using the variables in this data set. It may be informative, and indeed may be necessary; to compare the forecasting performance of forecast combination from individual indicators with factor models. Hence the bullet *v* above is added to the list. Since there is a huge literature on the forecasting with factor models, rather than aiming to make an encyclopedic review of the literature, key papers covering above topics are surveyed. After reviewing the papers, implications of them for this thesis are discussed.

2.6.1 Extracting factors

Forecasting with factor models requires extracting factors that summarize information in a data set. In this thesis, Stock and Watson (SW)'s static principal components approach and Forni, Hallin, Lippi and Reichlin (FHLR)'s dynamic principal components approach to estimate the factors are used. There are other methods for obtaining factors such as the one proposed by Kapetanios and Marcellino (2009). However, as the literature review by Eickmeier and Ziegler (2008) shows, SW and FHLR approaches are more commonly used in the forecasting literature. Hence, forecasting performance of these two mainstream methods are compared and contrasted. This section starts with the papers by Stock and Watson. Then, Forni et al. (2005) is reviewed. After covering the main papers of the two approaches, three papers that compare the forecasting performance of the SW and FHLR approaches are reviewed.

Stock and Watson (1999) forecast US inflation for one year ahead. They use simulated out of sample methodology to study the forecasting performance of the models that they test. They define their procedure as follows (page 302): consider forecasting the inflation rate from January 1980 to January 1981 in January 1980. All models are estimated, information criteria are computed and lag lengths are selected. From this model a forecast is obtained. Then, moving the information set by one month, all of the models are re-estimated, information criteria are computed and models are selected. From this model, forecast is obtained for the respective period. Authors find that some indicators outperform unemployment. They use a wide range of indicators from different data blocks for testing the forecasting power of 167 indicators. Relative performance of those indicators changes over time. For example, National Association of Purchasing Managers' new orders index results in lower forecast error in the first evaluation sample while it performs worse in the second evaluation sample. In the paper forecasts from a large data set are considered as well. They do this in two ways. In the first method, they combine bivariate forecasts. In the second method, they construct indices from the large data set. In this method, authors estimate factors from principal components using data up to time t for forecasting the inflation rate at t+12. They select the number of factors and lag length using BIC. It is found that factor models produce the best forecasts.

In two papers that were published in 2002, Stock and Watson formalized their approach to factor model forecasting. In Stock and Watson (2002a) entitled as *forecasting using principal components from a large data set*, authors show that feasible forecasts using principal components as factors are asymptotically efficient. Moreover, factors are consistently estimated even in slight time variation. They use both Monte Carlo simulations and empirical examples to study the proposed methodology. This paper is discussed in more detail in Chapter 3. Stock and Watson (2002b) is more empirically oriented. So in this chapter, key aspects of this paper are covered in more detail.

Stock and Watson (2002b) focus on h-step ahead forecasts using principal components as the factors. They note that in the case of multistep ahead forecasting, there are at least two approaches. First option is to use an iterative scheme, where one builds a VAR model and iterates this model. Second approach is direct approach. Stock and Watson (2002b) use direct approach for obtaining forecasts.

Since factors are unobserved, they are estimated by principal components. Authors study forecasting performance of the factor models for eight macroeconomic variables. Four of these are related to real economic activity and remaining four variables are about the prices. Stock and Watson (2002b) compute forecasts for each series for six, twelve and twenty-four month-ahead forecasting horizons. They use direct forecasting approach. In other words, they estimate separate equations for each forecast horizon in the following form (page 149):

$$\hat{y}^{h}_{T+h/T} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}'_{hj}\, \widehat{F}_{T-j+1} + \sum_{j=1}^{p} \hat{\gamma}_{hj} y_{T-j} \tag{2.2}$$

This is the most general form for a diffusion index (DI) model. Three versions of forecast equation are used to get forecasts:

a. DI-AR Lag: this type of equation includes lags of the factors and the lags of y_t. BIC is used to determine k, m and p.

b. DI-AR: this form of the forecast equation uses the contemporaneous values of the factors and lags of y_t.

c. DI: this speciation uses only the contemporaneous factors.

Theory used in the paper assumes that the data set that is used to extract factors are composed of stationary series. Hence, each of the 215 series that is used in the paper is transformed to ensure stationarity. Authors use an autoregressive model and multivariate leading indicators forecasts as the benchmark. They use simulated out-of-sample forecasting design to compare performance of the models.

Results show that relative performance of the factor models changes with the forecast equation, target variable and forecast horizon. For example, for twelve month-ahead forecasts, DI produces the lowest forecast error for industrial production, while DI-AR Lag outperforms other two specifications for personal income. For six month-ahead forecasts, DI-AR Lag produces marginally lower forecast errors for industrial production. Relative performance in the case of price variables also changes depending on the specification. DI type equation performs substantially worse than the other. This suggests that using lags of the price variables help reducing forecast errors substantially while for real variables they can even harm the forecast performance. Regarding the relative performance of DI-AR Lag and DI-AR, there is no clear winner for six and twelve month-ahead forecasts.

In summary, these papers are important contributions in formalizing how to do empirical application of factor models in the spirit of SW approach. Stock and Watson (1999) show the potential of factor models for reducing forecast errors by applying the method to inflation forecasts. Stock and Watson (2002a) develop the theoretical dimension of using principal components for factor extraction. Finally, Stock and Watson (2002b) offers a systematic way for the empirical methodology.

In this thesis, key aspects of these papers are applied. For example, recursive out of sample forecasting exercise is used to compare forecasting models. Factors are estimated, lag lengths and number of factors are determined at each iteration of the recursive exercise. Performance of combining forecasts from bivariate models is compared with that of the factor models. Additionally, DI-AR Lag and DI type forecast equations are used for testing the relative performance of factor models for forecasting inflation, industrial production and stock market. There are departures from some of the practices of these papers though. First of all, in the 2000s researchers come up with ideas for determining number of factors more formally. Also, it is shown that more data may not always be better for factor analysis. Hence, effect of data set structure and the number of factors extracted from this data set on forecast performance are evaluated as well.

Next paper that is going to be summarized is Forni et al. (2005) where they use dynamic principal components approach to factor extraction. This is the paper that formalizes the FHLR approach. After reviewing this paper, studies that compare the effect of factor extraction method on the forecast performance are discussed.

Forni et al. (2005) extend their previous work, Forni et al. (2000), on the factor models. In 2000 paper, authors use spectral density of the data which means a two-sided filter is used to get the factors. To be more clear, factors are linear combinations of past, present and future observations. For ex-post analysis, using a two-sided filter may not pose a significant limit on the use of the factors. But in the case of forecasting, this presents itself as a major problem. Forni et al. (2005) note that while SW's static principal component approach solves this issue by using the contemporaneous values to estimate the factors, they may lose valuable information, such as lead-lag relations in the data. Forni et al. (2005) aim to combine these two approaches. They explain their two-step methodology as follows:

i. Obtain common and idiosyncratic variance-covariance matrices at all leads and lags as inverse Fourier transformations of the corresponding estimated spectral density matrices.
ii. Use contemporaneous combinations of these estimates.
iii. h-step ahead projections are obtained as onto the obtained factors.

As Forni et al. (2005) emphasize, while both SW and FHLR approaches are based on one-sided linear combination, their weighting schemes are different. Relative performance is expected to improve by using a more sophisticated technique when the cross-sectional items differ significantly in the lag structure of factor loadings.

These two methods attract considerable interest of the researchers. Researchers compared the performance of these two methods for different countries and variables. Next, three papers that focus on comparing performance of these two methods are reviewed. First paper that is analyzed is Boivin and Ng (2005). In this paper, authors are interested in the effect of two modelling dimensions on forecast performance: effect of factor extraction method and the forecast equation method. They define the factor extraction step as step E (with SW and FHLR being the choices) and the obtaining forecast as step F (direct, iterative, unconstrained and nonparametric forecasts). They use simulations to understand the effect of different type of error structure on the forecast performance. Their main conclusion is that taking step F as given, choice of step E does not play a significant role on the forecast performance. However, preferences on the step F affect the conclusions about step E. Regarding the factor extraction methods, this paper shows that SW approach gives competitive forecasts. Since it is easier to implement, in day to day use it may be preferred over dynamic approach.

Second paper that is highlighted in this section is by Schumacher (2007) who compares alternative factor models for forecasting German GDP. He uses 124 quarterly series such as GDP, industrial production, labor market indicators, prices, interest rates, spreads, and survey data. Author transforms data to get stationary series. He uses three types of factor models. Two of the methods are same as the ones used in this thesis, namely SW and FHLR approaches. Third method uses subspace algorithm. Schumacher (2007) uses direct forecasting approach. His forecast evaluation is based on both recursive and rolling schemes. Number of factors is selected with certain information criteria and a performance based system is considered as well. His findings are as follows. FHLR and subspace algorithm approaches perform better than SW approach but choice of the auxiliary parameters may affect the forecast performance significantly. Since, in an out of sample forecasting context these parameters are unknown, forecast performance evaluations based on the past performance may not always be a good guide for future. Another finding is that criteria for the number of static factors proposed by Bai and Ng (2002) play non-negligible role on the forecasting performance. In summary, effect of factor extraction method on the forecast performance depends on the modelling choices such as choosing auxiliary parameters, number of factors and forecast equation. General principles of this paper are followed in this thesis and effects of different modelling choices on the forecast performance are evaluated.

Third paper that is examined regarding the effect of factor extraction method on forecasting performance is the paper by D'Agostino and Giannone (2012). This paper compares the forecasting performance of SW and FHLR approaches for the US economy. Authors use industrial production and CPI as the target variables and conduct a simulated recursive out-of-sample forecasting exercise. They find that until 1985, factor models beat autoregressive model. However after 1985, gain from using factor models disappears. For example, for twelve month-ahead forecasts, towards the end of the evaluation sample, AR model beats the factor models by a large margin, especially for the CPI. Authors note that poor performance relative to the benchmark does not only apply to factor models for this sample. Other forecasts, both model based and institutional, show a similar pattern. Comparing the performance of factor extraction approaches, before 1985, SW approach performs better than FHLR while after 1985 reverse is in general true. Despite differences for some periods, factor forecasts are highly collinear.

This paper is specifically interested in comparing SW and FHLR approaches 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. Additionally, number of factors may play an important role on the relative forecast performance. This thesis will shed more light on these issues for an emerging market economy data for different types of target variables. After reviewing these three papers comparing the relative performance of factor extraction approaches, papers studying the effect of the number of factors on the forecast performance are reviewed.

2.6.2 Number of Factors

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 considerably. 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. In this respect, this paper is discussed in more detail.

Barhoumi et al. (2013) analyze the effect of proposed methods for selecting the number of factors on the forecasting performance. For the number of static factors, they consider methods proposed by Bai and Ng (2002) and Alessi et al. (2010). For the number of dynamic factors they test several criteria. They consider criteria proposed by Stock and Watson (2005) and Amengual and Watson (2007). These two papers are the modifications of the Bai and Ng (2002) criteria for the number of static factors. Third method for the number of dynamic factors is due to Bai and Ng (2007) who estimate a VAR with the static factors and then take the spectral density of the residuals of this VAR. Breitung and Pigorsch (2013) propose another methodology that uses the correlation analysis of the principal components. Hallin and Liska (2007)'s methodology for the number of dynamic factors uses the eigenvalues of the spectral density matrix of the observations.

Barhoumi et al. (2013) analyze the effect of choosing the number of factors with one of the above information criteria for the French and German GDP growth for one to four quarter-ahead. Authors obtain factors from SW and FHLR approaches. Their results can be summarized as follows:

i. Estimated over five-year moving windows, number of factors change over time. For example, with Bai and Ng (2002) criterion for 1993-1998, number of static factors is estimated to be 4 and 1 for France and Germany, respectively. These figures change over time and for 2003-2008, number of factors is estimated to be 3 and 4 for France and Germany, respectively (page 72). Similar observations also hold for number of dynamic factors.
ii. Factor models perform better than an AR model for one quarter ahead forecasts but perform poorer than an AR model for the longer horizons.
iii. Using lags of the dependent variable helps to reduce the forecast errors marginally for French GDP forecasts while it increases forecast errors for German GDP forecasts.
iv. Authors conclude that information criteria of Bai and Ng (2002) for the static factor models, and Bai and Ng (2007) and Breitung and Pigorsch (2013) for the dynamic models appear as the most robust overtime (page 77).

v. They stress that from a forecasting point of view, adding an ad hoc number of factors is not necessarily a good choice. So, using tests can help to reduce forecast errors.

vi. They have an important caveat at the end of the paper. Since the number of observations is low, spectral density matrix may not be estimated accurately. So, when the sample size is not long, static factor models are more robust.

In short, number of factors may have an important role on the forecasting performance. Selecting the number of factors with an information criterion may produce better forecast results. In this thesis, 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.

2.6.3 Data Set

Ability to process large amount of information is praised as a key feature of the factor models. One can extract factors from thousands of variables in just a second. This opens up an important question. Should one use as many series as one can collect to get "more accurate" factors and hence hopefully get better forecasts? In this section, some key papers in this area are reviewed. Reading of the literature shows that more data is not always better!

Boivin and Ng (2006) is a seminal work on the effect of number of series that is used to get factors. They start by noting that in the Bai and Ng (2002) it is shown that factors can be estimated with reasonable precision with as small as 40 series for iid errors. Then they postulate the question that "can it be undesirable to increase the number of variables beyond a certain limit." The intuition is that unlike in a survey design where the aim is to represent the population, in the factor extraction a researcher collects *a* data set. Since there is not a population to target, different researchers may come up with different factors and hence different forecasts. Indeed, in the case of macroeconomic data our variables come from broad categories such as industrial production or prices. Consider the case that one uses headline series from this set. As he/she add more series from a given category, possibility of correlated errors will increase. Then, there may be situation where the cross-correlation of the errors will be higher than the threshold assumed by the theory. Authors take the work of Stock and Watson (2002b) where around 150 series are used to extract the factors. Boivin and Ng (2006) show that pre-selecting as few as 40 variables from this set may yield lower forecast errors than using the larger data set.

Caggiano et al. (2011) work on this issue with the European data. They analyze data for euro area, six largest euro area countries (Germany, France, Italy, Spain, the Netherlands and Belgium) and the UK. They consider datasets in the range of 105-133 indicators (page 740). Different pre-selection criteria are used for reducing the dimension of these datasets. They find that pre-selection improves forecast performance. For example, they find that best performance is obtained when as few as 12 variables for UK and 22 variables for Italy are used to get the factors (page 749).

To wrap up, these two papers show that using more data is not always better for factor analysis. In this respect, an aim of this thesis is to analyze the effect of data set structure on the forecast performance. Rather than doing this as the above papers, a different strategy is used based on another branch of this literature. Next; these papers are reviewed.

Data sets to extract factors are constructed from main data blocks such as production or prices. For example, consider the seminal work of Stock and Watson (2002b). They consider 215 (149 in the case of balanced panel) series from fourteen blocks. In the production block, they use total industrial production and subgroups such as durable and nondurable consumer goods. For other groups, a similar strategy is observed as well. But just as Boivin and Ng (2006) explain, adding subcategories from a group may cause high cross-correlation between errors. Barhoumi et al. (2010) start from this point and analyze whether using aggregate or disaggregate data affects the forecast performance of the factor models in a systematic way. They put up three different data sets; small, medium and large. Small data set includes a total of 20 series. In the medium dataset they use disaggregated version of soft data so that number of variables increase to 51. For the large dataset, they use sectoral disaggregation. For example industrial sector is divided into consumer goods, equipment goods, intermediate goods, agri-food goods and car industry.

Authors also look at the different factor extraction techniques. In addition to the static factor model of SW and dynamic factor model of FHLR that are analyzed in this thesis, they use other versions of dynamic factor models. In particular, they consider two-step approach of Doz et al. (2011) based on Kalman filtering and also quasi maximum likelihood approach of Doz et al. (2012). Their results suggest that forecasts from small data set and factors extracted with SW approach are not statistically worse than other methods and other datasets.

Methodology of Barhoumi et al. (2010) is followed in this thesis and three different data sets are used. These data sets are constructed by increasing the disaggregation level of the variables. In this approach focus is on the effect of disaggregation on forecasting performance. From another angle, one may be interested in the effect of different data blocks on forecasting performance. For example, whether soft data such as surveys or financial variables help to forecast key macroeconomic variables may be informative. Next; papers that are reviewed focus on this point.

Forni et al. (2003) analyze forecasting performance of SW and FHLR approaches using 447 monthly variables for euro area. These variables are constructed using six groups as financial variables, monetary indicators, industrial production, prices, surveys and others. From these variables, they build 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. Results of the paper points out that:

i. In general SW and FHLR approaches perform better than an AR model for one and three month ahead forecasts.
ii. Best results are not always obtained by the largest dataset.
iii. For inflation, for both SW and FHLR approaches excluding financial variables cause deterioration in the forecast performance.
iv. 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) and Eickmeier and Ng (2011) analyze the role of international variables for forecasting factor models. Schumacher (2010) analyzes forecasting performance with and without international data. He finds that adding international variables does not reduce forecast errors of factor models. But, if a pre-selection is applied to the dataset though LARS-EN method, then there is improvement. Eickmeier and Ng (2011) consider alternative data-rich methods. Strategy is again constructing datasets with and without international variables. They conclude that results change depending on the data rich method and parameter choices within these methods.

In conclusion, more data may not always be better for factor analysis. In this respect, in this thesis effect of data set structure is analyzed by using three data sets that differ by the level of disaggregation; small, medium and large. Also, forecasts are obtained from data sets excluding some blocks, such as European Union variables, to understand the role of certain data blocks on forecast performance.

2.6.4 Type of Forecast Equation

Type of the forecasting equation refers to the case of using lags of the dependent variable or the factors. An example of this would be DI, DI-AR and DI-AR Lag of Stock and Watson (2002b) as discussed above. If one had enough data and estimate the parameters for a well specified model then one would be able to estimate coefficients pretty accurately. However, in real life applications data sets are limited in terms of time series observations. Hence, using lags of variables may harm forecasting performance by increasing parameter estimation uncertainty. On the other hand, for persistent series, such as core inflation, using lags of the dependent variables may improve forecasting performance. In this respect, it is also an empirical question to test the effect of using lags of the variables on forecast performance. Since this issue is touched on in the papers discussed in the previous sections, a new paper is not introduced.

2.7 Pooling Information or Pooling Forecasts

In the factor model approach, one summarizes information in a large data set with few factors. Then he/she uses those factors in the second stage forecast regression. An alternative way to use possibly very large amount of data in the forecasting would be using forecast combination. It is well established by now, following the seminal work of Bates and Granger (1969) that forecast combination improves forecasting performance. For example, Timmermann (2006) and Hendry and Clements (2004) are theoretical contributions on this subject. Stock and Watson (2003, 2004) are two empirical examples showing the benefit of combining individual forecasts. Next, these papers from Stock and Watson are discussed.

Stock and Watson (2003) analyze the role of asset prices in forecasting output and inflation. They use data from seven countries, obtain forecasts with direct forecasting approach and conduct pseudo out-of-sample forecasting to get a measure of the performance of the indicators. A wide range of indicators are considered such as interest rates, real and nominal exchange rate, and commodity prices. In addition to those asset prices, they use indicators from economic activity such as GDP and employment, from prices such as deflator and wages, and from monetary block. In a nutshell, it is found that some indicators are helpful at forecasting at certain periods for some countries while they perform poorly in other cases. Their tabulation shows that for inflation out of the 211 four-quarter ahead forecasts for seven countries, 25 percent was better than an AR model for 1971-1984. However, only 6 percent of those were still better than AR model for 1985-1999 while the remaining 18 percent performed worse than an AR model. They note however that in the literature it has been shown that forecast combination may outperform individual forecasts. Moreover, relative performance of forecast combination may be more stable than individual forecasts. Indeed, they find that using combination forecasts bring substantial improvement over the individual forecasts. This observation holds across different periods as well which indicates the stability of the forecast performance.

Stock and Watson (2004) analyze the forecast combination for seven countries from a different angle. Their focus is on the alternative forms of combining forecasts. An important message from this paper is that simple average of the individual forecasts outperforms more complicated methods. In this respect, in this thesis simple averages of the individual forecasts for each of the three data sets are considered as well. In particular, forecast performance of the mean of the forecasts from bivariate models from small, medium and large data sets are presented.

Naturally, there are examples of comparing factor models and pooling individual forecasts. For example, the first paper that is reviewed in this chapter, Stock and Watson (1999), is a seminal example of this practice. They consider forecasts of Phillips curve using 167 variables. In addition to those bivariate forecasts, they consider forecast combination of the forecasts of these series. Also, they obtain factors from these series and use the factors in the forecasting. They find that factor model approach outperforms individual models.

To conclude, there are various dimensions in the case of forecasting with factor models. Each dimension may have a different effect on the forecast performance. Moreover, different dimensions may not be mutually independent. For example, using lags of a few factors for a small data set may have different effects than using lags of a high number of factors from a large data set. In the literature, different papers study the effect of some of those issues on the forecast performance. In this thesis, the issue is approached in a more comprehensive manner. In particular, different dimensions of modelling for different type of target variables are considered.

2.8 Literature on Forecasting Turkish Economy Variables

Use of factor models for forecasting Turkish macroeconomic variables is rather 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 estimation methodology of Giannone et al. (2008) is utilized where the system is cast in state space and the factors are obtained. They determine the number of factors to extract from this system using Bai and Ng (2007). Authors use two types of forecasting equation, a single equation and a VAR model augmented with the factors.

Their results show that factor models perform poorly. In the case of single equation model, only for the two quarter ahead horizon factor model beats the benchmark random walk. FAVAR 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 three quarter-ahead forecasts there are much 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, as the papers covered in this chapter show both theoretically and empirically, it is not clear ex-ante whether dynamic approach is preferable for the short sample used by the authors. Indeed, there are various dimensions for forecasting with factor models which Öğünç et al. (2013) do not take into account. Hence, this thesis contributes to the literature by analyzing the sensitivity of the forecasting performance of factor models for inflation for Turkish economy. Soybilgen (2015) thesis, which is not publicly available at the time this thesis is written, 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 use of the factor models for real sector variables. Günay (2015) reports some of the findings of this thesis. 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 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.

There are studies that forecast the variables that are considered in this thesis with other methods. For example, Değerli (2012) analyzes the forecasting performance of VAR models for forecasting industrial production. He finds that using combination of VAR models bring improvement over the benchmark. Altug and Uluceviz (2013) analyze forecasting models for industrial production and inflation for Turkish economy. They estimate single equation models with direct forecasting approach for multi-period ahead forecasts. Results indicate that performances of the indicators are not stable. Literature on forecasting stock market return for a long horizon is limited. There are various work focusing on modelling the volatility of the returns and forecasting daily returns.

In summary, literature on forecasting with factor models is limited for Turkish economy. There are papers using other methods. A recurring conclusion is that performances of the models are not stable and forecast combination may help to reduce forecast errors more consistently.

CHAPTER III

3 METHODOLOGY

3.1 Factor Models

This chapter is devoted to the technical discussion. Technical aspect of factor models is presented following the expositions in Bai and Ng (2008), Stock and Watson (2002a), Bai (2003) and Barhoumi et al. (2014).

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,…, N, t = 1,…,T, a static factor model is defined as

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

$$= C_{it} + e_{it}. \tag{3.2}$$

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. Assumptions about factors and idioscyratic terms change with the factor model approach. For example, strict factor model approach assumes that idiosyncratic terms are cross-sectionally independent. The challenge comes from the fact that variables on the right hand side of the Equation 3.1 are not observed. Hence it is necessary to estimate factors and factor loadings. An example from Bai (2003) for the use of the factor models is cited to give some more insight on the idea of expressing a series as the sum of a common component and a series-specific part. Bai (2003) brings attention the fundamental assumption of Arbitrage Pricing Theory where asset prices follow a factor structure. In the notation of Equation 3.1, x_{it} would be the return for asset *i* at time *t*, F_t is a vector of factor returns, e_{it} are the idiosyncratic returns.

According to factor model approach, asset prices will be determined in part by the common factors affecting the economy such as GDP growth or weather conditions. The part for each series that cannot be explained by those common factors are considered as specific to that asset, i.e. idiosyncratic.

Following Breitung and Eickmeier (2006) and Bai and Ng (2008), factor model representation is expressed using matrices in the following way. Let $X_t = (x_{1t}, x_{2t}, \dots, x_{Nt})$' be the observed data. In an r-factor model each element of the X_t can be expressed as in Equation 3.3 (the example is for the first variable in X_t):

$$\begin{pmatrix} x_{11} \\ x_{12} \\ \dots \\ x_{1t} \end{pmatrix} = \begin{pmatrix} \lambda_1 f_{11} + \lambda_2 f_{21} + \cdots + \lambda_r f_{r1} + e_{11} \\ \lambda_1 f_{12} + \lambda_2 f_{22} + \cdots + \lambda_r f_{r2} + e_{12} \\ \dots \\ \lambda_1 f_{1t} + \lambda_2 f_{2t} + \cdots + \lambda_r f_{rt} + e_{1t} \end{pmatrix} = \left[\begin{pmatrix} \lambda_1 & \dots & \lambda_r \end{pmatrix} \begin{pmatrix} f_{11} & \dots f_{1t} \\ f_{21} & \dots f_{2t} \\ & \dots \\ f_{r1} & \dots f_{rt} \end{pmatrix} \right]' + \begin{pmatrix} e_{11} \\ e_{12} \\ \dots \\ e_{1t} \end{pmatrix} \tag{3.3}$$

Relation given in the above equation given can be expressed as,

$$x_{it} = \lambda_{i1} f_{1t} + \cdots + \lambda_{r1} f_{rt} + e_{it}, \text{ t=1,…,T.} \tag{3.4}$$

$$= \lambda_{i\cdot}' f_t + e_{it} \tag{3.5}$$

where f_t is a vector of r common factors, $e_{it}$are the idiosyncratic components and $\lambda_{i\cdot}' = [\lambda_{i1}, \dots, \lambda_{ir}]$. When all of the elements in X_t are taken into account, following matrix representation can be used:

$$\begin{bmatrix} x_{11} & \dots & x_{N1} \\ x_{12} & \dots & x_{N2} \\ & \dots & \\ & \dots & \\ x_{1t-1} & \dots & x_{Nt-1} \\ x_{1t} & \dots & x_{Nt} \end{bmatrix} = \begin{bmatrix} F_{11} & \dots & F_{r1} \\ F_{12} & \dots & F_{r2} \\ & \dots & \\ & \dots & \\ F_{1t-1} & \dots & F_{rt-1} \\ F_{1t} & \dots & F_{rt} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1N} \\ \lambda_{21} & \dots & \lambda_{2N} \\ & \dots & \\ \lambda_{r1} & \dots & \lambda_{rN} \end{bmatrix} + \begin{bmatrix} e_{11} & \dots & e_{N1} \\ e_{12} & \dots & e_{N2} \\ & \dots & \\ e_{1t} & \dots & e_{Nt} \end{bmatrix} \tag{3.6}$$

So, defining $X = (x_1', \dots, x_N')$ as a TxN matrix of observations, matrix representation of the factor model can be expressed as

$$X = F\Lambda' + \text{e} \tag{3.7}$$

Here, $X\ is\ (T\ x\ N), F\ is\ (T\ x\ r), \Lambda' is\ (r\ x\ N)\ and\ e = (e_1', e_2', \dots, e_N')\ is\ (T\ x\ N)$.

In the above equation, general form of factor model representation is shown but depending on the assumptions on F and e, one deals with different types of factor models. In particular, conditional on the dynamics of F, one can have a *static* or a *dynamic* factor model whereas assumptions about *e* will make the model an *exact factor model* or *an approximate factor model*. Bai and Ng (2008) note that although above representation implies a static relation between x_{it} and F_t, F_t can follow a dynamic process that evolves according to

$$A(L)F_t = u_t \tag{3.8}$$

where A(L) is a polynomial, which can be infinite order, of the lag operator. In other words, one can have a static relation between *x* and *F* while *F* may be following a dynamic process. This can be achieved, for example, by stacking the factors along with their lags in a matrix. Dynamic factor model representation can be shown formally as follows:

$$x_{it} = \lambda_i'(L)f_t + e_{it} \tag{3.9}$$

where $\lambda_i(L) = (1 - \lambda_{i1}L - \cdots - \lambda_{i1}L^s)$ shows the dynamic factor loadings of order s. Here the factors are assumed to evolve according to

$$f_t = C(L)\varepsilon_t \tag{3.10}$$

where ε_t are iid errors.

Bai and Ng (2008) combine above expressions and reach the representation shown in Equation 3.11. They note that in the literature, $q = \dim(\varepsilon_t)$ is referred as the number of dynamic factors.

$$x_{it} = \lambda_i'(L)C(L)\varepsilon_t + e_{it} \tag{3.11}$$

When factor models are classified in terms of the assumptions on the errors, two categories can be defined:

i. exact factor models (also named as strict factor models)
ii. approximate factor models

In the case of exact factor models one assumes that idiosyncratic errors are independent while in the case of approximate factor models correlation, in cross-section and time-series dimension, is allowed to a certain extent. 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 the economic data. This is due to the fact that in general economic time series data are 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. Keeping these concerns in mind, first of all results for the classical factor analysis are derived mathematically since it will serve as a building block for our further analysis. Then it will be discussed how one can relax some of the restrictive assumptions.

Analysis starts by giving assumptions. Breitung and Eickmeier (2006, page 28) state that for the exact factor models it is assumed that

i. $E(e_t) = 0$

ii. $E(e_t e_t') = \Sigma = diag(\sigma_1^2, \dots, \sigma_N^2)$

iii. $E(f_t) = 0$

iv. $E(f_t f_t') = \Omega$.

It is further assumed that errors and the factors are uncorrelated with each other

v. $E(f_t e_t') = 0$

Note that assumptions $E(e_t) = 0$ and $E(f_t) = 0$ imply the assumption that $E(x_t) = 0$. This can be achieved by de-meaning the variables before doing the factor analysis. Also note that variance-covariance matrix of the idiosyncratic terms are assumed to be a diagonal matrix. As stated by Bai and Ng (2008, page 95) when the Σ is diagonal, one has a strict factor model. Naturally, loadings should be different from zero for having a relation between factors and the indicators.

Assumptions *i-v* are used to show how one can reduce the dimension of the data using factor model representation. Recalling that sample covariance matrix of X can be written as in the Equation 3.12, one can start by plugging-in the factor model representation for X and then do the matrix multiplication. Dimensions of the matrices are shown explicitly to make sure that matrices are conformable.

$$V = \frac{(X'X)}{N} \tag{3.12}$$

$$NV = (X'_{NxT} X_{TxN}) \tag{3.13}$$

$$= (\,F_{Txr}\Lambda'_{rxN} + \mathrm{e}_{TxN})'(F_{Txr}\Lambda'_{rxN} + \mathrm{e}_{TxN}) \tag{3.14}$$

$$= (e'_{NxT} + \Lambda_{Nxr}\mathrm{F}'_{rxT}) (F_{Txr}\Lambda'_{rxN} + \mathrm{e}_{TxN}) \quad (3.15)$$

$$= e'_{NxT}F_{Txr}\Lambda'_{rxN} + e'_{NxT}\mathrm{e}_{TxN} + \Lambda_{Nxr}\mathrm{F}'_{rxT}F_{Txr}\Lambda'_{rxN} + \Lambda_{Nxr}\mathrm{F}'_{rxT}\mathrm{e}_{TxN} \quad (3.16)$$

$$= 0 + N\Sigma + \Lambda\mathrm{N}\Omega\Lambda' + 0 \quad (3.17)$$

In the Equation 3.16, the first and the last terms are zero because factors and errors are assumed to be uncorrelated (assumption v above). Other terms follow from the definition of variance-covariance matrix for the idiosyncratic terms and the factors. Therefore, following expression is obtained:

$$V = \Sigma + \Lambda\Omega\Lambda' \quad (3.18)$$

Bai and Ng (2008) also assume $E(F_t F_t') = I_r$ which serves as an identification condition. This condition is necessary since the variance-covariance matrix is symmetric. So, number of parameters on the left hand side would be less than the unknowns on the right hand side. Therefore, one should impose some conditions to be able to identify the system. Intuitively, this assumption implies that factors are orthogonal to each other and the variances of each factor are normalized to unity. Using these assumptions one ends up with the following expression for the variance-covariance matrix of the data:

$$V = \Sigma + \Lambda\Lambda' \quad (3.19)$$

Note that factors do not appear in the above equation but loadings are used. However, one still does not observe the factors or the factors loadings. In the factor analysis approach, one obtains the factor loadings first and then find the estimated factors. This is done through an iterative algorithm. Since there are latent variables, one needs to use methods like Expectation Maximization. This approach comes with certain limitations for economic problems. In particular, one needs to have certain assumptions on the structure of the idiosyncratic errors and factors such as those coming from a normal distribution. Another limitation is that a large amount of series may not be used with this approach. This is due to the fact that as the number of variables increase, one will need to estimate more parameters. Convergence of the maximum likelihood estimation may be difficult and with the increasing size of the parameters it may be even impossible to estimate the factors. Hence, even though there are some applications of this approach in economics, in these thesis techniques that do not assume errors are iid from a normal distribution are utilized.

In conclusion, in the factor model approach each variable is decomposed into two components: common and idiosyncratic. When idiosyncratic terms are uncorrelated across time and cross-section one has the exact factor model. This assumption may be restrictive for economic time series data. When this assumption is relaxed, one enters to the world of approximate factor models. If the dynamics of the factors are modelled, it is called dynamic factor model approach. As Bai and Ng (2008, page 95) note, a dynamic factor model with q factors can be written as a static factor model with $r = q(s + 1) \geq q$ static factors where s is the lag length.

3.2 Approximate Factor Models for Large N

In this section methods for obtaining factors in the case of approximate factor models are discussed. Section starts by motivating why one needs to resort to approximate factor models. Discussion is based on Barhoumi et al. (2014, page 82).

i. Number of variables is often larger than the number of observations in economic data. For example, considering domestic and international data for production, trade, surveys and financial indicators one can collect hundreds of indicators. But, even if one uses monthly observations, number of time series dimension would be limited. Hence, unless one uses techniques to deal with large N, he/she will need to work in a constrained environment which may lead to loss of information.

ii. IID hypothesis and hypothesis on the diagonality of the variance-covariance matrix of the idiosyncratic components is too restrictive for economic data. For example, consider a researcher that uses detailed production data and he/she includes durable, non-durable, investment and intermediate goods production to his/her data set. Even if part of those series are explained with the factors and unexplained part is called as idiosyncratic, there may still be some correlation in those idiosyncratic parts. There may be various data groups, such as industrial production, prices and interest rates, in the data sets. Hence it is necessary to relax the assumption that there is no correlation across cross section and times dimension.

iii. Maximum Likelihood Estimation is challenging for large N as one needs to estimate a large number of parameters. In this case, a convergent solution may not even be found.

In their review of the dynamic factor models, Barhoumi et al. (2014, page 82) provide the evolution of the approximate factor models. They state that Chamberlain and Rothschild (1983) are the first to relax the assumptions of strict factor models. In particular, Chamberlain and Rothschild (1983) show that if N tends to infinity then principal component analysis is equivalent to factor analysis. In a series of papers, Connor and Korajczyk (1986, 1988, 1993) relax the assumption of Chamberlain and Rothschild (1983) that the variance-covariance matrix of the population is known. Forni et al. (2000, 2004) extend the approximate static factor models to approximate dynamic factor models. In particular, in these two works Forni et al. relax the assumptions of Brillinger (1981) that N is fixed and T tends to infinity.

In summary, Barhoumi et al. (2014) note that using approximate factor models over the strict factor models has a number of benefits for the economic time series data.

i. The idiosyncratic components can both be weakly mutually correlated and show little heteroscedasticity.
ii. It is possible to have a weak correlation between the factors (F_t) and the idiosyncratic components.

3.3 Estimation of Factors for Large N

In this section, techniques for estimating approximate static and dynamic factor models are presented. Sections starts by giving a general picture from the work of Schumacher (2007, pages 274-275). He analyzes forecasting performance of alternative factor models for German GDP growth. He explains how to get the factors in two mainstream methods in economics, namely Stock and Watson (2002a)'s principal component method and Forni et al. (2005)'s dynamic principal component analysis. In the Stock and Watson (2002a) (SW hereafter) methodology, factors can be estimated as

$$\hat{F}_t^{SW} = \hat{S}' X_t \tag{3.20}$$

where $\hat{S}_j$ corresponds to the r largest eigenvectors of the variance-covariance matrix of the data. So, factors are simply the eigenvectors times the data matrix. Forni et al. (2005)'s (hereafter FHLR) method ends up solving a generalized eigenvalue problem in the following formula.

$$\hat{\Gamma}_\chi(0)\hat{Z}_j = \hat{\mu}_j \hat{\Gamma}_\xi(0)\hat{Z}_j \tag{3.21}$$

where $\hat{Z} = (\hat{Z}_1, \dots, \hat{Z}_r)$ denotes the eigenvectors and solves the following generalized eigenvalue problem. Then factors can be obtained as

$$\hat{F}_t^{FHLR} = \hat{Z}' X_t \tag{3.22}$$

Note that above methods imply that cross-sectional average of the series is used. In their review of dynamic factor models, Stock and Watson (2011) present a discussion about this issue. They remark that (page 8 of the working paper version) cross-sectional averaging is nonparametric. For example, in the principal components approach one does not have a model for the factors or the idiosyncratic components. Indeed factors are treated as r-dimensional parameters to be estimated using N-dimensional data. Stock and Watson (2011) state that along the lines of Chamberlain and Rothschild (1983) weak assumptions are made. Conditions are that factors affect all of the series and factor loadings are heterogeneous. Another assumption is that correlation across idiosyncratic parts is limited. They note that there are many different cross-sectional weighting schemes that yield consistent estimator for the factors. After giving this broad view on the estimation of the factors, more detail on these techniques are presented in the following sections.

3.3.1 Estimating Factors with Principal Component Method

This section studies the approach that obtains factors with principal component method. Presentation is based on the Stock and Watson (2002a) and Bai and Ng (2008). In these papers, authors relax assumptions on the idiosyncratic terms being diagonal and show how one can estimate the factors using principal components. Stock and Watson (2002a) consider the following nonlinear least squares problem:

$$V(\tilde{F},\tilde{\Lambda}) = (NT)^{-1}\sum_i\sum_t(x_{it} - \tilde{\lambda}_i\,\tilde{F}_t)^2 \tag{3.23}$$

where $\tilde{F} = (\widetilde{F_1},\widetilde{F_2},\dots,\widetilde{F_T})'$ are the hypothetical values for the factors and $\tilde{\lambda}_i$ is the i^{th} row of the $\tilde{\Lambda}$ which are the hypothetical factor loadings. Let $\hat{F}$ and $\hat{\Lambda}$ denote the minimizers of $V(\tilde{F},\tilde{\Lambda})$. Stock and Watson (2002a) state that after concentrating out $\hat{F}$, minimizing the nonlinear least squares problem given above is equivalent to

$$maximize\ tr(\tilde{\Lambda}'X'X\tilde{\Lambda})$$

such that

$$\frac{\tilde{\Lambda}'\tilde{\Lambda}'}{N} = I_r$$

where $X$is the (TxN) data matrix with the t$^{\text{th}}$ row X_t' and $tr(\cdot)$ denotes the matrix trace.

Stock and Watson (2002a, page 1169) observe that this is the classical principal components problem. The solution to the problem can be obtained by setting $\hat{\Lambda}$ equal to the eigenvectors of $X'X$ corresponding to the r largest eigenvalues. Then the expression for the estimating the factors would be

$$\hat{F} = X'\hat{\Lambda}/N \tag{3.24}$$

Stock and Watson (2002a) note that for calculating the above expression one needs the eigenvectors of the $X'X$ which is an (N x N) matrix. However, when N>T, one can decrease the computational burden by looking at the problem from another angle. In particular, one can concentrate out $\hat{\Lambda}$ which in turn imply that minimizing nonlinear least squares problem given by

$$\text{Maximize } tr(\tilde{F}'(XX')\tilde{\mathrm{F}}) \text{ such that}$$

$$\frac{\tilde{F}'\tilde{F}}{T} = I_r.$$

Solution to the above problem will yield $\tilde{F}$ which is the matrix of the first r eigenvectors of XX'. Stock and Watson (2002a, page 1169) note that *column spaces* of $\hat{F}$ and $\tilde{F}$ are *equivalent*. Since interest in this thesis is on forecasting, getting space spanned by the factors is enough for applications. Therefore, one can use either approach, concentrating out factors or loadings, depending on the sample at hand.

In addition to showing that one can obtain factors with principal components, Stock and Watson (2002a) prove that principal component estimator is consistent and this approach can still be used even if there is slight time variation in the loadings over time. This is an important point as it is highly likely that there will be a time-varying structure in the economy.

3.3.2 Estimating Factors with Dynamic Principal Component Method

In this section FHLR method is explored in more detail. FHLR method of obtaining factors with dynamic principal component is introduced in Forni et al. (2003, 2005). Following mathematical presentation of FHLR method is based on Schumacher (2007), Barhoumi et al. (2014) and D’agostino and Giannone (2012).

D’agositono and Giannone (2012) compare forecasting performance of SW and FHLR approaches. They note that FHLR propose efficiency improvement in two ways: First of all, FHLR take into account the signal to noise ratio of the variables in the weighting step. This is achieved by using the generalized principal components analysis. Second improvement is done by taking into account the leading and lagging relations across series.

Mechanics of this method is introduced using the presentation of Schumacher (2007). He notes that the method works in two steps: Common component and idiosyncratic parts are estimated in the first step. And then in the second step static factors are obtained. In the next two subsections, these steps are shown in more detail.

3.3.2.1 Estimating the Covariances of the Common and Idiosyncratic Components

Schumacher (2007, page 274) summarizes steps for obtaining the autocovariances:

i. Let $\hat{\Gamma}(k) = T^{-1}\sum_{t=1}^{T} X_t X'_{t-k}$ be the k-lag estimated autocovariance of the series that is used for obtaining the factors. An estimator of spectral density of X_t is given by

$$\hat{\Sigma}(\theta_h) = \sum_{-M}^{M} w_k\, \hat{\Gamma}(k) e^{-ik\theta_h} \text{ at frequency } \theta_h = \frac{2\pi h}{2H} \text{ for h=0,…,2H}$$

and with Bartlett lag weights $w_k = 1 - \frac{|k|}{(M+1)}$.

ii. For each frequency, compute the dynamic eigenvalues and eigenvectors of $\hat{\Sigma}(\theta_h)$ and denote $\Lambda(\theta_h)$ as the (qxq) diagonal matrix with the largest q dynamic eigenvalues of the main diagonal and the (Nxq) matrix $\hat{P}(\theta_h) = (\widehat{P_1}(\theta_h), \dots, \widehat{P_q}(\theta_h)$ of the corresponding eigenvectors.

iii. Spectral density of the common component is given by:

$\hat{\Sigma}_\chi(\theta_h) = \hat{P}(\theta_h)\Lambda(\theta_h)\hat{P}^*(\theta_h)$, where asterisk denotes complex conjugates.

iv. Spectral density of the idiosyncratic components can be obtained by

$$\hat{\Sigma}_\xi(\theta_h) = \hat{\Sigma}(\theta_h) - \hat{\Sigma}_\chi(\theta_h)$$

v. The inverse discrete Fourier Transform provides time-domain autocovariances of the common components

$\hat{\Gamma}_\chi(k) = \frac{1}{2H+1}\Sigma_{h=0}^{2H}\hat{\Sigma}_\chi(\theta_h)e^{ik\theta_h}$ for k

3.3.2.2 Estimating the Factors

Schumacher (2007) states that aim at this step is to find the r linear combinations of the time series $\hat{Z}_j' X_t$ for j=1,…,r so to maximize the contemporaneous variance explained by the common factors $\hat{Z}_j'\hat{\Gamma}_\chi(0)\hat{Z}_j$. There is an additional restriction here following Forni et al. (2005) as

$$\hat{Z}_j'\hat{\Gamma}_\xi(0)\hat{Z}_i = 1 \text{ for } i = j \text{ and for } i \neq j. \quad (3.25)$$

This problem can be formulated as a generalized eigenvalue problem

$$\hat{\Gamma}_\chi(0)\hat{Z}_j = \hat{\mu}_j\hat{\Gamma}_\xi(0)\hat{Z}_j \quad (3.26)$$

where $\hat{\mu}_j$ denotes the jth generalized eigenvalue and $\hat{Z}_j$ is corresponding eigenvector. Then one can obtain the factors as

$$\hat{F}_t^{FHLR} = \widehat{Z}'X_t \quad (3.27)$$

where $\hat{Z}$ denotes the (Nxr) matrix of eigenvectors corresponding to the largest r eigenvalues.

After giving the mechanics of the system, comments on the steps are provided to bring to attention some key issues. D'agostino and Giannone (2012, page 312) note that the estimate of the covariance matrix of the idiosyncratic component, $\widehat{\Gamma}_{\xi}$, is ill-conditioned when the cross-sectional dimension is large. This can make the generalized principal component solution unstable. The solution of the FHLR approach to this issue is to set the off-diagonal elements to zero. So, one can interpret approach of FHLR as a modification of the static principal components approach. The modification is inversely weighting the data by the variance of the idiosyncratic components.

3.4 Determining the Number of Factors

Theoretical representation of factor models is discussed and two methods for obtaining the factors are illustrated. However, in real life applications there are several challenges to deal with before using the factors. In this section, focus is on how to determine the number of factors that one will extract from a given data set. First of all, number of static factors is discussed and then analysis moves to the number of dynamic factors that one needs for FHLR approach.

3.4.1 Number of Static Factors

Bai and Ng (2002) develop theory for the determining 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 AIC or BIC would not be appropriate for large panel of data.

Bai and Ng (2002, page 195) cite some of the alternatives for determining the number of factors. For example, number of factors can be found by a likelihood ratio test if the normality of the idiosyncratic terms is assumed. Other methods are also proposed assuming that one of the dimensions (N or T) is fixed. However, these methods perform poorly for large N and T (see Dhrymes, Friend and Glutekin (1984) and Cragg and Donald (1997)). Bai and Ng (2002) note that the problem with the previous approaches is that they do not apply when N and T tends to infinity. For example, when N>T, the rank of the sample covariance matrix is no more than T while the population covariance matrix can have rank N. Moreover, as previous discussion shows, in the case of using economic time series one needs to relax assumptions on idiosyncratic terms. Bai and Ng (2002) develop their theory by allowing heteroscedasticity in idiosyncratic terms and also some weak dependence between the factors and the errors.

The basic idea of Bai and Ng (2002, page 199) is that if one knew the factors but not the loadings, he/she could approach the issue as a model selection. So, a model with k+1 factors cannot do worse than a model with k factors. But since more parameters are estimated, there will be a tradeoff between fit and efficiency. So, following the exposition in Bai and Ng (2002), general form of the information criteria is presented.

Let F^k be a matrix of k factors and

$$V(k, F^k) = \min_{\Lambda} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^{k\prime} F_k^t)^2 \tag{3.28}$$

be the sum of squared residuals when one uses k factors. Bai and Ng (2002) note that following type of loss function can be used to determine the k:

$$V(k, F^k) + kg(N, T) \tag{3.29}$$

where $g(N, T)$ is the penalty for overfitting.

Let

$$\left(k, \hat{F}^k\right) = \min_{\Lambda} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^{k\prime} \hat{F}_k^t)^2 \tag{3.30}$$

is the sum of squared residuals when one estimates k factors from this data set (as opposed to knowing the true number of factors). 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:

$$PC(k) = V\left(k, \hat{F}^k\right) + kg(N, T) \tag{3.31}$$

$$IC(k) = \ln(V\left(k, \hat{F}^k\right)) + kg(N, T) \tag{3.32}$$

Given these forms, Bai and Ng (2002, page 201) offer the following criteria:

$$BN1: PC_{p1}(k) = V(k, \hat{F}^k) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right); \tag{3.33}$$

$$BN2: PC_{p2}(k) = V(k, \hat{F}^k) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln C_{NT}^2; \tag{3.34}$$

$$BN3: PC_{p3}(k) = V(k, \hat{F}^k) + k\hat{\sigma}^2 \left(\frac{lnC_{NT}^2}{C_{NT}^2}\right). \tag{3.35}$$

$$BN4: IC_{p1}(k) = \ln(V(k, \hat{F}^k)) + k\left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right); \tag{3.36}$$

$$BN5: IC_{p2}(k) = \ln(V(k, \hat{F}^k)) + k\left(\frac{N+T}{NT}\right) \ln C_{NT}^2; \tag{3.37}$$

$$BN6: IC_{p3}(k) = \ln(V(k, \hat{F}^k)) + k\left(\frac{lnC_{NT}^2}{C_{NT}^2}\right). \tag{3.38}$$

where $\hat{\sigma}^2$ is a consistent estimate of $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} E(e_{it})^2$. In applications one can replace this with $V(kmax, \hat{F}^{kkmax})$. And $C_{NT}^2 = \min(N, T)$. So, the criteria depend both on N and T. Bai and Ng (2002) discuss other possible information criteria which only depend on N or T. They show that those criteria will fail in certain dimensions. However, they highlight the following criterion as well (page 202 of Bai and Ng, 2002):

$$BN7: BIC_3(k) = V(k, \hat{F}^k) + \ k\hat{\sigma}^2 \left(\frac{(N+T-k)\ln(NT)}{NT}\right) \tag{3.39}$$

This criterion does not satisfy the condition of the theorem that Bai and Ng (2002) use to show the consistent estimation of factors. This is due to the fact that $g(N, T)$ does not always vanish. So, BIC_3 may perform well in some data sets but not in all data sets. Ultimately, it is an empirical issue to test the performance of this criterion.

Bai and Ng (2002) test their criteria with both simulated and actual data. 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)). 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.

3.4.2 Number of Dynamic Factors

In this thesis information criteria suggested by Bai and Ng (2007) is used to determine the number of dynamic factors. Bai and Ng (2007) start by considering a vector of observed stationary time series, F_t $(rx1)$ which follows the following VAR:

$$A(L)F_t = u_t \tag{3.40}$$

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

$$u_t = R\epsilon_t \tag{3.41}$$

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).

Define

$$\Sigma_u = E(u_t u_t') \tag{3.42}$$

Then,

$\Sigma_u = R\Sigma_\epsilon R'$ has rank $q \leq r$.

Let $c_1 > c_2 \geq \cdots \geq c_r \geq 0$ be the ordered eigenvalues and define (page 53 of Bai and Ng, 2007)

$$D_{1,k} = \left(\frac{c_{k+1}^2}{\sum_{j=1}^{r} c_j^2}\right)^{1/2} \tag{3.43}$$

$$D_{2,k} = \left(\frac{\sum_{j=k+1}^{r} c_j^2}{\sum_{j=1}^{r} c_j^2}\right)^{1/2} \tag{3.44}$$

Bai and Ng (2007) note that Σ_u and F are not observed. However, they can be estimated using $\hat{\Sigma}_u$. They show that $\hat{D}_{1,k}$ and $\hat{D}_{2,k}$ that are constructed from the eigenvalues of $\hat{\Sigma}_u$ converge to 0 $(k \geq q)$ asymptotically at a rate depending on the convergence rate of $\hat{\Sigma}_u$ to Σ_u.

In real life situations one does not know the actual values of the factors and hence they have to be estimated. In this respect, Bai and Ng (2007) discuss also the case where F_t is not observable. They offer two criteria in the case of using estimated factors. Bai and Ng (2007, page 56) note that the main contribution comes from the fact that it is not necessary to estimate dynamic factors to determine q.

3.5 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 thesis, the direct approach for multi-step ahead forecasting is used since it is the common method for the 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, page 149) focus on the multi-step ahead prediction. For industrial production they consider the following transformation for h step ahead direct forecasting:

$$y_{t+h}^{h} = \left(\frac{1200}{h}\right) \ln\left(\frac{IP_{t+h}}{IP_t}\right) \tag{3.45}$$

$$y_t = 1200 \ln\left(\frac{IP_t}{IP_{t-1}}\right) \tag{3.46}$$

Diffusion Index Forecasts (Page 149 of Stock and Watson (2002b))

$$\hat{y}_{T+h/T}^{h} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}_{hj}' \hat{F}_{T-j+1} + \sum_{j=1}^{p} \hat{\gamma}_{hj}\, y_{T-j} \tag{3.47}$$

where $\hat{F}_t$ is the vector of k estimated factors. Note that (1200/h) implies that annualized version of the h-period change is used. Stock and Watson (2002b) use three versions of the above equation.

i. DI-AR Lag: this version includes lags of the factors and lags of y with m and p estimated by the Bayesian Information Criterion. They set $1 \le m \le 3$ and $0 \le p \le 6$.

ii. DI-AR: This form of the forecasting equation uses the contemporaneous values of the factors while picks the lag length of the y over $0 \le p \le 6$.

iii. DI: This type of the forecast equation uses only the contemporaneous values of the factors so that m=1 and p=0.

iv. 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. Rather than using information criterion in the form of Bai and Ng (2002), Stock and Watson (2002b) use BIC information criterion to select the model. In particular, they consider $1\leq k \leq 4$ for the first case and $1\leq k \leq 12$ for the second and the third cases.

Stock and Watson (2002b) model the real variables as I(1). They find little difference for the inflation for modelling it as I(1) or I(2). In this thesis, following the practices in the literature and upon inspecting unit roots tests for each indicators, industrial production, inflation and the stock exchange are modeled as I(1).

CHAPTER IV

4 FORECAST ENVIRONMENT

4.1 Introduction

In this chapter, mechanics of the forecasting process used in the empirical applications are introduced. First of all, data sets are presented. After discussing which blocks are used in those data sets, how the disaggregation level increases for the indicators are explained. Then, information criteria suggested by Bai and Ng (2002) for determining the number of static factors are applied to data sets. This step shows how the number of factors and factors themselves from these data sets behave over time. In the thesis, empirical focus is on forecasts from three to twelve month ahead. Hence, equation for multi-step ahead forecasting is introduced. Forecasts are obtained from various specifications so it is necessary to use a metric for comparing all of these alternatives. Following the common practice in the literature, Root Mean Squared Error (RMSE) is used to compare models. RMSEs are obtained from simulated out-of-sample forecasting exercise. Role of different data blocks are analyzed by excluding certain indicators from master data sets. Next, construction of these restricted data sets are discussed. A lot of time and energy is invested for obtaining all of the forecasts and calculate respective RMSEs: many papers are read, codes with hundreds of lines are written and run which takes hours, data from both domestic and international economy are collected. However, there is no guarantee that this process actually helps at forecasting. Literature has shown over and over again that simple benchmarks such as AR models can beat very sophisticated forecasting models, including factor models. Therefore, forecasting performance of models are compared with a simple benchmark. After explaining this benchmark, in the last section computations used in the thesis are discussed.

4.2 Data

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. Even if the forecast model is a simple AR model, questions about data will pop up. To name just two simple examples, whether one should use seasonally adjusted or unadjusted data, or whether one should use series in first difference or in levels. In the case of more complex models, choice of the explanatory variables may play important role on the forecasts. Hence, issues related to the composition and structure of the data set may play important role on the forecasting performance.

Data set structure is 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 have considerable effect on forecasting performance.

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 in different aggregation levels. For example, one can collect data on industrial production as headline index. 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 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 head line series, five series from MIGS or around twenty from NACE or all of them at the same time. A similar picture arises for soft data as well. One can use consumer confidence as the headline index or subcomponents can be considered. These are questions about the recent state of the economy as well as about expectations. Aggregating subcomponents as simple averages, which is the method for calculating consumer confidence index, may result in loss of information. 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. Table 4.1 demonstrates the increasing level of detail for industrial production.

In the empirical section, data from following groups are used: 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 (Table 4.2). Details about the source of the data and which series are used in medium and large sets are provided in the Appendix (Table A.1 to Table A.3). Due to the technical requirements of principal components analysis it is necessary to work with stationary series. Following common practice in the literature the series are transformed by taking logs if appropriate and first differenced to ensure stationary. Target variables are treated as I(1). 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 is no consensus on data set design. One option is to follow Angelini et al. (2011) who use series from different aggregation levels in the same data set. For example, they use headline index and subcomponents for industrial production in the single data set of their paper. An alternative way is offered by Barhoumi et al. (2010) who use different data sets by gradually disaggregating the indicators. They compare three different data sets that are constructed by disaggregating indicators; small, medium and large. By considering different factor extraction approach and different forecast horizons 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.

Table 4.1. 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.

Following the approach of Barhoumi et al. (2010) for extracting factors, for the empirical exercise of this thesis 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.

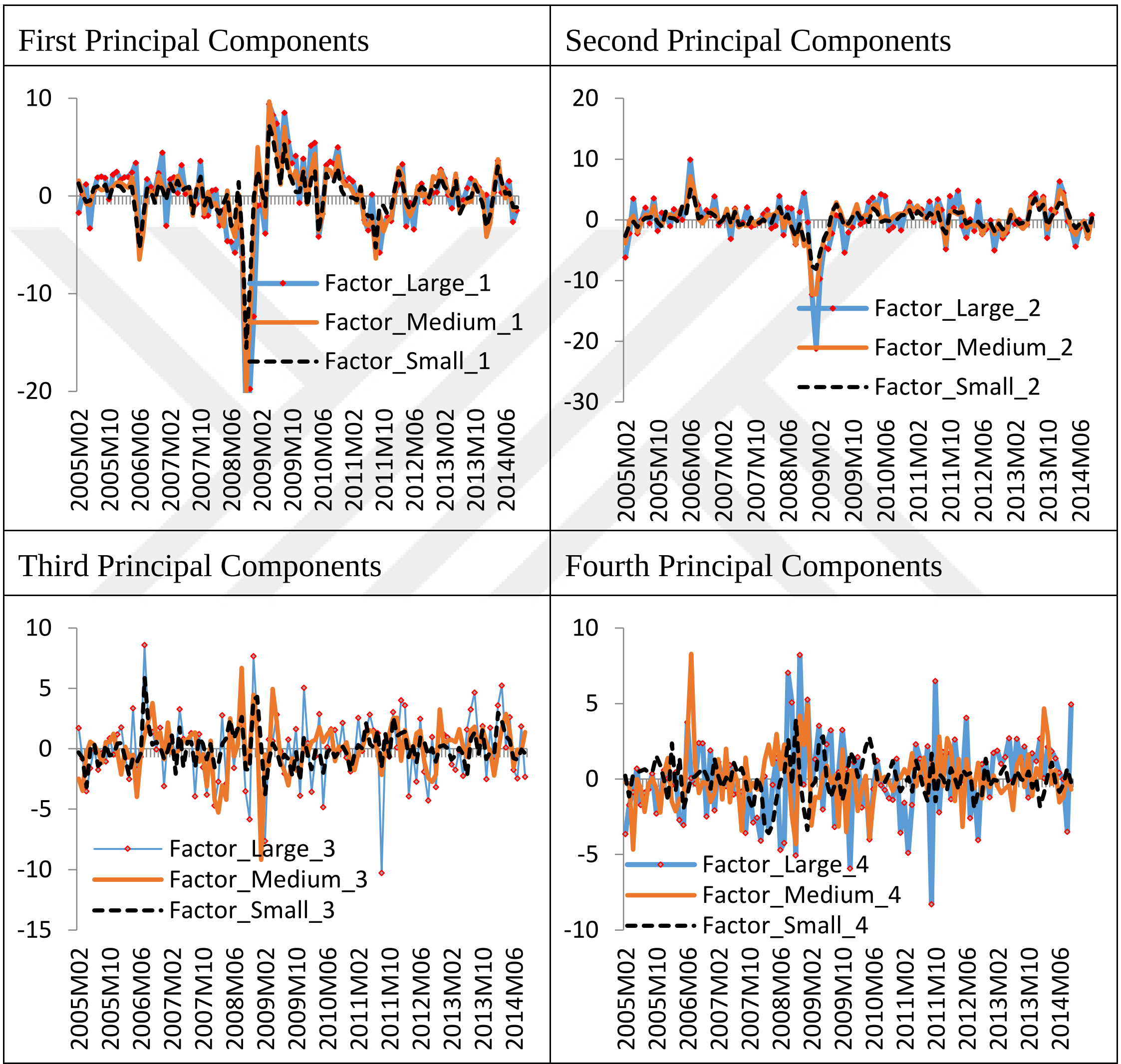
Table 4.2. Indicators Used in the Small Data Set

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

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.

4.3 Factors

Once data sets are constructed, one can extract factors that summarize information in these sets. In this section, the first four principal components from three data sets are plotted. These principal components are the factors that are used for SW approach in the forecasting.



Notes: Graphs show the principal components that are obtained from three different data sets namely, small, medium and large.

Figure 4.1.Principal Components from Small, Medium and Large Data Sets

Source: Author's own calculations

Figure 4.1 shows the first four principal component from each of the three data sets. It is seen that first principal components from each data set show similar patterns overtime. Second principal components are also fairly similar for medium and small data sets. When one moves in the analysis and consider third and fourth principal components it is seen that factors diverge from each other. The fact that one gets different factors from different data sets imply that forecasts from these data sets may be different. One of the aims in this thesis is to see whether there is a systematic pattern between forecast performance of factor models and data set size. For instance, if the best performing models use factors from large data set this implies that more data is better. So one can use disaggregated data for forecasting purposes. It should be noted though that number of factors and use of lags of the factors may affect forecast performance. Thus, focus should not be only on the disaggregation level but overall modelling choices should be taken into account.

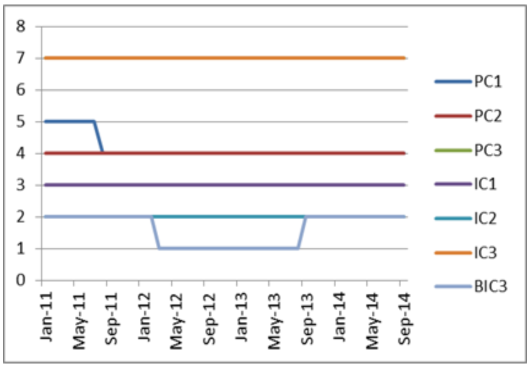
4.4 Number of Factors

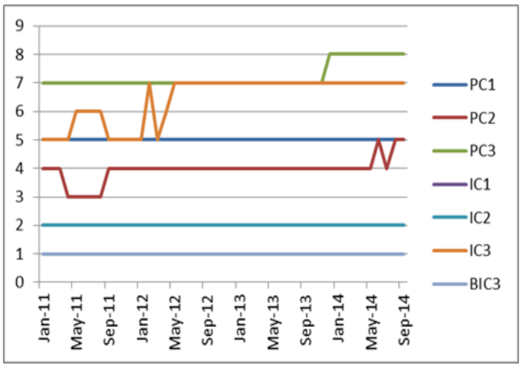
In the previous section, it is observed that factors become less similar as the number of factors increase. Correlation between the first factors are very high, so if one uses only one factor from each of the data sets he/she will get fairly close forecasts. However, fourth factors behave rather differently. Thus if one use four factors in forecast equations, it is highly likely that forecasts will differ by a large margin. These observations suggest that number of factors used in the forecasting applications may affect forecasting performance substantially. In this respect, in this section determining the number of factors is discussed.

Bai and Ng (2002) note that if one knows the true factors, he/she 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 section 3.4.1 in the methodology chapter). In the presence of cross-section correlations, BIC3 has very good properties. This criterion can be used despite not fulfilling all the conditions of Theorem 2 in their paper. Figure 4.2 and Figure 4.3 show number of factors that one gets by recursively expanding medium and large data sets. As evident in the figures below, 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 one adds more observations through time. To be more specific, PC3 and IC3 suggest around 7 factors while BIC3 suggest 1 factor.

In the thesis, recursive out–of-sample forecasting exercise is used to compare models. In the evaluation sample that is used to assess the performance of models, for each month number of factors proposed by different criteria from Bai and Ng (2002) are obtained. In the graphs, x-axis shows the last month of the data set that factors are extracted from. For example, by using the criteria of Bai and Ng (2002) number of factors suggested by each criterion are calculated from February 2005 to January 2011. Then, data set is expanded by one month and calculations for the number of factors suggested by the seven criteria are re-done. Results are shown in February 2011 point of the x-axis. Maximum number of factors (required for PC1, PC2 and PC 3) is set as for four, seven and nine for small, medium and large data sets, respectively.

Some authors do not use all of these criteria or do not even check the role of using a certain information criterion on forecasting performance. For instance, although Barhoumi et al. (2013) analyze effect of the number of factors used on forecasting performance, they only employ IC1 among the Bai and Ng (2002) criteria. Another example is Gupta and Kabundi (2011) where they 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 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, all of the seven criteria suggested by Bai and Ng (2002) are considered in this thesis.





Notes: In the thesis, recursive out–of-sample forecasting exercise is used to compare models. In the evaluation sample, at each month number of factors are obtained by different criteria from Bai and Ng (2002). X-axis shows the last date that factors are extracted. For small data maximum number of factors (required for PC1, PC2 and PC 3) is set as four, for medium data set it is set as seven and for large data set it is set as nine. Results for small data set are not shown since the pattern is similar with the medium and large data sets. To be more specific, BIC3 suggests the lowest number of factor (one factor) and PC3 suggests the highest number of factors (four factors).

Figure 4.2.Number of Factors Obtained from Different Information Criteria for Medium Data Set (left)

Figure 4.3.Number of Factors Obtained from Different Information Criteria for Large Data Set (right)

Source: Author’s own calculations

4.5 Forecast Horizon and Forecast Equation

In this thesis, results are analyzed for three, six, nine and twelve month-ahead forecasts. Attention is on these four horizons for the sake of clarity in presentation. Three month-ahead forecast performance is expected to be informative about the short run performance of models while twelve month-ahead forecasts can be informative for longer run performance. An important question emerges when one forecasts for more than one period ahead. Consider the month-on-month growth rate of industrial production as presented in Figure 4.4. One can define three month-ahead forecast as the month-on-month growth from three months from now. For example, in a case where one has access to the January figure as the last data point, he/she 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. If for some reason the March figure is unusually low, there may be correction in April and growth rate may be high. Hence, month-on-month growth rates from three and twelve month from now may not be very interesting from the user's perspective, be it a policy maker or a financial analyst. Rather, they may be interested in the over-all growth during these periods.

In this respect, following Stock and Watson (2002a) forecasts are obtained for the *cumulative* growth rates for three, six, nine and twelve month. In this approach, for the case that one has access to January data, he/she forecasts the growth rate in April relative to the level in January. In other words, one works with the cumulative growth in the horizon of interest. Three and twelve month-ahead cumulative growth rates in Figure 4.5 show that, as expected, twelve month-ahead cumulative growth rates are relatively more stable than three month-ahead growth rates. It is also observed that volatility of the three month growth decreased after around mid-2011. This observation suggests that forecast performance may change for different time periods. As a result, forecast performance will be evaluated for two consecutive samples (see the discussion on Equation 3.47).

$$\hat{Y}_{t+h/t}^{h} = \hat{\alpha}_h + \sum_{j=1}^{m} \hat{\beta}'_{hj}\hat{F}_{t-j+1} + + \sum_{j=1}^{p} \hat{\gamma}'_{hj} Y_{t-j}, \tag{4.1}$$

where Y is the variable that we want to forecast, $\hat{F}$ denote the vector of factors, Y_{t-j} shows the month on − month change of the dependent variable .

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

Equation 4.1 shows the direct forecasting equation. It is estimated with OLS. In this equation, dependent variable is the cumulative growth rate from time t to time t+h so that one is forecasting h-period ahead. Month-on-month change of industrial production and the estimated factors are used as independent variables. Using different letters in the notation of Equation 4.1 for the lag length (namely m and p) indicates that one can use different number of lags for the lag of dependent variable and for the factors. Cap on the F shows that estimated factors are used. This is due to the fact that one cannot observe actual factors.

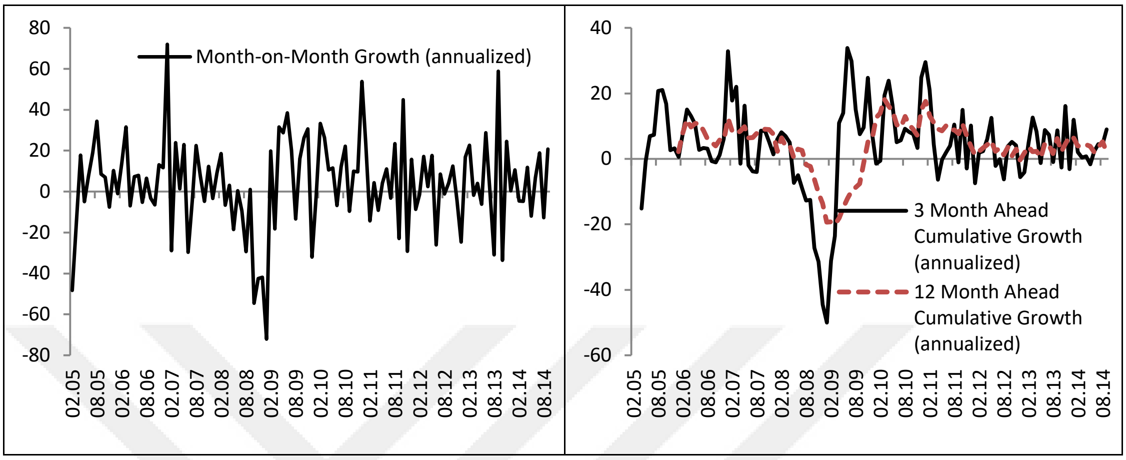


Figure 4.4.Month-on-Month Growth of Industrial Production

Figure 4.5.Three and Twelve Month Cumulative Growth of Industrial Production (Annualized))

Source: TURKSTAT

4.6 Forecast Evaluation

There are various alternative methods for obtaining forecasts. Factor model approach used in this work is just one type of the methodologies developed by theoretical and applied contributions over a long period of time. Yet, ex-ante it is not clear whether a method can produce informative forecasts. So, before using a method in real life applications, one needs to have an idea about the success of the models. One option for testing a model's success is to build a model today and follow its forecasting performance for several years to come. If it is considered to be successful then that model may be included in the toolkit in the future. Of course this is not a very practical way. In the end, that model may turn out to be a bad choice. So one needs a quicker way to gauge the success of a model. In the literature a method called "simulated-out-of-sample forecasting" is used as a quick and practical way to evaluate models. In this approach, one does the following though experiment. If I had used this model three years ago what would its forecasts have been for, say six month ahead? Model is estimated with the data that would be available three years ago and the forecast is obtained. Then, data set is expanded by one month and a new forecast is obtained for six month ahead. In the end, forecaster will end up with a vector of forecasts for six month ahead that would have been obtained if this model had been used in the past. He/she can compare these forecasts with the realizations. If the forecasts miss most of the realizations by a high margin, this would suggest that forecasts from this model may not be reliable in the future as well. Some modifications in the model can be done or an alternative model can be considered. Then evaluations can be re-done to see whether there is any improvement.

One way to do the evaluations would be using a graph. Plotting forecasts with realization will reveal a lot of information. However, looking at the graph one can think it is a good model for forecasting while somebody else might not agree. Moreover, two researchers forecasting the same variable with different methods cannot compare their methods in an objective way with the graphs. Also, a researcher may build a large number of methods, just like our thesis has 340 alternative specifications. Comparing all of them with a graph is practically impossible. So, a metric to evaluate models is necessary.

In this thesis evaluation criterion, the metric, is the Root Mean Squared Error (RMSE) that one gets from a pseudo out-of-sample forecasting exercise. RMSE is the most frequently used metric for comparing models. If one only considered average forecast errors, negative and positive errors would cancel. So even if the model makes very large forecast errors, there is a chance that average error would be close to zero. This might give the wrong signal that model makes very low forecast error. So, square of the errors are used in the evaluations. This would punish larger errors heavily. Formally RMSE can be expressed as follows:

$$RMSE = \sqrt{\frac{[(Realization\ at\ time\ t+h) - (Forecast\ for\ time\ t+h\ at\ time\ t)]^2}{Number\ of\ forecasts}} \quad (4.2)$$

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 (2013) 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.

In the graphical analysis of chapters six to eight, twelve month rolling RMSEs are presented. For the tables presented at the end of these chapters for showing the best and the worst performing specifications, evaluation is done for two 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 is available up until September 2014, and the longest horizon that the thesis is interest in is 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 had forecast at that point in time. At each step factors are obtained with data that would be available at that time, lag lengths in Equation 4.1 are calculated, appropriate equation for h step-ahead forecasting is estimated, and forecasts are obtained. Two versions of Equation 4.1 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. When equations for DI-AR Lag are estimated, lag lengths are determined using the Bayesian Information Criteria. After finding the appropriate model, using this model forecasts are obtained.

4.7 Excluding Data Blocks

In the forecasting process, indicators from different data blocks are considered such as industrial production and interest rates. For the analysis of the data set structure on the forecast performance, two dimensions are taken into account. First dimension is disaggregating data set and this issue is discussed in Section 4.2. There is another angle that can increase one's understanding of the data set structure on the forecasting performance. Namely, analyzing the effect of different data blocks on forecast performance.

Table 4.3. Indicators Excluded for the Construction of Data Sets for Analyzing the Effect of Data Blocks

Data sets excluding European Union variables

1. EU-Industrial Production
2. EU Consumer Confidence
3. EU-Business Confidence

Data sets excluding commodities and financial variables

1. Commodity Price Index
2. VIX
3. SP 500
4. Borsa Istanbul-30

Data sets 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 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 Table 4.3, 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 European Union indicators on the forecast performance, forecasts from two data sets are compared: 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 by excluding certain blocks one at a time are compared. In the tables 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 master data sets.

4.8 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 random walk. Intuition behind comparing with a benchmark is the fact that going over all the messy details of a forecasting model may not worth it if it cannot beat a simple benchmark. Choice of an AR model as benchmark suggests that just by using time series properties one can construct forecasts. If the proposed model cannot beat this simple AR model, it suggests that it cannot add value to the forecasting practice. In the literature a frequently observed finding is inability of the sophisticated models beating the simple benchmarks.

Benchmark model in this thesis 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 choses as the benchmark. In Chapter 5 to Chapter 7, in the tables where alternative specifications are compared, relative RMSE of the factor models compared to the simple benchmark are presented. A figure greater than 1 means that, on average, model makes higher forecast error than the simple benchmark.

4.9 Computation

In this thesis a comprehensive analysis of a large number of models are done by considering 340 specifications that one obtains by changing the factor extraction approach, number of factors, data set size and forecast equation. For the out of sample forecasting exercise, for each month, factors are obtained from each of these specification for three to twelve month ahead forecasts. DI-AR Lag type equations require setting the lag length of each of the indicators. Every possible combination of lag lengths of variables are taken into account. This means estimating thousands of equations for determining lag length of a given specification for each round of out of the sample exercise. So, computational burden is heavy. Estimations are done in Eviews 7.1 and Matlab 2013 with codes that are written for this thesis. For the Matlab code, codes of Schumacher (2007) are used as the starting point, in particular Matlab functions for obtaining the factors and estimating the number of factors suggested by the information criteria. For the FHLR approach for some functions it is also made of use of the codes by Mario Forni that are available in his website. In the end, codes are designed that are suitable to use in the future for studying the effect of modeling choices.

CHAPTER V

5 FORECASTING INDUSTRIAL PRODUCTION

5.1 Introduction

In this chapter, forecasting performance of factor models are analyzed for industrial production growth for a comprehensive set of specifications for three, six, nine and twelve month-ahead. Since it is a key indicator for monitoring the state of the economy, policy makers and market players closely follow developments in industrial production. Industrial production is affected both from external and domestic demand conditions. Weather conditions and occasional events like strikes can also have temporary but strong effects on industrial production. Another element that affects the industrial production is the inventory management of firms. They may build-up inventories anticipating strong demand in the future or vice versa. These kinds of determinants may affect industrial production growth at different magnitude for different horizons. Hence, industrial production growth may be quite volatile, hard to model and forecast. Extensive analysis in this chapter can guide forecasters in enlarging or modifying their forecasting toolkit.

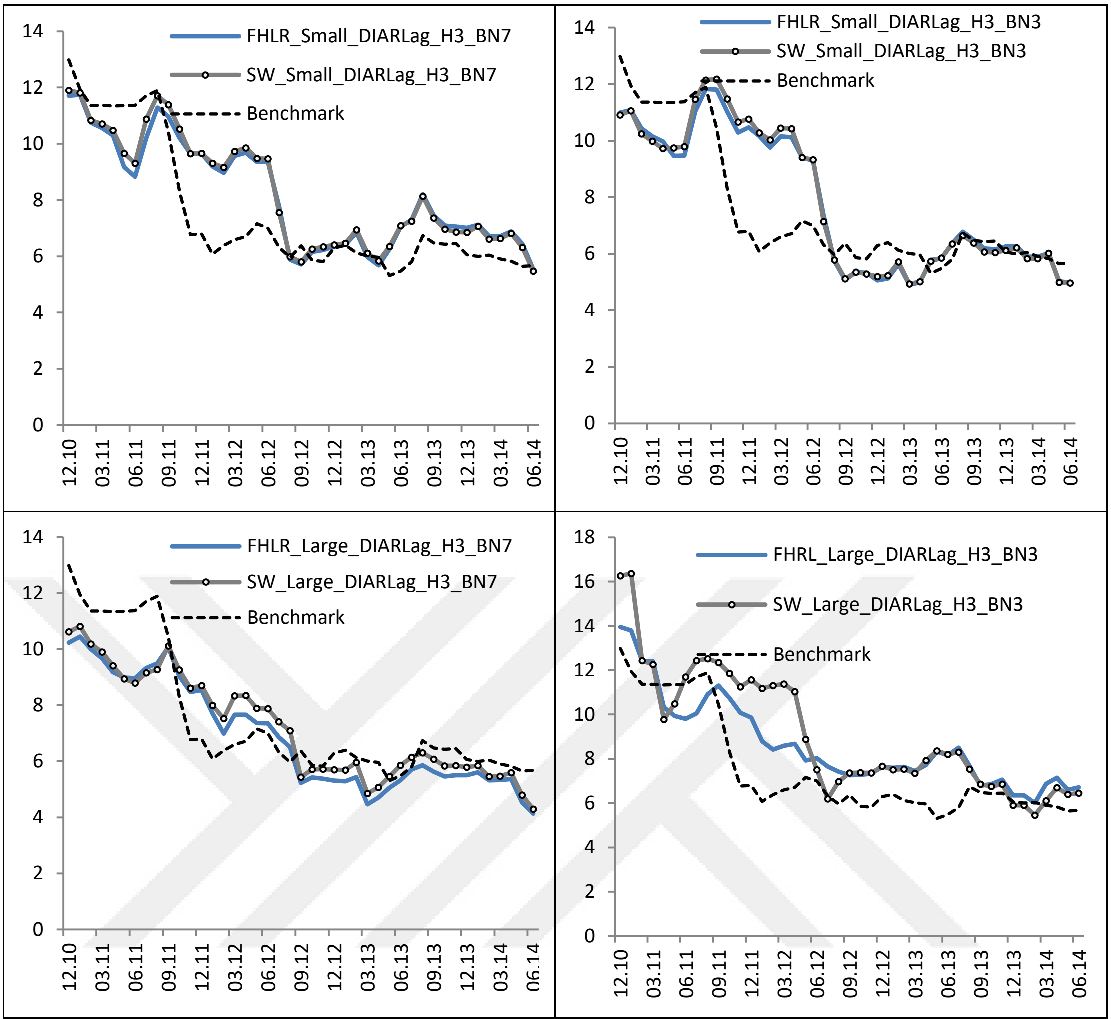
Papers reviewed in Chapter 2 show that core inflation and industrial production respond differently to the modelling decisions. In particular, using lags of factors and inflation helps at forecasting while this practice may even harm forecasting performance for industrial production. Effect of data set structure may also be different for core inflation, industrial production and stock market. This is due to the fact that data sets are in general quite heterogeneous. Data blocks, such as interest rates and European Union variables, may play different roles on the variables that one aims to forecast. In this respect, analyzing effect of modelling decisions for forecasting industrial production and contrasting findings of this exercise with core inflation and stock market is expected to contribute to the literature.

Chapter is organized as follows: First of all, analysis starts with factor models using three master data sets; small, medium and large. Then an alternative method for using big data sets in forecasting, namely pooling of bivariate forecasts, is considered. In the fourth section focus is on the effect of data blocks on forecasting performance. In these three sections, rolling twelve-month RMSEs are presented for the benchmark and factor models. Aim of this format, namely making comparison through figures, is to give a sense of the effect of modelling specifications on forecasting performance. Since relative performance may change over time, RMSEs are presented in twelve-month moving windows. This practice enables reader to see how a given set of specification affects forecast performance, compare the given specification with alternative ones and also with respect to the benchmark in different periods. There are various possible specifications, to be accurate 340, that can be obtained by changing some dimensions of the modelling process such as forecast extraction approach or building the forecast equation. In the last section, a horse-race type analysis is conducted to see whether there is a systematic pattern in the best and worst performers. For this aim, in the tables the five best and the five worst specifications are shown for three-, six-, nine- and twelve- month ahead horizons for two forecast evaluation samples. This is also the section that results are discussed and interpreted.

5.2 Forecasting with Factor Models

5.2.1 FHLR vs SW

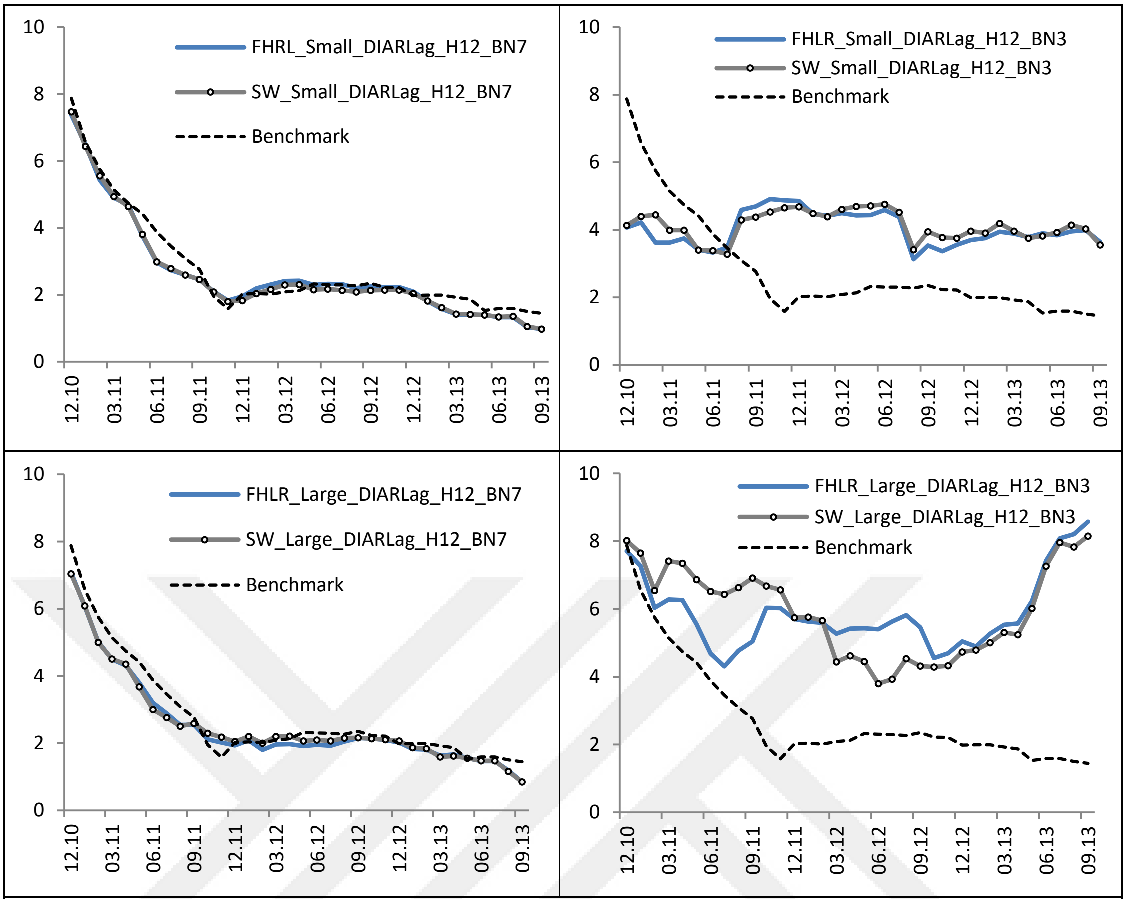
In this subsection, performance of FHLR and SW approaches for the factor extraction are compared. Discussion starts with results for industrial production for three-month-ahead forecasts with DI-AR Lag type forecast equation. Comparison of these approaches is made by changing the criterion for the number of factors and size of the data set (Figure 5.1). Note that for the clarity in the presentation only two information criteria out of the seven available are presented in the figures while all seven are considered in the last section of this chapter.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach of Stock and Watson (2002b) and Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 shows the forecast horizon for three month ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 5.1 Rolling RMSEs for Comparing Factor Extraction Approaches for Industrial Production: Three Month Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach for Stock and Watson (2002b) and Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H12 shows the forecast horizon for twelve month-ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 5.2. Rolling RMSEs for Comparing Factor Extraction Approaches for Industrial Production: Twelve Month Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

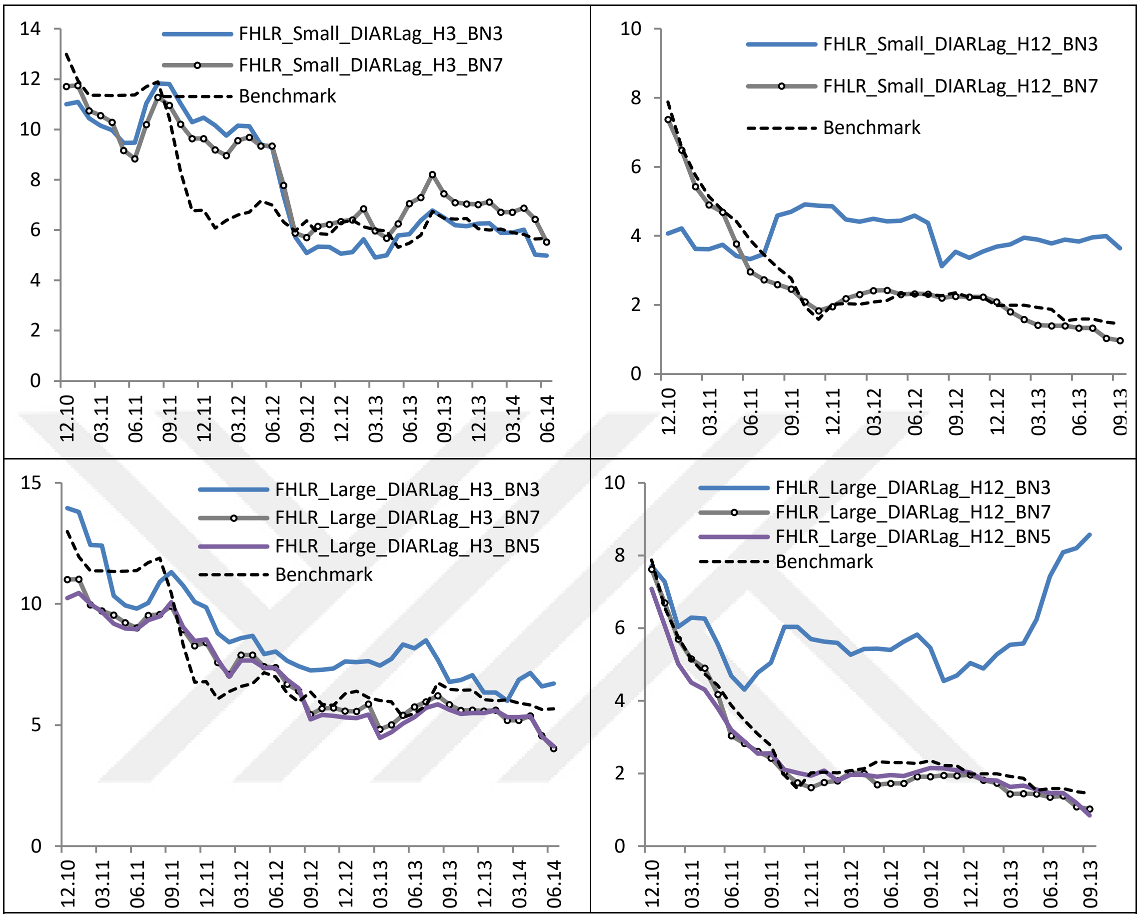
For the factors extracted from small data set, forecasts errors are close to each other for SW and FHLR for two information criteria, namely BN7 (north-west graph of the figure) and BN3 (north-east graph of the figure). Moving to the large data set, in the case of choosing number of factors with BN7, forecast errors are again close for SW and FHLR with FHLR performing slightly better (south-west graph of the figure). For BN3, which tends to give a high number of factor, for the large data set forecast errors diverge relatively more at the beginning of the sample (south-east graph of the figure). For this case, FHLR approach results in lower forecast error over time while there is a brief period where SW performs better. It should be noted however that factor forecasts are, in general, worse than the benchmark.

Analysis of three month-ahead forecast errors guides one about short-term forecast performance. As forecast horizon increases, however, relative performance of different specifications may change. In the next figure, forecast performance is presented for twelve month-ahead forecasts (Figure 5.2). Inspecting graphs shows that for the cases that one selects number of factors with BN7 for small (north-west graph in the figure) and large (south-west graph in the figure) data sets and small data set with BN3 criterion (north-east graph in the figure), forecast errors for SW and FHLR approaches are close to each other. On the other hand, for the case that one selects the number of factors with BN3 and use large data set (south-east graph in the figure), forecast errors diverge for SW and FHLR approaches. It is observed also that for both small and large data sets with BN3 criterion factor models perform worse than the benchmark. With BN7, at the beginning of the sample, factor models beat the benchmark.

5.2.2 Number of Factors

In the previous section, it is observed that effect of number of factors on forecast performance is stronger than the effect of factor extraction method. In this respect, in this section effect of the information criteria for choosing the number of factors is analyzed. Bai and Ng (2002) offer seven criteria for selecting the number of factors. Plotting RMSEs that would result in using all the seven criteria in a single graph will not be visually appealing. In this respect for small data set, results are presented for BN3 (named as PC3 in Bai and Ng, 2002) and BN7 (named as BIC3 in Bai and Ng, 2002). Choice is due to the fact that, in general BN3 gives the highest number of factor while BN7 suggest the lowest number of factors. In addition to these two criteria, for large data set results obtained from using information criterion BN5 (named as IC2 in Bai and Ng, 2002) is presented as well. Number of factors suggested by this criterion is in general between

BN3 and BN7. Since results for SW and FHLR approaches are close to each other for most of the cases, to save space it is reported below only the models with FHLR approach.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 5.3. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Industrial Production: DI-AR Lag

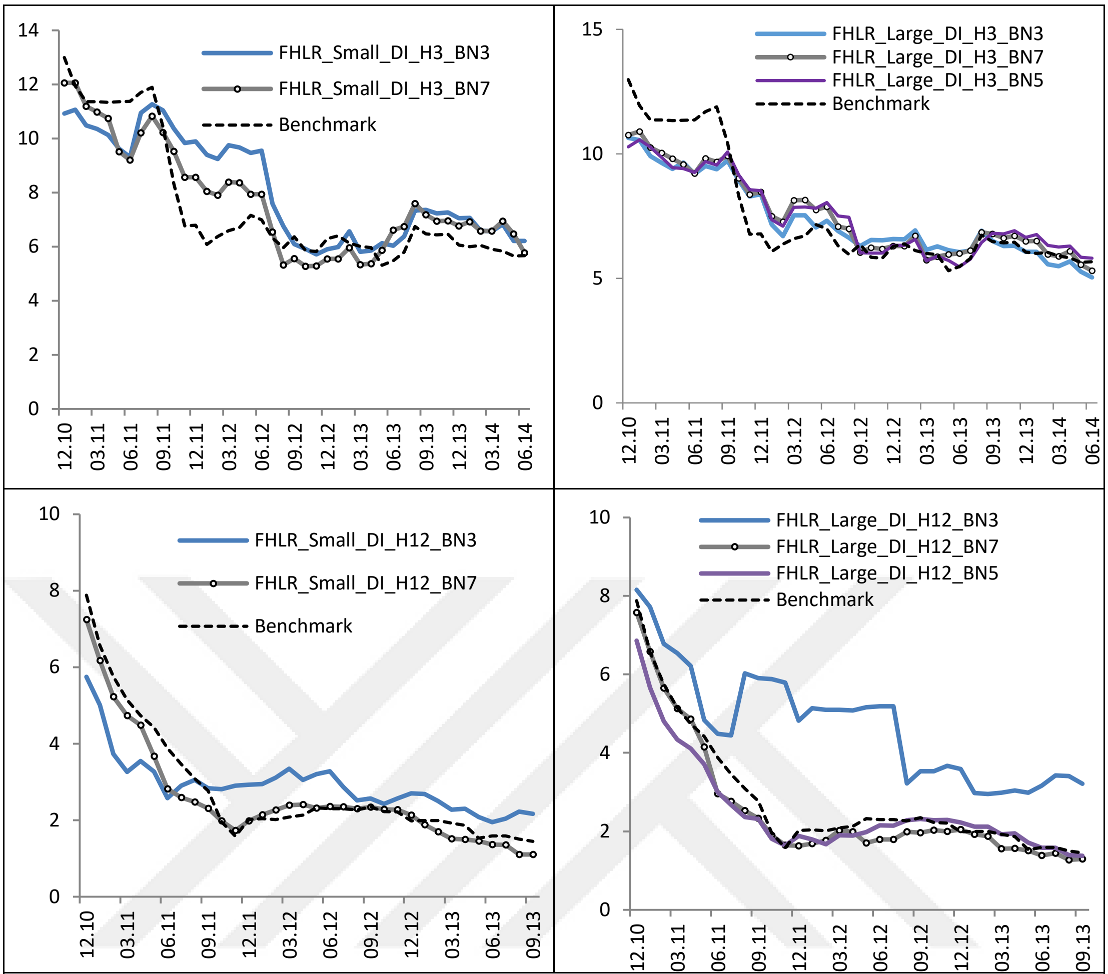
Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

It is seen in Section 4.4 that BN3 gives around eight factors for large data set while BN7 gives around one factor. Using lags of those factors increases parameter uncertainty.

This effect will be more severe for the case that one selects the number of factors with BN3. So, it is important to take into account the type of the forecast equation for analyzing the effect of the number of factors on the forecast performance. To that end, first of all discussion focuses on the case of DI-AR Lag type forecast equation and then attention is given to the DI type forecast equation. In general using BN3 increases forecast errors relative to BN7 or BN5 (Figure 5.3). An exception to this observation is the results for small data set for three month ahead forecasts (north-west graph of the figure). For large data set, using BN3 causes deterioration both in the short-run (south-west graph of the figure) and for long-run forecasts (south-east graph of the figure).

In the DI-AR Lag case, selecting number of factors with BN3 causes deterioration for three of the four cases. This may be due to the fact that BN3 selects a high number of factors and using lags of all of those factors may increase parameters uncertainty in the estimations. Hence, using a DI type forecast equation, where one uses only contemporaneous factors, may change the conclusion about the role of the number of factors on forecast performance (Figure 5.4). For three month-ahead forecasts using large data set with DI type forecast equation, it is seen that unlike DI-AR Lag case, using BN3 produces similar forecasts errors with BN7 (north-east graph of the figure). For twelve month-ahead forecasts, BN3 still results in poor forecasts albeit with some improvement over the DI-AR Lag case. Forecast errors from using BN5 criterion, which suggests higher number of factors from BN7 but lower than BN3 for large data set, are presented as well. In this case performance is close to specifications with BN7.

In summary, for industrial production growth forecasts, using BN7 which in general suggests only one factor produces lower forecast errors compared to cases where one uses an information criterion that picks a relatively larger number of factors.



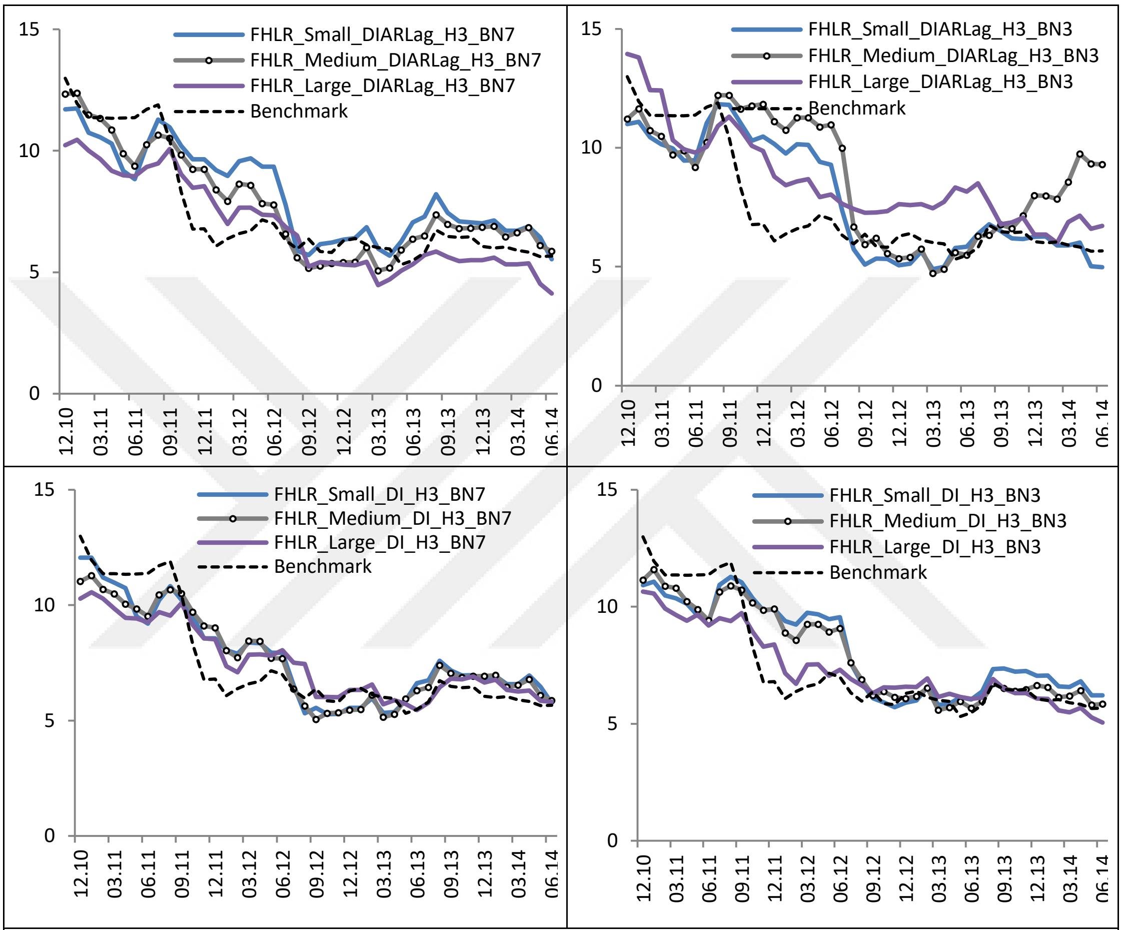
Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI refers to the forecast equation where one uses only contemporanous factors, H3 and H12 shows the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PIC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 5.4. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Industrial Production: DI

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

5.2.3 Data Set Size

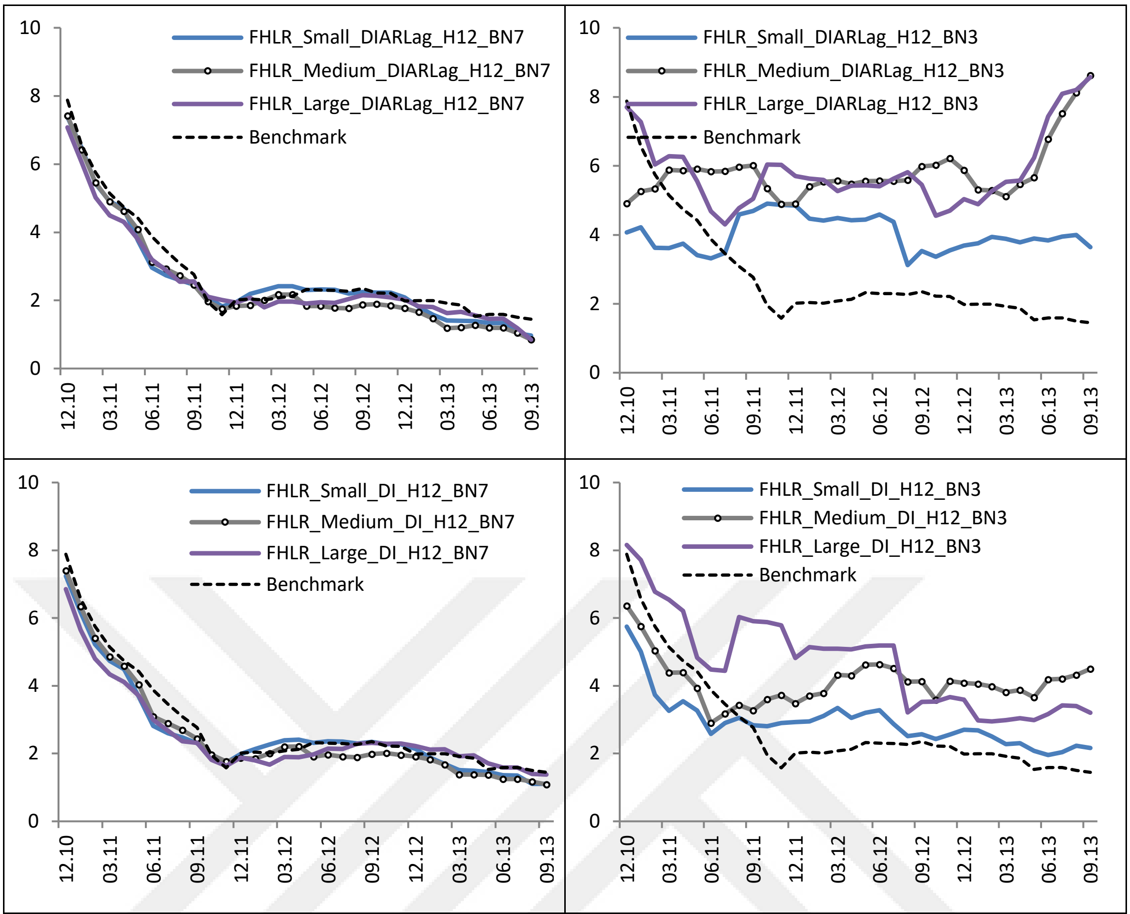
In this section, the focus is on comparing and contrasting forecasting performance for small, medium and large data sets. Analysis starts with three month-ahead forecasts (Figure 5.5).



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H3 shows the forecast horizon for three month-ahead forecast while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 5.5. Rolling RMSEs for Comparing Data Set Size for Industrial Production : Three Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H12 shows the forecast horizons for twelve month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 5.6. Rolling RMSEs for Comparing Data Set Size for Industrial Production: Twelve Month-Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

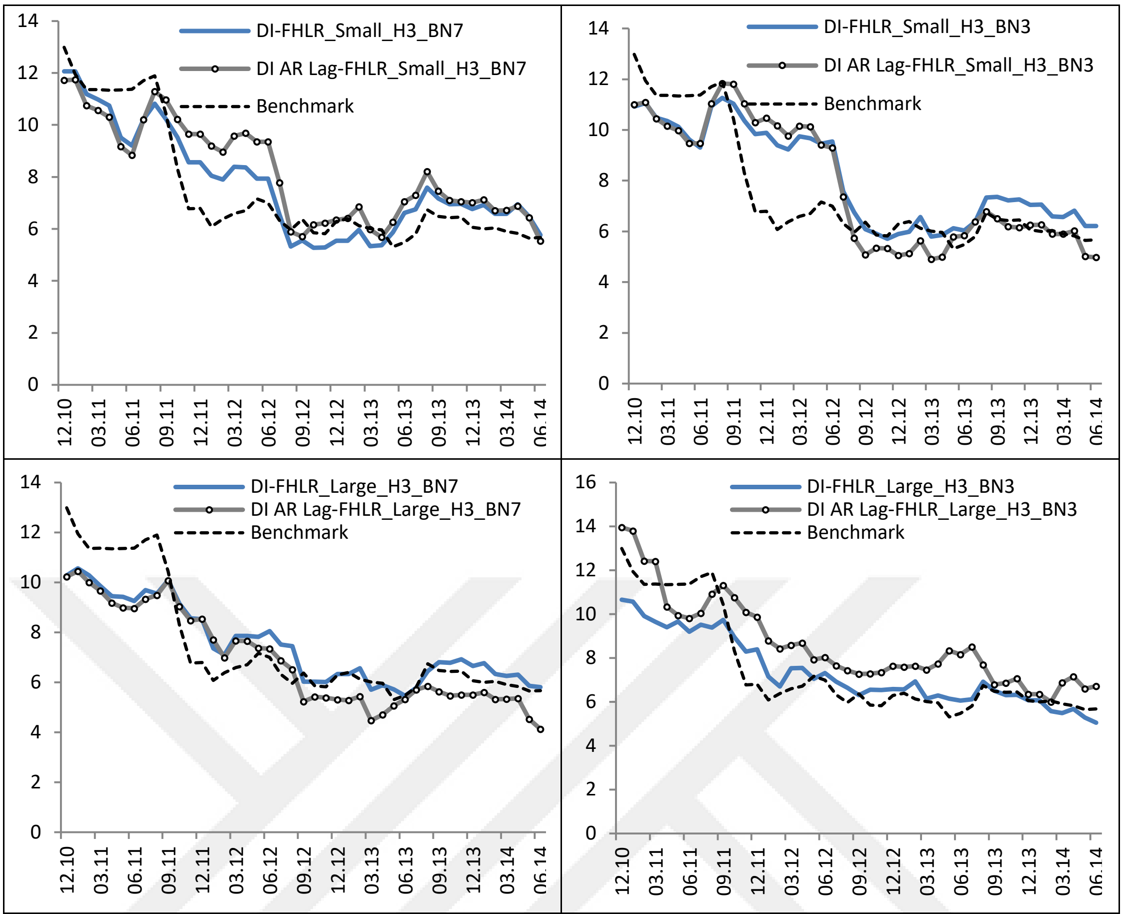
When one uses BN7 criterion with DI-AR Lag forecast equation, large data set improves forecasting performance over medium and small data sets (north-west graph of the 5.5). Changing the forecast equation to DI, large data set stills perform better in general but now worse than the benchmark (south-west graph of the figure).

Next, relative forecast performance of different specifications are evaluated for twelve month-ahead forecasts (Figure 5.6). Twelve month-ahead forecasts provide insight about the longer run performance of the factor models. For specifications with BN7, RMSEs are close to each other (north-west and south-west graphs of the figure). When one uses BN3, RMSEs are considerably worse than the benchmark. In the case of three month-ahead forecasts presented above, RMSEs were relatively more different with BN7 and closer to the benchmark with BN3. Hence, it is seen that the question that whether factor models can beat benchmark or whether a specification have a positive or negative effect on forecast performance is not independent from the forecast horizon.

5.2.4 Forecast Equation Type

In this section focus is on the effect of forecast equation type on forecast performance for three month-ahead forecasts (Figure 5.7). It is clearly seen once again that it is important to carefully analyze the sensitivity of the modelling choices on the forecast performance. In particular, for small data set with BN7 using DI performs better (north-west graph of the figure) while for BN3 using DI-AR-Lag performs better (north-east graph of the figure). For the specification with large data set and BN7 criterion, using DI-AR-Lag type forecast equation produces lower forecast error than DI (south-west graph of the figure) while for BN3 reverse is true (south-east graph of the figure).

In summary, analyzing the performance of the factor models require a comprehensive approach. In the process of forecasting with factor models, inputs are used from different dimensions for producing the output, forecasts. These inputs may be interacting with each other. Interestingly factor extraction approach is not the most important determinant of the forecasting performance. While FHLR approach uses a more complicated and theoretically more comprehensive technique, performance of SW and FHLR are close to each other.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI AR Lag refers to the forecast equation where one uses lags of the factors and the dependent variables, DI refers to the forecast equation that only the factors are used, H3 shows the forecast horizons for three month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

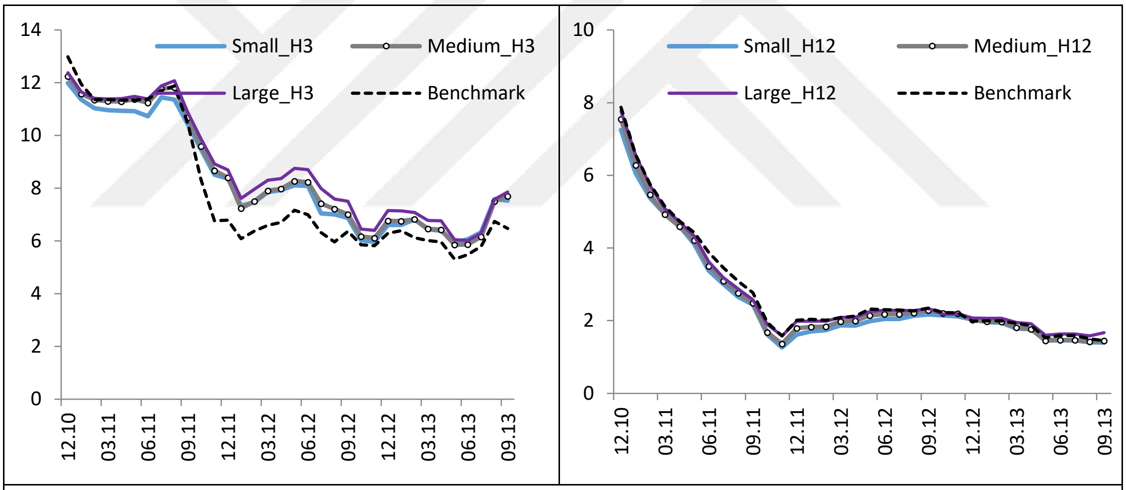
Figure 5.7. Rolling RMSEs for Comparing Forecast Equation Type for Industrial Production: Three Month-Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

Results indicate that number of factors, forecast equation type and the data set size may play significant role on the forecast performance. Care is needed to take into account all of these dimensions in setting up the forecasting process. One should be aware of the fact that using a large number of factors along with their lags may harm the forecasting process even if this process improves the fit of the forecast equation.

5.3 Pooling of Forecasts

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 are 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. Results reveal that from three data sets for both three month- and twelve month-ahead horizons, forecasts are close to each other (Figure 5.8). For three month-ahead forecast, pooling of forecasts does not beat the simple benchmark.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, Small/Medium/Large refer to the data sets that forecasts are obtained from. H3/H12 refer to the forecast horizon for three and twelve month-ahead forecasts.

Figure 5.8. Rolling RMSEs for Comparing Pool of Forecasts for Industrial Production

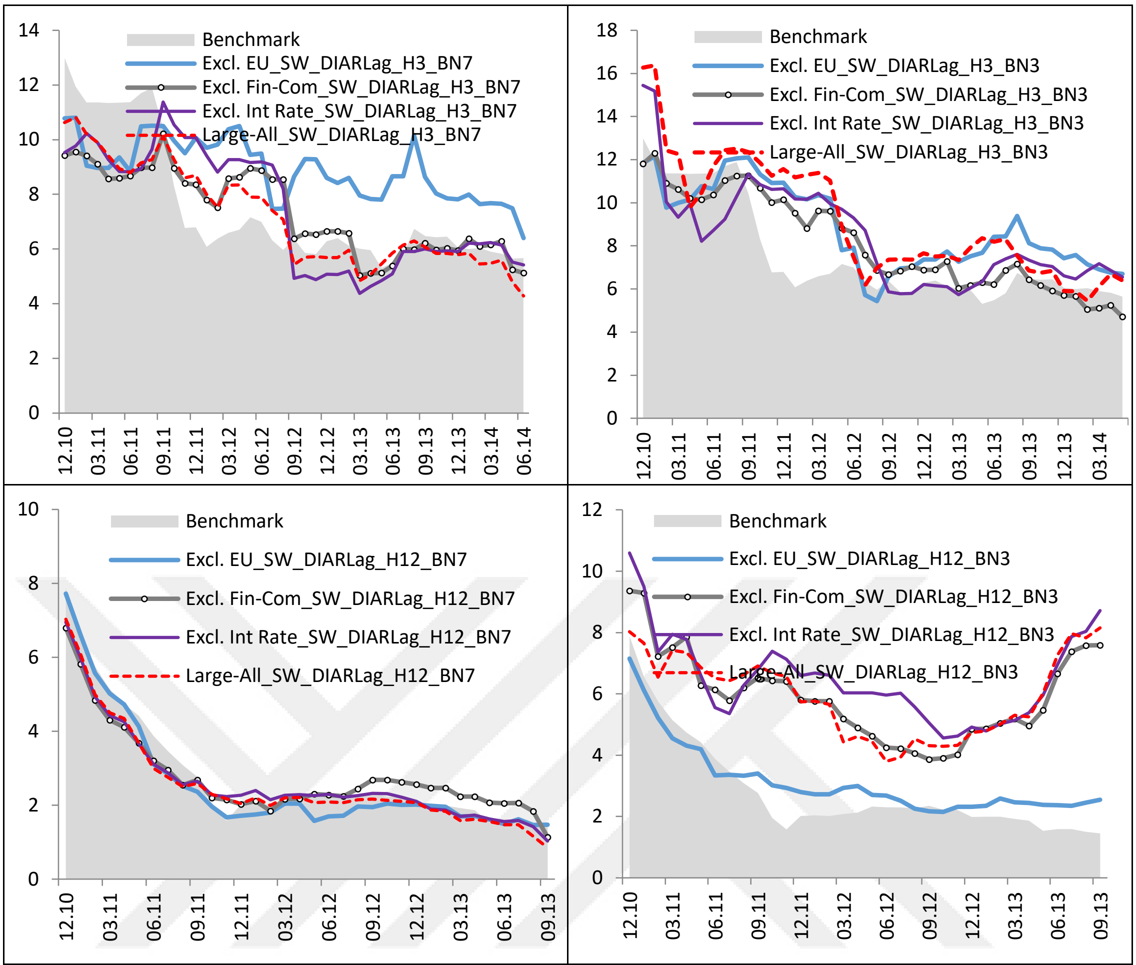
Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

5.4 Excluding Data Blocks

Factor models are powerful tools to deal with big data. One can summarize a large number of series with a few factors. Yet, there is no golden rule, rule of thumb or recipe for choosing the composition of the data set for macroeconomic applications. Hence, researchers construct data sets from different blocks such as real variables, prices and surveys. In the second section of this chapter it is discussed the effect of the data set structure by focusing on the aggregation level of the series that are used in the data sets. By increasing detail of the series within blocks, three data sets are constructed; small, medium and large. To some extent, this exercise enables one to see the effect of data set structure on the forecast performance. However, there are still important questions about the composition of the data set. Should one use a certain block at all and whether certain blocks are more important for forecasting than others? Following the practice in the literature to answer these questions, forecasts are obtained by excluding data blocks. In particular, four different data sets are constructed:

i. Master data sets (Large, Medium and Small data sets with all of the variables discussed in Chapter 4)
ii. Data sets excluding European Union variables (Excl. EU)
iii. Data sets excluding commodities and financial variables (Excl. Fin.)
iv. Data sets excluding interest rates (Excl. Int. Rates)

Sensitivity of the factor models' performance to different specifications are analyzed for the master data sets in the second section. For data sets that exclude certain blocks, one can redo the analysis for the effect of factor extraction methodology, number of factors, data set size and the forecast equation. To save space, for the graphical analysis all of these steps are not repeated. However in the next section, all of the specifications will be considered when the models are ranked. For instance small data set excluding European Union variables for alternative number of factors criteria will be in the competition list.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW refers to the factor extraction approach for Stock and Watson (2002b), Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 5.9. Rolling RMSEs for Comparing Factor Models for Industrial Production by Excluding Data Blocks: DI-AR Lag

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

In Figure 5.9 results are presented for the large data set (all series and excluding blocks) with DI-AR Lag type forecast equation. For three month-ahead forecasts, excluding European Union variables from the factor extraction data set causes deterioration with BN7 (north-west graph of the figure). With BN3, at the end of the sample excluding European Union variables increases RMSEs as well (north-east graph of the figure). On the other hand, excluding interest rates actually improves the forecast performance relative to the case where one uses all of the series, including interest rates.

For twelve month-ahead forecasts with BN7, forecast errors from using different data sets are close to each other (south-west graph of the figure). It is worth noting though that excluding financial variables results in higher forecast errors. Finally, for twelve month-ahead forecast with BN3, excluding interest rates increases the forecast error while excluding European Union variables decreases the errors (south-east graph of the figure).

In summary, in the short run European Union variables have forecasting power for industrial production while in the longer run interest rates and financial variables have forecasting power. These findings reveal that care is needed for constructing data sets while obtaining forecasts from factor models at different horizons.

5.5 Comparing Forecast Performance of Factor Models and Pool of Bivariate Equations

In the previous sections, some examples are presented from selected specifications that can be used for forecasting with factor model. In this type of graphical analysis, one modelling choice is changed at a time. While this strategy enables one to compare alternatives in a given set of modelling choices, effect of different sets of alternatives may not be mutually independent. For example, consider the question that whether aggregation level of the series used in data sets affects forecast performance. For this purpose, forecast errors from small, medium and large data sets are compared for different forecast equation types and number of factors. In the case of fixing the factor extraction approach, forecast equation type and criterion for the number of factors, it is presented how forecasts from small, medium and large data sets perform. In another section, one changes one of the choices that was fixed such as forecast equation type but this time fixing the data set size. However, effect of modelling choices may not be mutually independent. For example, for large data set with DI type forecast equation BN3, which suggest a relatively high number of factors, may produce lowest RMSE. However, for large data set using BN3 with DI-AR Lag one may get a relatively poor forecast performance. The intuition is that with DI-AR Lag and BN3 one needs to estimate a large number of parameters. This increase in parameter uncertainty may wipe out the benefit of using extra factors in the forecast equation. Hence, an exhaustive analysis taking into account different dimensions of modelling is necessary.

Bookkeeping for all of the available alternative specifications that can be used for forecasting reveals the importance of making a comprehensive comparison of forecast performance. There are 84 alternatives (7 criteria for the number of factors x 2 factor extraction approach x 2 forecast equation type x 3 data set size) for the factor models presented in the second section. For the pool of bivariate equations, there are 3 alternatives. For the factor models excluding data blocks one at a time there are 3x84=256 alternatives. In total, 84+3+256= 339 alternatives exist. Considering the simple benchmark, there are a total of 340 specifications. RMSEs for all of these specifications are calculated for three, six, nine, and twelve month-ahead horizons. In the previous sections it is seen from the graphs that relative performance is time-varying. In this respect, 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.

Four tables are presented. In the first two, top 5 specifications are shown (Table 5.1 and Table 5.1) while in the third and fourth the worst 5 are presented (Table 5.3 and Table 5.4). This exercise enables one to see whether there is a pattern in the best and the worst specifications. Following general points are worth highlighting:

- 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.
- 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.

- 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 specifications, 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.
- Modelling decisions affect forecasting performance of the factor models considerably. For example, in the best specifications 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.

Table 5.1. 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-Septr 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
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, Hallin, Lippi and Reichlin (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 2.4. 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 5.2. 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, Hallin, Lippi and Reichlin (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 2.4. 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

Table 5.3. 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: January 2010-September 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, Hallin, Lippi and Reichlin (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 2.4. 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 5.4 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, Hallin, Lippi and Reichlin (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 2.4. 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.

CHAPTER VI

6 FORECASTING CORE INFLATION

6.1 Introduction

In this chapter, focus is on analyzing performance of factor models for forecasting core inflation for a comprehensive set of specifications. Core inflation, which is the growth of the price level, is more persistent compared to industrial production growth. Papers reviewed in Chapter 2 show that core inflation and industrial production respond differently to the modelling decisions. In particular, using lags of the factors and the dependent variable helps at forecasting core inflation while this practice may even harm forecasting performance for industrial production. Effect of data set structure on forecasting performance may also change for core inflation, industrial production and stock market. This is due to the fact that data sets are quite heterogeneous in terms of data set. For example data blocks, such as interest rates and European Union variables, may play different roles on the forecasting of inflation and industrial production. In this respect, comparing effect of modelling decisions on forecast performance for core inflation and contrasting these findings with industrial production and stock market is expected to contribute to the literature.

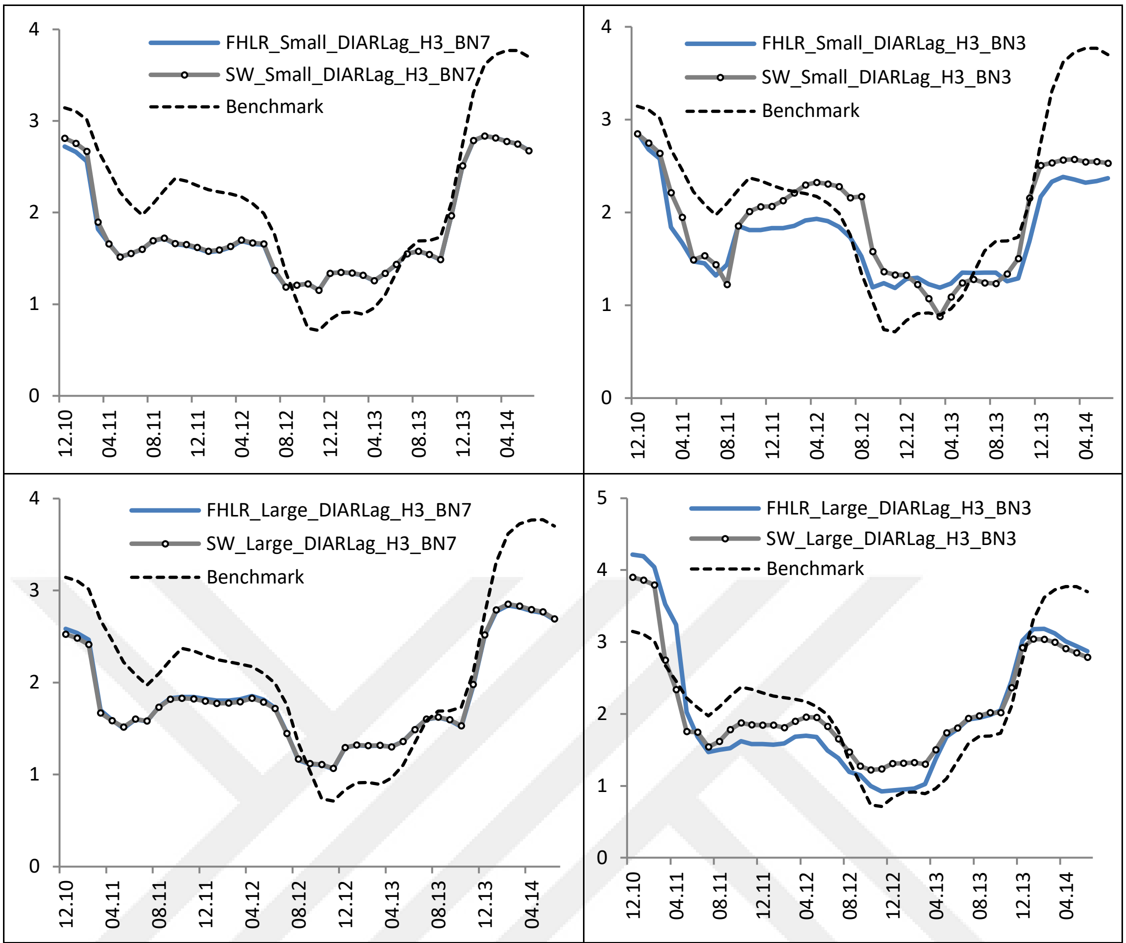
Chapter is organized as follows: First of all, forecasts are obtained with factor models using master data sets; small, medium and large. Then attention is devoted to an alternative method for using large data sets in forecasting, namely pooling of bivariate forecasts. Then, in the third section effect of data blocks on forecasting performance is analyzed. In these three sections, rolling twelve-month RMSEs are presented for the benchmark and factor models. Aim of this format, namely using graphical analysis, is to give a sense of the effect of modelling specifications on forecasting performance. Since relative performance may change over time, twelve month rolling RMSE are shown in the figures. This practice enables reader to see how a given set of specification affects forecast performance with respect to alternative specification and also with respect to the benchmark. There are various possible specifications, to be accurate 340 alternative specifications are used for forecasting the target variables. In the last section, results of a horse-race type analysis are presented. In particular, the five best and the five worst specifications are reported. Results are shown for three-, six-, nine- and twelve- month ahead forecasts for two forecast evaluation samples. This is also the section that results are discussed and interpreted.

6.2 Forecasting with Factor Models

6.2.1 FHLR vs SW

In this subsection, forecasting performance of FHLR and SW approaches to the factor extraction are compared. Analysis starts by discussing the results for core inflation for three-month-ahead forecasts with DI-AR Lag type forecast equation by changing the criterion for the number of factors and size of the data set (Figure 6.1).

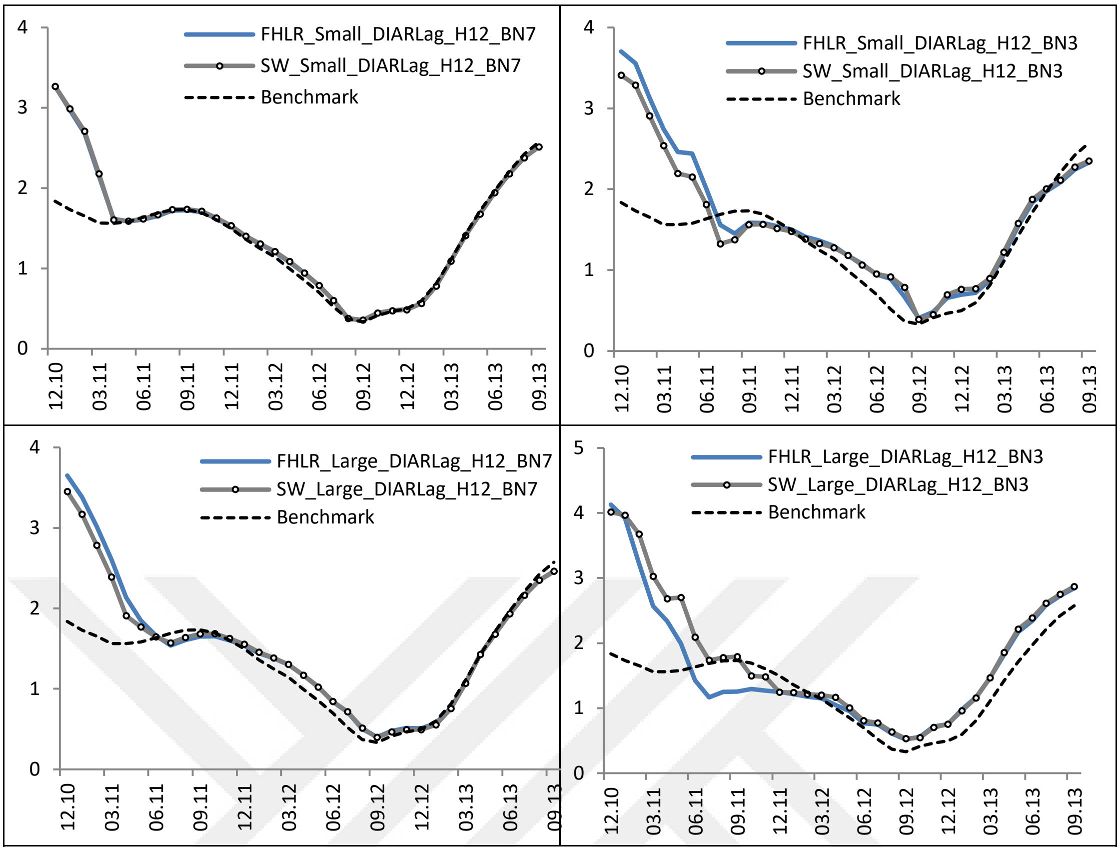
Graphs show that information criterion used for deciding the number of factors play a key role. In particular, forecast errors are close to each other for SW and FHLR approaches, for both small and large data sets when one uses the BN7 criterion for deciding the number of factors (north-west and south-west graphs in the figure). When one uses BN3 criterion for the number of factors, for both small and large data sets, relative performance of SW and FHLR approaches differ over time (north-east and south-east graphs of the figure). Compared to the benchmark, there is no clear winner. At the beginning and at the end of the sample, factor models perform better while in the middle of the sample they are unable to beat the benchmark.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach of Stock and Watson (2002b) and Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 shows the forecast horizon for three month ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 6.1. Rolling RMSEs for Comparing Factor Extraction Approach for Core Inflation: Three Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach for Stock and Watson (2002b) and Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H12 shows the forecast horizon for twelve month-ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 6.2. Rolling RMSEs for Comparing Factor Extraction Approach for Core Inflation: Twelve Month-Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

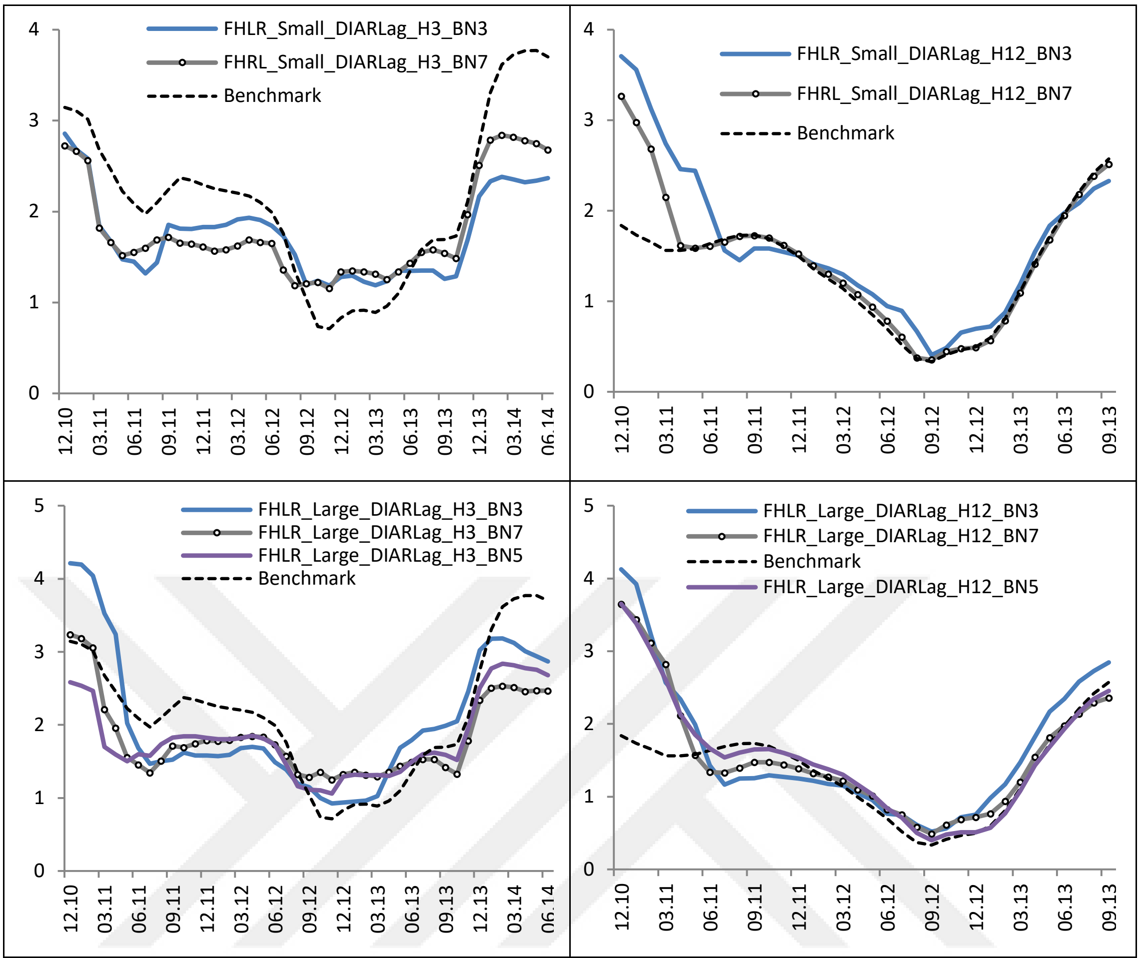
Analysis of three month ahead forecasts guides one about the short-term forecast performance. As the horizon increases, however, relative performance may change. Next, forecast performance is analyzed for twelve month-ahead forecasts (Figure 6.2). Inspecting the graphs show that in general performance of the factor models are similar to each other and RMSEs are also close to the benchmark. That being said, at the beginning of the sample for twelve month-ahead forecasts, FHLR approach performs worse than the SW approach for small data set (north-east graph of the figure) while performs better than SW approach for large data set (south-east graph of the figure). In summary, when alternative factor model approaches are compared forecast performances are close to each other.

6.2.2 Number of Factors

In this subsection, effect of information criteria for choosing the number of static factors on the forecasting performance is analyzed. Since the results for SW and FHLR approaches are close to each other for most of the cases, to save space, below only the models with FHLR approach are reported.

It is seen in Section 5.4 that BN3 gives around eight factors for large data set while BN7 gives around one factor. Using lags of those factors may increase parameter uncertainty especially for the BN3 case. So, it is important to take into account type of the forecast equation for analyzing effect of the number of factors on forecast performance. To that end, first of all the case of DI-AR Lag type forecast equation (Figure 6.3) and then the case of DI type forecast equation are studied (Figure 6.4). Comparing alternative information criteria, relative performance of BN3 and BN7 changes over time so there is no clear winner (Figure 6.3).

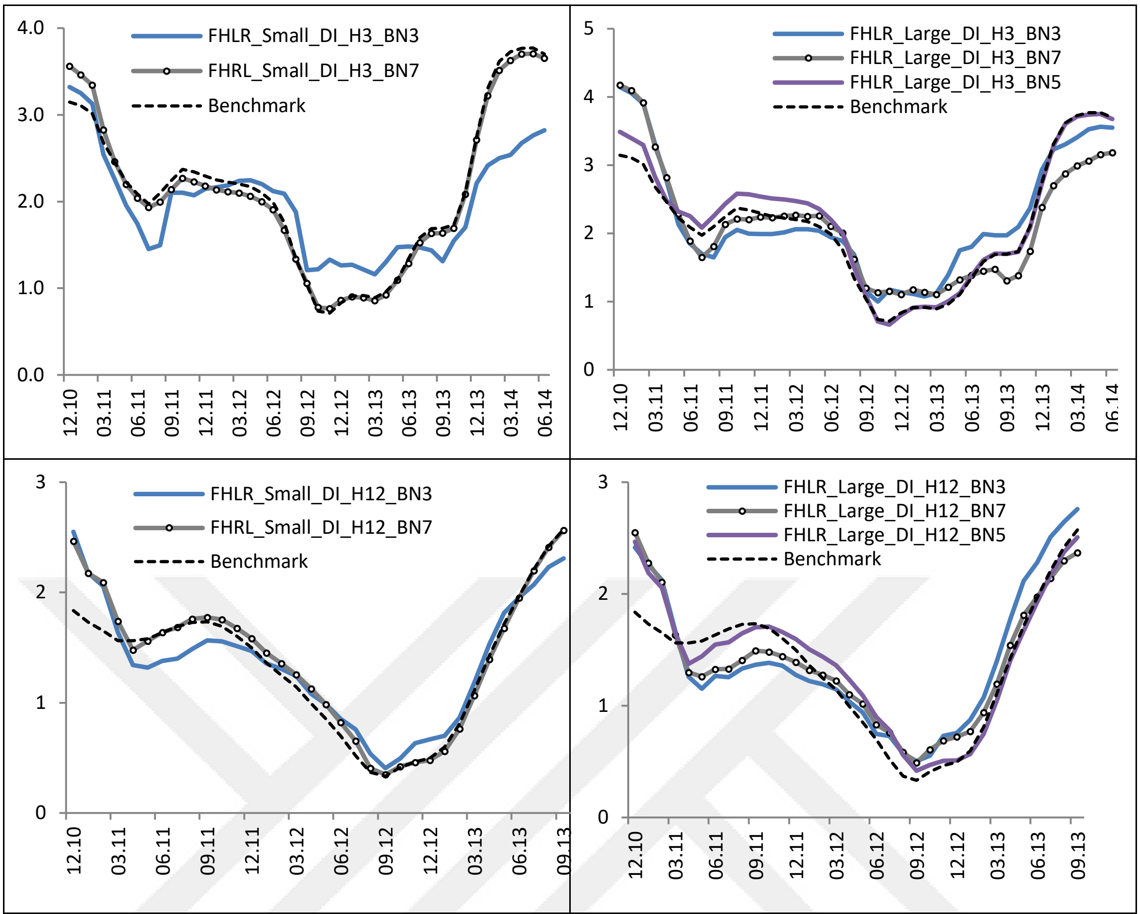
Next, analysis moves to the case with DI type forecast equations, where in the forecast equations one uses only contemporaneous factors (Figure 6.4). In this case, unlike DI-AR Lag case, more noticeable differences emerge for twelve month-ahead forecasts. This is due to the fact that in the DI-AR Lag case, forecast equation gets information from lags of the core inflation. Since, it is a persistent series once the lags of the core inflation is taken into account, effect of factors becomes less important. Overall analysis of the results does not imply a definitive winner. For example, with small data set for three month ahead forecasts at the beginning and at the end of the sample BN3 gives smaller forecast errors (north-west graph of the sample). However, in the middle of the sample using BN3 results in worse performance compared to both factor model with BN7 and compared to the benchmark.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 6.3. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Core Inflation: DI-AR Lag

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI refers to the forecast equation where one uses only contemporanous factors, H3 and H12 shows the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PIC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 6.4. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Core Inflation: DI

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

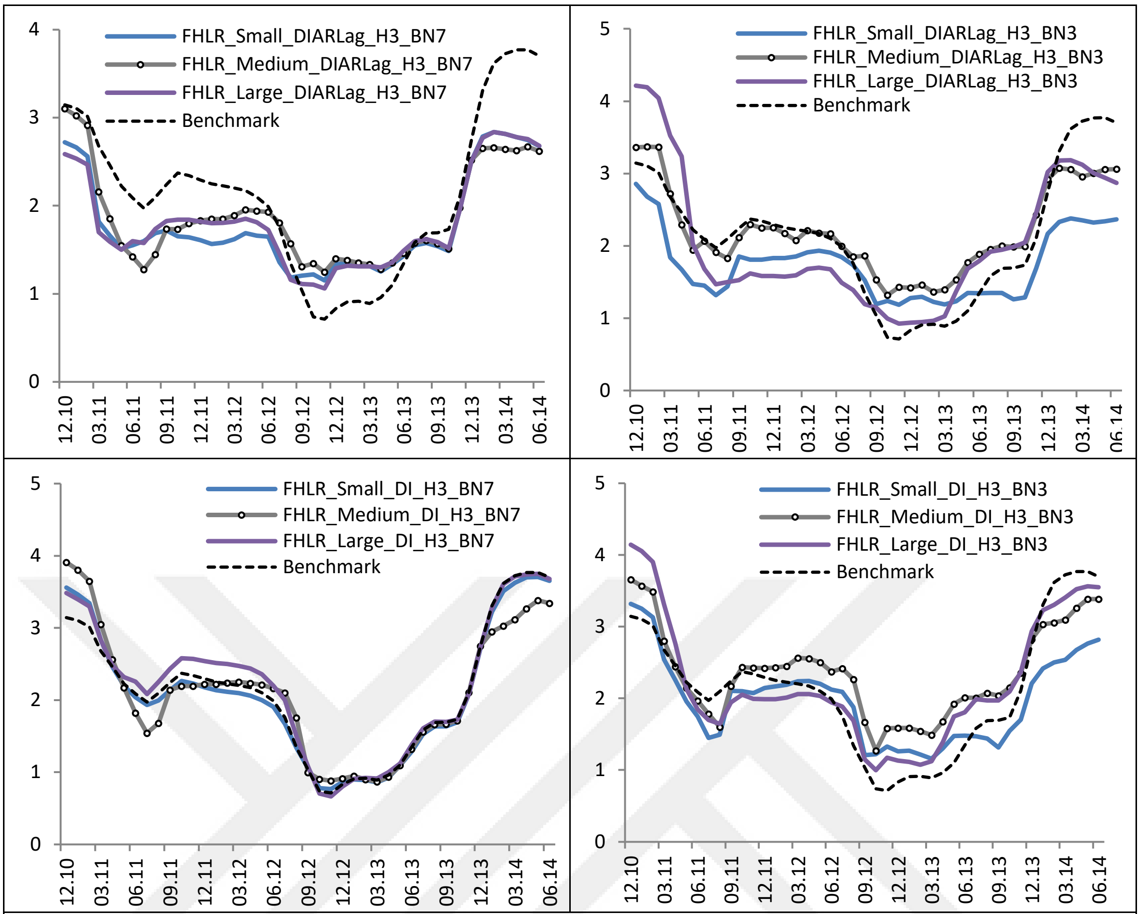
6.2.3 Data Set Size

In this section, focus is on comparing and contrasting forecasting performance for small, medium and large data sets, which are obtained by successively expanding the disaggregation level of the indicators.

Factor models are powerful tools for reducing dimension of the data. However, using a larger number of variables may not always result in lower forecast error. While it may not be possible to predict how forecast performance will react to the data set composition, increasing cross-correlation in the idiosyncratic terms after a certain threshold is not desirable in terms of theoretical considerations. So, results in this section is expected to give more insight about whether there exists a systematic pattern of forecast performance and the data set size.

First of all, three month-ahead forecasts are evaluated (Figure 6.5). When one uses BN7 criterion with small data set, both for DI-AR Lag (north-west graph of the figure) and DI (south-west graph of the figure) type forecast equations, RMSEs are close to each other compared to BN3 cases. For BN3 there are more noticeable differences both for DI-AR Lag (north-east graph of the figure) and DI (south-east graph of the figure). Similar to the cases that one compares alternative factor extraction approaches and criteria for the number of factors, there is no clear and consistent winner.

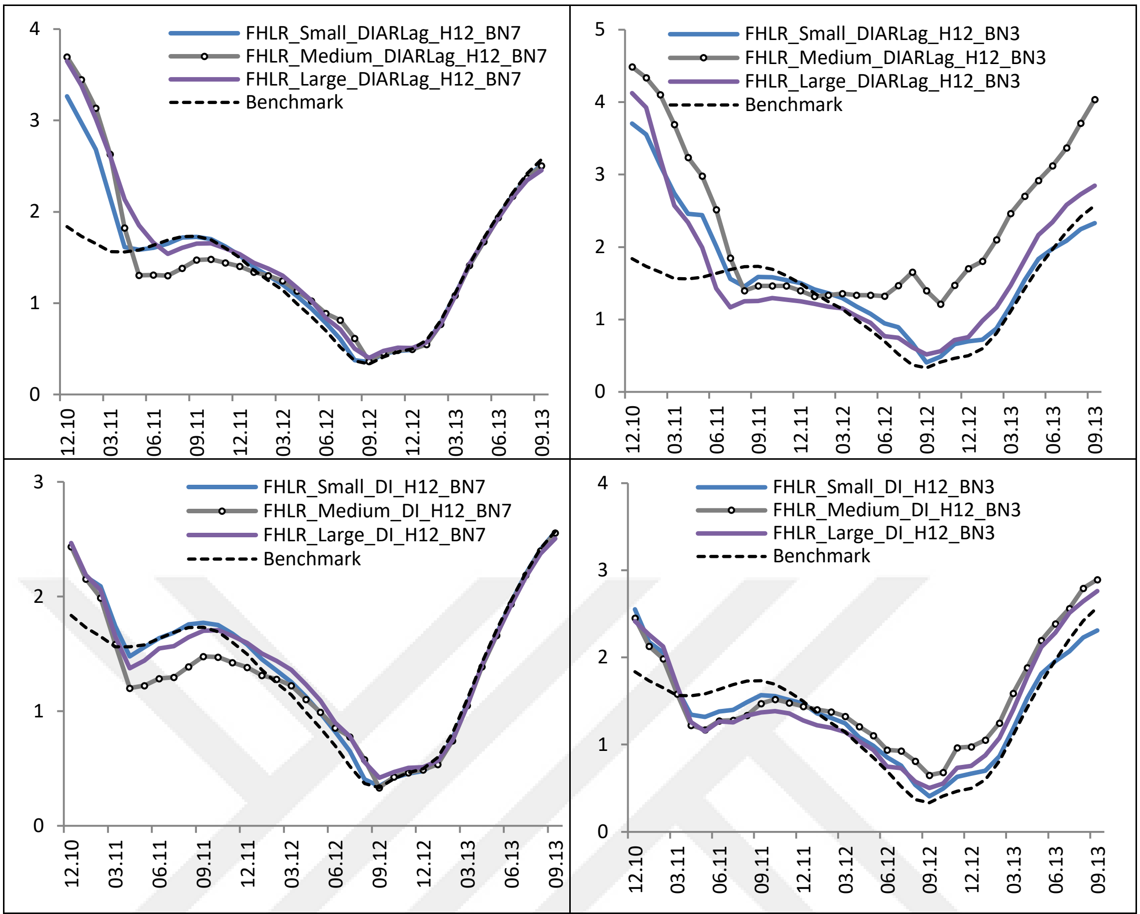
Next, relative forecast performance of different specifications are compared for twelve month-ahead forecasts (Figure 6.6). For this horizon with BN7 information criterion, forecasts errors are close to each other similar to the three month-ahead forecasts analyzed above. Compared to the cases with BN7, one gets more noticeable differences with BN3. For DI-AR Lag case (north-east graph of the figure), using medium data set results in higher forecast errors.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H3 shows the forecast horizon for three month-ahead forecast while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 6.5. Rolling RMSEs for Comparing Data Set Size for Core Inflation: Three Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



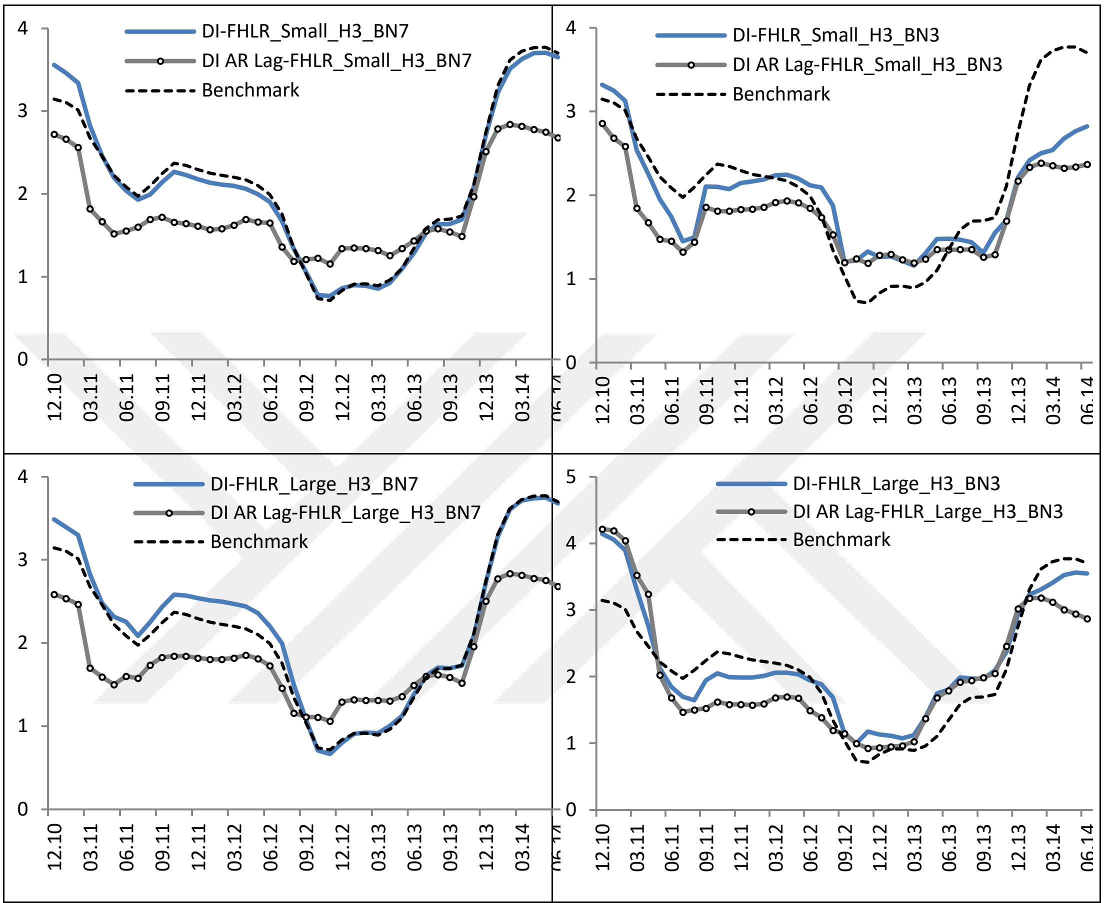
Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H12 shows the forecast horizons for twelve month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 6.6. Rolling RMSEs for Comparing Data Set Size for Core Inflation: Twelve Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

6.2.4 Forecast Equation Type

This section is devoted to the analysis of the effect of the type of forecast equation on forecast performance for three month-ahead forecasts.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI AR Lag refers to the forecast equation where one uses lags of the factors and the dependent variables, DI refers to the forecast equation that only the factors are used, H3 shows the forecast horizons for three month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

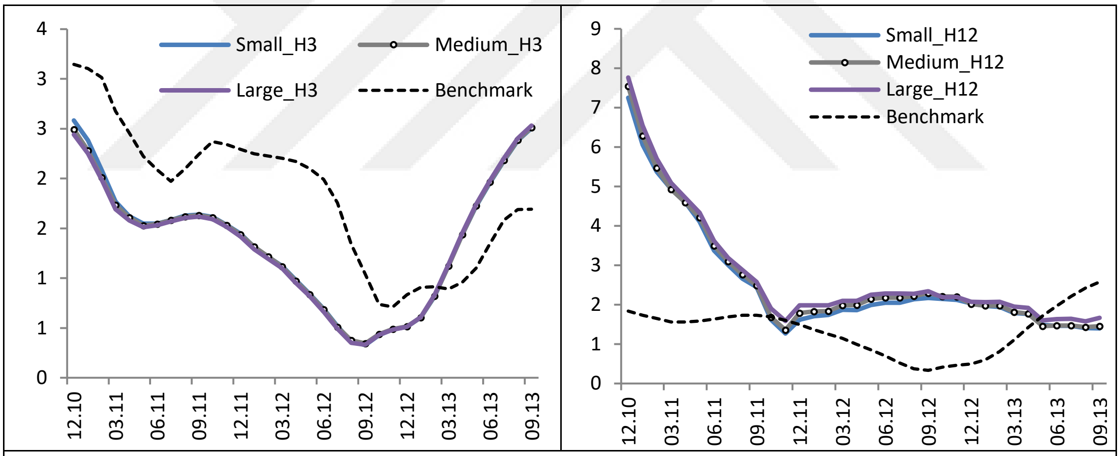
Figure 6.7. Rolling RMSEs for Comparing Forecast Equation Type for Core Inflation: Three Month-Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

Core inflation is more persistent relative to the industrial production growth. Hence, using lags of the core inflation may carry useful information. Comparing DI-AR Lag and DI type forecast equations for BN3/BN7 and small/large data sets, it is seen that at the beginning and at the end of the sample DI-AR Lag produces lower forecast errors than DI type forecast equation (Figure 6.7). However, in the middle of the sample using BN7 leads to lower forecast errors for DI.

6.3 Pooling of Forecasts from Bivariate Equations

Bivariate equations are estimated from each of the 22 series from the small data set, 63 series from the medium and 167 series from the large data set. Then, forecasts are obtained from these equations. Forecast combination is done by taking the average of the forecasts from these individual equations. This exercise enables one to compare the effect of pooling forecasts as opposed to factor models where one pools information.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, Small/Medium/Large refer to the data sets that forecasts are obtained from. H3/H12 refer to the forecast horizon for three and twelve month-ahead forecasts.

Figure 6.8. Rolling RMSEs for Comparing Pooling of Forecasts for Core Inflation

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

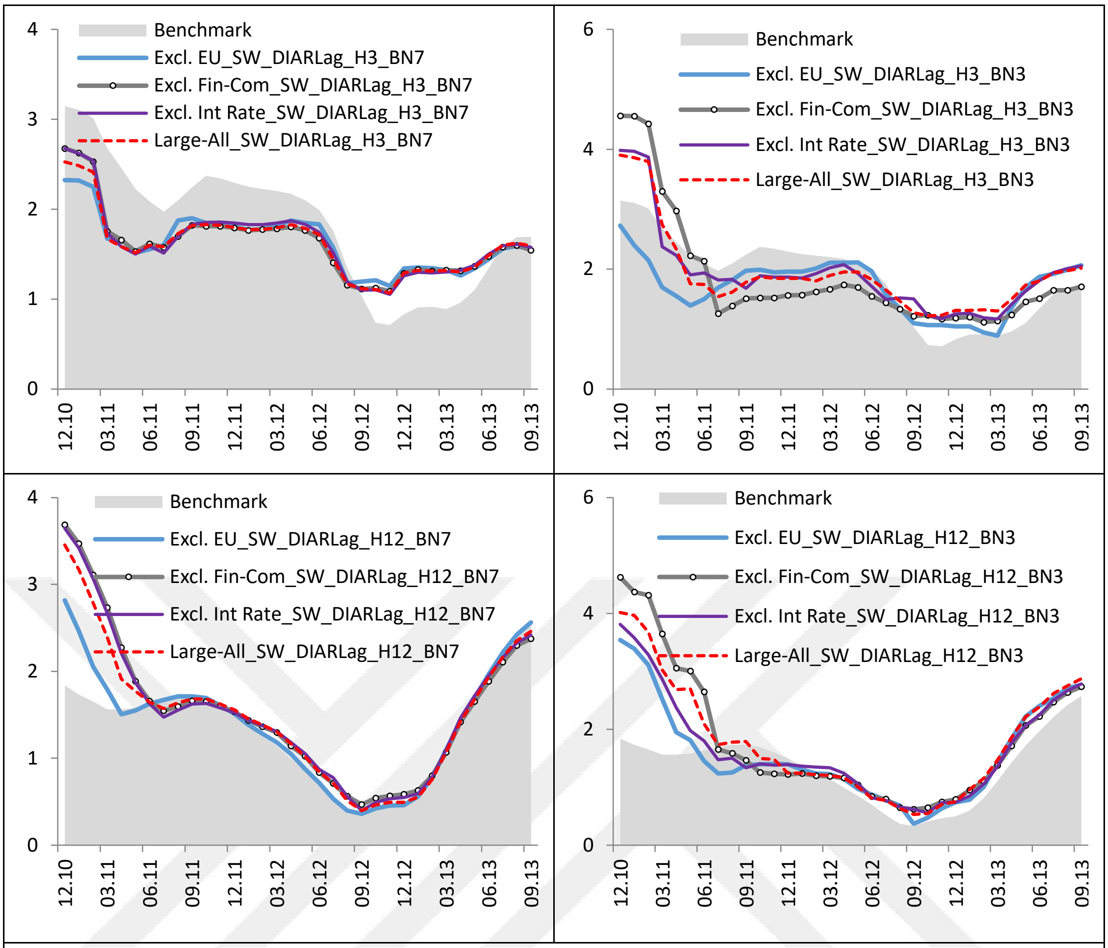
Results reveal that, for both three and twelve month-ahead horizons, forecast errors from three data sets are close to each other (Figure 6.8). In these graphs one cannot see whether pooling of forecasts results in lower forecast errors than using factor models. In Section 4, tables show the best and the worst specifications out of 340 specifications including factor models. If bivariate equations perform relatively well, they will appear in the tables.

6.4 Forecasts by Excluding Data Blocks for Factor Models

Factor models are powerful tools to deal with large number of variables. A few factors that summarizes information in the large data set can be used in the forecasting equation. Yet, there is no golden rule, rule of thumb or recipe for choosing the composition of the data set for macroeconomic applications. Hence, researchers construct data sets from different blocks such as real variables, prices and surveys. Moreover, it is not clear whether data set composition should change with the type of the series that one wants to forecast.

In the second section of this chapter, effect of the data set structure is discussed by focusing on the aggregation level of the series that are used in the data sets. By increasing the detail of the series within blocks, three data sets are constructed; small, medium and large. To some extent, this exercise enables one to see the effect of data set structure on the forecast performance. However, there are still important questions about the composition of the data set. Should one use a certain block at all and whether certain blocks are more important for forecasting than others? Following the practice in the literature to answer these questions, forecasts are obtained by excluding data blocks. In particular, four different data sets are constructed by excluding data blocks one at a time.

i. Master data sets (Large, Medium and Small data sets with all the variables discussed in Chapter 4)
ii. Data sets excluding European Union variables (Excl. EU)
iii. Data sets excluding commodities and financial variables (Excl. Fin.)
iv. Data sets excluding interest rates (Excl. Int. Rates)



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW refers to the factor extraction approach for Stock and Watson (2002b), Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 6.9. Rolling RMSEs for Comparing Factor Models with Excluding Data Blocks for Core Inflation

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

In the four subsections of Section 2, sensitivity of the forecast performance to the modelling choices is studied by considering different specifications for the master data sets. For the data sets that exclude certain blocks, one can redo the sensitivity analysis for the effect of factor extraction methodology, number of factors, data set size and the forecast equation. However, to save space, for the graphical analysis all these steps are not repeated. These specifications will be considered in the horse race of the models that will be presented in the next section.

In Figure 6.9 results are presented for the large master data and with the restricted data sets with DI-AR Lag type forecast equation. For three month-ahead forecasts, excluding European Union variables cause a noticeable decrease of forecast errors at the beginning of the sample for BN3 (north-east graph of the figure). With BN3, at the beginning of the sample, excluding financial variables increases forecast errors for both three- (north-east graph of the figure) and twelve month-ahead (south-east graph of the figure) forecasts. It should be noted however that relative performance changes over time. For example, for three month-ahead forecasts with BN3 (north-east graph of the figure), at the beginning of the sample excluding financial and commodity variables increases RMSE while at the end of the sample this practice decreases forecast errors. Hence, it is hard to come to a final verdict about definitive effect of certain data blocks on forecast performance.

6.5 Comparing Forecast Performance of Factor Models an Pool of Bivariate Equations

In the previous three sections, some examples are presented graphically for selected specifications that can be used for forecasting with factor models. In particular, some dimensions are fixed and it is analyzed how changes in certain modelling choices affect forecast performance. While this strategy enables one to compare alternatives in a given set of modelling choices, effect of modelling choices may not be independent. For example for large data set with *DI* type forecast equation using BN3, which suggest a relatively high number of factors, may produce the lowest RMSE. For large data set using BN3 with *DI-AR-Lag*, however, one may get a relatively poor forecast performance. The intuition is that with DI-AR-Lag and BN3 one needs to estimate a large number of parameters. This increase in parameter uncertainty may wipe out the benefit of using extra factors in the forecast equation.

Bookkeeping for all of alternatives reveal the importance of making a comprehensive comparison of models. There are 84 alternatives (7 criteria for the number of factors x 2 factor extraction approach x 2 forecast equation type x 3 data set size) for the factor models presented in the first section. For the pool of bivariate equations there are 3 alternatives. For the factor models excluding data blocks one at a time, there are 3x84=256 options. In total, there exists 84+3+256= 339 alternatives. Considering the simple benchmark, a total of 340 alternative models are considered. RMSEs are calculated for three, six, nine, and twelve month-ahead horizons for all of these specifications. It is seen in the graphs that relative performance is time-varying. In this respect, in the tables relative RMSEs are presented for two sub-periods. Namely, episode 1 is for January 2010-September 2011 and episode 2 is for Oct. 2011-Sept. 2013.

In this section, four tables are presented. In the first two, the top 5 specifications are shown (Table 6.1 and Table 6.2) while in the third and fourth, the worst 5 are presented (Table 6.3 and Table 6.4). 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 specifications while DI appears more frequently in the worst 5 specifications.
- Both in the best and worst five specifications 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.
- As shown in Chapter 5, PC3 and IC3 give a relatively large number of factors. They appear frequently in the worst 5 specifications. However, there are cases that they show up in the best models.

- 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.
- For the first evaluation sample, even the best specifications cannot beat the benchmark for twelve-month ahead forecasts.
- Modelling decisions affect forecasting performance of the factor models considerably. For example, in the best specifications one can get up to thirty percent improvement relative to the benchmark while for the same horizon deterioration up to 20 percent is observed.
- These points support the main hypothesis of this dissertation: before reaching a conclusion about the performance of factor models one needs to conduct a comprehensive analysis.

Table 6.1. 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	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, Hallin, Lippi and Reichlin (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 2.4. 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 6.2. 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: October 2011-September 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, Hallin, Lippi and Reichlin (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 2.4. 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 6.3 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: January 2010-September 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, Hallin, Lippi and Reichlin (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 2.4. 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 6.4 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: October 2011-September 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, Hallin, Lippi and Reichlin (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 2.4. 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.

CHAPTER VII

7 FORECASTING STOCK MARKET

7.1 Introduction

In this chapter, for a comprehensive set of specifications, forecasting performance of factor models are studied for the change in stock market index for three, six, nine and twelve month-ahead forecasts. Unlike industrial production or core inflation, regarding stock market forecasting there are more fundamental questions than which factor model specification would fit best to our forecasting needs. From the perspective of efficient market hypothesis, trying to forecast stock market better than a random walk for twelve month from now may not even make sense. Nevertheless, in addition to a real sector variable and a price indicator, applying empirical methodology described in Chapter 4 to a financial variable may provide useful insights. This comparison enables one to see how factor models behave for different type of series; real, price and financial. For the stock market growth, analysis based on the effect of data blocks on forecasting performance may be particularly informative.

Chapter is organized as follows: First of all, focus is on comparing and contrasting factor models using three master data sets; small, medium and large. Then an alternative method is analyzed for using big data sets in forecasting, namely pooling of bivariate forecasts. Then, in the fourth section effect of data blocks on forecasting performance is considered. In these three sections, rolling twelve-month RMSEs are presented for the benchmark and factor models. Aim of this format, namely using figures, is to give a sense of the effect of modelling specifications on forecasting performance. Since relative performance may change over time, RMSEs are presented in twelve month rolling windows. This practice enables reader to see how a given set of specification affects forecast performance with respect to alternative specifications and also with respect to the benchmark. Since there are various possible specifications, to be accurate 340 alternative specifications are used, it is not preferred to present all of the graphically. Rather, the best and the worst specifications are shown with tables. RMSEs relative to the benchmark are reported for three-, six-, nine- and twelve- month ahead horizons for two forecast evaluation samples. This is also the section that results are discussed and interpreted.

7.2 Forecasting with Factor Models

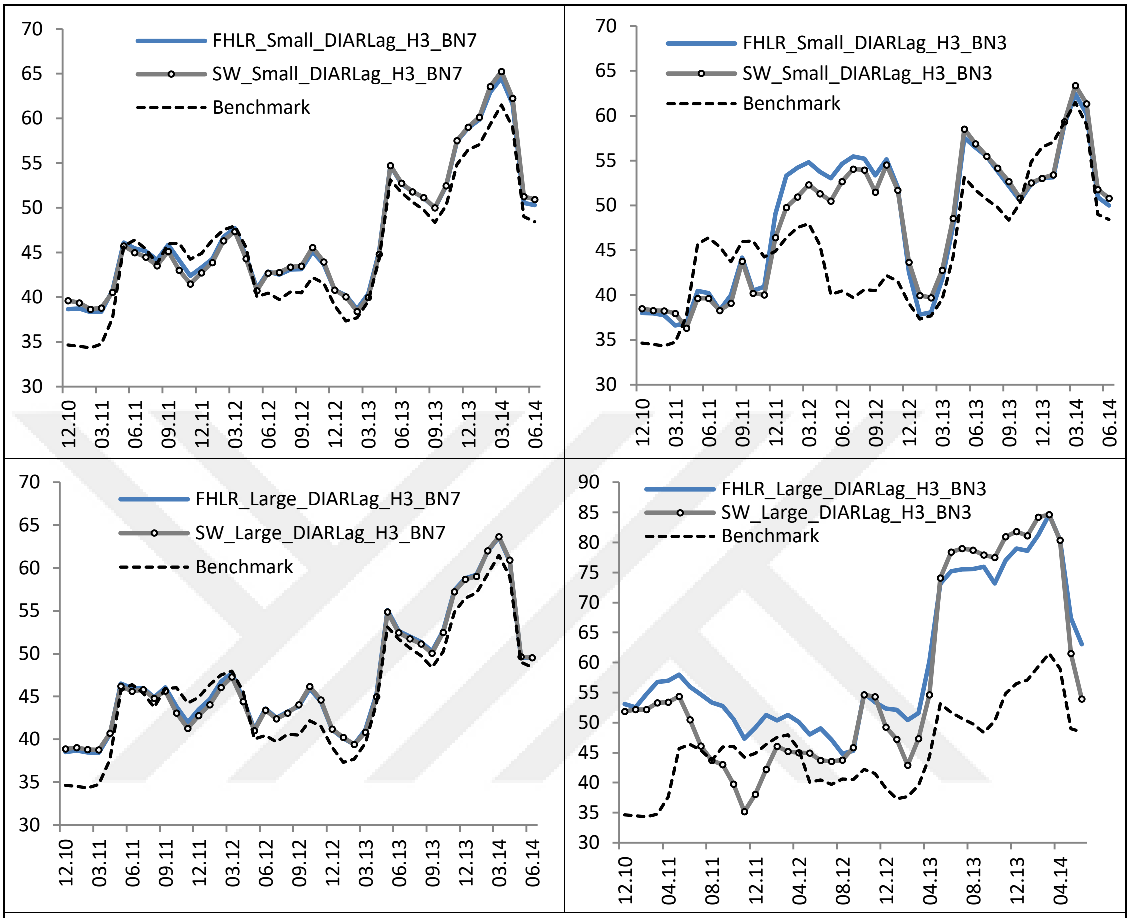
7.2.1 FHLR vs SW

In this subsection, FHLR and SW approaches to the factor extraction are compared. Discussion starts with the results for stock market for three-month-ahead forecasts with DI-AR Lag type forecast equation by changing the criteria for the number of factors and size of the data set (Figure 7.1).

Graphs show that except the specification using large data set and number of factors chosen with BN3 (south-east graph of the figure), SW and FHLR approaches result in similar forecast errors. For the specification with large data set and BN3, at the beginning of the sample SW approach performs better than FHLR and for a brief period also better than the benchmark. While there are periods that factor models beat the benchmark, careful analysis of the magnitude of the RMSEs reveal that forecast errors are very large. So, while one can beat the benchmark, forecasting performance is not promising for real life applications.

After analyzing three month-ahead forecasts, relative performance are presented for twelve month-ahead forecasts (Figure 7.2). Inspecting graphs shows that in general, specification with small data set and BN3 being an exception (north-east graph of the figure), performance of factor models are close to each other. For the specification with small data and BN3, at the beginning of the sample FHLR produces lower forecast errors while in the later part of the sample SW performs better than FHLR.

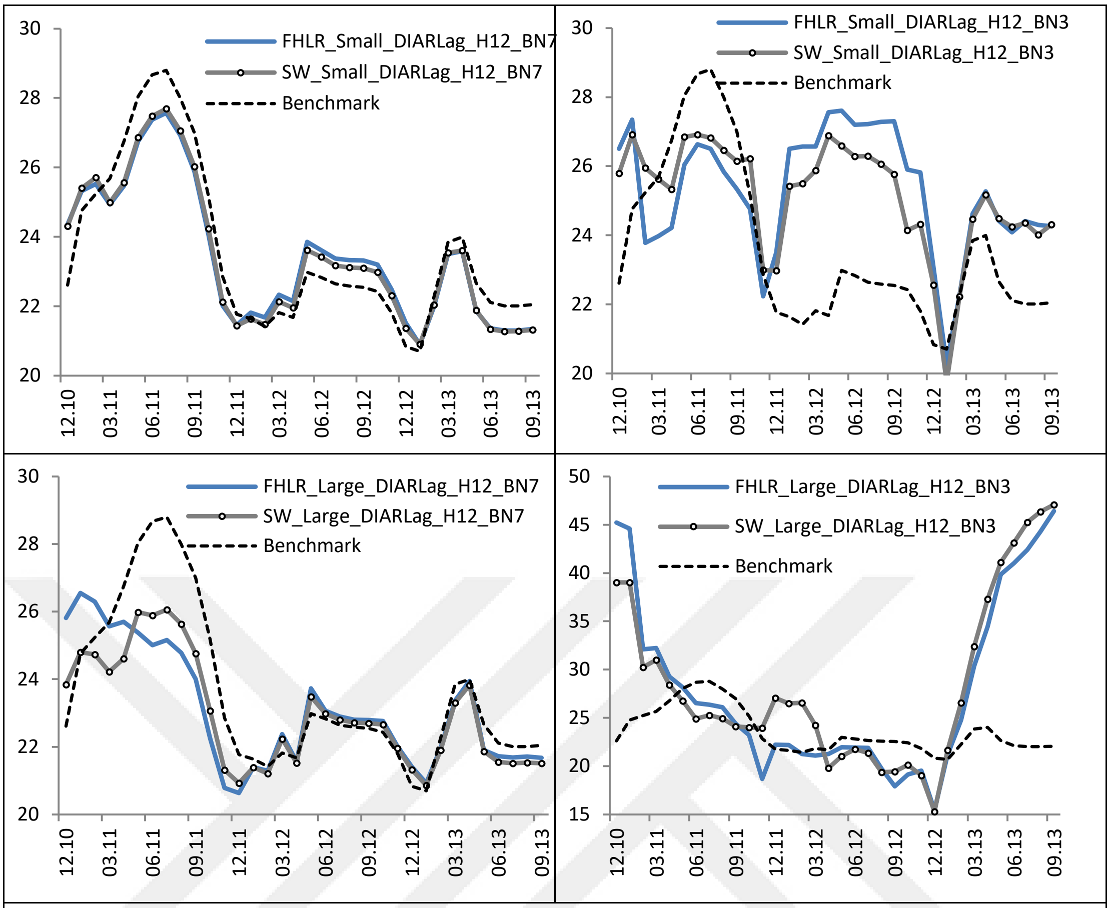
There are periods that factor models perform better than the benchmark. For the specifications that number of factors are chosen with BN7, more stable performance are observed compared to BN3. With large data set and BN7, at the beginning of the sample factor models beat the benchmark and then perform similarly. On the other hand, with large data set and BN3 after beating the benchmark both factor approaches perform considerable worse than it (south-east graph of the figure).



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach of Stock and Watson (2002b) and Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 shows the forecast horizon for three month ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 7.1. Rolling RMSEs for Comparing Factor Extraction for Stock Market: Three Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW and FHLR refer to the factor extraction approach for Stock and Watson (2002b) and Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H12 shows the forecast horizon for twelve month-ahead forecasts while BN3/BN7 show PC3/BIC3 information criterion from Bai and Ng (2002); for selecting number of static factors.

Figure 7.2. Rolling RMSEs for Comparing Factor Extraction Approach for Stock Market: Twelve Month-Ahead Forecasts

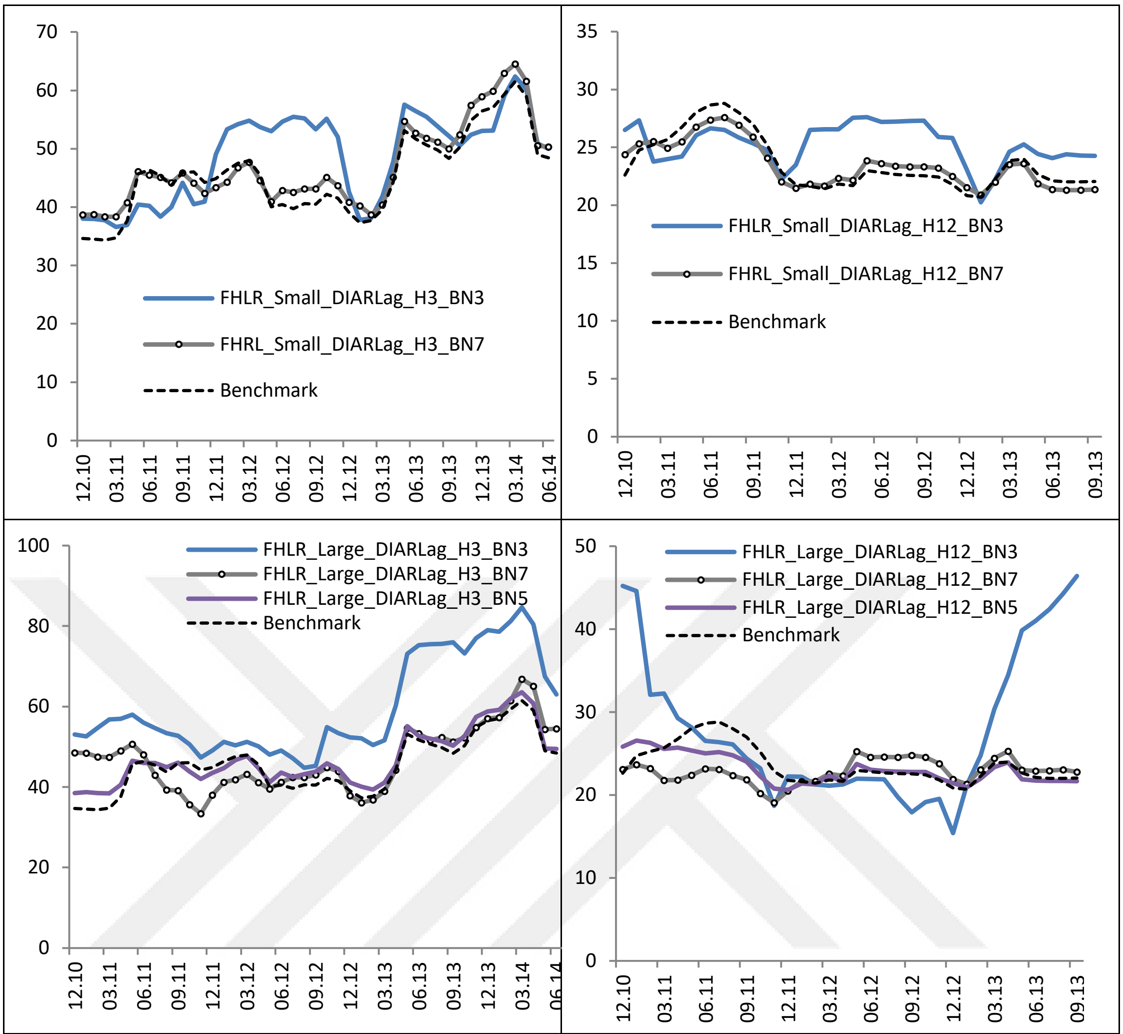
Source: Author's calculations based on pseudo-out-of sample forecasting exercise

7.2.2 Number of Factors

Number of factors is another input that one has to provide to the system for producing forecasts. This number can be specified in an ad-hoc manner. However, there are theoretical contributions that guide forecasters about the number of factors that should be extracted from the given data set. Bai and Ng (2002) offer seven criteria that can be used to decide the number of static factors. In this section focus is on analyzing the effect of information criteria for choosing the number of factors. Only models where factors are extracted with FHLR are reported to save space. Also, in the graphical analysis results are shown for BN3, BN5 and BN7. BN3 tends to deliver highest number of factors while BN7 tends to deliver lowest number of factors. First of all the case of DI-AR Lag type forecast equation is discussed and then the case of DI type forecast equation is studied.

Analysis of the forecast performance over time and across specifications show that using BN3 causes deterioration in the forecast performance (Figure 7.3). There are exceptions to this observation. However, these are short-lived improvements.

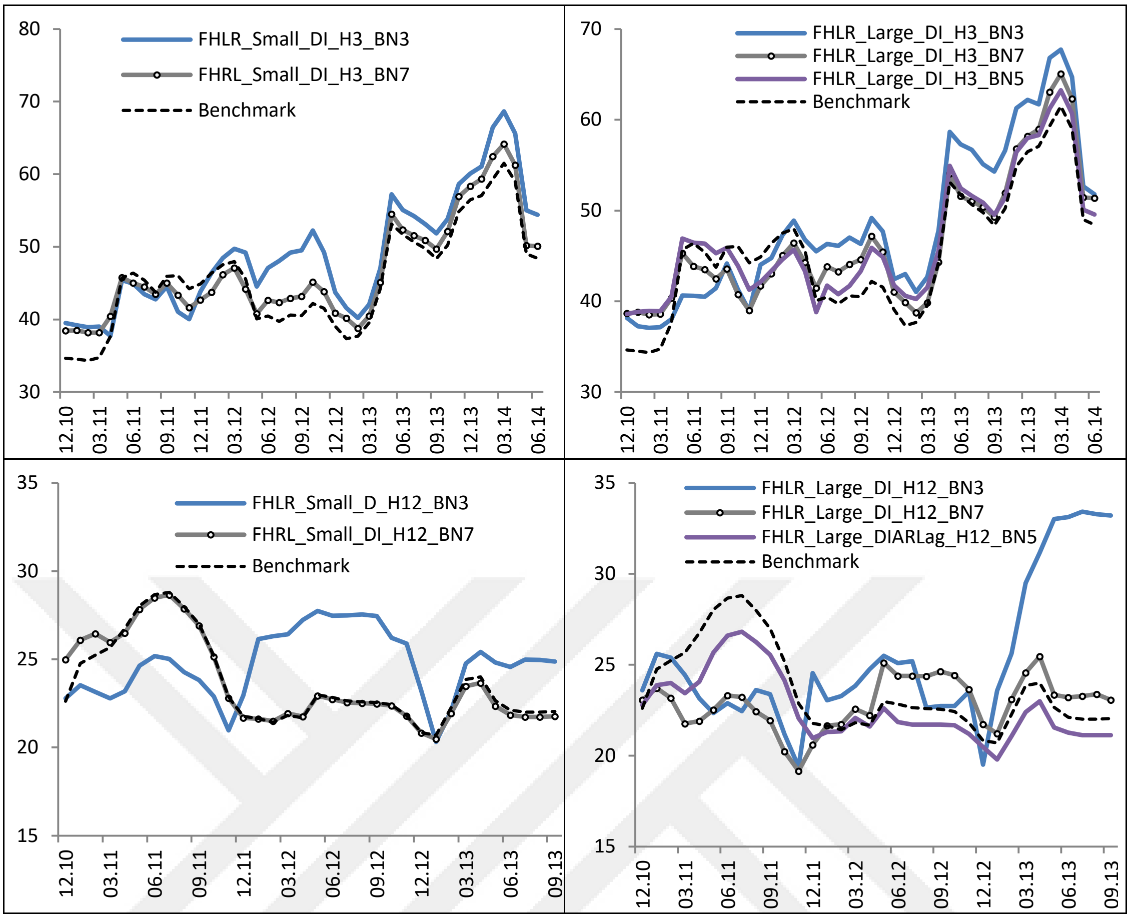
Next, specifications with DI type forecast equations which use only contemporaneous factors are analyzed (Figure 7.4). Most of the time, for three month-ahead forecasts BN3 still results in the worst performance both for small (north-west graph of the figure) and large (north-east part of the figure) data sets. For twelve month-ahead forecast, rather unstable results are obtained. For the specification with small data set and BN7, which suggest one factor, forecast errors are close to the benchmark (south-west graph of the figure). For twelve month-ahead forecasts with small data set, using BN3 criterion for deciding the number of factors results in a highly volatile picture. At the beginning of the sample, using BN3 reduces forecast errors relative to the factor forecasts with BN7 and relative to the benchmark (south-west graph of the figure). Yet, in the rest of the sample using BN3 results in considerably higher forecast errors. For twelve month-ahead forecasts with large data set, using BN3, BN5 and BN7 results in rather different forecast performance over time (south-east graph of the figure). At the beginning of the sample, BN7 results in lowest errors while towards the end of the sample BN5 performs best. It is worth noting that using BN3 increases forecast errors considerably at the end of the sample.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.3. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Stock Market :DI-AR Lag

Source: Author's calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refer to the factor extraction approach of Forni et al. (2005) , Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI refers to the forecast equation where one uses only contemporanous factors, H3 and H12 shows the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN5/BN7 show the information criteria PIC3/IC2/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.4. Rolling RMSEs for Comparing Information Criteria for the Number of Factors for Stock Market :DI

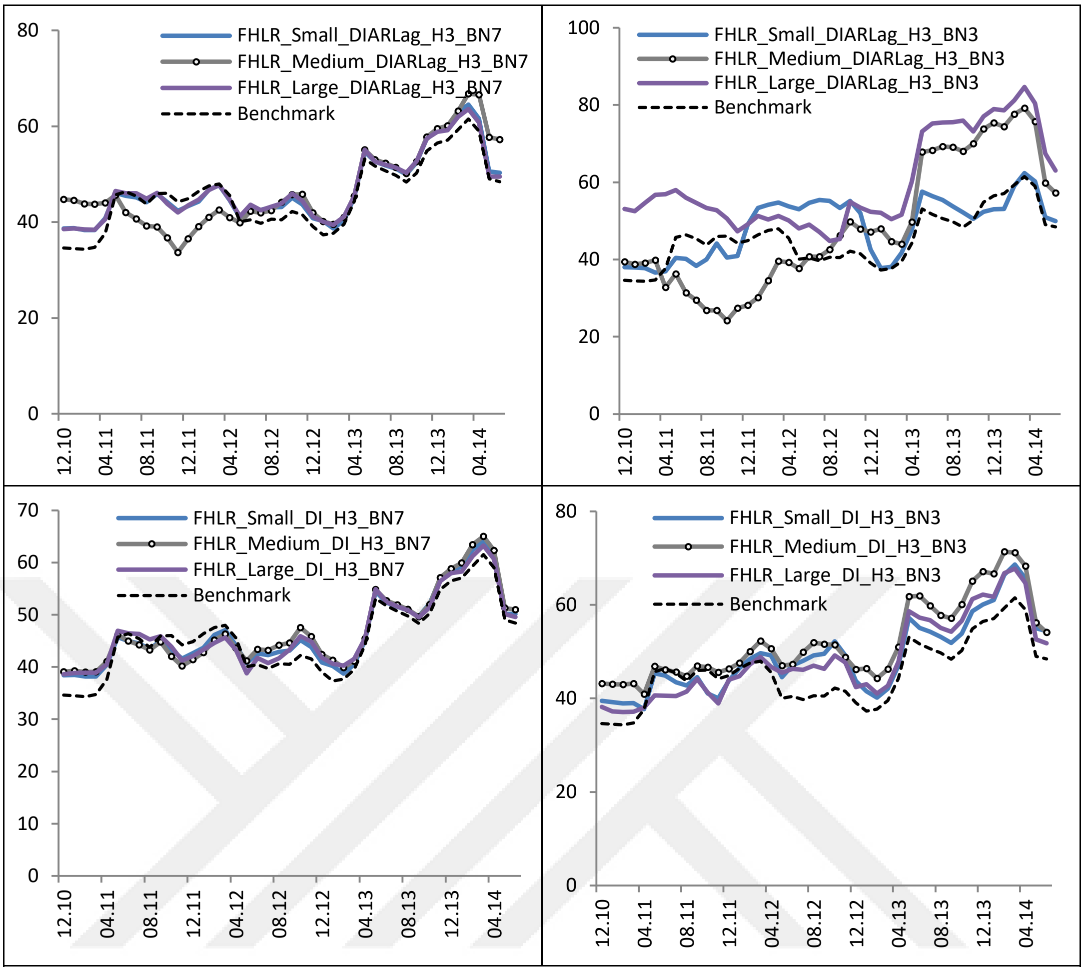
Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

7.2.3 Data Set Size

In this section, forecasting performances of specifications that use small, medium and large data sets are compared. Discussion starts with three month-ahead forecasts (Figure 7.5). When one uses BN7 criterion for selecting the number of factors, forecasts from medium data result in the lowest forecast error at the beginning of the sample while that specification is the worst performer at the end of the sample (north-west graph of the figure). For three month ahead forecasts with BN3 criterion, one gets rather unstable results (north-east graph of the figure). At the beginning of the sample, using medium data set results in the best forecast performance, which even beats the benchmark. However, at the end of the sample both forecasts from large and medium data sets perform considerably worse than the benchmark. At the end of the sample, it is observed that forecasts from small data set show the best performance.

Moving to the DI case, it is seen that with BN7 criterion, for all three data sets and the benchmark forecast performances are close to each other (south-west graph of the figure). With BN3 criterion, one still gets volatile and time-varying relative performance (south-east graph of the figure). It should be noted, though, that volatility is less than the DI-AR Lag case.

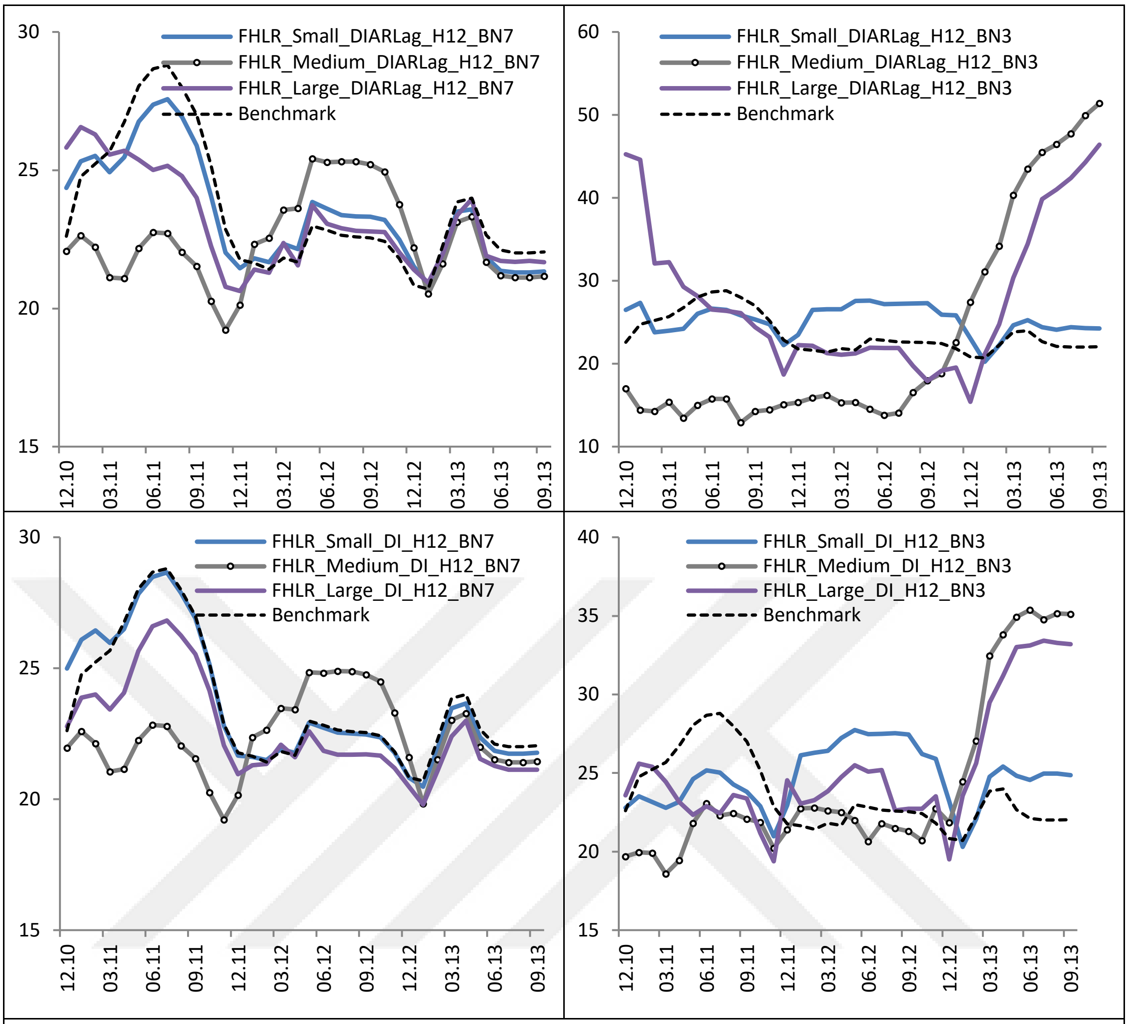
Next, relative performance of three data sets are evaluated with different specifications for twelve month-ahead forecasts (Figure 7.6). For this case, highly volatile relative performances are observed. With BN7 criterion, both for DI-AR Lag (north-west graph of the figure) and DI (south-west graph of the figure), at the beginning of the sample using medium data set results in the lowest forecast error. However, in the middle of the sample specification with medium data set performs worst. At the end of the sample, all three factor models and the benchmark perform similarly. With BN3 criterion, volatility of the relative performance increases for all three data sets (north-east and south-east graphs of the figure). It is noted that at the end of the sample no specification beats the benchmark with BN3 while with BN7 reverse is true.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H3 shows the forecast horizon for three month-ahead forecast while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.5. Rolling RMSEs for Comparing Data Set Size for Stock Market: Three Month-Ahead Forecasts

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Medium/Large refer to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, DI refers to the specifcation where one uses only the factors in the forecast equation, H12 shows the forecast horizons for twelve month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.6. Rolling RMSEs for Comparing Data Set Size for Stock Market: Twelve Month-Ahead Forecasts

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

7.2.4 Forecast Equation Type

In this section, focus is on analyzing the effect of forecast equation on forecast performance for three month-ahead forecasts (Figure 7.7).

Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, FHLR refers to the factor extraction approach of Forni et al. (2005), Small/Large refer to the aggregation detail of the data set that the factors are extracted from, DI AR Lag refers to the forecast equation where one uses lags of the factors and the dependent variables, DI refers to the forecast equation that only the factors are used, H3 shows the forecast horizons for three month-ahead forecasts while BN3/BN7 show the information criteria PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.7. Rolling RMSEs for Comparing Forecast Equation Type for Stock Market: Three Month-Ahead Forecasts

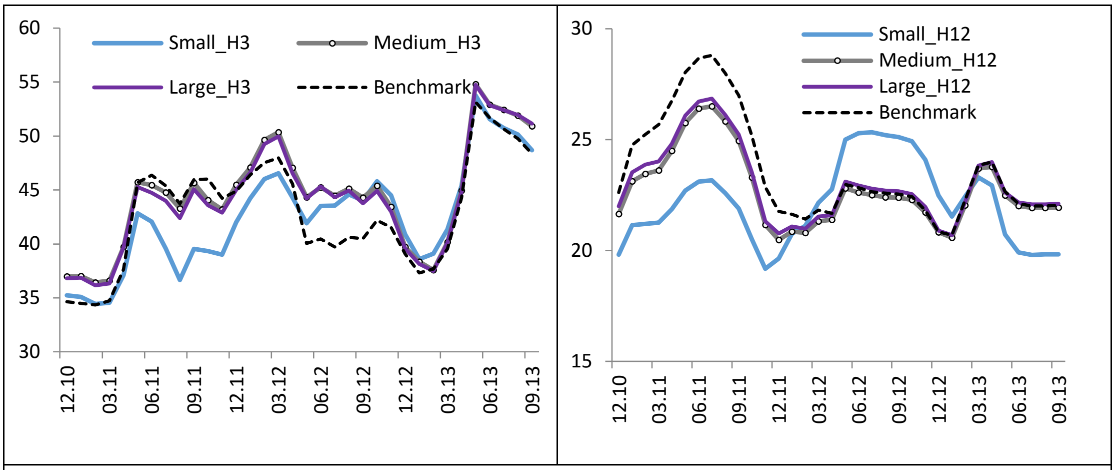
Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

With BN7 criterion, both for small (north-west graph of the figure) and large data sets (south-west graph of the figure), relative performances are close to each other. Using BN3 criterion, it is observed that relative performance for DI-AR-Lag and DI changes over time. In particular, for the small data set (north-east graph of the figure), DI-AR Lag produces the lowest forecast errors at the beginning and at the end of the sample while in the middle of the sample, it performs worst. For the large data set, picture is more homogenous (south-east graph of the figure). To be more concrete it is seen that using DI-AR Lag increases the forecast errors.

7.3 Pooling of Forecasts

Bivariate equations are estimated from each of the 22 series from the small data set, 63 series from the medium and 167 from the large data set. Then, forecasts are obtained from these equation in the fashion of out-of-sample forecasting. For the forecast combination, average of the forecasts are calculated from the forecasts of individual equations. This exercise enables one to compare the effect of pooling forecasts as opposed to factor models where pooling information is the main strategy.

Results show that unlike industrial production and inflation cases, pooling of forecasts from small data behave differently than other two data sets. Using small data set results in lowest forecast error at the beginning of the sample both for three and twelve month-ahead forecasts (Figure 7.8). For twelve month-ahead forecasts, in the middle of the sample it is the worst performer.



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, Small/Medium/Large refer to the data sets that forecasts are obtained from. H3/H12 refer to the forecast horizon for three and twelve month-ahead forecasts.

Figure 7.8. Rolling RMSEs for Comparing Pool of Forecasts for Stock Market

Source: Author’s calculations based on pseudo-out-of sample forecasting exercise

7.4 Excluding Data Blocks

Factor models are powerful tools to deal with large number of series. A few factors that summarizes information in a large data set can be used in the forecasting equation. Yet, there is no golden rule, rule of thumb or recipe for choosing the composition of the data set for macroeconomic applications. Hence, researchers construct data sets using variables from different blocks such as real sector, prices and surveys. Moreover, it is not clear whether data set composition should change with the type of the series that one wants to forecast.

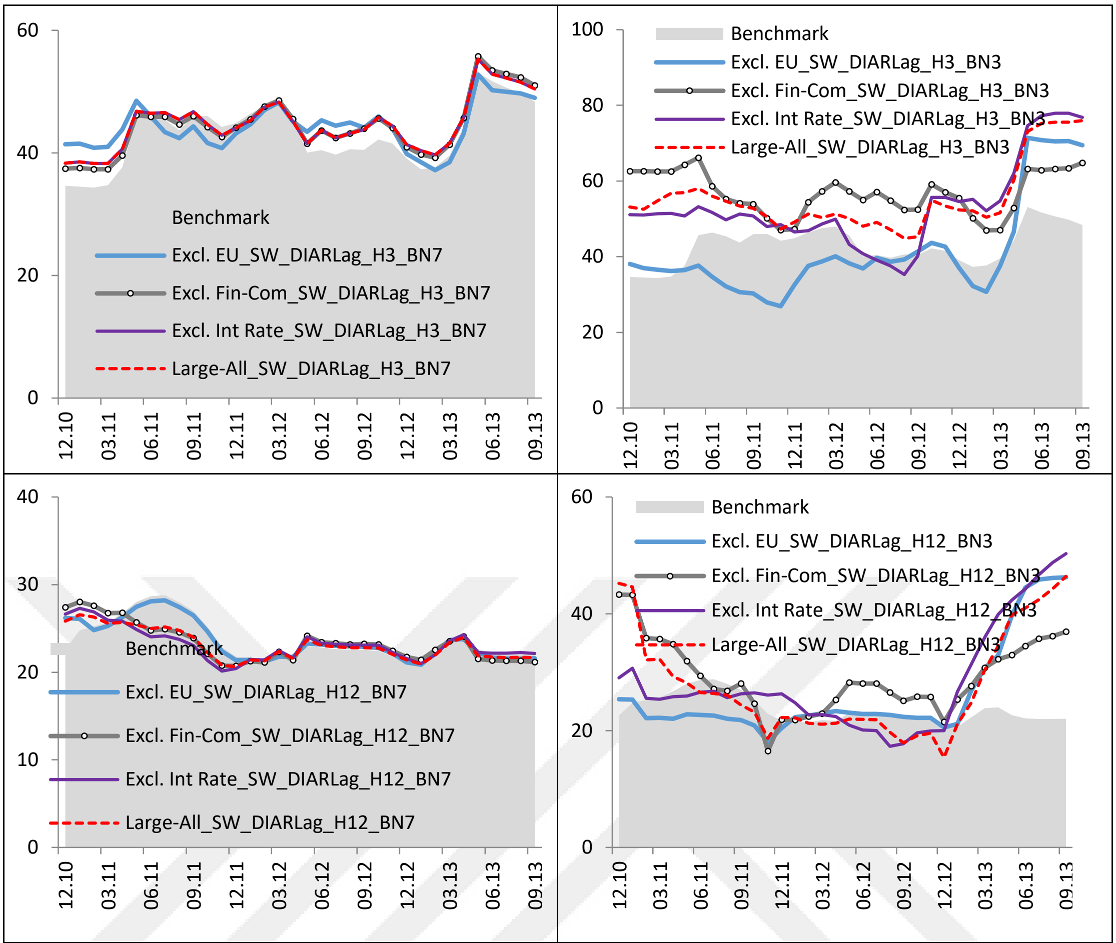
In the second section of this chapter, effect of the data set structure is discussed by focusing on the aggregation level of the series that are used in the data sets. By increasing the detail of the series within blocks, three data sets are constructed; small, medium and large. To some extent, this exercise enables one to see the effect of data set structure on forecast performance. However, there are still important questions about the composition of the data set. Should one use a certain block at all and whether certain blocks are more important for forecasting than others? Following the common practice in the literature, to answer these questions forecasts are obtained by excluding data blocks one at a time. In particular, four different data sets are constructed by excluding data blocks.

i. Master data sets (Large, Medium and Small data sets with all the variables discussed in Chapter 4)
ii. Data sets excluding European Union variables (Excl. EU)
iii. Data sets excluding commodities and financial variables (Excl. Fin.)
iv. Data sets excluding interest rates (Excl. Int. Rates)

In the second section of this chapter, sensitivity of the forecast performance of factor models are analyzed for different specifications for the master data sets. For data sets that are produced by excluding certain blocks, one can redo the analysis conducted in four subsections: effect of factor extraction methodology, number of factors, data set size and the forecast equation. To save space, for the graphical analysis, these steps are not repeated. Results are discussed for large data set and three data sets that are obtained by excluding certain blocks. In the last section, a more detailed analysis will be presented for the relative forecast performance of these specifications.

In Figure 7.9, results are presented for the large data set with DI-AR Lag type forecast equation. For three month-ahead forecasts, when BN7 is used for deciding the number of factors, by excluding European Union variables one gets lower RMSEs at the end of the sample compared to other data sets (north-west graph of the figure). Effect of European Union variables on the forecast performance is more striking when BN3 is used for selecting the number of factors (north-east graph of the figure). In this case, by excluding European Union variables one can even beat the benchmark. On the other hand until the end of the sample excluding financial variables harm the forecasting performance.

For twelve month-ahead forecasts, mixed results are obtained for the BN3 case. In particular, while excluding financial variables increase RMSE and excluding European Union variables decrease RMSE at the beginning of the sample, reverse is true at the end of the sample (south-east graph of the figure).



Notes: In the graphs, Root Mean Squared Errors (RMSE) are presented. RMSEs are obtained from pseudo out-of-sample forecasting exercise. 12-month moving RMSEs are plotted to see the performance over time. In the legends, SW refers to the factor extraction approach for Stock and Watson (2002b), Large refers to the aggregation detail of the data set that the factors are extracted from, DIARLag refers to the forecast equation where one uses lags of the factors and the dependent variable, H3 and H12 show the forecast horizons for three month- and twelve month-ahead forecasts while BN3/BN7 show the information criteria of PC3/BIC3 from Bai and Ng (2002) for selecting number of static factors.

Figure 7.9. Rolling RMSEs for Comparing Factor Models with Excluding Data Blocks for Stock Market: DI-AR Lag

Source: Author's calculations based on pseudo-out-of sample forecasting exercise

7.5 Comparing Forecast Performance Factor Models and Pool of Bivariate Equations

In the previous three sections, examples are presented for selected specifications that can be used for forecasting with factor models. In particular, some dimensions of the modelling process is fixed and how changes in certain modelling choices affect forecast performance is analyzed. While this strategy enables one to compare alternatives in a given set of modelling choices, effect of different sets of alternatives may not be mutually independent. Moreover, for a clear comparison in a graphical analysis one modelling choice is changed at a time. For example, consider the question that whether aggregation level of the series used in data sets affects forecast performance. Comparisons are made for small, medium and large data sets by fixing factor extraction approach, forecast equation type and criterion for the number of factors. In another section, a case is presented by changing one of the choices that was fixed such as forecast equation type but this time by fixing the data set size.

Effect of modelling choices may not be mutually independent. For example, for large data set with DI type forecast equation BN3, which suggest a relatively high number of factors, may produce the lowest RMSE. However, for large data set using BN3 with DI-AR-Lag one may get relatively poor forecast. The intuition is that with DI-AR-Lag and BN3, one needs to estimate a large number of parameters. This increase in parameter uncertainty may wipe out the benefit of using extra factors in the forecast equation. Hence, an extensive analysis taking into account different dimensions of modelling is necessary.

Bookkeeping for all of the available alternative specifications that can be used for forecasting reveal the importance of making a comprehensive comparison of forecasts. There are 84 alternatives (7 criteria for the number of factors x 2 factor extraction approach x 2 forecast equation type x 3 data set size) for the factor models presented in the first section. For the pool of bivariate equations, there are 3 alternatives. For the factor models excluding data blocks one at a time, 3x84=256 alternatives exist. In total, there are 84+3+256= 339 alternatives. Considering the simple benchmark, there are 340 alternative models. RMSEs are calculated for three, six, nine, and twelve month-ahead horizons. In the previous sections, it is observed from the graphs that relative performance is time-varying. In this respect, tabulations of relative RMSEs are presented for two sub-periods. Namely, episode 1 is for January 2010-September 2011 and episode 2 is for October 2011-September 2013.

In this section, four tables are presented. In the first two, the top 5 specifications are shown (Table 7.1 and Table 7.2) while in the second the worst 5 are shown (Table 7.3 and Table 7.4). This exercise enables one to see whether there is a pattern in the best and worst specifications. Following points are worth highlighting:

- Considering the best performing specifications, for the first evaluation sample it is observed that DI-AR-Lag type forecast equation appears relatively more frequently while for the second evaluation sample almost all of the specifications use DI type forecast equation. Analysis of the worst performing specifications show that DI-AR Lag type forecast equation appears relatively more frequently.
- Comparing SW and FHLR approaches, except for the first evaluation sample for the twelve month-ahead forecasts, two approaches perform similarly in terms of size of the RMSE.
- In the literature review, it is noted that IC1 and IC2 are more frequently used for deciding the number of factors. While Bai and Ng (2002) point out the promising performance of BIC3, its use in practice is rare. For the second evaluation sample, all these three criteria appear frequently while for the first evaluation sample PC3/IC3 are also seen. In the worst performing specifications PC3 and IC3 appear relatively more frequently.
- In the best specifications, it is observed that excluding European Union variables decrease RMSE. For twelve month-ahead forecasts, in the second evaluation sample excluding interest rates or financial variables cause a reduction in the forecast errors.
- For the best models, except three month-ahead forecasts in the second evaluation sample, factor forecasts beat the benchmark. However, as graphical analysis shows errors are very large, around 30 percentage points for twelve

month ahead forecasts. Hence, this result does not mean that one can use factor models for forecasting stock market in practical applications.

- Modelling decisions affect forecasting performance of the factor models considerably. For example, in the best equations one can get thirty-five percent improvement relative to the benchmark while for the worst specifications one may get seventy-five percent higher RMSE relative to the benchmark.
- These points support the main hypothesis of this dissertation: before reaching a conclusion about the performance of factor models we need to conduct a comprehensive analysis.

Table 7.1. Rankings of the Models for Stock Market (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: January 2010-September 2011
Three Month-Ahead						
1	DI_AR_Lag	FHLR	PC3	M=H=16	Large/Excl. EU	0.835
2	DI_AR_Lag	FHLR	PC1	M=H=16	Large/ Excl. Int. Rates	0.847
3	DI_AR_Lag	FHLR	PC3	M=H=16	Medium/All	0.864
4	DI_AR_Lag	FHLR	IC3	M=H=16	Medium/All	0.864
5	DI_AR_Lag	SW	PC1	-	Large/ Excl. Int. Rates	0.883
Six Month-Ahead						
1	DI	SW	PC1	-	Small/Excl. EU	0.792
2	DI	SW	PC2	-	Small/Excl. EU	0.792
3	DI	SW	PC3	-	Small/Excl. EU	0.792
4	DI	SW	IC1	-	Small/Excl. EU	0.792
5	DI	SW	IC2	-	Small/Excl. EU	0.792
Nine Month-Ahead						
1	DI_AR_Lag	FHLR	PC1	M=H=16	Large/ Excl. Int. Rates	0.655
2	DI_AR_Lag	FHLR	PC2	M=H=16	Large/ Excl. Int. Rates	0.664
3	DI_AR_Lag	FHLR	IC3	M=H=16	Large/ Excl. Int. Rates	0.690
4	DI_AR_Lag	SW	PC1	-	Medium/Excl. EU	0.720
5	DI_AR_Lag	SW	IC1	-	Medium/Excl. EU	0.728
Twelve Month-Ahead						
1	DI_AR_Lag	FHLR	PC1	M=H=16	Medium/Excl. EU	0.591
2	DI_AR_Lag	FHLR	PC3	M=H=16	Medium/All	0.739
3	DI_AR_Lag	FHLR	IC3	M=H=16	Medium/All	0.739
4	DI_AR_Lag	SW	PC1	-	Medium/Excl. EU	0.742
5	DI_AR_Lag	SW	PC3	-	Medium/All	0.778

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, Hallin, Lippi and Reichlin (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 2.4. 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 7.2. Rankings of the Models for Stock Market (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: October 2011-September 2013
Three Month-Ahead						
1	DI_AR_Lag	FHLR	PC1	M=H=16	Medium/Excl. EU	0.994
2	Benchmark	-	-	-	-	1.000
3	DI_AR_Lag	SW	IC1	-	Medium/Excl. EU	1.010
4	DI_AR_Lag	SW	IC2	-	Medium/Excl. EU	1.010
5	DI	FHLR	IC2	M=H=16	Large/Excl. EU	1.024
Six Month-Ahead						
1	DI	FHLR	IC2	M=H=16	Medium/All	0.944
2	DI	FHLR	IC1	M=H=16	Medium/All	0.947
3	DI	FHLR	IC1	M=H=16	Medium/Excl. EU	0.952
4	DI	FHLR	IC2	M=H=16	Medium/Excl. EU	0.952
5	DI	SW	IC2	-	Medium/All	0.954
Nine Month-Ahead						
1	DI	FHLR	BIC3	M=H=16	Large/Excl. Fin.	0.976
2	DI	FHLR	BIC3	M=H=16	Large/Excl. Int. Rates	0.980
3	DI_AR_Lag	SW	BIC3	-	Medium/ Excl. Int. Rates	0.982
4	DI	SW	BIC3	-	Large/Excl. Int. Rates	0.983
5	DI	SW	BIC3	-	Large/Excl. Fin.	0.985
Twelve Month-Ahead						
1	DI	FHLR	BIC3	M=H=16	Large/Excl. Fin.	0.958
2	DI	FHLR	BIC3	M=H=16	Large/All	0.961
3	DI	SW	BIC3	-	Large/Excl. Fin.	0.961
4	DI	SW	BIC3	-	Medium/Excl. Int. Rates	0.964
5	DI	FHLR	BIC3	M=H=16	Large/Excl. Int. Rates	0.964

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, Hallin, Lippi and Reichlin (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 2.4. 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 7.3. Rankings of the Models for Stock Market (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	PC3	M=H=16	Large/ Excl. Int. Rates	1.27
337	DI	SW	PC3	-	Medium/Excl. Int. Rates	1.29
338	DI	SW	IC3	-	Medium/Excl. Int. Rates	1.29
339	DI_AR_Lag	FHLR	PC3	M=H=16	Large/All	1.31
340	DI_AR_Lag-F	FHLR	PC3	M=H=16	Medium/Excl. EU	1.32
Six Month-Ahead						
336	DI	FHLR	PC3	M=H=16	Medium/Excl. Int. Rates	1.26
337	DI	FHLR	IC3	M=H=16	Medium/Excl. Int. Rates	1.26
338	DI_AR_Lag	FHLR	PC1	M=H=16	Medium/ Excl. Fin	1.27
339	DI_AR_Lag	FHLR	PC3	M=H=16	Large/ Excl. Int. Rates	1.29
340	DI_AR_Lag	FHLR	IC1	M=H=16	Large/ Excl. Int. Rates	1.29
Nine Month-Ahead						
336	DI	SW	IC3	-	Medium/Excl. Int. Rates	1.16
337	DI	FHLR	PC3	M=H=16	Large/Excl. Int. Rates	1.18
338	DI_AR_Lag	SW	PC1	-	Small/All	1.20
339	DI_AR_Lag	SW	PC2	-	Small/All	1.20
340	DI_AR_Lag	SW	PC3	-	Small/All	1.20
Twelve Month-Ahead						
336	DI_AR_Lag	FHLR	PC1	M=H=16	Medium/All	1.23
337	DI_AR_Lag	SW	PC2	-	Large/ Excl. Int. Rates	1.24
338	DI_AR_Lag	FHLR	IC3	M=H=16	Large/ Excl. Fin	1.24
339	DI_AR_Lag	SW	PC1	-	Medium/ Excl. Int. Rates	1.28
340	DI_AR_Lag	FHLR	PC3	M=H=16	Large/ Excl. Int. Rates	1.30

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, Hallin, Lippi and Reichlin (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 2.4. 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 7.4. Rankings of the Models for Stock Market (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	Data Set	Evaluation Sample: October 2011-September 2013
Three Month-Ahead						
336	DI_AR_Lag	SW	IC3	-	Large/ Excl. Int. Rates	1.34
337	DI_AR_Lag	FHLR	PC1	M=H=16	Large/ Excl. Int. Rates	1.35
338	DI_AR_Lag	SW	PC2	-	Large/All	1.35
339	DI_AR_Lag	SW	PC1	-	Large/ Excl. Int. Rates	1.36
340	DI_AR_Lag	SW	PC1	-	Large/All	1.36
Six Month-Ahead						
336	DI_AR_Lag	FHLR	PC1	M=H=16	Large/ Excl. Int. Rates	1.62
337	DI_AR_Lag	SW	PC2	-	Large/ Excl. Int. Rates	1.62
338	DI_AR_Lag	SW	IC3	-	Large/ Excl. Int. Rates	1.62
339	DI_AR_Lag	FHLR	PC2	M=H=16	Large/ Excl. Int. Rates	1.67
340	DI_AR_Lag	SW	PC3	-	Large/All	1.69
Nine Month-Ahead						
336	DI_AR_Lag	SW	PC1	-	Large/All	1.64
337	DI_AR_Lag	FHLR	IC3	M=H=16	Large/ Excl. Int. Rates	1.69
338	DI_AR_Lag	FHLR	PC1	M=H=16	Large/All	1.69
339	DI_AR_Lag	SW	IC3	-	Large/All	1.74
340	DI_AR_Lag	SW	PC3	-	Large/All	1.75
Twelve Month-Ahead						
336	DI_AR_Lag	SW	PC1	-	Large/ Excl. Int. Rates	1.59
337	DI_AR_Lag	SW	IC3	-	Large/ Excl. Int. Rates	1.59
338	DI_AR_Lag	SW	IC3	-	Medium/ Excl. Int. Rates	1.60
339	DI_AR_Lag-F	FHLR	IC3	M=H=16	Large/Excl. EU	1.60
340	DI_AR_Lag	SW	PC2	-	Large/ Excl. Int. Rates	1.61

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, Hallin, Lippi and Reichlin (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 2.4. 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.

CHAPTER VIII

8 CONCLUSION

Decision makers need to form expectations about future course of various variables before developing their strategies and acting based on those strategies. To this end, forecasting is conducted by different actors and institutions using various techniques. In the meantime, technological advances make it easier to construct data sets with hundreds of domestic and international variables easily. At first sight, it may be thought that with increasing data availability it gets easier to decrease forecast errors as one can get information about a wide range of areas. However, standard techniques cannot handle a large number of indicators due to degrees of freedom problem. Therefore, the trend toward 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 a 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. They have been used in disciplines such as psychology for a long time but their use had been limited in economics due to mismatch between assumptions made in the classical factor analysis and the nature of economic problems and data sets. After seminal contributions in the last decade that made it possible to use factor models for large data sets to answer economic problems in a theoretically consistent manner, they have become popular in the economics and especially in the forecasting profession.

Factor models can be a useful tool in various branches of the economics. One can use factors in an atheoretical way by just taking the first principal component of a data set and interpreting it as the business cycle conditions. On the other hand, one can go a more structural path and use FAVARs to study the monetary policy. In this thesis, factor models are used for forecasting economic and financial variables.

Forecasting performance of factor models are evaluated taking into account different inputs required for forecasting with factor models. There are different dimensions for evaluating the relative forecasting performance of 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 thesis takes a broader approach and make a comprehensive analysis of the sensitivity of forecasting performance of factor models to inputs used in the models. Empirical exercises analyze 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. In addition to factor models, an alternative method to utilize large amount of data, namely combining forecasts, is considered as well. This systematic and comprehensive approach provides useful insights for practical applications as forecasters can become more familiar about how forecasting performance changes with different modelling decision. In the end, this effort may help to optimize model selection for forecasting with factor models.

Findings indicate that factor extraction approach plays a minimal role on the forecast performance. Rather, combination of the decisions on the number of factors, forecast equation and the data set structure plays the pivotal role. Using a high number of factors from a large data set and then using the lags of these factors in the forecast equation deteriorates the forecast performance for most of the cases. Using disaggregated data or using more data do not necessarily improve the forecast performance.

It is also found that depending on the type of the target variable best and worst performing specifications change. For industrial production, using variables from the

European Union helps to improve forecast performance especially in the short term forecast. On the contrary, for core inflation and stock market forecasts many of the best specifications exclude European Union variables. Forecast equation set up also has non-negligible effect. Depending on the whether one uses lags of the factors in the forecast equation, relative performance of the information criteria for selecting the number of factors change. It is also worth noting that in the case of core inflation, worst performing models are observed for the cases that only contemporaneous factors are used in the forecast equation. So, data set structure is a key determinant of forecasting performance.

Forecast performances are analyzed for two consecutive evaluation periods. Relative performance is time varying. For example, for industrial production for the first evaluation sample forecast equation type that uses only contemporaneous factors appears relatively more frequently in the top 5 specifications. However, for the second evaluation sample forecast equation that uses also lags of the factors dominates the top 5 best performing specifications list. Depending on the source of demand, external or internal, private or public, it is natural that different models perform relatively better. While this point may seem obvious, it has important implications for setting up the forecasting system for practical use. This result shows the importance of regularly monitoring the forecast performance of models with the data flow and adjust the forecasting system accordingly.

All in all, a specification with a carefully minted small or medium size data set with a few factor may perform relatively successfully. It is not easy to know which specification will be relatively more successful for the target variable one is interested in. And as shown in the applied part of the thesis, modelling decisions are not mutually independent. There is no systematic pattern to prescribe a recipe for the inputs of factor models that produces relatively successful forecasts at all times. So, continuous analysis of the performance of the alternative specifications for the variables that is to be forecast is necessary.

There are fruitful avenues for further research. One obvious path is to focus on the stability of the performance of the factor models. To better model this dimension, using time varying parameter approach may improve the forecast performance. This can be done by allowing time varying loading or time varying estimation of the parameters of the forecas equation. Another possible route would be trying to integrate a Markov Switching type analysis to the system so that modelling different regimes. It may be the case that rather than small and continuous change in the model parameters, depending on the state of the economy forecasting models may need to be differentiated.

Another dimension that can be focused on to improve the model performance would be working on the model selection. It may be the case that rather than using all of the first few factors decided by the information criteria, one may work with say the first and the third factor in the forecast equation. These are left for further research.

9 REFERENCES

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

Alessi, L., Barigozzi, M., & Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. Statistics & Probability Letters, 80(23), 1806-1813.

Altug, S., & Uluceviz, E. (2014). Identifying leading indicators of real activity and inflation for Turkey, 1988-2010. OECD Journal: Journal of Business Cycle Measurement and Analysis, 2014(1), 1-37.

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., & Veronese, G. (2010). New Eurocoin: tracking economic growth in real time. The review of economics and statistics, 92(4), 1024-1034.

Amengual, D., & Watson, M. W. (2007). Consistent estimation of the number of dynamic factors in a large N and T panel. Journal of Business & Economic Statistics, 25(1), 91-96.

Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., & Rünstler, G. (2011). Short-term forecasts of euro area GDP growth. The Econometrics Journal, 14(1), C25-C44.

Angelopoulou, E., Balfoussia, H., & Gibson, H. D. (2014). Building a financial conditions index for the euro area and selected euro area countries: what does it tell us about the crisis?. Economic Modelling, 38, 392-403.

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. Journal of Business & Economic Statistics, 52-60.

Bai, J., & Ng, S. (2008). Large Dimensional Factor Analysis. Foundations and Trends® in Econometrics, 3(2), 89-163.

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

Barhoumi, K., Darné, O., & Ferrara, L. (2013). Testing the number of factors: An empirical assessment for a forecasting purpose. Oxford Bulletin of Economics and Statistics, 75(1), 64-79.

Barhoumi, K., Darné, O., & Ferrara, L. (2014). Dynamic factor models: A review of the literature. OECD Journal: Journal of Business Cycle Measurement and Analysis, 2013(2), 73-107.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. Journal of the Operational Research Society, 20(4), 451-468.

Baumeister, C., Liu, P., & Mumtaz, H. (2010). Changes in the transmission of monetary policy: evidence from a time-varying factor-augmented VAR. Bank of England Working Paper, No. 401.

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

Boivin, J., & Ng, S. (2005). Understanding and comparing factor-based forecasts. International Journal of Central Banking.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis?. Journal of Econometrics, 132(1), 169-194.

Breitung, J., & Eickmeier, S. (2006). Dynamic factor models. Allgemeines Statistisches Archiv, 90(1), 27-42.

Breitung, J., & Pigorsch, U. (2013). A canonical correlation approach for selecting the number of dynamic factors. Oxford Bulletin of Economics and Statistics, 75(1), 23-36.

Brillinger, D. R. (1981). Time Series: Data Analysis and Theory. New York, NY: Holt, Rinehart and Winston.

Caggiano, G., Kapetanios, G., & Labhard, V. (2011). Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK. Journal of Forecasting, 30(8), 736-752.

Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. Econometrica: Journal of the Econometric Society, 1305-1323.

Chamberlain, G., & Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. Econometrica: Journal of the Econometric Society, 1281-1304.

Connor, G., & Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. Journal of financial economics, 15(3), 373-394.

Connor, G., & Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. Journal of Financial Economics, 21(2), 255-289.,

Connor, G., & Korajczyk, R. A. (1993). A test for the number of factors in an approximate factor model. The Journal of Finance, 48(4), 1263-1291.

Connor, G., & Korajczyk, R. A. (1995). The arbitrage pricing theory and multifactor models of asset returns. Handbooks in operations research and management science, 9, 87-144.

Cragg, J. G., & Donald, S. G. (1997). Inferring the rank of a matrix. Journal of econometrics, 76(1), 223-250.

D'Agostino, A., & Giannone, D. (2012). Comparing alternative predictors based on large-panel factor models. Oxford bulletin of economics and statistics, 74(2), 306-326.

Değerli, A. (2012). Short-term Industrial Production Forecasting for Turkey (Master's thesis, Middle East Technical University).

Dhrymes, P. J., Friend, I., & Gultekin, N. B. (1984). A critical reexamination of the empirical evidence on the arbitrage pricing theory. The Journal of Finance,39(2), 323-346.

Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Journal of Econometrics, 164(1), 188-205.

Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi–maximum likelihood approach for large, approximate dynamic factor models. Review of economics and statistics, 94(4), 1014-1024.

Eickmeier, S. (2007). Business cycle transmission from the US to Germany—A structural factor approach. European Economic Review, 51(3), 521-551.

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

Eickmeier, S., & Ng, T. (2011). Forecasting national activity using lots of international predictors: An application to New Zealand. International Journal of Forecasting, 27(2), 496-511.

Favero, C. A., Marcellino, M., & Neglia, F. (2005). Principal components at work: the empirical analysis of monetary policy with large data sets. Journal of Applied Econometrics, 20(5), 603-620.

FED, C. (2001). CFNAI background release. Discussion paper, http://www. chicagofed. org/economicresearchanddata/national/pdffiles/CFNAI bga. pdf.

Forni, M., & Reichlin, L. (1998). Let's get real: a factor analytical approach to disaggregated business cycle dynamics. The Review of Economic Studies,65(3), 453-473.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. Review of Economics and statistics, 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?. Journal of Monetary Economics, 50(6), 1243-1255.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2004). The generalized dynamic factor model consistency and rates. Journal of Econometrics, 119(2), 231-255.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The Generalized dynamic factor model: One sided estimation and forecasting, Journal of the American Statistical Association, 100, 830-840.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics,55(4), 665-676.

Günay, M. (2015). Forecasting Turkish Industrial Production Growth With Static Factor Models. International Econometric Review (IER), 7(2), 64-78.

Gupta, R., & Kabundi, A. (2011). A large factor model for forecasting macroeconomic variables in South Africa. International Journal of Forecasting,27(4), 1076-1088.

Hallin, M., & Liška, R. (2007). Determining the number of factors in the general dynamic factor model. Journal of the American Statistical Association,102(478), 603-617.

Hatzius, J., Hooper, P., Mishkin, F. S., Schoenholtz, K. L., & Watson, M. W. (2010). Financial conditions indexes: A fresh look after the financial crisis, NBER Working Paper (No. w16150).

Hendry, D. F., & Clements, M. P. (2004). Pooling of forecasts. The Econometrics Journal, 7(1), 1-31.

Kapetanios, G., & Marcellino, M. (2009). A parametric estimation method for dynamic factor models of large dimensions. Journal of Time Series Analysis,30(2), 208-238.

Kose, M. A., Otrok, C., & Whiteman, C. H. (2008). Understanding the evolution of world business cycles. Journal of international Economics, 75(1), 110-130.

Modugno, M., Soybilgen, B., & Yazgan, E. (2016). Nowcasting Turkish GDP and news decomposition. Finance and Economics Discussion Series Board of Governors of the Federal Reserve System.

Öğünç, F., Akdoğan, K., Başer, S., Chadwick, M. G., Ertuğ, D., Hülagü, T., ... & Tekatlı, N. (2013). Short-term inflation forecasting models for Turkey and a forecast combination analysis. Economic Modelling, 33, 312-325.

Roll, R., & Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. The Journal of Finance, 35(5), 1073-1103.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. Journal of economic theory, 13(3), 341-360.

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

Schumacher, C. (2007). Forecasting German GDP using alternative factor models based on large datasets. Journal of Forecasting, 26(4), 271-302.

Schumacher, C. (2010). Factor forecasting using international targeted predictors: the case of German GDP. Economics Letters, 107(2), 95-98.

Soybilgen, B. (2015). Three essays on forecasting, (Doctoral dissertation, İstanbul Bilgi Üniversitesi).

Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. Journal of Monetary Economics, 44(2), 293-335.

Stock, J. H., & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. Journal of the American statistical association, 97(460), 1167-1179.

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

Stock, J. H., & Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. Journal of Economic Literature, 41(3), 788-829.

Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. Journal of Forecasting, 23(6), 405-430.

Stock, J. H., & Watson, M. W. (2005). Implications of dynamic factor models for VAR analysis (No. w11467). National Bureau of Economic Research.

Stock, J. H., & Watson, M. W. (2011). Dynamic factor models. Oxford handbook of economic forecasting, 1, 35-59.

Timmermann, A. (2006). Forecast combinations. Handbook of economic forecasting, 1, 135-196.

APPENDICIES

APPENDIX A. INDICATORS USED IN THE DATA SETS

Table A.1. Small Data Set		
	Data (Abbreviations Used in the Table A.2 and Table A.3 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.2. Medium 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 restaruants	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.3. Large 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

APPENDIX B. CIRRICULUM VITAE

PERSONAL INFORMATION
Surname, Name: Günay, Mahmut.
e-mail: mahmutgunay@gmail.com

EDUCATION

Degree	Institution	Year of Graduation
MS	The University of Texas at Austin	2005
BA	Boğaziçi University	2003

PROFESSIONAL EXPERIENCE

Year	Place	Position
2005-…	Central Bank of Turkey	Economist

FOREIGN LANGUAGE
English

PUBLICATIONS
Günay, M. (2016). Forecasting Turkish GDP Growth: Bottom-Up vs Direct? (No. 1622). Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Günay, M. (2016). "Forecasting Turkish GDP Growth with Financial Variables and Confidence Indicators," CBT Research Notes in Economics 1614, Research and Monetary Policy Department Central Bank of the Republic of Turkey.

Ozel, Ö. & Günay, M. (2015). "Sanayi Uretim Endeksi Yillik Degisiminin Tahminine Istatistiksel Yaklasim," CBT Research Notes in Economics 1521, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Günay, M. (2015). I Just Ran Four Million Regressions" for Backcasting Turkish GDP Growth," Working Papers 1533, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Günay, M. (2015). Forecasting Turkish Industrial Production Growth With Static Factor Models. International Econometric Review (IER), Econometric Research Association, vol. 7(2), pages 64-78, September.

Akkoyun, H. Ç., Şen-Doğan, B., & Günay, M. (2014). Business Cycle Synchronization of Turkey with the Eurozone and the United States: What Has Changed Since 2001?. Emerging Markets Finance and Trade, 50(4), 26-41.

Akkoyun, H. Ç. & Günay, M. (2013). Milli Gelir Buyume Tahmini: IYA ve PMI Gostergelerinin Rolu. CBT Research Notes in Economics 1331, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Akkoyun, H. Ç. & Günay, M. (2012). Nowcasting Turkish GDP Growth. Working Papers 1233, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Aldan, A., Bozok, İ. & Günay, M. (2012). Short Run Import Dynamics in Turkey. Working Papers 1225, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Günay, M. (2011). Milli Gelir Guncellemeleri. CBT Research Notes in Economics 1121, Research and Monetary Policy Department, Central Bank of the Republic of Turkey

Günay, M. (2010). Sanayi Uretimine Iliskin Piyasa Beklentilerindeki Sapmalar ve Takvim Etkisi. CBT Research Notes in Economics 1003, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

Aldan, A. & Günay, M. (2008). Entry to Export Markets and Productivity: Analysis of Matched Firms in Turkey. Working Papers 0805, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.

APPENDIX C. TURKISH SUMMARY

GİRİŞ:

Bilgi işlem teknolojilerindeki hızlı gelişimin de etkisiyle reel ve finansal değişkenlere ulaşmak ve bu değişkenleri depolamak her geçen gün kolaylaşmaktadır. Çok uluslu şirketlerin ürünlerinin üretim aşamalarını farklı ülkelere dağıtmalarının da katkısıyla özellikle 90'lı yıllarda küresel ticaret küresel ekonomideki büyümeden çok daha hızlı artmıştır. Reel taraftaki bu harekete ek olarak, sermaye akımlarının seyri kredi kanalı ve kurlar üzerinden özellikle cari açık veren gelişmekte olan ülkelerin büyüme ve talep görünümünde önemli roller oynayabilmektedir. Bu çerçevede, yerel verilerin yanında uluslararası alandan verilerin de analiz ve tahmin süreçlerine dâhil edilmesi gerekmektedir. Ülkemiz özelinde ise Gümrük Birliği anlaşması, coğrafi yakınlık, doğrudan yabancı yatırımlarla otomobil gibi sektörlerde üretimin artması gibi sebeplerle Avrupa Birliği talebinin ihracat için önemi artmıştır. Bir başka açıdan bakıldığında ise, petrol ithalatçısı olan ülkemiz cari açık vermektedir. Bu durum, emtia fiyatlarının seyrine ve sermaye akımlarına hassasiyeti de artırmaktadır. Sonuç olarak, Türkiye ekonomisine dair sağlıklı ekonomik analizler ve isabetli tahminler için, ülkemize ilişkin değişkenlerin yanında Avrupa Birliği gibi dış pazarlar için reel ve finansal verilerin, anket göstergelerinin ve emtia fiyatlarının takibi faydalı olmaktadır.

Veriye ulaşımda artan bu kolaylığa karşın En Küçük Kareler veya Vektör Otoregresif gibi klasik yöntemlerle katsayıların tahmin edildiği bir modelde kullanılabilecek değişken sayısı sınırlıdır. Bu nedenle, çok sayıdaki değişkeni kullanabilmek için uygun yöntemlerden yararlanmanın önemi artmaktadır. Faktör modelleri çok sayıdaki değişkeni kullanmaya imkân sağlayan bir yaklaşım olarak bu noktada oldukça işlevsel olabilmektedir. Bu yaklaşımda, bir veri setindeki bilgi birkaç faktör ile özetlenmekte ve tahmin süreçlerinde bu faktörler kullanılmaktadır. Faktör modellerinin bu özellikleri nedeniyle son yıllarda iktisat yazınının birçok alanında kullanılmaya başlanmıştır. Faktör modellerinin makroekonomik değişkenlere ilişkin tahminlerde kullanımı son yirmi yılda artmıştır. Değişik kurum ve kuruluşların tahmin gereksinimleri farklı olmaktadır. Örneğin, merkez bankaları para politikası kararlarını alırken parasal aktarım mekanizmasının gecikmeli etkileri nedeniyle bir-iki yıllık vadede enflasyonun ne olacağı konusunda tahminler yapmaktadırlar. Söz konusu tahminlerin oluşturulmasında hikâye anlatmaya imkân veren yapısal ya da yarı-yapısal modeller tercih edilmektedir. Bu tahminlere ek olarak, yakın dönemde ekonominin ne durumda olduğu da politika yapımında önemli bir girdi olmaktadır. Enflasyondaki kısa vadede gözlenebilecek bir yükseliş beklentileri bozabileceğinden, para politikası karar aşamasında ve bu kararlara dair iletişim süreçlerinde enflasyona ilişkin daha yakın vadeli tahminler de üretilmekte ve kamuoyu ile paylaşılmaktadır. Merkez bankaları enflasyonun yanında reel ekonomideki gelişmeleri de takip etmektedir. İktisadi faaliyete dair sanayi üretimi ve milli gelir büyümesi gibi verilere ilişkin daha kısa vadeli tahminler de üretilmektedir. Görece daha kısa vadeli bu tahminler için yapısal modellerden ziyade ekonometrik modeller daha isabetli tahminler verebilmekte ve verileri en etkin şekilde işlemeye izin vermektedir. Kamu maliyesine ilişkin politikalarda ise vergi gelirleri tahmininde kilit girdiler olan enflasyon ve büyüme oranlarının yıllık frekanstaki tahminleri önemli parametrelerdir. Uzun vadeli böyle bir bakış açısının yanında yatırım bankaları gibi kuruluşlar kur ve borsa için çok kısa dönemli, bazen günlük nispetinde, tahminlere ihtiyaç duyabilmektedir. Sonuç olarak, değişik kuruluşlar farklı değişkenler için farklı vadelerde tahminlere ihtiyaç duymaktadır. Tahmin üretmek için çeşitli yöntemler kullanılmaktadır. Faktör modelleri çok sayıda veriyi değerlendirmeye ve kullanmaya imkân vermesinin yanında gerek yakın dönemli gerekse orta-vadeli tahminlerde kullanılabildiği için tahmini bir girdi ya da çıktı olarak kullanan kuruluşlarda popüler hale gelmiştir.

Bu tezde faktör modelleri kullanılarak bir reel değişken olan sanayi üretim endeksindeki büyüme, bir fiyat değişkeni olan çekirdek enflasyon ve finansal bir değişken olan borsa endeksindeki değişim tahmin edilmektedir. Böylece, faktör modellerinin farklı tip değişkenler için tahmin performansı değerlendirilebilecektir.

YÖNTEM

Faktör modellerini anlatan genel ifade aşağıdaki iki denklemdeki gibi ifade edilebilir. Burada, x veri setimizdeki değişkenleri göstermektedir. "i" ifadesi ile o değişkenin tanımı yapılmaktadır. Örneğin, i=1 ve i=2 durumlarında, x_1 faiz oranlarına ilişkin bir değişkeni, x_2 ise tüketici güvenine dair başka bir değişkeni gösteriyor olabilir. "i" ifadesi, 1'den N'ye kadar gidecek şekilde tanımlanmıştır. İlk nesil faktör modellerinde N sınırlı sayıda olmaktaydı. Örneğin, 10'dan fazla değişken kullanılması bu teknikte zordu. 90'lı yılların sonunda geliştirilen ve 2000'li yılların başında yayımlanan çalışmalardaki tekniklerle (Stock ve Watson (2002) ve Forni vd (2005)) artık N çok fazla sayıda olabilmektedir.

"t" ise zamanı göstermektedir. İlk denklemin sağ yanında iki ifade bulunmaktadır. Bu ifadelerin ilki "i" değişkeninin "t" zamanındaki değerinin F faktörü ile açıklanan kısmını göstermektedir. Burada, lamda "faktör yüklemesi" olarak adlandırılan değişkenlerle faktörler arasındaki ilişkiyi göstermektedir. Dikkat edilirse lamda'nın alt-indisi "i" ifadesi iken F'in alt indisi "t" ifadesidir. Böylece, her bir değişken için ilgili faktöre ilişkin bir lamba değeri bulunmakta ancak aşağıdaki tanımda bu değer zamanla değişmemektedir. Faktörleri gösteren F ise zamana göre değişmektedir. Bu gösterimde F bir matris olarak farklı faktör serilerini içeren genel bir ifadedir. "F_t" ifadesinin kritik bir anlamı vardır: mevcut sistem statik bir faktör modeli yaklaşımını göstermektedir. x değişkeninin t zamanındaki değeri faktörlerin de t zamanındaki değerine bağlı olmaktadır. Ancak şu belirtilmelidir ki faktörlerin gecikmeli değeri yeniden tanımlanarak bu matriste yer alabilecektir.

$$\begin{aligned} x_{it} &= \lambda_i' F_t + e_{it} \\ &= C_{it} + e_{it}. \end{aligned}$$

i= 1,…, N, t = 1,…,T,

İlk denklemdeki ikinci ifade ise faktör ile açıklanamayan hareketi, seriye-özgü hareket, göstermektedir. İki alt-indis kullanılarak gösterilen bu terim her bir değişken için her bir zamanda farklı terimlerin olabileceğine işaret etmektedir.

Yukarıdaki denklem sistemi oldukça genel olduğundan statik faktör modelinin yanında kesin ve yaklaşık faktör modelleri olarak ifade edilen iki durumu göstermek için de kullanılabilmektedir. Kesin faktör modeli (exact factor models) yaklaşımında seriye özgü terimlerin birbirlerinden bağımsız olduğu varsayılmaktadır. Yaklaşık faktör modellerinde (approximate factor models) ise değişkenler arası ve bir değişken için zamanlar-arası, belirli sınırlar içinde, korelasyona izin verilmektedir. Faktör modellerinin ilk kullanım alanları olan psikoloji ve finansal verilerde, kesin faktör modeli yaklaşımı varsayımları makuldü. Ancak, bu varsayımın iktisadi analizlerde de kullanılması gerçeklikten oldukça uzaklaşılan bir duruma yol açabileceğinden yaklaşık faktör modelleri daha yaygın şekilde kullanılmaktadır. Psikoloji literatüründeki bir örneğe göre, bir sınıftaki öğrencilerin başarı düzeylerini belirleyen temel değişken zekâ seviyeleri olarak modellendiği bir durumu ele alalım. Zekâ dışında başarıyı etkileyen faktörler ise seriye-özgü değişken içinde yer almaktadır. Bir sınav öncesi bir öğrenci hastalanıp zekâsının ima ettiği düzeyin altında bir not alabilir. Hastalığın bir başka öğrenciyi etkilememe ihtimali yüksektir, bir başka deyişle öğrencileri etkileyen şoklar birbirinden bağımsızdır. Bu şokların zamanlar arası bağımsız olması da makul bir varsayımdır. Bu çerçevede, bir sınavda iki öğrencinin aldığı notların faktörlerle açıklanamayan kısmı zamanlar arası ve öğrenciler arası birbirinden bağımsız olabilir.

İktisat literatüründe karşılaşılan değişkenlerin ise faktörlerle açıklanan dışında kalan seriye-özgü zamanlar ve seriler arası bağımsız olduğu çok güçlü bir varsayımdır. İlk olarak, zaman serisi analizindeki değişkenler yatay-kesit analizinden farklı olarak zamanlar arası korelasyon sahip olabilmektedir. Buna ek olarak, veri setlerinde birbirleriyle ilişkili değişkenler yer almaktadır. Örneğin, sanayi üretimi bloğunda tekstil üretimi ve giyim üretiminin kullanıldığı durumu ele alalım. Faktörler ile açıklanan kısım dışında kalan seriye-özgü bileşenler birbirleriyle halen ilişki olabilecektir. Bu nedenle, iktisadi uygulamalar için seriye özgü terimlerdeki varsayımların esnetilmesi gerekebilecektir. Yaklaşık faktör modelleri bunu yapmaktadır.

Faktör modellerinin ifade edildiği yukarıdaki ilk denklemde bir değişken faktörün bir fonksiyonu olan “ortak terim” ve faktör tarafından açıklanamayan hareketi gösteren seriye özgü kısım olarak ifade edilmiştir. Ancak, bu denklemde yer alan ifadelerden lamda, faktörler ve seriye özgü bileşenler gerçek hayatta gözlenememektedir. Diğer yandan, faktör elde edilecek veri setinin yapısına dair bir standart bulunmamakta ve bu veri setinden kaç tane faktör oluşturulacağı da bilinmemektedir. Son olarak, elde edilen faktörlerin tahmin süreçlerinde nasıl kullanılacağına dair de farklı seçenekler bulunmaktadır, örneğin faktörlerin gecikmeli değerlerinin kullanılıp kullanılmamasına göre tahmin performansı değişebilecektir. Sonuç olarak, faktör modelleri çok sayıdaki değişkeni kullanmak için güçlü bir araç olsa da tahmin performansını etkileyebilecek model tercihleri yapılmaktadır. Diğer yandan, fiyat değişkenlerindeki değişim reel değişkenlere göre daha yapışkan olabilirken, finansal değişkenlerin etkin bir piyasada tahmini mümkün olmayabilecektir. Bu çerçevede, faktör modellerindeki model tercihlerinin tahmin performansına etkisi sanayi üretimi büyümesi, çekirdek enflasyon ve borsa endeksindeki değişim için ayrı ayrı değerlendirilmektedir.

FAKTÖR ELDE ETME YÖNTEMLERİ

Bir veri setinden faktör elde etmek için farklı yaklaşımlar bulunmaktadır. Stock ve Watson (2010) tarafından ilk nesil faktör modelleri olarak adlandırılan yöntem En Çok Olabilirlik (Maximum Likelihood) yöntemidir. Bu yöntemde değişken sayısı sınırlı olmaktadır. Bu durumun oluşturduğu kısıt, Stock ve Watson (2002) ve Forni vd. (2005) çalışmalarında önerilen yöntemlerle aşılabilmiştir. İkinci nesil olarak adlandırılan bu tür faktör modellerinde parametrik olmayan ortalamalar metotları kullanılmaktadır. Bu yaklaşımlarda temel olarak veri setindeki değişkenlerin ağırlıklı ortalaması faktör olarak kullanılmaktadır. Ağırlık olarak kullanılacak değerlerin elde edilmesinde öne çıkan iki yaklaşım aşağıda özetlenmektedir.

Stock ve Watson (2002) aşağıdaki problemden yola çıkmaktadır. Bu problemde, x değişkenindeki hareketin faktör ile açıklanamayan kısım minimize edilmektedir. Stock ve Watson (2002) bu problemin temel bileşenler analizi yöntemleri kullanılarak çözülebileceğini göstermiştir.

$$V(\tilde{F},\tilde{\Lambda}) = (NT)^{-1}\sum_{i}\sum_{t}(x_{it} - \tilde{\lambda}_i \tilde{F}_t)^2$$

Stock ve Watson (2002)'nın yöntemi çok sayıda değişkeni kolay bir şekilde kullanmaya izin verse de ekonomik ve finansal değişkenlerdeki dinamik yapıyı tam olarak yansıtamamaktadır. Bunun nedeni, faktörlerin elde edildiği temel bileşenler analizinde "t" zamanındaki faktör değeri için "t" zamanındaki değişkenlerin değerleri kullanılmaktadır. Ancak, faiz oranı gibi bir değişken ile güven endeksi ile gecikmeli bir ilişki olabilir. Forni vd (2005) çalışması bu çerçevede, faktörleri dinamik temel bileşenler yöntemi ile elde etmektedir. Bu yöntemde, Fourier dönüşümü kullanılarak frekans alanında faktörler elde edilmektedir. Yöntemin ilk uygulamalarında iki taraflı filtre kullanıldığından tahmin uygulamalarında kullanımında bir kısıt vardı. Bu kısıtı aşan çalışmalarla desteklenen yöntem ile birlikte tahmin süreçlerinde kullanılmak üzere faktörler elde edilebilmektedir. Söz konusu yöntemin kullanılmasındaki yol haritasını Schumacher (2007, sayfa 274) çalışması özetlenmektedir. İlk olarak, otokovaryans değerleri elde edilmektedir. Spektral yoğunluk matrisi hesabında Bartlett pencere genişliği kullanılmaktadır. Her bir frekans için dinamik özdeğer ve özvektör hesaplanmaktadır. Ardından, ters-Fourier dönüşümü kullanılarak zaman-alanındaki (frekans –alanı yerine) otokovaryans değerleri elde edilmektedir. Böylece, aşağıdaki genelleştirişmiş özdeğer problemi çözülmektedir.

$$\hat{\Gamma}_{\chi}(0)\hat{Z}_j = \hat{\mu}_j\hat{\Gamma}_{\xi}(0)\hat{Z}_j$$

Yukarıdaki denklemde ilk ifade ortak terimlerin otokovaryansları, sağdaki bölümde ortadaki ifade ise seriyr-özgü bileşenler için bu değerleri göstermektedir. Böylece, aşağıdaki formül ile Forni vd. (2005) yöntemine göre faktörler elde edilebilmektedir.

$$\hat{F}_t^{FHLR} = \hat{Z}'X_t$$

FAKTÖR SAYISI

Stock ve Watson (200) temel bileşenler ile Forni vd. (2005) dinamik temel bileşenler yöntemleri çok sayıda veriden istenilen sayıda faktör almayı çok hızlı bir şekilde yapabilecek olsa da faktör sayısının çok fazla olduğu bir durumda yöntemin temel erdemlerinden biri olan boyut-küçültme işlevi zayıflayacaktır. Ancak, çok az sayıda faktör kullanıldığında ise ver setindeki bilgiler yeterince etkin bir şekilde özetlenemeyecektir. Faktörler gözlenemediği için kaç tane faktör olması gerektiği de bilinmemektedir. Faktör sayısını belirlemek için çeşitli yöntemler kullanılmaktadır. Sabit sayıda faktör kullanarak bunların performansına göre faktör sayısına karar vermek bir yöntem iken Akaike Bilgi Kriteri gibi yöntemlerin faktör modellerine uygulanmasıyla bilgi kriteri bazlı faktör sayısı seçimi de bir başka yöntemdir. Bu tezde, statik faktör sayını bulmak için Bai ve Ng (2002) çalışmasında önerilen yedi kriter de ayrı ayrı kullanılmıştır. Bu kriterlere ilişkin formüller aşağıda gösterilmektedir. N, değişken sayısını, T ise gözlem sayısını ifade etmektedir.

$$BN1: PC_{p1}(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right);$$

$$BN2: PC_{p2}(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln C_{NT}^2 ;$$

$$BN3: PC_{p3}(k) = V\left(k, \hat{F}^k\right) + k\hat{\sigma}^2 \left(\frac{lnC_{NT}^2}{C_{NT}^2}\right).$$

$$BN4: IC_{p1}(k) = \ln(V\left(k, \hat{F}^k\right)) + k \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right);$$

$$BN5: IC_{p2}(k) = \ln(V\left(k, \hat{F}^k\right)) + k \left(\frac{N+T}{NT}\right) \ln C_{NT}^2 ;$$

$$BN6: IC_{p3}(k) = \ln\left(V\left(k, \hat{F}^k\right)\right) + k\left(\frac{lnC_{NT}^2}{C_{NT}^2}\right);$$

$$BN7: BIC_3(k) = V\left(k, \hat{F}^k\right) + \ k\hat{\sigma}^2 \left(\frac{(N+T-k)\ln(NT)}{NT}\right).$$

Dinamik faktör sayısı için ise, Bai ve Ng (2007) çalışmasında önerilen yöntem kullanılmıştır.

TAHMİN DENKLEMİ

Faktör sayısını belirleyip faktörleri elde ettikten sonra tahmin modellerinin oluşturulma aşamasında da alternatifler mevcuttur. İlk olarak, bir dönemden uzun vadeli tahminlerde bağımlı değişkenin yapısına ilişkin iki alternatif mevcuttu. Örneğin, tahmin yapılan dönemden üç ay sonrasındaki aylık büyümeler tahmin edilebilecektir. Ancak, aylık değişimler dalgalı olacağından üç veya on iki ay sonrasındaki aylık dalgalanmaları yakalamak mümkün olmayabilir. Ayrıca, politika yapıcılar ve karar alıcılar açısından ileriki bir zamandaki aylık değişim yerine bu dönemdeki birikimli değişimler daha anlamlı olabilir. Bu çerçevede, bağımlı değişken olarak, aşağıdaki formülde gösterildiği üzere, h dönem sorası için birikimli büyüme oranları tahmin edilmiştir.

$$y_{t+h}^{h} = \left(\frac{1200}{h}\right) \ln\left(\frac{IP_{t+h}}{IP_t}\right)$$

Tahmin edilmek istenilen değişkenin yapısına karar verildikten sonra belirlenmesi gereken bir sonraki konu ise tahmin denkleminin sağ tarafında hangi açıklayıcı değişkenlerin kullanılacağıdır. Aşağıdaki denklem tahmin modelinin genel ifadesini göstermektedir. Bu denklem üç şekilde kullanılmıştır. DI-AR Lag olarak ifade edilen yapıda hem faktörlerin hem de bağımlı değişkenin gecikmeli değerleri kullanılmaktadır. DI olarak ifade edilen yapıda ise faktörlerin cari değeri kullanılırken bağımlı değişkenin aylık değişmesinin gecikmeli değerleri de yer almamaktadır, bir başka deyişle denklem sadece faktörlerden bilgi almaktadır.

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

Tezdeki temel soru, faktör modellerinin performansına model tercihlerinin nasıl yansıdığıdır. Tahmin denklemi yapısı bu konunun önemini açıkça ortaya koymaktadır. Bai ve Ng (2002) tarafından önerilen kriterlerin bazıları yüksek sayıda faktör önermektedir. Bu kriter, DI-AR Lag tipi bir tahmin denklemi ile birlikte kullanıldığında, denklemde çok sayıda katsayı yer almaktadır. Bu durum, katsayılara ilişkin belirsizliği artırmaktadır. Böylece, model tercihlerinde tek bir boyutu dikkate almak sağlıklı sonuç vermeyebilecektir. Faktör sayısının tahmin performansına etkisi tahmin denklemi yapısından bağımsız olmayacaktır.

VERİ SETİ

Faktör modeli yaklaşımında kullanılacak veri setlerine ilişkin genel geçer bir kural yoktur. İstatistikler, farklı detaylarda açıklanmaktadır. Bu durum da veri seti oluşumundaki karar sürecini daha da karmaşık hale getirmektedir. Zira, sanayi üretiminin faktör elde edilecek veri setine dâhil edildiği bir durumda, bunun hangi detayda yapılacağı sorusu ortaya çıkmaktadır. Bu çerçevede bu çalışmada serilerin farklı detay seviyelerini dikkate alarak üç farklı veri seti oluşturulmuştur: küçük, orta ve büyük. Aşağıdaki tablo, sanayi üretimi için duruma bir örnek teşkil etmektedir.

Tablo. Göstergelerde Artan Detay: Sanayi Üretimi Örneği

Küçük Veri Seti	Orta Veri Seti	Büyük Veri Seti
Sanayi Üretimi	Ara Malları	Madencilik
	Sermaye Malları	Gıda
	Dayanıksız Mallar	İçecek
	Dayanıklı Mallar	Tütün
	Enerji	Tekstil
		Giyim
		Deri
		Ağaç
		Kağıt
		Basım
		Rafine petrol
		Kimya
		Eczacılık
		Kauçuk
		Diğer mineral
		Ana metal
		Fabrikasyon metal
		Elektronik ve optik
		Elektrikli cihazlar
		Makine ve teçhizat
		Taşıt
		Diğer ulaşım
		Mobilya
		Diğer imalat
		Makine-ekipman kurulum
		Elektrik-gaz-su

Yukarıdaki örnekten gidilirse küçük veri seti olarak isimlendirilen grupta sanayi üretiminin toplamı yer almaktadır. Orta büyüklükteki veri setinde ise Ana Sanayi Grupları Sınıflandırması kapsamında sanayi üretimi için açıklanan beş alt kalem kullanılmaktadır.

Büyük veri setinde ise, sanayi üretimi toplamının yirmi altı alt sektör detayında açılandığı NACE sınıflaması kullanılmaktadır.

Göstergelerin hangi detayda kullanılacağına ek olarak bir göstergenin faktör elde edilen veri setinde yer alıp almaması da sonuçları etkileyebilecektir. Bu doğrultuda, faiz, Avrupa Birliği ve emtia fiyatları-finansal değişkenler blokları ayrı ayrı veri setlerinden çıkarılarak tahmin performansı karşılaştırılmıştır.

Sonuç olarak, faktör elde etme, faktör sayısını belirleme, tahmin denklemi ve veri seti yapısı konusunda alternatifler vardır. Bu alternatifler birbirlerinden bağımsız değildir. Büyük bir veri setinden çok sayıda faktör elde edilerek bunların gecikmeli değerleri de tahmin süreçlerinde kullanıldığında artan katsayı belirsizliği nedeniyle istikrarsız ve isabetsiz tahminler ortaya çıkabilecektir. Bu tezde, bu unsurların tahmin performansına etkileri kapsamlı bir şekilde değerlendirilmektedir.

YAZIN

Faktör modelleri iktisat yazınında birçok alanda kullanılmaktadır. Varlık fiyatları modellemesinde faktör modelleri hisse senetleri gibi finansal değişkenlerdeki hareketler faktörlerle açıklanan kısım ve varlığa özgü kısım olarak ayrıştırılmaktadır. İş çevrimi modellerinde ise milli gelir ya da sanayi üretimi gibi tek bir değişkeni ekonominin durumunu yansıtan gösterge olarak kullanmak yerine birçok göstergenin ortak faktörü kullanılarak daha geniş bir veri seti değerlendirilebilmektedir. Bu yöntemde ayrıca farklı ülkelerin iş çevrimlerinin birbirleriyle olan ilişkileri de incelenebilmektedir. Para politikası analizinde faktör modelleri özellikle faydalı olabilmektedir. Enflasyon, işsizlik/büyüme ve faiz oranlarının kullanıldığı klasik bir Vektör Otoregresif modelde faiz şokları sonrası enflasyon yükselebilmektedir. “Fiyat bilmecesi” olarak ifade edilen bu durumun bir sebebinin merkez bankalarınca kullanılan geniş veri setlerinin küçük ölçekli Vektör Otoregresif modellerde dikkate alınmaması olduğu ileri sürülmüştür. Eski FED başkanı Bernanke’nin diğer bazı araştırmacılarla yaptığı çalışmalarda reel ve finansal alandan verilerden oluşan faktörler Vektör Otoregresif modellere entegre edilerek (FAVAR yaklaşımı) fiyat bilmecesinin neden kaynaklanmış olabileceğine dair bir açıklama getirilmiştir. Faktör modellerinin bir diğer kullanım alanı ise ekonominin ve finansal koşulların durumuna ilişkin endeksler oluşturmaktadır. Chicago FED tarafından açıklanan ve temel bileşenler analizi kullanılarak ABD ekonomisi için ekonominin durumunu gösteren bir endeks her ay güncellenmektedir.

Faktör modellerinin değişik alanlardaki kullanıma ek olarak, tahmin üretmek için yaygın bir kullanımı vardır. Araştırmacılar faktör modellerinin yapısının tahmin performansına etkisini değişik boyutlardan ele almışlardır. Örneğin bir çalışma veri setindeki detayın tahmin sonuçlarına etkisini incelerken bir diğer çalışma faktör elde etme yönteminin etkisini incelemektedir. Bazı çalışmalar belirli veri bloklarının tahmin performansına etkilerine bakarken bir diğer kısmı ise faktör sayısını belirlemede kullanılan yöntemlerin performansa etkisini incelemektedir. Bu çalışmalarda genel olarak reel bir değişken olan milli gelir büyümesi ve sanayi üretimi ile fiyatlara ilişkin bir değişken olan tüketici fiyatları enflasyonu tahmin edilmektedir. Bu tezde ise bu boyutlar aynı sistemde değerlendirilmektedir.

BULGULAR

Sanayi üretimi, enflasyon ve borsa verileri için 336 farklı faktör modeli tahmin modeli spesifikasyonu üç, altı, dokuz ve on iki ay sonrası birikimli değişimler için incelenmiştir. Tahmin performansı zamanla değişebileceği için örneklem dışı tahminler iki ayrı dönem için değerlendirilmiştir. Bilgiyi birleştiren bir yöntem olan faktör modeli yaklaşımına ek olarak, tahminleri birleştirme yaklaşımı da dikkate alınmıştır. Bunun için her bir veri setinde kullanılan değişken ile tahminler üretilmiş ve bu tahminlerin ortalaması alınmıştır. Yazında standart bir uygulama olan tahmin performansının bir ölçüt modele göre değerlendirilmesi yapılarak göreli tahmin performansı tablolar haline sunulmuştur.

Genel olarak bakıldığında sonuçlar, tahmin modeline da tercihlerin performansı önemli derece etkilediğini göstermektedir. Bir değişken için ölçüt modellere göre yüzde 20 civarında daha iyi tahmin üretilebildiği gibi aynı değişken için farklı bir tahmin modeli spesifikasyonu kullanılarak görece büyük tahmin hataları yapılabilmektedir. Özetle,

i. Faktör elde etme yönteminin tahmin performansı üzerindeki etkisi sınırlıdır.

ii. Daha detaylı veri kullanmak her zaman tahmin performansını iyileştirmemektedir.
iii. Büyük bir veri setinden çok sayıda faktör elde edip tahmin modelinde bunların gecikmelerini de kullanmak tahmin performansını olumsuz etkileyebilmektedir.
iv. Farklı değişken tiplerinde ve farklı vadelerde en iyi modeller farklı olmaktadır. Bu nedenle, her bir değişken ve vade için farklı modeller kullanılması faydalı olacaktır.
v. Modellerin performansı zamanla değişmektedir. Bu çerçevede, tahmin modellerinin performansının dinamik bir şekilde değerlendirilerek uygun modellerin belirlenmesi tahmin performansı açısından faydalı olacaktır.

SANAYİ ÜRETİMİ

Sanayi üretimi özelinde bakıldığında,

i. Tahminlerin değerlendirildiği ikinci döneme bakıldığında, bütün tahmin ufuklarında DI-AR-Lag [faktörlerin hem cari dönem hem de gecikmelerini, bağımlı değişkenin gecikmeli değerleri ile birlikte kullanan denklem] tipi tahmin denklemi kullanılmıştır. İlk dönem için ise DI [sadece faktörlerin cari değerlerini kullanan denklem] tipi denklemler daha yaygın kullanılmaktadır. Diğer taraftan, DI-AR-Lag tipi denklemler en kötü performans gösteren modellerde de kullanılmaktadır. Bu sonuç, tahmin performansında denklem tipi dışındaki model tercihlerinin önemini göstermektedir.
ii. Faktör elde etme yöntemleri olan SW ve FHLR yaklaşımları karşılaştırıldığında ise ilk tahmin değerlendirme döneminde SW yaklaşımının, ikincisinde ise FHLR yaklaşımının daha iyi sonuçlar verdiği görülmektedir. Bununla birlikte, iki yöntemden elde edilen tahmin sonuçları birbirine oldukça yakındır.
iii. Yazında, faktör elde etmek için IC1 ve IC2 yaklaşımlarının yaygın olarak kullanıldığı görülmektedir. Bai ve Ng (2002) ise BIC3'ün de ümit verici

bir kriter olduğunu belirtmektedir. En iyi performans gösteren modellerde, IC1 ve IC2 yanında BIC3 yaklaşımının da kullanıldığı görülmektedir. En kötü performans gösteren modellerde ise PC3 ve IC3 listeyi domine etmektedir.

iv. En iyi modellerde, faiz oranlarını ve finansal değişkenleri, ayır ayrı, veri setinden çıkarmak tahminleri iyileştirmektedir. Ancak, en çok tahmin hatası yapan modeller için de bu değişkenlerin olmadığı modeller kullanılmaktadır. Bu gözlem de model tercihlerinin tahmin performansına etkisinin birbirinden bağımsız olmadığını göstermektedir.

v. Sonuç olarak, model tercihleri tahmin performansını önemli derecede etkilemektedir. En iyi denklemlerde ölçüt modele göre yüzde kırklık bir iyileştirme görülürken, en kötü denklem, ortalamada, ölçüt modelden dört kat daha fazla tahmin hatası yapmaktadır.

ENFLASYON

Sonuçlara enflasyon özelinde bakıldığında,

i. Üç aydan dokuz aya kadar olan tahminlerde, DI-AR Lag tipi modeller en iyi denklemlerde daha yaygın şekilde kullanılırken, en kötü performans gösteren modellerde DI tipi denklem daha yaygın şekilde kullanılmaktadır.

ii. Hem en iyi performans gösteren beş modelde hem de en kötü performans gösteren beş modelde FHLR yaklaşımı daha sık görülmektedir. Bununla, birlikte SW yaklaşımı ile ortaya çıkan tahmin hataları FHLR yaklaşımı ile olana oldukça yakındır. Bu durum, FHLR yaklaşımı kullanmaktan elde edilen kazanımın sınırlı olduğunu göstermektedir.

iii. En iyi modelleri gösteren tablolardan ortaya çıkan sonuç, faktör sayısını belirlemek için kullanılan Bai ve Ng (2002) kriterlerinden IC1 ve IC2 yanında, BIC3'ün de kullanıldığıdır.

iv. Veri setlerinden çok sayıda faktör kullanmak gerektiği yönünde tavsiye veren PC3 ve IC3 kriterlerinin en kötü performans gösteren denklemlerde yaygın şekilde kullanıldığı görülmektedir. Bununla birlikte, en iyi modellerde de bu kriterlerin kullanıldığı görüldüğünden kapsamlı bir analiz yapmadan model tercihleri hakkında bir yargıya varmak mümkün olmayabilir.

v. En iyi modellerde, Avrupa Birliği verilerini çıkarmak tahmin hatalarını düşürürken finansal değişkenleri veya faiz oranlarını veri setinden çıkarmak tahmin hatalarını artırmaktadır.

vi. On iki ay sonrası için yapılan tahminlerde, en iyi modeller dahi ölçüt modele göre daha iyi sonuç vermemektedir.

vii. Model tercihleri, faktör modellerinin başarısını önemli ölçüde etkileyebilmektedir. Ölçüt modele göre, en iyi denklemlerde, yüzde 30 daha az tahmin hatası yapılabilirken en kötü modellerde yüzde 20 daha fazla hata yapılabilmektedir.

BORSA

Borsa özelinde sonuçlara bakıldığında,

i. En iyi performans gösteren modellerde, ilk tahmin değerlendirme döneminde DI-AR-Lag tipi tahmin denklemi daha yaygın olarak kullanılırken, ikinci tahmin değerlendirme döneminde DI tipi denklemler daha yaygın olarak kullanılmaktadır. En yüksek tahmin hatası yapan modellerde ise DI-AR-Lag tipi denklem daha yaygın olarak kullanılmaktadır.

ii. Faktör elde yöntemlerinin tahmin performansına ektisi karşılaştırıldığında ise, SW ve FHLR yaklaşımlarının birbirlerine yakın tahmin hataları yaptığı görülmektedir.

iii. Faktör sayısı belirleme yöntemlerinden IC1 ve IC2 yanında yine BIC3 iyi tahmin performans sergilemektedir. Yüksek tahmin hatası yapan modellerde ise PC3 ve IC3 kullanılmaktadır.

iv. En iyi modeller genel olarak ölçüt modelden daha iyi performans göstermektedir. Bu durum, ölçüt modelin tahmin hatalarının yüksek olduğu bilgisi ile birlikte değerlendirilmelidir. Böylece, her ne kadar ölçüt modelden daha iyi tahminler üretilebilse de bu durum faktör modelleri kullanılarak hisse senedi piyasasının başarılı bir şekilde tahmin edilebileceği şeklinde yorumlanmamalıdır.

v. Model tercihlerine bağlı olarak, ölçüt modele göre yüzde 35 iyileşme sağlanabilirken, en yüksek tahmin hatası yapan modellerde ölçüt modele göre yüzde 75 daha yüksek tahmin hataları görülmektedir.

SONUÇ:

Bu tezde, reel, fiyat ve finansal bloktan birer değişken için faktör modelleri ile tahmin elde edilmiş ve tahmin performansına etki eden unsurlar karşılaştırılmıştır. Çok sayıda değişkeni tahminlerde kullanmaya imkân veren bir diğer yaklaşım olan tahmin birleştirmesi ile de bir karşılaştırma yapılmıştır. Sonuçlar, faktör elde etme yönteminin sonuçlar üzerinde etkisinin sınırlı olduğuna işaret etmektedir. Faktör sayısı, veri seti detayı ve tahmin denklemi tipine dair seçimlerin performansı önemli ölçüde etkileyebildiğini göstermiştir. Ayrıca, modellerin tahmin performansı zamanla değişebilmektedir. Bu çerçevede, tahmin modellerinin sürekli bir şekilde takip edilerek değişen şartlara göre modellerin güncellenmesinin faydalı olacağı değerlendirilmiştir.