Essays on Applied Time Series and Macroeconometrics

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig–Maximilians–Universität München
2012

vorgelegt von
Xiaowen Jin

Referent: Professor Dr. Gerhard Illing
Korreferent: Professor Dr. Gebhard Flaig
Promotionsabschlussberatung: 15. Mai 2013

To My Family

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Dr. Gerhard Illing for his continuous teaching, guidance and encouragement of academic work and personality development. I am also grateful to my secondary supervisor Prof. Dr. Gebhard Flaig, for his countless critical advice of time series econometrics and dissertation writing.

I thank my colleagues from Munich Graduate School of Economics and from Seminar for Macroeconomics at LMU: Desislava Andreeva, Agnès Bierprigl, Jin Cao, Lei Hou, Sebastian Jauch, Sebastian Missio, Monique Newiak, Angelika Sachs, Heike Schenkelberg, Sebastian Watzka, Michael Zabel and Jiarui Zhang. They all helped the progress of my project in many ways.

I am indebted to my research internship supervisor: Sebastian Willutzky from Allianz Global Investors, for opening my eyes of Kalman Filter and mix frequency time series analysis. In particular, I am grateful to my co-authors of my second chapter: Jasper de Winter and Jos Jansen from De Nederlandsche Bank. It is my great pleasure to work with you. The second chapter belongs to De Nederlandsche Bank's research programme 2012.

Last but not least, many thanks to Zhejiang University, my mentors Prof. Dr. Yuexiang Jiang and Yanmin Qian, my parents and husband.

Xiaowen Jin

Acronyms

ERPT	Exchange Rate Pass-Through
NEER	Nominal Effective Exchange Rate
REER	Real Effective Exchange Rate
FPI	Foreign Price Index
BDI	Baltic Dry Index
MFVAR(/MA)	Mix frequency Vector Autoregressive (/Moving Averaging Model)
FA-MFVAR	Factor augmented MFVAR
RMS(F)E	Root Mean Square (Forecast) Error
MIDAS(-AR)	(Autoregressive) Mixed Data Sampling Regression Model
FA-MIDAS	Factor augmented MIDAS
QVAR	Quarterly VAR Model
(AR-)BEQ	(Autoregressive) Bridge Equation
SFM	Static Factor Model
DFM	Dynamic Factor Model

Preface

A time series is a sequence of data observed over successive time, such as weekly, monthly, quarterly and yearly data. Time series analysis is the statistical models to analyze or represent characteristics of time series data. As a rapidly evolving field, it exerts profound influence for empirical macroeconomic and monetary policy studies. For instance, the co-integration and structural models are widely utilized to detect the association among macroeconomic variables, autoregressive model is deemed to be the benchmark of time series forecasting.

The primary purpose of this dissertation is to apply modern time series technique on macroeconomic issues. It consists of three self-contained essays. Chapter 1 estimates exchange rate pass-through (ERPT) in China by both linear and VAR models, further investigates its relationship with exchange rate regime, and explores the reasons for divergent ERPT for different price indices. Chapter 2 and 3 have the identical focus of Gross Domestic Production Index, large information set and mixed frequency data. Chapter 2 conducts a systematic comparison of the short-term forecasting performance of linear statistical models in a pseudo-real time setting. We also evaluate subjective GDP forecasts by professional forecasters. Chapter 3 lays out a factor augmented mix frequency VAR/MA model (FA-MFVAR/MA) to construct a GDP based monthly coincident index for German.

In Chapter 1, we shed light on the transmission movement from exchange rate to price level for China, namely ERPT, especially in the background of exchange rate policy reform. In the wake of dramatically increasing current account surplus, the revaluation of Chinese Yuan causes heated controversy in international community. Albeit Chinese Yuan appreciated almost 20% between 2005 and 2008, the adjustment of exchange rate appears to be the result of be political pressure. For a long time, China is on the edge of being labeled as currency manipulation country. Despite all of those, each step of exchange rate reform is extremely sluggish, due to the anxiety of economic growth stagnation, high unemployment and inflation.

We concentrate on the potential impact of exchange rate on inflation. If a low ERPT is measured, it reflects the fact that exchange rate adjustment has minor influence on price stability. The implication behind is that, China's central bank is qualified for pursing a relatively independent monetary policy.

In the context of fixed exchange rate regime, the meaning of measuring ERPT casts much doubt. Virtually, the exchange rate data employed for ERPT measurement covers wide range of trade partners. US follows Japan, Korea and Taiwan, is merely the fourth largest importing region for China. Researchers universally adopt weighted average exchange rate index, instead of exchange rate against one single country, i.e. the nominal effective exchange rate (NEER), which takes all main trading partners' currencies into account and fluctuates all the time. Henceforth the pegging to US dollar is not an obstruction for our estimation.

To assess ERPT in China, we primarily sketch out an autoregressive distributed lag model in accordance with Devereux and Yetman (2010). The linear regression is theoretically approaching the law of one price and producer pricing behavior. The dependent variables include foreign price index, inflation rate, broad money supply, distribution cost and output gap. In order to alleviate the endogeity issue of linear model, vector autoregressive model with the identical variables is executed, by the mean of Cholesky decomposition and impulse response function.

Different from the existing literatures, we firstly estimate China's ERPT systematically by both linear and VAR models. The concern about the connection between ERPT and exchange rate regime is also a contribution. We introduce the dummy variables in the linear model and adopt VAR model simultaneously. In contrast with the simple division of sample as Wang and Li (2009), our treatment is more rigorous and reliable. Moreover, we do a novel cause survey of disparate pass-through for CPI and PPI.

The linear model shows that 1% appreciation of NEER declines 0.132% CPI inflation rate and 0.495% PPI inflation rate over the long run. The VAR model supports results of the linear model, in terms of a fairly low CPI pass-through and relatively higher PPI pass-through. With considering the exchange rate regime, we find that it matters for the CPI pass-through, but not for the PPI. Inter alia, CPI pass-through in fixed exchange rate period is higher. Combined with the fact that appreciation declines inflation rate, it implies that Chinese government could pursue more flexible exchange rate policy. In addition, the reasons for the low ERPT for CPI are also discussed. We consider price control, basket and weight of Chinese price indices, distribution cost, non-tradable share and imported input for analysis.

Chapter 2 and 3 revolve around the estimation and forecast of GDP, using a large-scale datasets of monthly and quarterly data. Information on economic activity and its short-term prospects is of great importance to decision makers in governments, central banks, financial markets and non-financial companies. The GDP, which refers to the market value of all officially recognized final goods and services produced in a given period, is regarded as the most comprehensive and reliable macroeconomic indicator. Nevertheless, two drawbacks hinder its function for timely monitoring and precisely forecasting the state of economy. At first, as a token of macroeconomic state, a monthly GDP indicator is unavailable but requisite for a great amount of empirical economic studies. Economists have to seek for substitution index, e.g. industrial production and composite index. Additionally, an initial estimate of quarterly real GDP is published around six weeks after the end of the quarter. The substantial publication lags precipitate a front-edge topic for GDP forecasting, namely nowcasting[1] and short-term predicting. The estimation of monthly GDP and forecast of quarterly GDP is distinguishing but also highly correlated issues. In certain models, both problems can be solved simultaneously.

When the simple models are fully exploited, one branch of researchers deliberate the added value of mixed frequency and large scale data. Generally, the mixed frequency models here refer to the combination of monthly indicators and quarterly GDP data in the same framework, the estimation relies on Kalman filter. Large scale data denotes the large information sets which cover full range of economic activities (industrial production, prices of goods and services, expenditures, unemployment, financial market prices, loans and consumer and business confidence), the total number of series can be over 100. Former studies adopted small scale series due to the feasibility of model estimation. Fortunately, factor model creates an explicitly specified and statistically meticulous dimension reduction scheme for large scale information sets.

Specifically, the superior performance of two state of art methodologies arises our attention. Mariano and Murasawa (2010)'s mix frequency VAR model (MFVAR) displayed a desirable performance of monthly GDP construction, while Angelini et al. (2011) and Bańbura et al. (2011) put forward a dynamic factor model, which exhibited empirical tractability and typically good fit forecasting performance, in the use of large and mixed frequency datasets. As a consequence, the improvement and comparison of mixed frequency and large scale approaches for GDP estimation and short-term forecasting become our natural point of entry.

In Chapter 2, we investigate the optimal GDP short-term prediction mechanism handling large information set, in a pseudo-real time setting. During the preceding financial crisis, forecaster across the world failed to forecast the depth and duration of the crisis.

[1] Forecasting data in the current quarter or month.

This phenomenon fuels concern about the forecasting performance of mechanical models during crisis. There is a generally argument that it is hard to beat simple model (e.g. random walk and autoregressive model). However, they did a bad job for the past crisis. Therewithal, we doubt the status of simple model by more sophisticated models, like factor model, mixed frequency model and professional forecasters, etc.

What we reinforce here is the pseudo-real time forecast design. Real time data stands for the information delivered immediately after collection, often adjusted as better estimates become available. In this definition, two problems concerning real time data are the publication lag and data revision. Pseudo-real time setting means the mimic of real-time flow of information. Owing to the incomplete revision histories for all monthly series, we have to ignore the data revision issue. Luckily, strand of literatures support the argument that, a truly real time data would most likely display minor impact of forecasting practise[2]. Finally, we only aim at the publication lag issue, via replicating the availability of the data at the time the forecast was made.

In all, we implement eleven models for the broad range of comparison. Concerning the handling of a large information set, we contrast pooling with factor method. Pooling approach is to estimate the relationship between each indicator and GDP growth individually and somehow average or weight these forecast to come up with one central forecast for GDP growth. While factor model in the contrary, extracts the common factor out of all indicators and regress that factor on GDP. Concerning the treatment of variables at different frequencies, we contrast quarterly with mixed frequency model. Quarterly model simply aggregates monthly indicators, while mixed frequency model incorporates monthly and quarterly indicators in a framework. We also make comparison between before-mentioned mechanical model with professional forecast. The intuition behind is that expert judgement of the respondents could improve the forecast of the mechanical models, since panelists seems to add a sizeable judgmental element to the forecasts. The three categories of comparisons (i.e. pooling versus factor approach, quarterly versus mixed frequency model, mechanical model versus professional forecast) comprise the linchpin of our forecast evaluation.

This chapter could be partly viewed as an extension and update of Rünstler et al. (2009). They paralleled a continuum of factor based method for short-term GDP forecasting of euro area. We add some novel methodologies, such as MFVAR, mixed data sampling regression (MIDAS), and their factor augmented version. Our sample period (1996.I-2011.III) allows us to compare the models' forecasting abilities in the period before

[2]See Bernanke and Boivion (2003), Diron (2008), Foroni and Marcellino (2012), Schumacher and Breitung (2008)

the financial crisis of 2008 (Great Moderation) and the much more volatile subsequent period (financial crisis and its aftermath). Our second contribution comes from the survey of the potential usefulness of forecasts made by professional analysts (published by Consensus Forecasts on a quarterly basis). They may be deemed to "wisdom of the crowds", reflect much more information than the statistical information set. However, the quarterly consensus forecasts have only been used once, for a forecasting "horserace" for GDP growth in the Netherlands (de Winter, 2011). König et al. (2003) also stated that it is the most accurate professional forecaster, in contrast with other individuals. We investigate for our six countries to what extent subjective forecasts by analysts contain information beyond that generated by the best mechanical statistical model.

The empirical work has concern on the Eurozone and its five largest countries. We find that it pays off to model monthly indicators, in particular as the horizon shortens and more monthly information is processed. The static and dynamic factor model consistently outperform all other approaches. There is a significant difference between forecasting methods that extract factors and pool single indicator models. Factor based models' superiority is most demonstrated in times of crisis. Statistical models significantly differ in the rate at which they are able to absorb monthly information as time goes by. Contrast with mechanical models, the professional forecasts displays poor predicting performance on their own, but contain valuable extra information from the view of encompassing test.

In Chapter 3, we verify whether the large scale and mixed frequency datasets benefit for generating GDP based monthly indicator, in the framework of FA-MFVAR/MA model. To compensate the lack of monthly GDP, one idea is to generate prominent Stock and Watson coincident index. Nonetheless, this index is constructed via a large factor model purely on monthly indices, abandoning the GDP data completely. Another idea is the aforesaid MFVAR approach pursued by Mariano and Murasawa (2010), merely limited number of series can be included notwithstanding. Angelini et al. (2011) is one among the few papers that estimates the monthly GDP indicator through both large scale and mixed frequency datasets. However, the results are disappointing.

In view of aforementioned weakness, our FA-MFVAR/MA method is a promising attempt for monthly GDP construction. It proceeds in two steps. At first, the single factor is extracted by either standard principal component or the latest two-step method. Afterward, bivariate MFVAR/MA model is established between quarterly GDP and the single factor. The estimation of the parameters and monthly GDP growth is realized by Kalman filter and maximum likelihood estimation.

To the best of our knowledge, it is the first time that FA-MFVAR/MA model is exploited for monthly GDP construciton. Our methodology sufficiently accommodates information from both the quarterly GDP and numerous monthly indicators, broadly absorbs essence from both factor and mix frequency VAR model. Notably, the fitness of our model depends on specific datasets.

Taking Chow and Lin method, single indicator MFVAR model and dynamic factor model as counterpart, we investigate the pros and cons of our FA-MFVAR/MA model by employing German data. Our result renders considerable superiority of FA-MFVAR model over the remaining methodologies, via the schematic analysis and out of sample root mean square error comparison. It implies that large datasets might improve the performance of index construction, but it relies on a plausible model setup.

References

Angelini, E., Canba-Mendez, G., Giannone, D., Reichlin L. and Rünstler G., 2011. Short-term Forecasts of Euro Area GDP Growth. *The Econometrics Journal* 14, 25-44.

Angelini, E., Bańbura, M., Rünstler G., 2008. Estimating and Forecasting Euro Area Monthly National Accounts from a Dynamic Factor Model. *ECB Working Paper Series*, No 953.

Bańbura, M., Giannone, D., Reichlin, L., 2011. Nowcasting. In *The Oxford Handbook of Economic Forecasting*, Clements M. and Hendry D (eds). Oxford University Press: Oxford, 193-224.

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

Devereux, M., Yetman, J., 2010. Price adjustment and exchange rate pass through. *Journal of International Money and Finance* 29(1), 191-200.

Diron, M., 2008. Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data. *Journal of Forecasting* 27(5), 371-390.

Foroni, C., Marcellino, M., 2012. Comparison of mixed frequency approaches for modelling Euro areae macroeconomic variables. *ECO Working Paper* 2012/07.

König, E., Dolmas, S., Piger,J., 2003. The Use and Abuse of Real-Time Data in Economic Forecasting. *The Review of Economics and Statistics* 85(3), 618-628.

Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., den Reijer, A., Jakaitiene, A., Jeloniek, P., Rua, A., Ruth, K., van Nieuwenhuyze,C. 2009. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting* 28, 595-611.

Schumacher, C., Breitung, J., 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting* 24, 368-398.

Wang, J., Li, N., 2009. Exchange Rate Pass-through: The Case of China. *Journal of Economic Research* 4, 17-27.

Winter, J., 2011. Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models. *DNB Working Papers* No. 320.

Chapter 1

An Empirical Study of Exchange Rate Pass-Through in China

1.1 Introduction

The debate about revaluation of Chinese Yuan has been a hot topic for long time. Although Chinese Yuan appreciated almost 20% during 2005 and 2008, the degree of appreciation is not satisfied by international community. The exchange rate war restarted since subprime crisis was eased. Krugman's impossible trinity pointed out, in the framework of an open economy, the exchange rate stability, domestic monetary policy independence and free capital movement could not be obtained in the same time. It is no doubt that China would open its capital market gradually. Therefore the reform of the exchange rate is inevitable. Nevertheless, when and how would the reform be operated is still a major issue for China's central bank. After all, they would not like to sacrifice economic growth, suffer from high unemployment and inflation.

The transmission movement from exchange rate to price level, namely exchange rate pass-through (ERPT), is one of the concerns. In the new open economy macroeconomics, the degree of ERPT is crucial for appropriate monetary policy. A low ERPT implies that government need not worry about the price instability or inflation, when adjusting exchange rate policy. It is thought to provide great freedom for pursuing an independent monetary policy. In this chapter, we examine the movement from exchange rate to domestic price, which may assist Chinese government to consider the revaluation issue of Chinese Yuan in an manner. Additionally, due to the fact that Chinese Yuan experienced three years' revaluation, does this appreciation influence ERPT?

The meaning of measuring ERPT is suspicious, since most of time Chinese Yuan is purely pegging to US dollar. Whereas, Figure 1.1 illustrates the major importing countries and regions for China in 2000, which demonstrated that exchange rate should not only concern US dollar, but also other main currencies. Nominal effective exchange rate (NEER), which fluctuates all the time, is universally used. It takes all main trading partners' currencies into account.

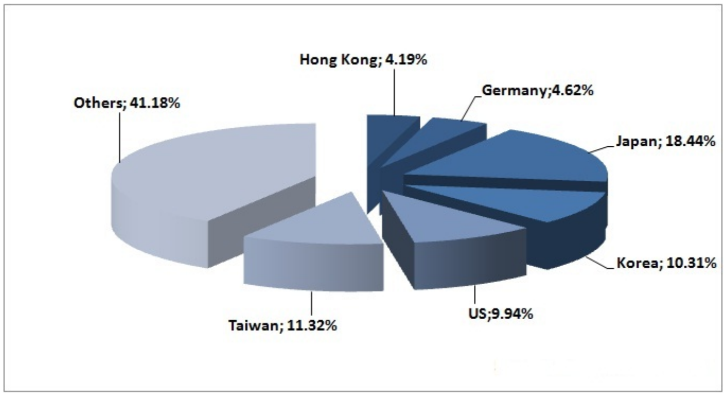


Figure 1.1: Main Importers for China, IMF IFS Data

In the empirical literature, three approaches are allied to gauge ERPT, single equation method, vector Autoregressive (VAR) model and cointegration model. However, few studies calculated China's EPRT. Ca' Zorzi et al. (2007) analyzed plenty of emerging countries' ERPT, including China. Shu and Su (2009), Wang and Li (2009), Chen and Liu (2007) calculated China's ERPT for a batch of price indices.

One contribution of this chapter is to exploit both single equation and VAR methods. Each method has its own drawback, the systematically estimation of the ERPT by two methods benefits from overcoming the deficiency in single approach. The second intriguing contribution is that it sheds light on the association between exchange rate regime and ERPT comprehensively. Although Wang and Li (2009) mentioned the increase of the ERPT during appreciation period by simple division of sample, their method is not convictable owing to the limited availability of data. We involve the dummy variables in the single equation method and use VAR model additionally, which is more reliable. Last but not least, we investigate the cause of pass-through discrepancy for various price indices (CPI and PPI), while previous work only stated the discrepancy, but not the reasons.

To assess the degree of pass-through for CPI and PPI, we firstly adopt the ordinary least square model (OLS). The results demonstrate that appreciation of Chinese Yuan could decrease the CPI inflation rate, but the degree is statistically insignificant. The effect of PPI is more noticeable than CPI's in the long run. Nevertheless, the shortage of the single equation regression is that it neglects the impact of inflation on exchange rate. As a result, VAR analysis with Cholesky decomposition could strengthen the robustness of previous OLS measure, which regards all the variables endogenously. The impulse response function illustrates a limited impact of NEER on CPI, and relative more conspicuous response in the case of PPI. This finding is in line with OLS estimation.

Besides, the factors which influence pass-through are discussed. The linear regression results show that exchange rate regime indeed impacts on CPI pass-through. The CPI pass-through is higher in fixed exchange rate regime. Inspired by Bustein et al. (2003), Campa and Goldberg (2006), we supply evidence which interprets the discrepancy between CPI and PPI pass-through. The evidence covers the divergent composition of price indices, the share of imported inputs and non-tradable goods, distribution cost.

This chapter is organized as follows. Section 1.2 is the literature review. In Section 1.3, the stylized facts on post 90's China economy is illustrated. Section 1.4 presents two analytical framework, namely OLS and VAR model. The results of ERPT measurement of CPI and PPI are reported. Section 1.5 further discusses the pass-through topic, investigates the relation between exchange rate regime and ERPT, and the reason for a fairly low CPI pass-through. Section 1.6 concludes.

1.2 Literature Review

1.2.1 Theoretical Work

In the traditional open economy macroeconomic models, the purchasing power parity holds all the time, which implies that pass-through is complete and immediate. Nonetheless, a strand of empirical works have demonstrated that neither the purchasing power parity nor the law of one price holds in real world, see Ihrig et al. (2006), Bailliu and Fujii (2004), etc.

One of the theoretical explanations is the export producers' pricing to market behavior (Krugman 1987). In an incomplete competition market, a firm will set different price for its goods across segmented national markets, to compete with local firms. It will adjust its mark-up in accordance with exchange rate shock additionally. A similar framework stemming from pricing behavior is local currency pricing, suggested by Devereux and Engel (2003). Instead of setting price in producer currency, a firm could choose local currency pricing, when exporting to countries with relative low exchange rate variability or stable monetary policies. Consequently, in new open economy macroeconomics, the incomplete ERPT is explained by the fact that some firms follow local currency pricing, while some others practise producer currency pricing.

Another cause for the incomplete ERPT is the distribution cost of tradable goods, which is proposed by Bustein et al. (2003). They stated that the costs of distributing tradable goods such as transportation, wholesale and retail also contribute to the incomplete ERPT. They defined distribution margin and calculated it for America and Argentina market. The results reflect the huge scale of distribution sectors. The growing globalization of production (Campa and Goldberg 2006) is another explanation. If the proportion of imported components is higher, then the domestic prices will be more sensitive to exchange rate. Because it provides another channel for exchange rate transmission except importing final goods. There is also argument related to slow adjustment of goods prices (Devereux and Yetman 2010). Due to the endogenous sticky price and menu cost, the ERPT is not complete.

1.2.2 Empirical Study

In empirical work, three methods are mainly applied to measure ERPT. Firstly is linear model. From the law of one price, Ihrig et al. (2006) added interpretation terms like trade barriers, lags of independent variables and other control variables in order to analyze the association between import price, consumer price and exchange rate in G7 countries. Especially, they employed a general to specific algorithm from Hendry and Krolzing (2001) to select the appropriate specifications for independent variables. They found a fairly low long-run ERPT of CPI in the period 1990-2004, from 0.002 to 0.083.

Campa and Goldberg (2005) started from the law of one price, and treated export prices of foreign producer by a monopolistic mark-up over exporter marginal cost. They estimated ERPT into import prices across OECD countries, the cross-country ERPT is around 46% in short run and 65% in long run.

Based on monopolistic producers' profit maximization, Bailliu and Fujii (2004) implemented dynamic panel data GMM model to measure ERPT for 11 industrial countries. On average, the estimated short and long-run ERPT of PPI are 0.202 and 0.301, while the one for CPI are 0.080 and 0.160 respectively.

The second approach is the Vector Autoregressive model (VAR), see McCathy (2000). This technique regards ERPT as the price response to the structural shock of exchange rate. He examined the impact of exchange rate and import price shock on CPI and PPI for industrial economies. By utilizing impulse response function and variance decomposition, the results demonstrated that the aggregate consumer price pass-through is moderate in most of the countries. Hahn (2003), Faruqee (2006) employed this method for euro area, while Ito and Sato (2008) applied it for east asian.

The third method is the Cointegration and Error Correction Model (ECM). Its advantage is that the ERPT could be simply estimated by the stable cointegration relationship among exchange rate, domestic price, foreign price. Kim (1998) adopted ECM on US data, while Beirne et al. (2009) employed it on eastern Europe. Whereas, the conditions for a cointegration estimation are much stronger than the ones of VAR model.

Concerning the empirical works of China, Ca' Zorzi et al. (2007) executed VAR model to plenty of emerging countries, including China. They found that 1% of exchange rate shock leads to 0.08% response of consumer price over 4 quarters, and 0.77% over 8 quarters. The corresponding results for China are higher in comparison with those in the other literatures. Referring to similar approach, Chen and Liu (2007) advocated that 1% of NEER shock only maximally causes 0.0076% response on the CPI. Shu and Su (2009) applied single equation regression combined with general to specific method. The results illustrated that 1% of appreciation causes 0.065% and 0.104% fall of consumer price and producer price respectively in short run, 0.201% and 0.367% decrease over long run. Meanwhile, the estimation of long-run ERPT for CPI is 0.24 in Wang and Li (2009).

Furthermore, the relation between ERPT and monetary policy regime is often discussed. The idea that inflation environment could affect ERPT, was firstly advocated by Taylor (2000). He established a simple sticky price setting model with market power. Moreover, he claimed that the low inflation may be associated with less persistent changes in the costs and the prices in other firms in the economy, which subsequently results in a low ERPT. Bailliu and Fujii (2004) empirically supported Taylor (2000)'s argument. They suggested that ERPT declined during inflation stabilization period for industrialized countries. Choudhri and Hakura(2006) provided powerful evidence that there is a positive relation between ERPT and average inflation rate across 71 countries. Ihrig et al. (2006) also compared ERPT in subsamples, and found a significant decline for almost all G7 countries.

For emerging and transition countries, Coricelli et al. (2006) stated that the reform towards more flexible exchange rate disconnects the link between exchange rate and price, under an inflation targeting framework. Beirne et al. (2009) claimed that countries in fixed regime across nine central and eastern European EU member states have higher ERPT. Barhoumi (2005) applied the Panel Cointegration model for 24 developing countries. He suggested that countries with fixed exchange rate, lower tariff barriers and higher inflation rate are associated with a higher long-run import price pass-through. Wang and Li (2009) found the increase of ERPT during appreciation period in China.

1.3 Stylized Fact on Post-1990's China

Due to the particularity of China economy, either exchange rate or inflation rate is influenced by specific policy regulation. Especially for the exchange rate policy, the Krugman's trilemma has claimed its close connection with monetary policy and capital market control. In this section, the evolution of China's policy regulations after 1990 is briefly summarized.

1.3.1 Capital Market Deregulation

Heavy control is the principal characteristic of China's capital market. As a result of the reform and open-door policy, foreign direct investment (FDI) was liberalized in the mid 90's. Other restriction of remaining inflows and outflows are still fairly strict nowadays. For local financial institutions and those aboard, Chinese government introduced a program called qualified foreign institutional investors (QFII) and qualified domestic institutional investors (QDII) in 2003. The institutions should satisfy the requirements on the scale of registered capital, financial status etc., then the special invest quota will be authorized. With the gradually deregulation, the number of QFII and QDII reaches 98 and 93 respectively. Non-QDII institution and individuals are forbidden to invest in the foreign capital market. Although the government controls so strictly in the capital market, the efficiency is not ideal. Large amount of hot money continuously flows in China, especially around 2007.

Moreover, the domestic capital market is incomplete as well. Stock market is the primary capital market in China. In 1990, Shenzhen and Shanghai stock exchanges were set up successively. Accompanied with the establishment of China Securities Regulatory Commission in Oct 1992, the period between 1993 and 1998 is the forming and initial developing stage. Along with the promulgation of Securities Law in 1998, the evolution of stock market stepped in the regularization stage. The reform of non-tradable shares is another milestone in this stage. Up to the end of 2009, there are 1718 listed companies, and the total market value reaches 24.40 trillion (see Figure 1.2). Stock becomes one of the most important households of financial assets.

The bond market is relatively lagged behind. The treasury bills are the only issued bond for a long time. In 2003 and 2004, the commercial paper and certificates of deposit were introduced. Meanwhile, the corporate bond market is insignificant.

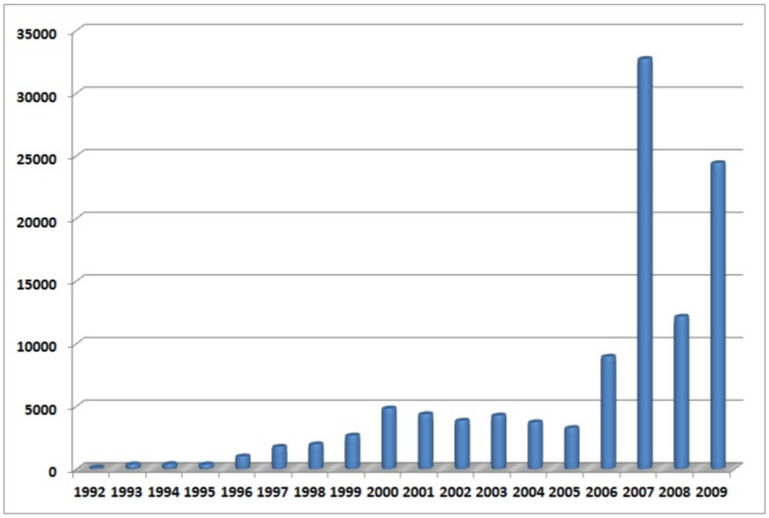


Figure 1.2: Total Market Value of Stock, IMF IFS Data

1.3.2 Monetary Policy

The major monetary policy instruments of the PBC are interest rates, bank required reservation ratio and money supply. Compared to market-based interest rate in the other countries, the strict control of the PBC declined the efficiency prominently. The situation was improved by the interest rate liberalization from 1996 (see Figure 1.3). The reform started in the money and bond markets. The PBC marketized the interbank market rate, the bond market rate and policy financial bonds. For the foreign currencies, the lending rate and the large deposit rate were liberalized in 2000, afterwards the small deposit rate of major foreign currency could also be determined by the commercial banks. In the subsequent two years from 1998, the PBC has expanded three times of the floating band of commercial lending rates, and attempted the liberalization of long-run large deposit rate. Accompanied with the launch of Shanghai Interbank Offered Rate, the PBC's commitment of pursuing market-based interest rate is fulfilled gradually.

Required reverse ratio exerts much more influence than those from interest rate instrument (see Figure 1.4). The original aim of reserve ratio is to maintain the liquidity of banks, make sure the stability of banks in case of large withdrawals. Later, it is widely employed as an instrument of liquidity management, in order to control the credit growth and indirectly affect money supply. The first significant reserve ratio change occurred in 1998. After that, the PBC adjusted the reverse ratio 30 times from mid 2006 to 2008. Since Sep 25th 1998, the adjustment started to classify small and large financial institutions, in the sense that large institutions required with a high reserve ratio. Meanwhile, excess reverse

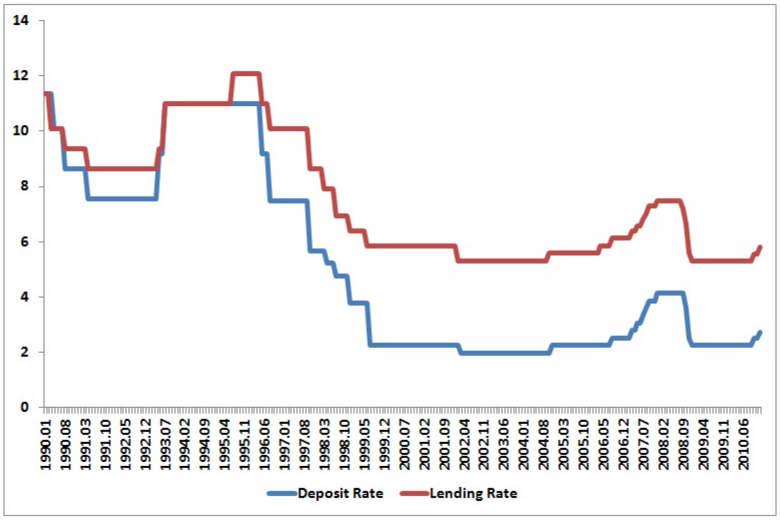


Figure 1.3: One Year Interest Rate, IMF IFS Data

ratio for commercial banks also exerts subsidiary function.

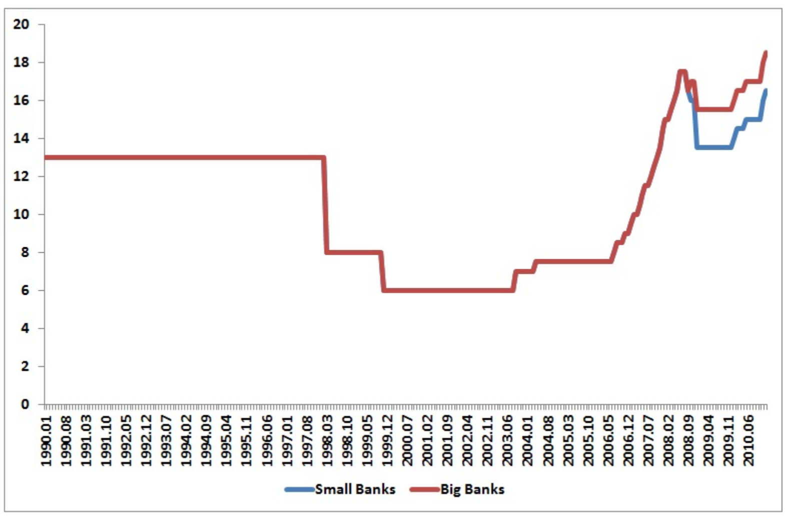


Figure 1.4: Reserve Ratio, IMF IFS Data

The control of base money supply is implemented through open market, currency and bill issuance. The base money and the money multiplier (monitored by reverse ratio) constitute the money supply. As China's normal anchor, money supply is crucial for execution of loosened or tightened monetary policy, although there is always a gap between actual and target broad money supply (M2). An intriguing phenomenon is the close correlation between inflation rate and M2 growth (see Figure 1.5), which implies that M2 played a key role in price stability.

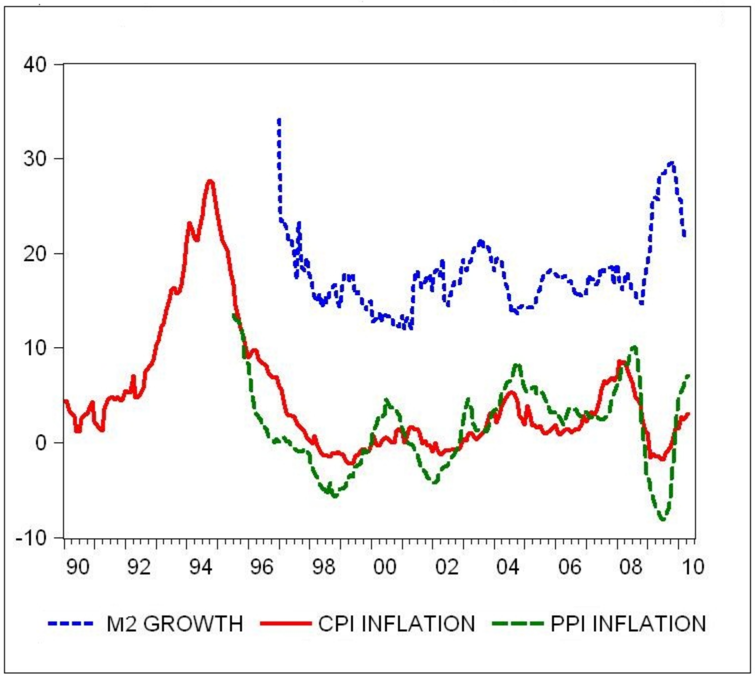


Figure 1.5: Money Growth and Inflation(12 months change), IMF IFS Data

1.3.3 Foreign Reserve and Exchange Rate Policy

In the 21^{st} century, attributed to the growing current account surplus and FDI inflows, reserve accumulation increases rapidly (see Figure 1.6). The current account surplus was growing especially fast from 2005, due to the trade surplus. The first time, the current account surplus decreased in 2009 since 2000. Following a sudden raise in 1997, the FDI's growth is relatively smooth afterwards. Although China owns the largest reserve accumulation all over the world, how to manage it is a severe task for PBC. Foreign reverse is the major component, while the proposition of position in IMF, Special Drawing Right (SDRs), and the other reserve is minor.

Oversize foreign reserve causes controversy about the exchange rate policy, consequently precipitates the policy reform. The reform is propelled step by step. Before 1994, there were two different exchange rates: the official rate and swap market rate. On Jan 1994, government merged the two exchange rates, implemented a so-called market-based, single and managed floating exchange rate system. However, it was still a pure US Dollar pegging system, the nominal USD/CNY rate was adjusted suddenly from 5.8 to 8.7 (see Figure 1.7).

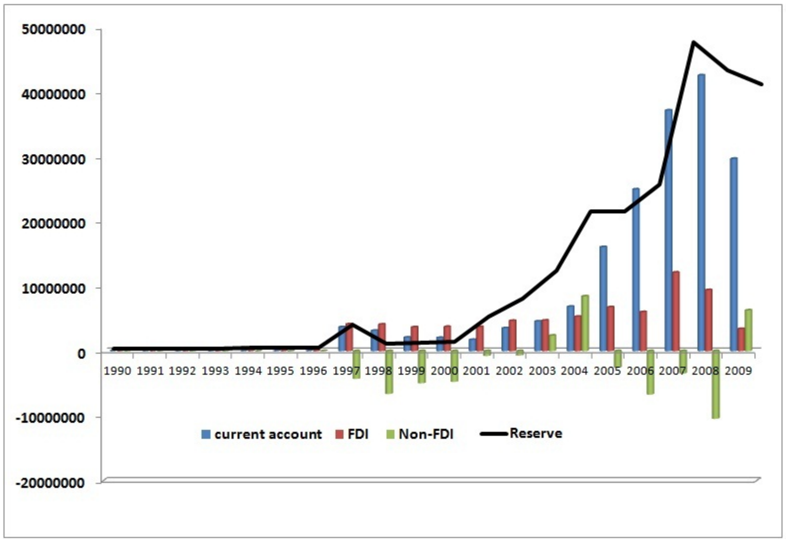


Figure 1.6: Reserve and Component (Million Dollar), IMF IFS Data

Since then until 2005, the nominal value of USD/CNY fluctuated in a rather narrow range around 8.28, even during the Asian financial crisis. Due to the imbalance of international payments, the challenge of independent monetary policy, overheating concern and pressure from western countries, China adopted a new manageable floating exchange rate policy on July 21, 2005. This new policy was based on market supply and demand, and with reference to a basket of currencies. The reference currencies consisted of US Dollar, Euro, Japanese Yen and Korea Won, but the operation was not transparent yet. In the subsequent 3 years, the nominal USD/CNY rate decreased from 8.28 to 6.84, appreciated nearly 20%. Along with the breakout of subprime crisis, the step of appreciation was ceased, and the nominal USD/CNY rate maintained in this level until 2010.

1.3.4 Inflation and Exchange Rate

Between 1992 and 1993, China experienced a serious inflation period. The CPI inflation rate reached the peak around 28% in Oct 1994. The subsequent tightening policy lowered the inflation rate successfully, without apparently negative impact on output. Whereas, the Asian financial crisis brought the CPI to another direction: deflation. The deflation was mild (the peak value was only 2.2% occurred in June 1999), but lasted until 2002. Afterwards, China's inflation rate was relatively low and stable except in 2007, when the asset price bubble was serious. For the PPI inflation, the general trend is analogous to

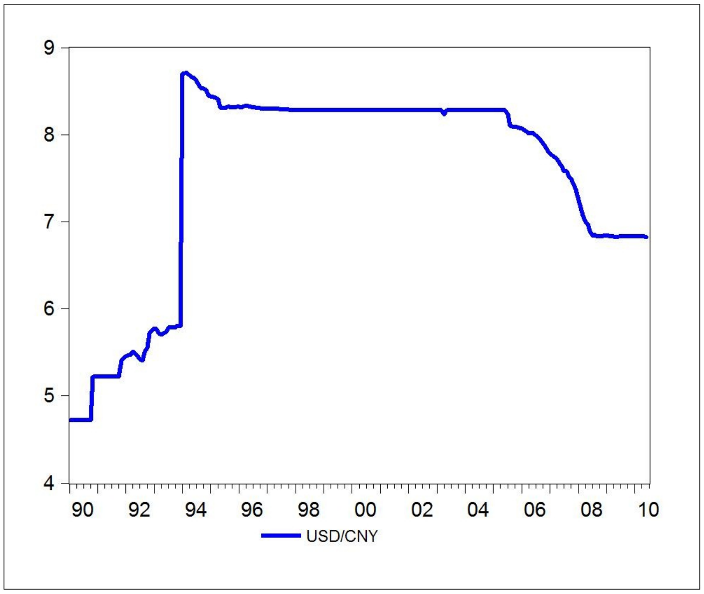


Figure 1.7: Nominal Exchange Rate vis-a-vis the U.S.Dollar, IMF IFS Data

that of CPI's, merely with more conspicuous variation. Figure 1.5 already illustrates the trend of CPI and PPI inflation rate based on yearly changes.

We has mentioned the importance of NEER in Section 1, Figure 1.8 describes the co-movement between domestic price (CPI, PPI) and the NEER. The trend of the CPI and PPI is similar. In comparison with NEER, they moved in the same direction during 1995-1997 and 2005-2010, while they varied in opposite direction during 1998 and 2004.

1.4 Empirical Analysis

1.4.1 Single Equation Analysis

Framework

Many literatures use single equation to regress the domestic price as a function of the exchange rate, foreign price and control variables such as output gap, e.g. Ihrig et al. (2006), Campa and Goldeberg (2005), Bailliu and Fujii (2004).

Following Devereux and Yetman (2010), we suppose a large number of import firms, which purchase a differentiated consumer good from foreign countries and sell it to domestic consumers. Demand function for firm i is as follows:

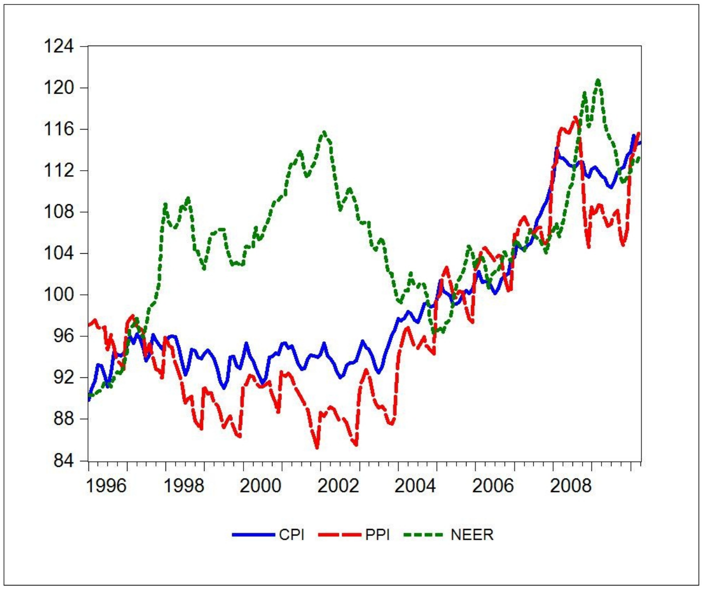


Figure 1.8: CPI, PPI and NEER (2005=100, base year), IMF IFS Data

$$C_t(i) = (\frac{P_t(i)}{P_t})^{-\lambda} C_t \tag{1.1}$$

where $C_t(i)$ is the demand of firm i, C_t is the total demand, $P_t(i)$ is the firm i's price, P_t is the composite price index for imported goods, λ is the elasticity of substitution.

The profit function for firm i is defined as:

$$\Pi_t(i) = P_t(i)C_t(i) - \frac{P_t^*}{S_t}\Theta_t C_t(i) \tag{1.2}$$

where S_t is the exchange rate foreign currency per unit of domestic currency, P_t^* is the all differentiated imported goods' foreign currency price, Θ_t is the per unit distribution cost. The firm i's import price setting is defined as:

$$P_t(i) = \frac{\lambda}{\lambda - 1}\Theta_t \frac{P_t^*}{S_t} \tag{1.3}$$

Assuming all import firms are identical, $P_t(i) = P_t$. The logarithm version of this equation can be written as:

$$p_t = \zeta_t + \theta_t + p_t^* - s_t \tag{1.4}$$

where $\zeta_t = \ln(\frac{\lambda}{\lambda - 1})$, the small letters represent natural logarithm.

As a result, the import price could be expressed as a function of the markup, distribution cost, exchange rate and foreign price. In order to control monetary policy's influence for domestic price, monetary policy instrument is introduced.

Data Description and Econometric Specification

We access two ERPTs respectively for consumer price index (CPI) and producer price index (PPI). The import price index is ruled out due to the limited data length. The NEER is used as the exchange rate index. Taking advantage of the definition of NEER and the real effective exchange rate (REER), the foreign price index (FPI) is calculated by FPI=NEER*CPI/REER. Through applying Hodrick-Prescott filter to gross industrial production data, we get an approximation of output gap. The monthly average Baltic Exchange Dry Index (BDI) is utilized to represent distribution cost (mainly transportation cost). The monetary policy variable adopts broad money supply (M2)[1].

Considering the price stickiness, the lagged variables for p_t, p_t^* and s_t are involved. In what follows, the ERPT could be estimated by the following Autoregressive Distributed Lag Model. Denoting the short-run ERPT by β_0, the long-run ERPT can be written as $(\sum_{i=0}^{5} \beta_i)/(1 - \sum_{i=1}^{5} \alpha_i)$[2][3].

$$\begin{aligned} dlncpi = \sum_{i=1}^{5} \alpha_i dlncpi_{t-i} + \sum_{i=0}^{5} \beta_i dlnneer_{t-i} + \sum_{i=0}^{5} \gamma_i dlnfpi_{t-i} + \zeta outputgap \\ + \eta dlnbdi + \delta dlnm2 \end{aligned} \tag{1.5}$$

$$\begin{aligned} dlnppi = \sum_{i=1}^{5} \alpha_i dlnppi_{t-i} + \sum_{i=0}^{5} \beta_i dlnneer_{t-i} + \sum_{i=0}^{5} \gamma_i dlnfpi_{t-i} + \zeta outputgap \\ + \eta dlnbdi + \delta dlnm2 \end{aligned} \tag{1.6}$$

[1] Detail description of data see Appendix 1.8.1.

[2] dln denotes the corresponding variable is in the first difference level of natural logarithm. lncpi and lncpi are represented for p_t, lnfpi for p_t^*, lnneer for s_t, outputgap for ζ_t, lnbdi for θ_t, lnm2 for broad money supply.

[3] Different versions of independent variables are taken into account. For instance, lagged variables only consider lncpi, lnneer or lnfpi, and different numbers of lags. According to the information criteria such as Akaike information criterion and adjusted R^2, this form is more suitable than the others.

Table 1.1: Unit Root Test

Variable	Deterministic	ADF test stat	PP test stat	5% Crit Value	Decision
lncpi	c,t	-1.312	-0.949	-3.438	I(1)
dlncpi	none	-2.962	-10.301	-1.943	
lnppi	c,t	-1.807	-1.632	-3.436	I(1)
dlnppi	none	-5.746	-5.681	-1.943	
lnneer	c,t	-2.364	-2.169	-3.436	I(1)
dlnneer	none	-8.347	-9.070	-1.943	
lnfpi	c,t	-3.214	-2.904	-3.436	I(1)/I(0)
dlnfpi	none	-2.351	-5.758	-1.943	
outputgap	none	-4.786	-8.963	-1.943	I(0)
lnbdi	c,t	-3.458	-2.150	-3.438	I(1)/I(0)
dlnbdi	none	-5.729	-7.910	-1.943	
lnm2	c,t	-0.503	- 2.773	-3.437	I(1)
dlnm2	c	-7.088	-18.022	-2.879	

Note: I(1) means the variable is integrated by order 1, I(0) presents that the corresponding variable is stationary.

We adopt monthly data from Jan 1996 to April 2010, taking 2005 as the base year (index equal to 100 in 2005), and all the data series is seasonally adjusted by X12-ARIMA. To exam the stationarity of these variables, Augmented Dickey-Fuller (ADF) Test and Phillips-Perron (PP) Test are executed in Table 1.1. The results indicate that most of the variables except output gap are integrated by order 1. Therefore in the OLS estimation, all the variables except output gap are taken the logarithm-difference form.

Results

All the OLS estimation results are reported in Table 1.2. We observe that the lagged CPI inflation has a positive influence on the current inflation. The coefficients of $lnneer_{t-1}$ and $lnneer_{t-5}$ are significant at the level of 10%, which implies that the movement from NEER to inflation is not immediate. For FPI, it seems that both the current and lagged FPI influence CPI inflation rate. The short and long-run CPI pass-through is 0.016 and -0.132 respectively. It suggests that 1 percent increase in NEER (1% appreciation) leads to 0.016% incline of CPI inflation rate in the first month and 0.132% decline in the long run. The hypothesis tests indicate neither short nor long-run ERPT is significantly different from 0 at the significance level of 5%[4].

[4]For the hypothesis test of long-run EPRT, delta method is utilized to compute the mean and variance, and establish statistic based on normal distribution.

Table 1.2: ERPT by OLS Estimation

CPI		PPI	
Variable	Coefficient	Variable	Coefficient
dlncpi(-1)	0.104 (0.082)	dlnppi(-1)	0.311***(0.081)
dlncpi(-2)	0.116 (0.082)	dlnppi(-2)	0.147*(0.085)
dlncpi(-3)	0.070 (0.079)	dlnppi(-3)	0.015 (0.088)
dlncpi(-4)	0.108 (0.078)	dlnppi(-4)	0.120 (0.086)
dlncpi(-5)	0.138*(0.080)	dlnppi(-5)	0.051 (0.076)
dlnneer	0.016 (0.027)	dlnneer	0.012 (0.028)
dlnneer(-1)	-0.058*(0.029)	dlnneer(-1)	-0.058*(0.031)
dlnneer(-2)	0.041 (0.029)	dlnneer(-2)	-0.069**(0.030)
dlnneer(-3)	-0.036 (0.029)	dlnneer(-3)	-0.011 (0.031)
dlnneer(-4)	0.026 (0.028)	dlnneer(-4)	-0.031 (0.030)
dlnneer(-5)	-0.050*(0.027)	dlnneer(-5)	-0.019 (0.030)
dlnfpi	0.613***(0.199)	dlnfpi	0.999***(0.212)
dlnfpi(-1)	0.035 (0.206)	dlnfpi(-1)	0.523 (0.233)
dlnfpi(-2)	-0.392*(0.203)	dlnfpi(-2)	-0.112(0.224)
dlnfpi(-3)	0.260 (0.204)	dlnfpi(-3)	-0.150 (0.222)
dlnfpi(-4)	-0.022 (0.214)	dlnfpi(-4)	-0.125 (0.229)
dlnfpi(-5)	-0.527***(0.194)	dlnfpi(-5)	-0.578***(0.215)
outputgap	0.000 (0.000)	outputgap	0.000 (0.000)
dlnbdi	-0.003(0.002)	dlnbdi	0.002 (0.002)
dlnm2	0.015 (0.034)	dlnm2	0.052 (0.036)
R^2	0.284	R^2	0.647
Short Run		Short Run	
ERPT	0.016 (-)	ERPT	0.012 (-)
Long Run		Long Run	
ERPT	-0.132 (-)	ERPT	-0.495 (-+)

Note: The figures in parentheses are standard errors, the figures in square parenthesis are p value. * significance at the 10% level, ** significance at the 5% level, *** significance at the 1% level. +(-) implies an ERPT elasticity is significantly different from 0(1) at the 5% level. dlncpi(-i) denotes $dlncpi_{t-i}$.

To keep the comparability of outcome, the same regression is applied on PPI, although it is not the best choice in the sense of information criteria and adjusted R^2. R^2 improved significantly compared to the same regression for CPI. According to the significance of coefficients, the PPI inflation rate is affected by previous PPI inflation rate, previous NEER change, and both current and prior FPI variation. The short and long-run PPI pass-through is 0.012 and -0.495 respectively. Initially, the PPI response to NEER is weaker than CPI, but over long run PPI pass-through is almost fourfold that of CPI's. The long-run ERPT is significant different from 0, while the short-run ERPT is not.

The measured ERPT from single equation illustrates that appreciation of domestic currency has negative impact on the inflation rate, especially significant in the case of PPI. The negative association between exchange rate and inflation is consistent with theory. In principal, the appreciation of domestic currency reduces the price of import goods directly, and declines the CPI and PPI partially. The impact on CPI and PPI is proportional with the percentage of import goods in these indices. The composition of price index is one of the main reason for the difference between CPI and PPI's ERPT. This will be discussed concretely in Section 1.5.2 .

In comparison with the industrial countries' estimation (mainly concentrates on long-run EPTR), the estimation of China's CPI pass-through in this chapter is a bit higher. A potential explanation is the divergence of economy structure. In view of other measurement for China's data, the long-run CPI pass-through is a little lower in our analysis. But the significant gap between CPI and PPI is widespread phenomenon.

1.4.2 Vector Autoregressive Model

The single equation regression neglects the fact that inflation rate could impact on exchange rate. Thereby, VAR model with Cholesky decomposition is supplement to measure ERPT. This approach regards all variables endogenously, strengthens the robustness of the previous OLS measurement.

The variables for VAR model include CPI and PPI, NEER, output gap (control supply shock), FPI and M2. All variables adopt the same form as OLS estimation. The baseline VAR model involves CPI and PPI separately, contains five variables. Under the framework, the results could be compared with the former OLS measurement.

VAR Approach and Cholesky Decomposition

A reduced-form VAR(p) model is as follows,

$$Y_t = c + A(L)Y_t + \mu_t \tag{1.7}$$

$$E(\mu_t \mu_t^{'}) = \Omega \tag{1.8}$$

where c is the deterministic, A(L) is a polynomial matrix of degree p in the lag operator L, μ_t is the vector of reduced-form VAR residuals.

In order to generate structural shock, namely orthogonalized innovations ε_t, a unique lower triangular Cholesky matrix C is applied,

$$C\epsilon\varepsilon_t = \mu_t \tag{1.9}$$

where $E(\epsilon_t \epsilon_t^{'}) = I$, $CC^{'} = \Omega$ and C= $\begin{bmatrix} C_{11} & 0 & 0 & 0 & 0 \\ C_{21} & C_{22} & 0 & 0 & 0 \\ C_{31} & C_{32} & C_{33} & 0 & 0 \\ C_{41} & C_{42} & C_{43} & C_{44} & 0 \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} \end{bmatrix}$.

According to the structure of Cholesky matrix C, the order indicates which shocks are not allowed to contemporaneously affect which variables. Hence, selecting the appropriate order of the endogenous variables through economic interpretation is crucial. Foreign price index is ordered first because that its residual is hardly influenced by any other shocks. Afterwards, we order the output gap. It is only affected by foreign price index and affects all other shocks except foreign price index. Assuming that monetary policy does not react to current inflation and NEER change, but to expected inflation (see Ito and Sato 2008), M2 can be settled prior to NEER. NEER is ordered before domestic price. Strand of empirical literature claimed that NEER granger causes price, but the opposite is not true. To begin with, the vector for VAR model is (dlnfpi,outputgap,dlnm, dlnneer, dlnp), where lnp_t could be lncpi or lnppi. The number of lags is determined by Akaike Information Criterion. Both CPI and PPI models are VAR(7).

Impulse Response Functions

In this section, we implement two baseline models with five variables, in order to check the effects of domestic price to different structural shocks over two-year (24 months horizon). The results for NEER shock to CPI and PPI are in Figure 1.9. The vertical axis in figures is the accumulative percentage change in domestic prices responding to 1% shock. The dotted line denotes a 2 standard-error 95% confidence bands of the estimates. It is computed by Monte Carlo method.

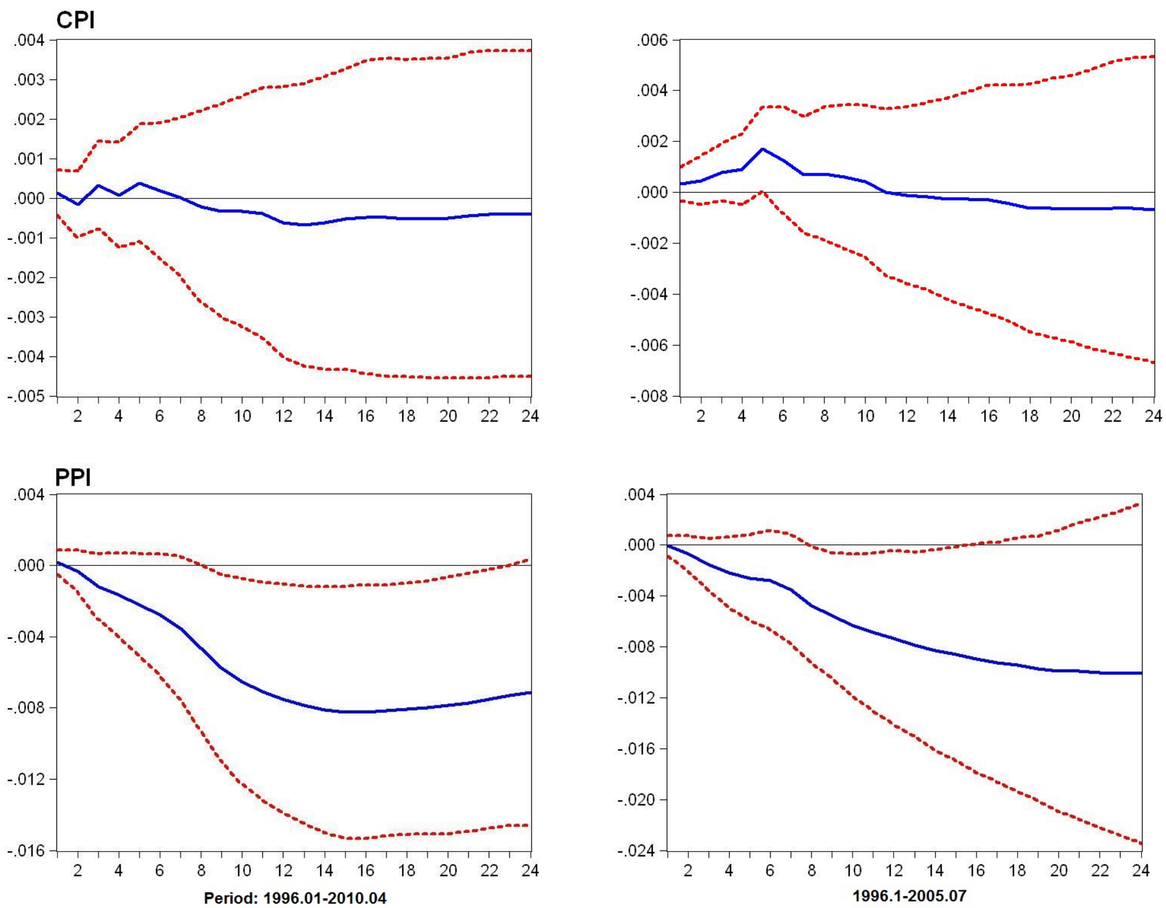


Figure 1.9: Impulse Response to NEER Shocks

The response of CPI to NEER shock is positive initially, posterior to slight fluctuation, it remains around -0.040%. It means that an appreciation of NEER will reduce CPI inflation in the long run. In addition, the response is statistically insignificant, as two standard error confidence bands surround the zero line almost symmetrically.

In contrast, PPI response is much larger and lasts for a longer horizon, approximately 13 months. The initial impact of NEER shock to PPI is identical as to CPI. The distinct appears one year later, the response of PPI reaches -0.786% which is 18 times larger than the one of CPI [5].

[5]The order of variables is crucial for impulse response analysis. For robustness consideration, we apply

1.5 Discussion

1.5.1 Exchange Rate Regime and Pass-Through

In Section 1.2.2, a flood of papers support that the monetary policy regime affects the ERPT. It could be a change of inflation environment, such as an adoption of inflation targeting or a structural break in inflation rate. It also could be a shift in exchange rate regime from the fixed to the floating exchange rate policy. China has a relative stable inflation environment, the concern comes from the influence of exchange rate regime.

To investigate the relation between ERPT and exchange rate regime, we still begin with the single equation framework. The stability Chow test might be the most intuitive method to judge whether the exchange rate policy reform causes a structural change. Regarding 2005 July as the potential breakpoint, the linear regression of CPI (equation 1.5) rejects the null hypothesis of no breaks at the 10% significance level, while PPI's regression (equation 1.6) accepts the above hypothesis.

Table 1.3: Chow Test for Structural Break

CPI			
F Statistic	1.538	Prob.F(21,124)	0.077
Log likelihood ratio	38.426	Prob.Chi2(21)	0.012
Wald Statistic	32.298	Prob.Chi2(21)	0.055
PPI			
F Statistic	0.584	Prob.F(21,124)	0.923
Log likelihood ratio	15.645	Prob.Chi(21)	0.789
Wald Statistic	12.255	Prob.Chi(21)	0.933

Note: The null hypothesis is there is no breaks in 2005M7 in the OLS regression.

With the preliminary evidence from Chow test, a new econometric equation is laid out for examining the association between exchange rate regime and ERPT. The dummy variables EX is added to equation 1.5 and 1.6 for this distinction. EX equals to 1 during nominal appreciating exchange rate period from July 2005 to July 2008, otherwise 0. The ERPT is calculated as in Table 1.4, while the new econometric equation as follows[6].

different orders. The results suggest there is little difference.

[6] Chen and Li (2009) estimated ERPT before and after the exchange rate reform. Even so, they simply separated the sample size, the estimation based on only 44 observation data could not be so accurate. My approach followed Edward (2006) could overcome this drawback, as the discrepancy during different monetary policy regimes can not only be tested by gauging pass-through in different periods, but also by the significance of dummy variables.

$$\begin{aligned}
dlncpi(dlnppi)_t = \sum_{i=1}^{5} \alpha_i dlncpi(lnppi)_{t-i} + \sum_{i=0}^{5} \beta_i dlnneer_{t-i} + \sum_{i=0}^{5} \gamma_i dlnfpi_{t-i} \\
+ \zeta outputgap + \eta dlnbdi + \delta dlnm2 + \sum_{i=1}^{5} \varphi_i dlncpi(dlnppi)_{t-i} EX \\
+ \sum_{i=0}^{5} \chi_i dlnneer_{t-i} EX
\end{aligned} \tag{1.10}$$

Table 1.4: Long Run ERPT and Monetary Policy

	Expression	CPI	PPI
Whole period		-0.132	-0.495
Fixed	$(\sum_{i=0}^{5} \beta_i)/(1-\sum_{i=1}^{5} \alpha_i)$	-0.140	-0.502
Flexible	$(\sum_{i=0}^{5} \beta_i + \sum_{i=0}^{5} \chi_i)/(1-\sum_{i=1}^{5} \alpha_i - \sum_{i=1}^{5} \varphi_i)$	-0.031	-0.524

Note: the long run ERPT for different regime is calculated through the new econometric equation, according to the expression part of table.

The regression results are presented in the Appendix 1.8.2. Table 1.4 lists the ERPT estimation for different periods. Short-run ERPT coefficient is always insignificant, hence more attention is paid to the long-run ERPT. There are two evidence supports that the shift of exchange rate regimes affects the CPI pass-through, but not the PPI pass-through. First, there is an obvious distinction of ERPT between two exchange rate regimes. In the fix exchange rate regime, long-run CPI pass-through is -0.140, while it is -0.031 in the flexible exchange rate regime. Nevertheless, one shortcoming of this study is the limited sample size, especially for the flexible exchange rate regime. Due to the short period of appreciation (only three years), the estimated ERPT in the flexible period can not be completely trusted. Consequently, we compare CPI pass-through for the whole period with the one in the fix exchange rate regime (-0.132 vs -0.140 respectively). It seems CPI pass-through is slightly higher in the period when the flexible exchange rate period is ruled out. Second, in Appendix 1.8.2 we can see that some coefficients of CPI equation for interaction term (χ_i, φ_i) are significant.

In terms of PPI, the ERPT among divergent regime is consistent (-0.502 vs -0.524 respectively), and none of the interaction terms coefficients are significant. This conclusion of higher CPI pass-through in the fix exchange rate regime is in line with Beirne et al.

(2009) and Barhoumi (2005).

Analogously, VAR model is implemented to eliminate the endogeneity doubt from OLS. The new VAR model is estimated within the reduced datasets from Jan 1996 to June 2005. From Figure 1.9, the impulse response of subsample displays slight distinction with full period analysis. Nevertheless, relatively significant response of both price indices within fix exchange rate regime are found in long run[7]. The VAR analysis partially supports the evidence from single equation, but failed to prove the divergent influence of exchange rate regime for two price indices[8].

The revaluation period between 2005 and 2008 is extraordinary for China's economy. A series of policies are implemented to response to the challenge. Figure 1.2-1.6 in Section 1.3 reflect the special volatility of policy and economy status from 2005 to 2008. Owing to the intensive appreciation expectation in 2005, capital inflow inclined rapidly regardless of the strict capital control (Figure 1.5), meanwhile the funds outstanding for foreign exchange rose. This type of funds increased the base money supply, produced bubble in stock and real estate market [9], finally led to a higher inflation rate. In order to control money supply and liquidity, PBC heightened interest rate and bank reverse ratio vigorously (Figure 1.2 and 1.3). Nonetheless, with enlarged interest difference and appreciation expectation, hot money even flew into more fiercely. The inflation rate was fairly high between 2006 and 2008, since it failed to be controlled. The challenge of the efficiency of monetary policy is severe. With the chain reaction following the appreciation, we see that the inflation degree here actually depends on the efficiency of monetary policy against the excess liquidity risk in the beginning of the appreciation period.

1.5.2 Explanation for Pass-Through Difference

From the previous empirical work, the results indicate that the CPI pass-through is much lower than PPI pass-through. China's pass-through is obviously larger than those of industrial countries. This section concentrates on explaining ERPT difference from various aspects.

One explanation of CPI and PPI pass-through might be different definitions and composition of price indices. PPI reflects the price change when industrial product goes into circulation for the first time, while CPI traces the price of consumer goods and service for

[7]By the end of two years, the responses of CPI and PPI before exchange rate reform are -0.068% and -1.00%, compared with -0.040% and -0.713% for whole period.

[8]One explanation could be that impulse response function only consider the response of price indices to NEER, while the computation of ERPT in linear model consider the response of NEER to itself.

[9]See Figure 1.1, the market value of stock market reached the peak in 2007.

the final consumers. About the composition of price indices, different baskets and weights of price indices could be the reason. Albeit Chinese government does not publish the detailed construction of price indices, the general construction for CPI is as follows, food 33.2%, tobacco and alcohol 3.9%, clothes 9.1%, household equipment and maintenance services 6.0%, medical care and personal products 10.0%, transportation and communications 10.4%, entertainment, education and culture 14.2%, residence 13.2%. In the composition of PPI, means of production such as raw materials and machinery, electronics, chemical, textile products are the majority, the weight of means of subsistence for resident final consumption is less than 30%. Hence, the cross term is that CPI contains 50% of the industrial consumer goods, while PPI includes 30% of the means of subsistence.

Furthermore, Bustein et al. (2003), Campa and Ihrig (2006) explained this puzzle by considering imported inputs, the existence of non-tradable goods and distribution cost through information from input-output table. This section follows these work and provide the evidence from China [10].

Imported inputs and non-tradable sector

In Campa and Ihrig (2006), the calibration of ERPT into CPI demonstrated that imported inputs can improve CPI pass-through, as higher imported inputs contribute to the price of non-tradable goods and home produced tradable products. Imported inputs ratio refers to the ratio of total value of imported intermediate inputs to the value of the total intermediate inputs. In Table 1.5, the third column reports the share of imported inputs for several countries. The ratio of imported inputs for China rises gradually from 0.087 in 1995 to 0.109 in 2002, due to the heightening degree of openness. Paralleling with the other countries, the share of China is only a little lower.

There is a large consensus in literature that exchange rate only influences the price of tradable goods. Although Campa and Ihrig (2006) pointed out that the price of non-tradable goods can also be affected by exchange rate through imported inputs, the pass-through is still much lower than for tradable goods. Therefore, the larger the scale of non-tradable sector is, the lower CPI pass-through is, where CPI combines the price of both tradable and non-tradable goods. In contrst, PPI is deemed to a proxy of domestic price index for tradable goods, it can not be affected by the scale of non-tradable sector.

Following the method from Campa and Ihrig (2006), we compute the share of tradable goods in consumption, as the ratio of the value of consumption by households in tradable products relative to the value of total consumption by households. The tradable goods

[10] The resource of input-output table is OECD.

is selected from the OECD input-output table (category 1 to 25) [11], while the rest of categories are consider as non-tradable goods. The fourth column of Table 1.5 indicates that the share of non-tradable consumption in China increased from 0.26 in 1995 to 0.468 in 2002. Particularly in the year 2000, the ratio is nearly half of US, France, Japan's. Meanwhile India and Indonesia have the identical composition of consumption, implies that the degree of development in a country is proportional to the share of non-tradable goods.

Table 1.5: Share of Imported Input, Non-tradables

Country	Year	Imported Inputs	Non-tradables to Consumption
China	1995	0.087	0.260
	2000	0.095	0.304
	2002	0.109	0.468
US	2000		0.797
France	2000	0.144	0.676
Japan	2000	0.070	0.779
UK	2000	0.160	0.682
Germany	2000	0.201	0.689
Indonesia	2000	0.210	0.432
Korea	2000	0.235	0.698
India	1998	0.117	0.396

Note: Some data is identical to Campa and Ihrig(2006), due to the identical calculation method

Distribution cost

The concept of distribution cost comes from Bustein et al. (2003). They considered distribution cost as the cost of distributing tradable goods, such as transportation, wholesaling and retailing. Moreover, they defined distribution margin as the ratio of the difference between retail price and producer price over retail price. It is obvious that high distribution cost leads to low ERPT. In addition, distribution cost could also partially explain that why PPI pass-through is higher than CPI. The definition of CPI involves more intermediate distribution links, which lead to more distribution cost.

Whereas, there is no suitable data to calculate such type of margin[12]. Hence we use

[11] Category detail refers to OECD's input-output table.

[12] In the linear model, we only consider one kind of distribution cost - transportation cost, as it is easier to measure, compared with other distribution costs.

input-output table to calculate the distribution margin for three final demand sections, which is the final consumption expenditure by households, gross fixed capital formation and exports. The distribution cost for service and non-tradable goods is assumed to be zero, the whole cost is allocated to tradable goods. The distribution sectors are Wholesale & Retail trade, Repairs, Land transport, Transport via pipelines, Water transport and Air transport (category 31, 33-35).

The distribution margin for tradable goods is calculated as final demand divided by the total inputs of tradable goods. Table 1.6 reports distribution margin for three sectors. In all the countries, the distribution margin for final consumption expenditure by households is larger than those of the export and gross fixed capital formation sections. For final household consumption section, China's distribution service consists of 10.2%, 10.3% and 16.5% of final consumer price in the year 1995, 2000 and 2002 respectively. In comparison with other countries, the distribution sector is not an important element for tradable consumption, since the ratio of industrial countries is above 60% in 2000. This differences probably comes from distinct structure between developed and developing countries. As we observe from Table 1.6, the distribution margins for Indonesia, Korea and India are also smaller.

Table 1.6: Distribution Margin

Country	Year	Final consumption expenditure by households	Gross fixed capital formation	Exports
China	1995	0.102	0.051	0.032
	2000	0.103	0.046	0.094
	2002	0.165	0.083	0.174
US	2000	0.904	0.189	0.263
France	2000	0.593	0.120	0.111
Japan	2000	1.052	0.332	0.181
UK	2000	0.616	0.221	0.278
Germany	2000	0.600	0.103	0.094
Indonesia	2000	0.253	0.240	0.134
Korea	2000	0.377	0.120	0.144
India	1998	0.279	0.140	0.280

To summarize, low imported inputs, considerable scale of non-tradable and distribution sectors for China could partially explain the low CPI pass-through in China. However, the impact from distribution and non-tradable sectors is much smaller than that in industrial countries. Hence, there must be some other reasons, such as price regulation. Chinese government controls price control mainly in the field of energy industry, agricultural commodities, land and resource. Particularly, the energy industry, retail price of petrol, diesel oil and electricity are seriously administrated. Taking product oil as an example, the corresponding policy is that the price of domestic produced oil is made adjustment, while the average price of crude petroleum in international market changes more than 4% in consecutive 22 working days. This policy implies that the product oil price in China lacks of elasticity and is hysteretic to international oil price fluctuation [13]. Concerning electricity industry, coal price is determined by market demand and supply, but not the electricity price. Agricultural commodities price is also regulated, especially grain price. Paralleling with other agricultural commodities, such as pork, vegetables, the variation of grain price is always the lowest.

1.6 Conclusion

In this chapter, ERPT for China's CPI and PPI is estimated by single equation regression. The OLS model finds that 1% devaluation of exchange rate leads to 0.016% CPI deflation in the short run and 0.132% CPI inflation over the long run. The ERPT for PPI amounts to 0.495 in the long run, which is much higher. It is in conformity with prediction. The VAR model supports the results of linear model, about a fairly low CPI pass-through and relatively higher PPI pass-through. The negative correlation between exchange rate variation and long-run inflation implies that appreciation could reduce inflation rate.

In light of the relation between exchange rate regime and pass-through, the exchange rate regime impacts on pass-through to CPI, but not to PPI. In the long run, the periods within fix exchange rate regime have higher CPI pass-through.

Different definitions, baskets and weights of price indices could be a reason for the difference of CPI and PPI pass-throughs. Furthermore, the low ERPT could be partially explained by low imported inputs, considerable magnitude of non-tradable and distribution sector. Other potential reasons are left for the future research.

The estimation results suggest that government could pursue more flexible exchange rate policy and maintains the independence of monetary policy. The unfavorable impacts of exchange rate fluctuation on price stability is not a problem. Whereas, this conclusion is not absolute. The variation of pass-through should be considered, as the ERPT may vary in accordance with different macro factor, such as inflation performance (Taylor 2000), and monetary policy regime (Devereux and Yetman 2010). More extremely, ERPT might be

[13] Appendix 1.8.3 gives details for product oil price adjustment.

endogenous in the open economy macroeconomic framework (Devereux et al. 2004).

1.7 References

Bailliu, J. and Fujii, E., 2004. Exchange rate Pass-Through and the Inflation Environment in Industrialized Countries:An Empirical Investigation. *Bank of Canada Review*, Spring.

Barhoumi,K., 2006. Differences in Long Run Exchange Rate Pass-Through into Import Prices in Developing Countries: an Empirical Investigation. *Economic Modelling* 23, 926-951.

Beirne, J. and Bijsterbosch, M., 2009. Exchange Rate Pass-Through in Central and Eastern European Member States. *European Central Bank Working Paper* No.1120.

Burdekin,R., Siklos, P.L., 2008. What has driven Chinese monetary policy since 1990? Investigating the People's bank's policy rule. *Journal of International Money and Finance* 27, 847-859.

Burstein, A., Neves, J., Rebelo, S., 2003.Distribution costs and real exchange rate dynamics. *Journal of Monetary Economics* 50(6), 1189-1214.

Campa, J. M., Goldberg, L. S., 2005. Exchange rate pass-through into import prices. *Review of Economics and Statistics* 87, 679-690.

Campa, J. M., Goldberg, L. S., 2006. Distribution Margins, Imported Inputs, and the Sensitivity of the CPI to Exchange Rates. *NBER Working Paper Series* 12121.

Ca' Zorzi, M., Hahn, E., Sanchez, M., 2007. Exchange Rate Pass-Through in Emerging Markets. *ECB Working Paper* No. 739.

Chen, Liubo., Liu, Houjun., 2007. The Effect of RMB Exchange Rate Pass-through: An Empirical Analysis Based on VAR. *Journal of Fincial Research* 4, 1-13.

Choudhri, E., Hakura, D., 2006. Exchange rate pass-through to domestic prices: does the inflationary environment matter? *Journal of International Money and Finance* 25, 614-639.

Coricelli, F., Jazbec, B., Masten, I., 2006b. Exchange Rate Pass-Through in Acceding Countries: Empirical Analysis and Policy Implications. *Journal of Banking and Finance* 30(5), 1375-1391.

Devereux, M., Engle, C., 2003. Monetary policy in the open economy revisited: price setting and exchange rate flexibility. *Review of Economic Studies* 70, 765-783.

Devereux, M., Yetman, J., 2010. Price adjustment and exchange rate pass through. *Journal of International Money and Finance* 29(1), 191-200.

Devereux, M., Engle, C., Storgaard P., 2004. Endogenous exchange rate pass-through when nominal prices are set in advanced. *Journal of International Economics* 63, 263-291.

Edwards, S., 2006. The Relationship between Exchange Rates and Inflation Targeting Revisited. *NBER Working Paper Series* 12163.

Faruqee, H., 2006. Exchange Rate Pass-Through in the Euro Area. *IMF Staff Papers* 53, 63-88.

Frankel, J.A., Parsley, D.C., 2005. Slow Pass-Through around the World: a new Import for Developing Countries? *NBER Working Paper Series* No.11199.

Forssbeck, J., Oxelheim,L., 2007. The Transition to Market-based Monetary Policy:What can China Learn from the European Experience. *Journal of Asian Economics* 18, 257-283.

Gagnon, J.E., Ihrig, J., 2004. Monetary policy and exchange rate pass-through. *International Journal of Finance and Economics* 9, 315-338.

Glick, R., Hutchison, M., 2009. Navigating the Trilemma: Capital Flows and Monetary Policy in China. *Journal of Asian Economics* 20, 205-224.

Goldberg, P., Knetter, M., 1997.Goods prices and exchange rates: What have we learned? *Journal of Economic Literature* 35, 1243-1292.

Gu, Jinlin., 2008. On Price Control and Inflation in China Material Market. *Reformation and Strategy* 11, 23-28.

Hahn, E., 2003. Pass-Through of External Shocks to Euro Area Inflation. *European Central Bank, Working Paper* No.243.

Ihrig, J., Marazzi, M., Rothenberg, A., 2006. Exchange-Rate Pass-Through in the G-7 countries. *International Finance Discussion Papers* No.851.

Ito, T., Sato, K., 2008. Exchange rate changes and inflation in post-crisis Asian economies: VAR analysis of the exchange rate pass-through. *Journal of Money, Credit and Banking* 40, 1407-1438.

Kim, K.H., 1998. US Inflation and the Dollar Exchange Rate: A Vector Error Correction Model. *Applied Economics* 30(5), 613-619.

Krugman, P., 1986. Pricing to Market when the Exchange Rate changes. *NBER Working Papers* No.1926.

McCarthy, J., 2000. Pass-Through of Exchange Rates and Import Prices to Domestic Inflation in Some Industrialized Economics. *Eastern Economic Journal* 33(4), 511-537.

Shu, C., Su, X., 2009. Exchange Rate Pass-through in China. *China & World Economy* 17(1), 33-46.

Taylor, J.B., 2000. Low inflation, pass-through, and the pricing power of firms. *European Economic Review* 44, 1389-1408.

Wang, J., Li, N., 2009. Exchange Rate Pass-through: The Case of China. *Journal of*

Economic Research 4, 17-27.

1.8 Appendix

1.8.1 Date Resource and Special Treatment

CPI/PPI: China's CPI and PPI data published in international database are mainly based on CPPY=100 (current period previous year=100), e.g. International Financial Statistics (IFS) from IMF. Ideal data for our estimation is based on one specific year, afterwards the monthly growth could be obtained easily. In order to adjust these indices to base on 2005=100, which is consistent with other index data, We use CPI's chain index data and CPI data based on CPPY=100 for CPI adjustment, and use PPI data based on CPPY=100 and producer goods' chain price index data for PPI adjustment.

Foreign Price Index: Undoubtly, the most ideal index presents the importing cost would be the weighted average of export price index. Whereas, there are not enough data to support the calculation of this index. Some literatures direct replaced export price index by CPI. Nonetheless, in order to keep the consistency of weights within NEER, we adopt Campa and Goldberg (2005)'s treatment. In line with IMF's methodology, $REER_i = \prod_{j \neq i}[\frac{CPI_i R_i}{CPI_j R_j}]^{W_{ij}}$, where j is an index that runs over country i's trade partners, W_{ij} is the competitiveness weight put by country i on country j, R_i, R_j is the nominal exchange rate of country i and j's currencies in US dollars. $NEER_i = \prod_{j \neq i}[\frac{R_i}{R_j}]^{W_{ij}}$. Therefore, $FPI = NEER_i * CPI_i / REER_i = \prod_{j \neq i}[CPI_j]^{W_{ij}}$, FPI is regards as an appropriate weighted average of foreign countries' CPI.

Per unit distribution cost: Per unit distribution cost is difficult to measure. BDI is a daily average of prices to ship raw materials. It represents the cost paid by an end customer to have a shipping company transport raw materials across seas on the Baltic Exchange, which is the global marketplace for brokering shipping contracts. It actually measures the transportation cost.

Data resource: CPI, PPI (CPPY=100), NEER, REER, Industrial Production,and M2 collect from IMF IFS, downloaded through Data Stream. While Baltic Exchange Dry Index make monthly sum of the daily data from Baltic Exchange, downloaded also from Data Stream. CPI and producer goods chain price index from WIND info (authoritative database in China).

1.8.2 ERPT by OLS Estimation with Dummy Variable EX

CPI		PPI	
Variable	Coefficient	Variable	Coefficient
dlncpi(-1)	0.045 (0.097)	dlnppi(-1)	0.296*** (0.086)
dlncpi(-2)	0.020 (0.095)	dlnppi(-2)	0.151* (0.091)
dlncpi(-3)	0.004 (0.092)	dlnppi(-3)	0.003 (0.095)
dlncpi(-4)	0.108 (0.091)	dlnppi(-4)	0.084 (0.092)
dlncpi(-5)	0.170* (0.093)	dlnppi(-5)	0.094 (0.082)
dlnfpi	0.614*** (0.206)	dlnfpi	0.976*** (0.220)
dlnfpi(-1)	0.013 (0.214)	dlnfpi(-1)	0.486* (0.248)
dlnfpi(-2)	-0.247 (0.206)	dlnfpi(-2)	-0.113 (0.235)
dlnfpi(-3)	0.188 (0.205)	dlnfpi(-3)	-0.106 (0.232)
dlnfpi(-4)	-0.058 (0.218)	dlnfpi(-4)	-0.241 (0.243)
dlnfpi(-5)	-0.523*** (0.194)	dlnfpi(-5)	-0.673** (0.230)
dlnneer	0.009 (0.028)	dlnneer	0.014 (0.032)
dlnneer(-1)	-0.055* (0.031)	dlnneer(-1)	-0.054 (0.034)
dlnneer(-2)	0.013 (0.030)	dlnneer(-2)	-0.068** (0.033)
dlnneer(-3)	-0.036 (0.031)	dlnneer(-3)	-0.034 (0.034)
dlnneer(-4)	0.039 (0.030)	dlnneer(-4)	-0.028 (0.034)
dlnneer(-5)	-0.062** (0.028)	dlnneer(-5)	-0.015 (0.033)
dlncpi(-1)* EX	0.146 (0.194)	dlnppi(-1)* EX	-0.005 (0.236)
dlncpi(-2)* EX	0.477** (0.188)	dlnppi(-2)* EX	0.057 (0.266)
dlncpi(-3)* EX	0.222 (0.183)	dlnppi(-3)* EX	0.119 (0.257)
dlncpi(-4)* EX	-0.241 (0.186)	dlnppi(-4)* EX	0.282 (0.268)
dlncpi(-5)* EX	-0.306* (0.183)	dlnppi(-5)* EX	-0.290 (0.237)
dlnneer* EX	0.125 (0.083)	dlnneer* EX	-0.050 (0.093)
dlnneer(-1)* EX	-0.061 (0.082)	dlnneer(-1)* EX	-0.037 (0.093)
dlnneer(-2)* EX	0.152** (0.076)	dlnneer(-2)* EX	0.017 (0.084)
dlnneer(-3)* EX	-0.013 (0.083)	dlnneer(-3)* EX	0.120 (0.084)
dlnneer(-4)* EX	-0.199** (0.089)	dlnneer(-4)* EX	0.038 (0.096)
dlnneer(-5)* EX	0.077 (0.090)	dlnneer(-5)* EX	-0.011 (0.099)
dlnm2	0.008 (0.034)	dlnm2	0.021 (0.035)
outputgap	0.000 (0.000)	outputgap	0.000 (0.000)
dlnbdi	-0.004* (0.002)	dlnbdi	-0.002 (0.002)
R^2	0.367	R^2	0.666

Note: The figures in parentheses are standard errors, the figures in square parenthesis are p value. * significance at the 10% levels, **-significance at the 5% levels, ***- significance at the 1% levels. dlncpi(-i) denotes $dlncpi_{t-i}$.

1.8.3 Adjustment of Product Oil Price In China

Time	Adjustment
Apr,13, 2010	Domestic petrol and diesel oil price increase 320 Yuan per ton
Nov, 11, 2009	Domestic petrol and diesel oil price increase 480 Yuan per ton
Sep, 30, 2009	Domestic petrol and diesel oil price decrease 190 Yuan per ton
Sep, 2, 2009	Domestic petrol and diesel oil price increase 300 Yuan per ton
July, 29, 2009	Domestic petrol and diesel oil price decrease 220 Yuan per ton
June, 29, 2009	Domestic petrol and diesel oil price increase 600 Yuan per ton
May, 31, 2009	Domestic petrol and diesel oil price increase 400 Yuan per ton
Mar, 25, 2009	Domestic petrol and diesel oil price increase 290 Yuan
	Diesel oil price increase 180 Yuan per ton
Jan, 15, 2009	Domestic petrol and diesel oil price decrease 140 Yuan
	Diesel oil price decrease 160 Yuan per ton
Dec, 12, 2008	Ex-factory price of petrol declines 900 Yuan per ton
	Ex-factory price of diesel oil declines 1100 per ton
June, 20, 2008	Domestic petrol and diesel oil price raises 1000 Yuan per ton
	Aviation kerosene price raises 1500 Yuan per ton
Nov, 1, 2007	Price of petrol, diesel oil increase 500 Yuan perton
Jan, 14, 2007	Ex-factory price of petrol decreases 220 Yuan per ton
	Ex-factory price of aviation kerosene decreases 90 Yuan per ton
Mar, 26, 2006	Ex-factory price of petrol increases 300 Yuan per ton
	Diesel oil by 200 Yuan per ton
	Aviation kerosene by 300 Yuan per ton
July , 23, 2005	Increase ex-factory price of petrol by 300 Yuan per ton
	Diesel oil by 250 Yuan per ton
	Aviation kerosene by 300 Yuan per ton
June ,25, 2005	Increase ex-factory price of petrol by 200 Yuan per ton
	Diesel oil by 150 Yuan per ton
	Aviation kerosene by 300 Yuan per ton
May, 23, 2005	Decrease ex-factory price of petrol by 150 Yuan per ton
May, 10, 2005	Increase ex-factory price of diesel oil by 150 Yuan per ton
Mar, 23, 2005	Ex-factory price of petrol raises 300 Yuan per ton

Chapter 2

Number crunching or the wisdom of the crowds? Forecasting GDP Growth in the Eurozone and its five largest Countries

2.1 Introduction

Information on economic activity and its short-term prospects is of great importance to decision makers in governments, central banks, financial markets and non-financial companies. Monetary and economic policy makers and economic agents have to make decisions in real time with incomplete and inaccurate information on current economic conditions. A key indicator of the state of the economy is the growth rate of real GDP, which is available on a quarterly basis only and also subject to substantial publication lags. In many countries an initial estimate of quarterly real GDP is published around six weeks after the end of the quarter. Moreover, real GDP data are subject to sometimes substantial revisions, as more source data becomes available to statistical offices over time.

Fortunately, there is a lot of statistical information related to economic activity that is published on a more frequent and timely basis. This information includes data on industrial production, prices of goods and services, expenditures, unemployment, financial market prices, loans and consumer and business confidence. The forecasting literature has recently developed several statistical approaches to exploit this potentially very large information set in order to improve the assessment of real GDP growth in the current quarter (nowcast) and its development in the near future. Examples are bridge models, factor models, mixed-data sampling models (MIDAS) and mix frequency vector autoregressive (MFVAR) models[1]. These models differ in their solutions to the practical problems of how to handle a large-scale information set and the fact that the auxiliary variables are observed at different frequencies and with different publication lags.

For practitioners there is now a wealth of statistical models to choose from. So which one should they use? As each model has strengths and weaknesses it is difficult to make a choice on purely theoretical grounds. The ranking of the models in terms of forecasting ability and how it may vary with the prediction horizon or the economic circumstances has to be determined by empirical analysis. On these issues the jury is still out, however, as large-scale comparative studies are scarce. The empirical work in many papers refers to a single country and usually includes only a limited number of models. Furthermore, papers differ in the size of the information set and the sample period[2].

This chapter is motivated by this gap in the empirical literature. We undertake a systematic comparison of a broad range of statistical linear models - eleven models in all - that have been applied in the recent literature. To improve comparability and robustness, we include six countries in our analysis (Germany, France, Italy, Spain, the Netherlands and the euro area), utilizing the same information set across countries. Moreover, our sample includes the volatile episode of the financial crisis of 2008 and its aftermath, which may make it easier to discriminate between the various models. We contrast the models' forecasting performance before 2008 and that during the crisis period. This may be of great interest to policy makers, financial analysts and economic agents alike, as information on where the economy stands and where it is heading in the immediate short run is particularly valuable in times of great uncertainty.

Providing cross-country evidence on the relative performance of eleven different statistical forecasting models is our first contribution to the literature. Model forecasts are the result of purely mechanical recipes and do not incorporate subjective elements. Our second contribution concerns the potential usefulness of forecasts made by professional analysts (published by Consensus Forecasts on a quarterly basis). From a practical point of view, such forecasts are very cheap and easy to use. Moreover, they may, as expression of the "wisdom of the crowds", reflect much more information than the statistical information

[1]See Baffigi et al. (2004) Rünstler et al. (2009).

[2]Rünstler et al.(2009) is an important exception, comparing three factor models, a bridge model and a quarterly VAR model for ten European countries, but this study does not include the recent crisis episode. Kuzin et al. (2012) analyze the relative forecasting performance of MIDAS models versus dynamic factor models, including part of the crisis years (2008-2009). Liebermann (2012) analyzed the relative forecasting performance during the period December 2001 - December 2011, of a range of models, but only for the United States

set, which is inevitably limited. A questionnaire by the European Central Bank (ECB) among the participants of the ECB Survey of Professional Forecasters found that the panelists regard forty percent of their short-term GDP forecasts to be judgment-based (ECB, 2009). We investigate for our six countries to what extent subjective forecasts by analysts contain information beyond that generated by the best mechanical statistical model.

The remainder of the chapter is structured as follows. Second 2.2 discusses the challenges that large-scale and irregularly shaped datasets pose to the forecasting process and describes how the eleven statistical models attempt to overcome these problems. Section 2.3 describes the data, models and our pseudo real-time forecast design. Section 2.4 and Section 2.5 present the results for the mechanical models and the professional forecasts, respectively. Section 2.6 summarizes our findings and concludes.

2.2 Linear statistical models for short-term GDP growth forecasting

2.2.1 Overview

Utilizing auxiliary information for forecasting of real GDP in the immediate short run in practice poses several challenges. The first challenge is posed by the large size of the information set. There are countless potentially useful variables for forecasting GDP. The size of the datasets in the empirical literature varies from 70 to more than 300 (e.g. Rünstler et al.,2009 and Giannone et al., 2008). The second problem relates to the fact that indicator variables are more frequently (monthly, weekly, daily) observed than GDP. Moreover, the dating of the most recent observation may vary across indicators because of differences in publication lags. This is known as the "ragged edge" problem, see Wallis (1986). The various statistical approaches in the literature deal with these challenges in different ways. To facilitate the discussion, Figure 2.1 depicts a schematic representation of the process of translating a large dataset into a single final forecast along with several crucial modeling choices. Figure 2.1 shows that a forecasting procedure involves two transformations of the dataset of indicators to produce a final forecast: an aggregation and the application of a forecasting tool, which links auxiliary variables to real GDP growth. The two transformations can be executed in a different order, representing two fundamentally different strategies. The "factor strategy" takes the aggregation step first by summarizing the large dataset by a small number of series. This strategy exploits the fact that the auxiliary variables are correlated. Factor analysis is used to replace a

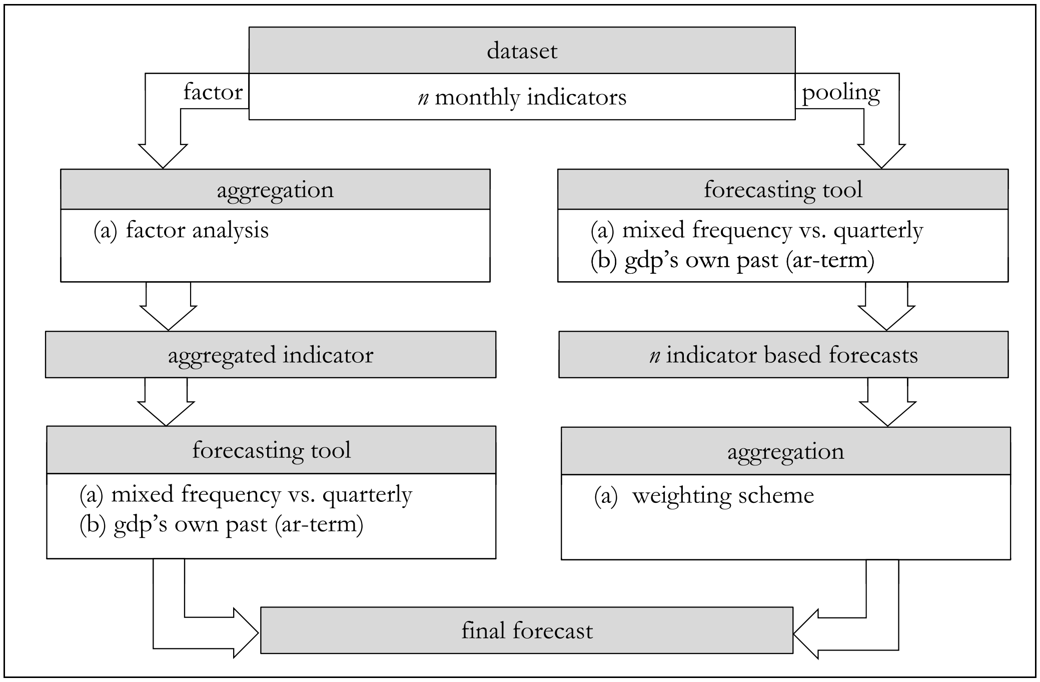


Figure 2.1: Schematic representation linear models for short-term GDP forecasting

large number of correlated time series by a limited number of uncorrelated (unobserved) factors or principal components representing the common information component of the original data series. The implicit weights (factor loadings) are determined by the correlation patterns in the original dataset. The factors serve as input for the forecasting procedure in the next step. Examples of this modeling strategy are static and dynamic factor models. By contrast, the second strategy first computes for each variable an indicator-specific forecast, which are then aggregated into a single final forecast in the second step. We call this strategy the "pooling strategy" as it pools a large number of individual indicator-based forecasts. In this approach it is necessary to specify the weighting scheme of the individual forecasts. A simple scheme is the simple average, which gives each forecast an equal weight, but weights may also be recursively computed depending on the indicators' (recent) forecasting performance. Examples of the pooling strategy are bridge equations and VAR models.

The specification of the forecasting tool is the second distinguishing feature of the approaches. The traditional approaches, such as the bridge models and VAR models, rely on forecasting equations that are solely cast in quarterly terms. That means that monthly indicator variables first have to be aggregated to quarterly averages, before they can be used for forecasting. As this may not be an efficient use of the available information, recently developed approaches accommodate both quarterly and monthly data within the same equation or system of equations. These approaches take publication lags into account. The mix frequency VAR (MFVAR) model treats GDP as an unobserved monthly variable in a state-space framework. Monthly GDP is transformed to the growth rate of quarterly GDP via an identity. The quarterly GDP growth rate is only observed in the third month of every quarter. The mixed-data sampling (MIDAS) design relates quarterly GDP directly to a large number of lags of monthly data series using a parsimonious specification of the lag structure.

A third, more practical, specification issue is whether or not to include GDP's own past in the forecasting tool. In general, forecasting equations can easily be augmented by autoregressive (AR) terms. Several authors have found that the AR-versions of models tend to result in modest improvements of forecasting performance (e.g. Foroni and Marcellino, 2012).

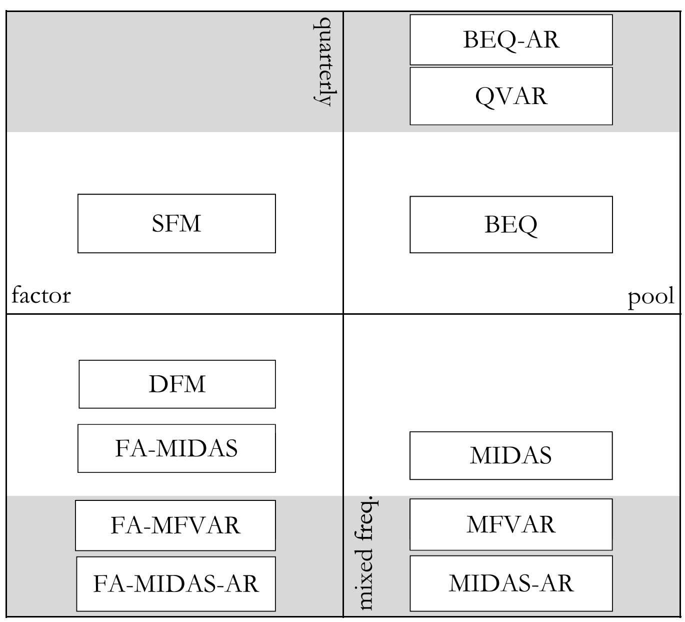


Figure 2.2: Classification of linear models for short-term GDP forecasting

In this chapter we analyze eleven statistical models. They are denoted as follows: (1) bridge model (BEQ), (2) BEQ with AR terms (BEQ-AR), (3) quarterly VAR model (QVAR), (4) diffusion index or static factor model (SFM), (5) dynamic factor model (DFM), (6) mix frequency VAR model (MFVAR), (7) factor-augmented MFVAR (FA-MFVAR), (8) mixed-data-sampling model (MIDAS), (9) MIDAS with AR terms (MIDAS-AR), (10) factor-augmented MIDAS (FA-MIDAS) and (11) FA-MIDAS with AR terms

(FA-MIDAS-AR). Figure 2.2 classifies the eleven models according to the choices made on the three issues discussed above. The horizontal axis puts factor strategies versus pooling strategies. The vertical axis puts purely quarterly forecasting equations versus forecasting models that combine monthly and quarterly data (labeled mixed frequency). Finally, shaded areas signify models augmented with auto-regressive terms. The remainder of this paragraph discusses the models depicted in Figure 2.2.

2.2.2 Quarterly models for GDP growth

This section describes the quarterly models we estimated.

Pooled bridge equations (BEQ)

The quarterly bridge equation is perhaps the most widely used method for forecasting GDP using monthly indicators. For applications see Kitchen and Monaco (2003) and Baffigi et al. (2004). Bridge equations are linear regressions that "bridge" the monthly variables, such as industrial confidence and retail sales to quarterly real GDP growth. Usually the monthly indicators are not known over the entire projection horizon. Therefore we proceed in two steps. Firstly, consider a vector of monthly series $x_t = (x_{1,t}, \ldots, x_{n,t})$, t=1,.... The monthly indicators are predicted over the forecasting horizon h on the basis of univariate autoregressive models. Secondly, the monthly forecasts of x are aggregated to the quarterly frequency (x_t^Q) and used to predict quarterly GDP (y_t^Q). The bridge model to be estimated is:

$$y_t^Q = \alpha_i + \sum_{s=0}^{p_i} \beta_{i,s} x_{i,t-s}^Q + \varepsilon_{i,t}^Q, \qquad \varepsilon_{i,t}^Q \sim N(0, \sigma^2_{\varepsilon^Q}) \tag{2.1}$$

where α_i is a constant term, p denotes the number of lags in the bridge equation and ε_i^Q is a normally distributed error-term.

Pooled vector autoregressive models (QVAR)

The pooled VAR approach has much similarity to the bridge equation approach. In contrast to bridge equations, the VAR models use the information content of GDP itself to produce forecasts of GDP (e.g. Camba-Mendez et al., 2001). More precisely, we consider the set of n monthly indicators x_t aggregated to the quarterly frequency x_t^Q and run a quarterly bivariate VAR including the indicator and GDP:

$$z_{i,t}^{Q} = \alpha + \sum_{s=1}^{p_i} A_s z_{i,t-s}^{Q} + \varepsilon_{i,t}^{Q}, \qquad \varepsilon_{i,t}^{Q} \sim N(0, \Sigma_{\varepsilon^Q}) \tag{2.2}$$

where $i = 1, \ldots, n$ with $z_{i,t}^{Q} = (y_t^Q, x_{i,t}^Q)'$, where α_i is a constant term and $\varepsilon_{i,t}^{Q}$ denotes the residual which is bivariate normally distributed. From each VAR, we produce forecasts $y_{t+h|t}^{Q}$ for GDP.

Static Factor Model (SFM)

A different approach to extracting information from the dataset by running regression for each individual indicator and pooling the forecast afterwards is to summarize the information in the vector if $x_t's$. The basic idea of this approach is to exploit the collinearity of the n series in our panel by summarizing all the relevant information in a few common factors r (r<< n).

The static factor model (or diffusion indices) due to Stock and Watson (2002) is one of the simplest versions of factor models, as the dynamics of the factors are not explicitly modeled. The forecasts of GDP growth are obtained in two steps. Firstly, we extract the static principal components are from the covariance matrix of the complete list of monthly indicators x_t to obtain estimates $\hat{f}_t$ of the r common static factors $f_t = (f_{1,t}, f_{2,t}, \ldots, f_{r,t})$. Formally:

$$x_t = \Lambda f_t + \xi_t, \qquad \xi_t \sim N(0, \Sigma_\xi) \tag{2.3}$$

which relates the n×1 vector of monthly observations x_t to the monthly factors f_t via a matrix of factor loadings Λ and an idiosyncratic component $\xi_t = \xi_{1,t}, \ldots, \xi_{n,t}$. In the second step the monthly factors f_t are aggregated to quarterly frequency f_t^Q and the GDP forecast can be derived from the "bridge" equation:

$$y_{t+h|t}^{Q} = \mu + \ \beta' f_t^Q + \varepsilon_t \tag{2.4}$$

Note that in this equation GDP appears with a lead of h quarters. Hence, the h-step ahead forecast $y_{t+h|t}^{Q}= \beta' f_t^Q$ is found directly and there is no need to forecast the monthly variables.

2.2.3 Mixed frequency models

This section describes the mixed frequency models we estimated. Recently the interest in these models amongst academics and policy makers has increased because of the general failure of simple quarterly models to forecast the sharp downturn at the onset of the financial crisis.

Dynamic factor model (DFM)

Dynamic factor models are all siblings of the static factor model. The dynamic factor model approach has been shown to provide relatively accurate forecast in the United States (Giannone et al., 2008), the euro area (see Bańbura et al., 2011; Rünstler et al., 2009), Spain (Camacho and Perez-Quiros, 2010) and the Netherlands (den Reijer, 2005). The main difference with the static factor model is that the dynamics in the factors are explicitly modeled as a VAR(p) process. In this chapter we use the dynamic factor by Bańbura and Rünstler (2011), that is used within several central banks within the euro area.

The first model equation of the DFM equals equation 2.3 of the SFM model and extracts the static factors from the covariance matrix of x's. Different from the SFM model the factors are assumed to follow a vector autoregressive process of order p:

$$f_t = \sum_{s=1}^{p} A_s f_{t-s} + \zeta_t \tag{2.5}$$

Moreover, the dynamic factor model setup assumes the covariance matrix of the VAR (σ_ζ) is driven by a q dimensional standardized white noise η_t, where B is a r× q matrix and A is a square r× r matrix:

$$\zeta_t = B\eta_t, \qquad \eta_t \sim N(0, I_q) \tag{2.6}$$

Combining the monthly factor model with a forecast equation for mean-adjusted quarterly GDP growth completes the model. For this purpose the unobserved monthly GDP growth rate y_t is introduced. The monthly GDP growth rate y_t is related to the monthly factors f_t via the equation:

$$y_t = \beta' f_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{2.7}$$

The model is estimated in four steps. Firstly, the factors loadings Λ and estimated static factors $\hat{f}_t$ are obtained as described in section 2.2.2. Secondly the coefficient matrix

A_i in equation 2.5 and β in equation 2.7 are estimated by Ordinary Least Squares using $\hat{f}_t$. Thirdly, the covariance matrix is decomposed into one or more principal components. Matrix B in equation 2.6 contains these components. By construction η_t is a white noise term. In the final step, the model is cast in state space and the Kalman filter and smoother are used to re-estimate the estimated factors ($\hat{f}_t$) and GDP growth.

Besides the dynamics in the factors the key feature of this approach is the use of the Kalman filter. The filter allows for an efficient handling of the unbalancedness of the data and mixing of the monthly frequency of the predictors x_i and quarterly GDP y_i^Q. The state-space setup of the dynamic factor model we used is presented in section 2.8.2. For a more elaborate description of the dynamic factor model see Bańbura and Rünstler (2011)[3].

Mix frequency vector autoregressive models (MFVAR)

The basic idea of the Mix Frequency VAR Model (MFVAR) is to establish a VAR model that efficiently mixes the monthly frequency of the indicator variables with the quarterly frequency of GDP. In contrast to the quarterly VAR model we use all the information from the monthly variables, and do not only use information when the monthly information over the full quarter is known.

The econometric technique used is comparable to the dynamic factor model. That is, we use the state space framework and the associated Kalman filter and smoother to estimate all coefficients. In contrast to the dynamic factor model we do not extract factors to summarize the signal from all indicators, but rather estimate bivariate MFVAR models and afterwards combine this information by pooling. Our estimation is based on Kuzin et al. (2011), but we modified the temporal aggregation scheme in order to be consistent with the scheme employed in the dynamic factor model.

Let $z_{i,t} = (y_t, x_{i,t})'$ be a vector of the latent monthly GDP variable and the variable $x_{i,t}$. The vector follows a VAR model:

$$z_{i,t} - \mu_{z_i} = \sum_{s=1}^{p} A_s(z_{i,t-s} - \mu_{z_i}) + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim N(0, \Sigma_\varepsilon) \tag{2.8}$$

where μ means the expectation of corresponding variable. As documented by Kuzin et al. (2011) the mean parameters μ_{z_i} are often quite difficult to estimate. Therefore, we work with demeaned GDP and monthly indicator series in estimation. After estimation we add the mean back to get the final estimation.

[3]See Durbin and Koopman (2001) for a treatment of state space models and the use of the Kalman filter and smoother.

The model is cast in state space (detailed in section 2.8.2) for estimation. We estimate the model by expectation-maximization algorithm as detailed in Mariano and Murasawa (2010). Another possibility would be to estimate the model by maximum likelihood, but we found this method had problems finding an optimal solution, especially for higher lag orders and when aiming for relatively strict convergence criteria.

Mixed data sampling regression models (MIDAS)

The Mixed-Data Sampling Model (MIDAS) is a single equation regression connecting monthly data (high frequency) to quarterly GDP (low frequency) (Ghysels et al., 2007), and could forecast the low frequency data directly. It can be seen as the mixed frequency version of the BEQ approach, as it is a single indicator model, mixing quarterly and monthly frequencies in an elegant way. MIDAS is a single-equation approach, whereas the MFVAR is a system approach that explains both GDP and the indicator. As such, a misspecification in one equation can affect the estimation and forecast accuracy of the other model equations. MIDAS has a sparse parameterization and does not have this problem. On the other hand, the MIDAS restrictions on the lag polynomial (see below) that relates the monthly indicator to quarterly GDP could be invalid, whereas the coefficients of the MFVAR polynomials are estimated unrestrictedly. The MIDAS model is defined by the following equation:

$$y^Q_{t+h} = \beta_0 + \beta_1 B(L^{(1/3)};\theta)x^{(3)}_{i,t+w} + \varepsilon_{i,t+h} \tag{2.9}$$

where w is the gap of final data available between x_t and y^Q_t[4] and $B(L^{(1/3)};\theta)$ is some lag polynomial. $x^{(3)}_t$ is skipped sampled from the monthly observations x_t. Every third observation, starting from the tth, one is included in the regressor $x^{(3)}_t$; thus, $x^{(3)}_{i,t} = x_{i,t}$ $\forall$ t=…,T-6,T-3,T. $B(L^{(1/3)};\theta) = \sum_{k=0}^{K} c(k,\theta)L^{k/3}$ denotes a weighting function, and $L^{k/3}x^{(3)}_{i,t-1} = x^{(3)}_{i,t-1-k/3}$ represents a fractional lag operator. $c(k,\theta)$ is a specific lag polynomial, that could be chosen from different lag polynomials. We follow Kuzin et al. (2011) and choose the exponential Almon lag polynomial[5], defined as:

[4]The last period of y^Q_t is transformed to the corresponding monthly period.

[5]We compared this polynomial to the recently proposed unrestricted lag polynomial (Marcellino and Schumacher, 2010) but this polynomial turns out to have a higher average RMSFE for most countries in our sample. Details available upon request with the authors.

$$c(k,\theta) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=0}^{K} \exp(\theta_1 k + \theta_2 k^2)} \tag{2.10}$$

k in equation 2.10 is fixed at 12. The parameters include $\theta_1, \theta_2,\ \beta_0,\ \beta_1$ are estimated by nonlinear least square, subject to $\theta_1 < 5, \theta_2 < 0$.

2.2.4 Factor and AR augmented models

Factor augmented models

Finally, we consider mixtures of the MIDAS and MFVAR models and the factor model described above, by considering factor-augmented version of the MIDAS (FA-MIDAS) as well as the factor augmented MFVAR (FA-MFVAR). While the basic MIDAS framework consists of a regression of quarterly GDP on the monthly indicators, the factor-augmented MIDAS and MFVAR approach exploits estimated factors rather than single indicators as regressors. The Factor MIDAS model for forecast horizon h is:

$$y^Q_{t+h} = \ \beta_0 + \ \beta_1 B(L^{(1/3)};\theta)\hat{f}^{(3)}_{t+w} + \varepsilon_{t+h} \tag{2.11}$$

The estimation procedure equals the procedure outlined in section 2.2.3. Similarly, the specification of the FA-MFVAR only needs a minor adjustment of equation 2.8 by replacing x_t by $\hat{f}_t$, then $z^Q_{i,t} = (y^Q_t, \hat{f}_t)'$ and $z_{i,t} = (y_t, \hat{f}_t)'$.

We follow the factor extraction algorithm proposed by Marcellino and Schumacher (2010). We directly adopt the same framework of equation 2.3, 2.5 and 2.6. The extraction of the monthly factors equals the factor extraction method used in the DFM, as described in Marcellino and Schumacher (2010). The three equation system can be cast in state space, where the Kalman filter and smoother allow for handing of the ragged edged dataset and forecasting of GDP growth.

AR-augmented models

To compensate the absence of GDP's own past in BEQ and MIDAS, we introduce AR-term into those two models. For simplification, we only consider AR(1), AR-BEQ can be written as

$$y^Q_t = \alpha + \lambda y^Q_{t-1} + \sum_{s=0}^{p_i} \ \beta_{i,s} x^Q_{i,t-s} + \varepsilon^Q_{i,t} \tag{2.12}$$

λ is estimated simultaneously with α and $\beta_{i,s}$.

The MIDAS with AR term (MIDAS-AR) is specified as proposed in Clements and Galvão (2008), the multi-step analogue of the normal MIDAS-AR and factor augmented MIDAS-AR (FA-MIDAS-AR) are defined as:

$$y^{Q}_{t+h} = \beta_0 + \lambda y^{Q}_{t} + \beta_1 B(L^{(1/3)};\theta)(1-\lambda L^{h})x^{(3)}_{i,t+w} + \varepsilon_{i,t+h} \tag{2.13}$$

$$y^{Q}_{t+h} = \beta_0 + \lambda y^{Q}_{t} + \beta_1 B(L^{(1/3)};\theta)(1-\lambda L^{h})\hat{f}^{(3)}_{t+w} + \varepsilon_{t+h} \tag{2.14}$$

2.3 Data and forecast design

This section describes the dataset we used to estimate the mechanical linear models (section 2.3.1), the pseudo real-time setup (section 2.3.2), the weighting scheme we used for pooling the QVAR, BEQ, BEQ-AR, MFVAR, MIDAS and MIDAS-AR models (section 2.3.3). Finally, section 2.3.4 describes the selection of the number of lags in the models as well as the factor selection criterion in the DFM and SFM model.

2.3.1 Dataset

Our monthly dataset consists of 72 monthly time-series variables, using harmonized definitions across countries. The indicator variables fall into four groups: production & sales, prices, monetary & financial indicators and surveys. Moreover, we added three composite indicators from the OECD[6]. Table 2.6 in the appendix provides an overview of all variables, the applied transformations and the starting date of the monthly series for each country in our sample. Monthly data are usually available on a seasonally (and calendar effects) adjusted basis at the source. When necessary, raw data series are seasonally adjusted by the US Census X12-method. All monthly series are made stationary by differencing or log-differencing (in case of trending data, such as industrial production, retail sales and monetary aggregates). For factor based models, all variables are standardized by subtracting the mean and dividing by the standard deviation. This normalization is necessary to avoid overweighting of large variance series in the determination of common factors.

[6] The primary source of the data is the ECB Statistical Datawarehouse (see `http://sdw.ecb.europa.eu/`). World trade and world industrial production are from the CPB World trade monitor (see `http://www.cpb.nl/en/world-trade-monitor`). Commodity price and most financial market indicators were taken from Thomson Reuters Datastream and most of the survey data from the European Commission (see `http://ec.europa.eu/economy_finance/db_indicators/surveys/index_en.htm`).

Quarterly GDP data for France, Italy, the Netherlands and Spain were taken from the OECD release data and revisions database[7]. Quarterly GDP data for Germany are from the Deutsche Bundesbank. Data refer to re-unified Germany from 1991.I onwards and to West Germany before 1991.I[8]. We constructed a synthetic GDP series for the euro area using the database underlying the ECB's Area-Wide Model, supplemented with data from the OECD[9].

2.3.2 Pseudo real-time design

The forecast design aims to replicate the availability of the data at the time forecasts are made in order to mimic as closely as possible the real-time flow of information. To this end, we used a data set downloaded on January 16, 2012 and combined this with the typical data release calendar to reconstruct the available dataset on the 16^{th} of each month during the period July 1995 - January 2012. All monthly indicator series start in January 1985, while the quarterly GDP series start in 1985.I. We thus employ a pseudo real-time design, which takes data publication delays into account, but ignores the possibility of data revisions for GDP and some indicators, such as industrial production. The latter implies that we might overestimate the forecasting accuracy of statistical models. However, the effects of data revisions on the final forecast may largely cancel out, since statistical methods typically attempt to eliminate noise in the process by either extracting factors from a large data set or pooling a large number of indicator-based forecasts. For example, Schumacher and Breitung (2008), using real-time data vintages for Germany, did not find a clear impact of data revisions on the forecasting accuracy of factor models. Moreover, the effect on the relative performance of models, which is the main focus of this chapter, is likely to be quite small (see also Bernanke and Boivin 2003). Abstracting from data revisions may affect the comparison of mechanical forecasts and forecasts by professional analysts to a greater extent, because GDP data are subject to substantial revisions. However, there is no obvious, feasible way to correct for this.

[7] The OECD release data and revisions database is publicly available at`http://stats.oecd.org/mei/default.asp?rev=1` For France we used the January 2012 vintage, for Italy the January 2012, December 2011 and April 2006 vintages, for the Netherlands the January 2012 and July 2005 vintages and for Spain the January 2012, November 2011, May 2005 and January 1999 vintages. The series for Italy, the Netherlands and Spain were constructed by backdating the January 2012 GDP-series by applying the quarter-on-quarter growth rates from the most recent GDP vintage.

[8] See `http://www.bundesbank.de/statistik/statistik_zeitreihen.en.php?lang=en&open=&func=row&tr=JB5000`.

[9] See `http://www.eabcn.org/data/awm/index.htm`.

Table 2.1: Timing of forecast exercise (example: forecast for third quarter)

Quarter to be forecast		Forecast made on 16^{th} day of
Two quarter ahead	1	January
	2	February
	3	March
One quarter ahead	1	April
	2	May
	3	June
Nowcast	1	July
	2	August
	3	September
Backcast	1	October
	2	November

We estimate the parameters of all models recursively using only the information available at the time of the forecast. See Rünstler et al. (2009), Giannone et al. (2008) and Kuzin et al. (2011), among others, for a similar approach. We construct a sequence of eleven forecasts for GDP growth in a given quarter, obtained in consecutive months. Table 2.1 explains the timing of the forecasting exercise, taking the forecast for the third quarter of 2011 as an example. We make the first forecast in January 2011, which is called the two-quarter-ahead forecast in month one. We subsequently produce a monthly forecast for the next ten months through November. The last forecast is made just before the first release of GDP in mid-November. Following the usual naming convention (see Bańbura et al., 2011), *forecasts* refer to one or two quarter ahead forecasts, *nowcasts* refer to current quarter forecasts and *backcasts* refer to forecasts for the preceding quarter, as long as official GDP figures are not yet available. In case of our example 2011.III, we make two quarter ahead forecasts from January to March, one quarter ahead forecasts from April to June, nowcasts from July to September, and backcasts in October and November.

2.3.3 Weighting scheme of indicator-based forecasts

The models BEQ, BEQ-AR, QVAR, MFVAR, MIDAS and MIDAS-AR construct a large number of different indicator-specific forecasts in the first stage, which have to be aggregated in the second stage to obtain the final forecast. Taking a weighted average of a large number of forecasts may ameliorate the effects of misspecification bias, parameters instability and measurement errors in the data, that may afflict the individual forecasts (Timmerman 2006). We have investigated three different weighting schemes:

(i) equal weights (simple mean); (ii) weights that are inversely proportional to the Root Mean Squared Forecast Error (RMSFE) measured from the start of the sample period until the previous quarter (recursive RMSFE scheme); and (iii) weights that are inversely proportional to the RMSFE measured over the past four quarters (moving window RMSFE scheme). Equal weights have been proven to work quite well as pooling mechanism (e.g. Stock and Watson, 2004 and Clark and McCracken, 2010). The latter two methods assign weights to the indicators based on their forecasting performance in the (recent) past. Formally, defining $\text{RMSFE}_{i,t}$ and $\text{RMSFE(4Q)}_{i,t}$ as the forecasting model's recursive RMSFE and moving window RMSFE using x_i as a dependent variable respectively,

$$RMSFE_{i,t} = \left[\tfrac{1}{t-1} \sum_{j=1}^{t-1} e_{j,i}^2 \right]^{\frac{1}{2}} \tag{2.15}$$

$$RMSFE(4Q)_{i,t} = \left[\tfrac{1}{4} \sum_{j=t-4}^{t-1} e_{j,i}^2 \right]^{\frac{1}{2}} \tag{2.16}$$

Then, the recursive and moving window RMSFE scheme are:

$$y_{t+h|t}^Q = \sum_{i=1}^{n} \omega_{i,t} y_{i,t+h|t}^Q, \tag{2.17}$$

where $\omega_{i,t} = \frac{(1 \backslash RMSFE_{i,t})}{\sum_{i=1}^{n} (1 \backslash RMSFE_{i,t})}$ for recursive RMSFE scheme, and $\omega_{i,t} = \frac{(1 \backslash RMSFE(4Q)_{i,t})}{\sum_{i=1}^{n} (1 \backslash RMSFE(4Q)_{i,t})}$ for moving window RMSFE scheme.

Tables 2.7 - 2.12 in the Appendix give an overview of the RMSFE of the three weighting schemes by horizon and country for the six relevant models. The overall picture is that the moving window RMSFE weighting scheme, which emphasizes performance in the recent past, has the smallest RMSFE on average, although the difference with the recursive RMSFE weighting scheme is quite small. In the rest of the chapter we therefore apply the moving window RMSFE weighting scheme for all relevant models and all countries.

2.3.4 Selection of maximum number of lags and number of common factors

Across models, countries and samples, the maximum number of lags in forecasting equations is determined recursively by the Schwartz information criterion (SIC). The maximum number of lags that allowed is 4 for quarterly data and 6 for monthly data. There is one exception. We find that the MFVAR for lag > 3 produces instable results, both when using the expectation-maximization algorithm as with maximum likelihood maximization. We conclude the specification with three lags results in the smallest RMSFE for most periods and most countries, so we choose this model as the base model[10].

Estimation of the static and dynamic factor model requires the specification of the number of static and dynamic common factors, denoted by r and q respectively. We base the choice of r and q on the combination that minimizes the RMSFE, evaluated over the entire sample 1996.I-2011.III[11]. We limited the search for r to the interval [1,6]. The upper bound of 6 was derived from the scree test of Cattell (1966). A grid search resulted in the following number of static factors: euro area: $r = 2$; Germany: $r = 1$; France: $r = 3$; Italy: $r = 4$; Spain: $r = 3$; Netherlands: $r = 6$.

We followed a similar procedure for the selection of the value of r and q in the dynamic factor model, imposing the restrictions r$\leq$6 and q$\leq$r. The second restriction is motivated by the finding of D'Agostino and Giannone (2012) that restricting the number of dynamic factors to be smaller than the number of static factors does not hurt predictive accuracy. Moreover, we applied the SIC to determine the maximum number of lags p in (equation 2.5) for p$\leq$6 (given r and q). The specification search led to the following numerical values: euro area: $r = 6$, $q = 5$, $p = 4$; Germany:$r = 2$, $q = 2$, $p = 3$; France: $r = 5$, $q = 2$, $p = 6$; Italy: $r = 6$, $q = 4$, $p = 2$; Spain: $r = 6$, $q = 2$, $p = 5$; Netherlands: $r = 6$, $q = 4$, $p = 2$.

2.4 Empirical results for statistical models

2.4.1 Forecasting performance

Table 2.2 presents data on the forecast performance of the eleven statistical models for our six countries for the complete sample period 1996.I-2011.III (63 quarters). The

[10]Detailed results are available upon request with the authors.

[11]Alternatively one could choose the number of factors r and q on the basis of in-sample criteria, as described in Bai and Ng (2002,2007). Our experience is that these criteria tend to indicate a relatively large number of factors, leading to volatile and less accurate forecasts (higher RMSFE). Detailed results are available upon request with the authors.

underlying empirical analysis has been carried out on a monthly basis for eleven horizons. To save space Table 2.2 (and the other tables in this chapter as well) reports results for the two and one-quarter ahead forecasts, the nowcast and the backcast, which have been calculated as the average of the corresponding monthly data. We measure forecast performance by the root mean square forecast error (RMSE). The first column of Table 2.2 reports the RMSE of the benchmark model (AR model). For the other statistical models the entries refer to their RMSE relative to that of the benchmark model in order to improve the comparability of the results across countries and horizons. Shaded entries indicate the model with the lowest RMSE in a row (for a particular horizon). Bold entries indicated models that have an RMSE that is less than 10 per cent larger than that of the best model and also smaller than the RMSE of the benchmark model[12]. The 10 per cent threshold is meant as a rough assessment of the economic significance of differences in forecasting ability. We will call models that meet this condition "competitive models" as in terms of forecasting performance they do not differ "too much" from the best model[13].

The outcomes in Table 2.2 point to several interesting results. First, incorporating monthly information in statistical forecasting procedures pays off in terms of forecasting accuracy, in particular for nowcasts and backcasts. The large majority of the relative RMSEs are smaller than 1 and they also tend to fall if the horizon shortens and more monthly information is absorbed. Second, for many models the gain is rather limited when truly forecasting. For the 2-quarter ahead forecast, the best models have on average an RMSE that is only 5% lower than the benchmark. Except for Spain, even the best statistical model does not deliver an economic significant improvement. For the 1-quarter ahead forecast the average improvement by the best models is 15% on the benchmark, but the other models generally post gains that are less than 10% on the benchmark. For the nowcast and backcast the average gain in accuracy is around a third for the best performing models. This pattern suggests that statistical models have greater value added when they can use information that pertains to the relevant quarter. Their relative strength is to improve the assessment of the current state of the economy. Third, the static and dynamic factor models display the best performance overall. Looking across countries and horizons, either the static or the dynamic factor model performs best. The only exception is the bridge model which is the best model in case of the 2-quarter ahead forecast for the Netherlands. The dynamic factor model works better for nowcasts and backcasts, while the static factor model has the edge for the 1-quarter ahead forecast. Fourth, many models

[12]If the best model has an RMSE of 0.6, the cut-off point is an RMSE of 0.66.

[13]Like other authors we refrain from doing conventional statistical tests as these are not discriminating in practice.

Table 2.2: Comparison RMSFE, euro area and five largest EA countries

Frequency	Quarterly					Mixed frequency						
Pool/Factor	Benchmark	Pooling			Factor	Factor	Pooling					
Model	AR	BEQ	BEQ-AR	QVAR	SFM	DFM	MIDAS	FA-MIDAS	MIDAS-AR	FA-MIDAS-AR	MFVAR	FA-MFVAR
RMSFE	(Absolute)	(Relative to AR-model)										
euro area												
2Q ahead	*0.64*	**0.96**	**0.97**	1.01	0.95	1.01	**0.97**	1.00	**0.97**	1.01	**0.97**	1.08
1Q ahead	*0.63*	**0.90**	**0.91**	1.00	0.83	**0.88**	0.92	0.98	0.93	0.97	0.92	0.99
nowcast	*0.60*	0.85	0.83	0.96	0.67	**0.68**	0.88	0.82	0.83	0.80	0.89	0.78
backcast	*0.53*	0.87	0.81	0.92	**0.61**	0.56	0.83	0.72	0.73	0.70	0.87	0.66
Germany												
2Q ahead	*0.96*	**0.94**	**0.94**	**0.96**	0.92	**0.93**	**0.94**	**0.98**	**0.98**	**1.00**	**0.94**	1.01
1Q ahead	*0.94*	**0.93**	**0.93**	0.97	**0.87**	0.85	0.93	0.97	0.94	0.97	0.94	0.98
nowcast	*0.93*	0.90	0.88	0.97	**0.77**	0.72	0.89	0.84	0.90	0.84	0.92	0.84
backcast	0.93	0.83	0.83	0.94	**0.73**	0.66	0.78	**0.71**	0.81	0.74	0.91	0.74
France												
2Q ahead	*0.54*	**0.95**	**0.95**	**0.99**	**0.94**	0.92	**0.94**	**1.00**	**0.95**	1.01	**0.94**	1.11
1Q ahead	*0.52*	0.92	0.91	1.01	0.80	**0.84**	0.94	0.97	0.92	0.95	0.92	1.02
nowcast	*0.47*	0.90	0.87	0.98	**0.75**	0.69	0.93	0.89	0.88	0.84	0.86	0.90
backcast	*0.44*	0.90	0.87	0.95	0.63	**0.64**	0.91	0.72	0.85	0.75	0.89	0.76
Italy												
2Q ahead	*0.78*	**0.95**	**0.95**	**0.97**	0.94	**0.96**	**0.94**	**0.96**	**0.95**	**0.96**	**0.95**	1.01
1Q ahead	*0.73*	**0.97**	**0.98**	1.01	0.89	**0.93**	**0.98**	1.01	**0.97**	1.00	0.98	1.05
nowcast	*0.70*	0.92	0.91	0.99	**0.76**	0.72	0.94	0.86	0.92	0.85	0.92	0.86
backcast	*0.68*	0.90	0.87	0.92	0.64	**0.65**	0.90	0.73	0.87	0.72	0.87	0.73
Spain												
2Q ahead	*0.58*	**0.98**	1.00	1.00	**0.94**	0.89	1.00	1.01	0.98	1.02	1.02	1.03
1Q ahead	*0.52*	1.02	1.02	1.03	**0.86**	0.80	1.04	1.06	0.94	0.97	1.14	1.06
nowcast	*0.46*	1.01	1.05	1.01	0.87	0.72	1.01	1.04	0.91	0.98	1.17	1.08
backcast	*0.49*	0.86	1.00	0.90	0.77	0.64	0.89	0.87	0.96	1.02	1.09	1.09
Netherlands												
2Q ahead	*0.71*	0.96	**0.97**	**1.00**	1.03	1.00	**0.97**	1.02	**0.97**	1.03	**0.98**	1.10
1Q ahead	*0.71*	0.91	0.91	0.99	0.81	**0.88**	0.95	0.98	0.92	0.98	0.96	1.02
nowcast	*0.68*	0.86	0.85	0.97	**0.74**	0.73	0.91	0.83	0.88	0.84	0.92	0.84
backcast	*0.64*	0.90	0.86	0.93	**0.71**	0.71	0.90	0.82	0.85	0.81	0.97	0.86

Notes: AR: Autoregressive model, BEQ: Pooled bridge equations, BEQ-AR: Pooled bridge equation with AR term, QVAR: Pooled vector autoregressive models, SFM Static Factor Model, DFM: Dynamic factor model, FA-MIDAS: Factor augmented MIDAS, FA-MIDAS-AR: Factor augmented MIDAS with AR term, MFVAR: Mixed frequency vector autoregressive model, FA-MFVAR: Factor augmented mixed frequency vector autoregressive model, MIDAS: MIDAS model, MIDAS-AR: MIDAS with AR-term.
Grey cells indicate model has lowest RMSFE. Bold figures indicate models that have relative RMSFE 10% or more lower than best model.

are competitive at the 2-quarter ahead horizon in most of the countries, but their number quickly falls as the horizon shortens. For the majority of the countries there typically is only one (other) competitive model for nowcasts and backcasts, usually the static or dynamic factor model. The first result is another piece of evidence that predictions by statistical models incorporate little information at the 2 quarters ahead horizon. The second result suggests that the static and dynamic factor model display a significantly larger ability to absorb monthly information than the other models. Within the latter group, factor-augmented models (FA-MIDAS, FA-MIDAS-AR and FA-MFVAR) are the best of the rest, while the quarterly and MFVAR models are the clear underperformers. Fifth, within our sample of countries Spain is an exceptional case as all statistical models do badly for all horizons, except for the static and dynamic factor model.

2.4.2 The marginal value of statistical models

Ranking models by their RMSE gives a first perspective on their relative usefulness. This subsection focuses on the marginal value of models by investigating whether forecasts generated by different models differ in their information content. As the various statistical approaches follow different strategies of extracting monthly information, it is conceivable that some models are complementary. In that case taking a weighted average of their respective forecasts may improve forecast accuracy. Even a badly performing model may have a positive marginal value provided it is able to pick up specific useful information. We establish the marginal value of the models versus the best statistical model (lowest RMSE) by running the encompassing test proposed by Stekler (1991). The test regression is

$$y^Q_{t+h|t} = \lambda \hat{y}^Q_{\mathrm{a}(t+h|t)} + (1-\lambda)\hat{y}^Q_{\mathrm{b}(t+h|t)} + \varepsilon_t \tag{2.18}$$

where y^Q_t is GDP growth in t, $\hat{y}^Q_{\mathrm{a}(t+h|t)}$ and $\hat{y}^Q_{\mathrm{b}(t+h|t)}$ are the forecasts for quarter t+h on time t of the alternative and best model respectively, λ is the weight of the alternative model and $(1-\lambda)$ is the weight of the best model. In order to get interpretable results, we impose the restriction that λ lies between 0 and 1. The alternative model contains additional information compared to the best model if $\lambda > 0$. We estimate λ and its standard error on the interval [0,1] by Maximum Likelihood (ML) and perform a one-sided (asymptotically valid) test of the hypothesis $\lambda = 0$ at the 5% level of significance. All calculations refer to the complete sample period 1996.I-2011.III (63 quarters).

Table 2.3 reports the results of our encompassing test. Entries depict the RMSE of the forecast combination relative to the RMSE of the best model as a measure of the potential gains from using forecast combinations. The estimated weight λ itself is not reported; bold entries signify λ estimates that are statistically greater than zero. A blank entry means that de ML algorithm returned the corner solution $\lambda = 0$.

The main message of Table 2.3 is that the gains from combining forecasts by different statistical models are limited in economic terms. Moreover, no model emerges as a clear winner, the best model in terms of marginal value is country-specific. It thus appears that the various approaches do not greatly differ with respect to the type of information they extract from large-scale monthly datasets. In the majority of the cases there is no gain in accuracy at all for horizons up to the nowcast. The best opportunities are for improving backcasts, when models have absorbed the maximum amount of monthly information. Except for Germany and Italy, the majority of the models offers some scope for improving backcasts. For the euro area the maximum possible reduction in the RMSE is 9%, for France 4%, for Spain 9% and for the Netherlands 6%. For nowcasts the maximum reduction in the RMSE does not exceed 4%. At the 2 quarter ahead horizon a comparatively large number of models appear to offer additional information, but the associated gains are very small (typically 1% reduction in RMSE). Finally, Table 2.3 shows that statistical significance and economic importance are different concepts. Most non-zero entries reflect a significant test result for the encompassing test, while most of the gains in forecast accuracy are very small.

2.4.3 Splitting the sample: Great Moderation versus Financial crisis

Our sample includes the financial crisis when real GDP went through a particularly volatile phase across the industrialized countries. An obvious question is whether and to what extent the performance of statistical forecasting models differs between the financial crisis period and the period before the financial crisis which was characterized by a large degree of macroeconomic stability. The latter period has been labeled as the Great Moderation. Most of the existing literature on short term forecasting is based on data from the Great Moderation period. Forecasting in volatile times poses of course greater challenges, so the results of a comparative analysis will be more informative on the issue which models are particularly apt at absorbing monthly information. Moreover, good forecasts and nowcasts are of greater importance to economic agents and policy makers in a volatile environment.

Table 2.3: Comparison RMSFE, euro area and five largest EA countries

Frequency	Quarterly					Mixed frequency						
Pool/Factor	Benchmark	Pooling			Factor	Factor	Pooling					
Model	AR	BEQ	BEQ-AR	QVAR	SFM	DFM	MIDAS	FA-MIDAS	MIDAS-AR	FA-MIDAS-AR	MFVAR	FA-MFVAR
RMSFE	(Reduction in relative RMSFE from combining model with best model)											
euro area												
2Q ahead	1.00	**0.99**	**0.99**	1.00	.	.	0.99	.	0.99	.	0.99	.
1Q ahead	.	.	.	.	.	.	.	.	.	.	.	.
nowcast	.	.	.	.	.	**0.96**	.	.	.	.	.	.
backcast	**0.98**	**0.97**	**0.97**	**0.97**	**0.96**	.	**0.96**	**0.96**	**0.94**	**0.96**	**0.98**	**0.91**
Germany												
2Q ahead	.	.	.	.	.	1.00	.	.	.	.	**0.99**	.
1Q ahead	.	.	.	.	.	.	.	.	.	.	.	.
nowcast	.	.	.	.	.	.	.	.	.	.	.	.
backcast	.	.	.	.	**0.99**	.	.	.	.	.	.	.
France												
2Q ahead	**0.99**	**0.99**	**0.99**	**0.99**	.	.	**0.98**	.	**0.98**	.	**0.98**	.
1Q ahead	.	.	.	.	.	.	.	.	.	.	.	.
nowcast	.	.	.	.	**0.99**	.	.	.	.	.	.	.
backcast	**0.99**	**0.99**	**0.99**	**0.98**	.	**0.96**	0.99	.	**0.99**	.	**0.99**	.
Italy												
2Q ahead	.	**0.99**	.	.	.	.	**0.99**	.	**0.99**	.	.	.
1Q ahead	**0.99**	.	.	.	.	.	1.00	.	**0.99**	.	.	.
nowcast	.	.	.	.	.	.	.	.	.	.	.	.
backcast	.	.	.	.	.	**0.98**	.	.	.	.	.	0.98
Spain												
2Q ahead	**0.97**	**0.99**	**0.99**	**0.98**	.	.	**0.98**	.	**0.98**	.	**0.99**	.
1Q ahead	**0.98**	.	1.00	**0.99**	**0.97**	.	.	.	**0.98**	.	.	.
nowcast	**0.96**	.	.	**0.99**	.	.	.	.	**0.99**	.	.	.
backcast	**0.91**	**0.98**	**0.99**	**0.95**	.	.	**0.98**	.	**0.99**	.	.	.
Netherlands												
2Q ahead	.	.	.	.	**0.99**	.	.	.	.	.	.	.
1Q ahead	.	.	.	.	.	.	.	.	.	.	.	.
nowcast	.	.	.	.	**0.96**	.	.	.	.	.	.	.
backcast	**0.96**	**0.95**	**0.94**	**0.95**	.	**0.96**	**0.95**	**0.97**	**0.95**	**0.96**	**0.95**	**0.96**

Notes: AR: Autoregressive model, BEQ: Pooled bridge equations, BEQ-AR: Pooled bridge equation with AR term, QVAR: Pooled vector autoregressive models, SFM Static Factor Model, DFM: Dynamic factor model, FA-MIDAS: Factor augmented MIDAS, FA-MIDAS-AR: Factor augmented MIDAS with AR term, MFVAR: Mixed frequency vector autoregressive model, FA-MFVAR: Factor augmented mixed frequency vector autoregressive model, MIDAS: MIDAS model, MIDAS-AR: MIDAS with AR-term.
Grey cells indicate model has lowest RMSFE. Bold figures indicate encompassing test indicates difference is statistically significant at the 5% level.

We divide the sample period into two parts: 1996.I - 2007.IV (Great Moderation) and 2008.I - 2011.III (Financial crisis). We discuss the performance of the models on the basis of their learning curve, which shows the relative decline in the RMSE as the forecasting horizon shortens, averaged over five countries.[14] We calculate a model's learning curve as the RMSE standardized by the RMSE for the first month of the 2 quarter ahead forecast. Figure 2.3 shows the learning curves of selected models for the complete sample period and the two subperiods (in the rows). The graphs on the left refer to models that aggregate indicator-specific forecasts, the graphs on the right refer to models that rely on factor analysis to summarize indicators. For presentational reasons we restrict the comparison to the AR versions of the MIDAS, BEQ and FA-MIDAS models (as they perform better than the non-AR versions) and leave out the QVAR model (as it learns in quarterly steps) and FA-MFVAR (as its forecast performance is meager).

For the complete sample period we find that the DFM model displays the steepest learning curve, with the SFM model being a close second. In addition, models involving factor analysis have steeper learning curves than models that aggregate indicator-specific forecasts. This is a stable pattern that holds both during the Great Moderation episode and the crisis episode (and also across countries). It thus appears that the DFM and SFM model are the fastest learning models in volatile as well as tranquil environments.

Predicting GDP is much more difficult in the crisis period. The RMSE of the benchmark model during the crisis period is two to three times as large as that during the Great Moderation. Part of this deterioration can be offset as the scope for improving forecasts by utilizing monthly information appears to be larger in volatile times, in particular for nowcasting and backcasting. For example, the RMSE of the dynamic factor model falls by 22% on average over the course of 11 months in the period before the crisis as compared to 58% in the crisis period. Differences in forecast accuracy across models are considerably larger after the crisis than before the crisis. This also means that the number of competitive models during the Great Moderation is much larger (about twice as many) than after the financial crisis, even for the nowcasting and backcasting horizons. This outcome support the outcome of d'Agostino and Giannone (2012) who show that the gain from using factor models is substantial, especially in periods of high comovements, as was the case during the Financial crisis. The crisis episode poses a more demanding test to models and consequently fewer models manage to pass. This finding also implies that the cost of employing a suboptimal model has increased after the crisis. Finally, the potential gains of combining

[14] We leave out Spain, because all statistical models fail to beat the benchmark model in the period 1996.I-2007.IV. Country details can be found in appendix which contains versions of Tables 2.2 and 2.3 for both subperiods.

statistical models (marginal value) tend to be markedly smaller during de financial crisis compared to the preceding period[15].

2.4.4 Assessing the building blocks of the models

The fact that our analysis includes eleven models and six countries allows us to shed some light on the issue which model features are especially valuable for forecasting. Referring to Figure 2.2, we focus on the following modeling choices: (1) employing factor analysis to summarize monthly information; (2) allowing for mixed frequencies in the forecasting equation or system; (3) exploiting GDP's own past by adding an autoregressive term to the forecasting equation. To assess a modeling feature's effect on the RMSE we compare (sets of) models that only differ in that aspect. Moreover, we take the average over five countries (once again we exclude Spain) to average out the country-specific component.

To measure the impact of utilizing factor analysis for aggregating monthly information rather than aggregating indicator-specific forecasts we can compare three pairs of models: (FA-MIDAS, MIDAS), (FA-MIDAS-AR, MIDAS-AR) and (FA-MFVAR, MFVAR). For the AR effect we can also exploit three pairs: (BEQ-AR, BEQ), (MIDAS-AR, MIDAS) and (FA-MIDAS-AR, FA-MIDAS). We assess the mixed frequency effect in two different ways. The first method involves the pairs (MIDAS, BEQ) and (MIDAS-AR, BEQ-AR), which relates quarterly GDP data on the left-hand side to monthly data (as opposed to quarterly averages of monthly data) on the right-hand side in a single forecasting equation setting. The second method is based on the pair (MFVAR, QVAR) and includes the effect of making GDP a monthly latent variable in a system.

Table 2.4 reports the impacts of the three modeling features (averaged over five countries) for the complete sample period and the two subperiods. Starting with the effect of utilizing factor analysis, we find that this improves the accuracy of nowcasts and backcasts, but for true forecasting factor augmented models show a weaker performance. This suggests that summarizing information of monthly data is primarily helpful when the information pertains to the quarter of interest itself. When truly forecasting the inevitable loss of information that summarizing implies appears to dominate any gains from the removal of noise. For the complete sample the average gain is 6% for nowcasts and 12% for backcasts. During tranquil times these gains are smaller, while in the crisis period the gains are larger, up to 17% for backcasts. Allowing for mixed frequencies in a single equation setting (MIDAS of BEQ) only modestly lowers the RMSE for backcasts, while for

[15]Moreover, the encompassing test is significant in only a few cases but this can partly be attributed to the low number of observations.

the other horizons the RMSE deteriorates slightly. In the crisis these effect are somewhat more pronounced, but still small. Treating GDP as monthly latent variable in a system has positive effects on all horizons for the whole sample, but this result appears to be completely driven by the crisis period. Lastly, exploiting GDP's own past by adding an AR term tends to improve the accuracy of forecasts a little bit for most horizons, especially for nowcasts (3%) and backcasts (5%) during the crisis episode.

2.5 Analysis of forecasts by professional analysts

The views of professional forecasters are an alternative and convenient source of information for policy makers and market participants. Currently, several surveys on the economic outlook exist and are regularly updated. The European Central Bank undertakes a quarterly survey among professional forecasters to get information on inflation expectations and growth prospects for the euro area. In the US, the Federal Reserve Bank of Philadelphia runs a well-known survey. Moreover, the private sector firm Consensus Economics collects and publishes economic forecasts on a monthly basis in the publication *Consensus Forecasts*. Consensus Forecasts offers an overview private analysts' expectations for a set of key macroeconomic variables for a broad range of countries. Consensus Forecasts is best known for its expectations on annual GDP growth for the current and next year. However, it also provides quarterly forecasts for GDP, which we will use in this chapter[16]. The panelists supply their forecasts for six consecutive quarters, starting from the first unpublished quarter. The number of respondents varies somewhat over time, but on average about nine institutions participate in the poll for the Netherlands, fifteen for Italy and Spain, twenty for France and thirty for Germany and the euro area.

This section investigates two issues. The first issue is the quality of Consensus forecasts as a separate forecasting device compared to the best statistical model. The second issue is the marginal value of Consensus forecasts based on an encompassing test versus the best model. In forming their expectations, analysts include subjective assessments on potentially a multitude of relevant factors. If a mixture of model-based and (subjective) Consensus forecasts improves the accuracy of forecasts, this can be viewed as evidence that forecasts by analysts indeed embody a different type of valuable information (subjective judgments).

We use the mean quarterly forecast as the measure of private sector expectations in our

[16] The annual Consensus forecasts have been analyzed in several papers (e.g. Ager et al. (2009), Batchelor (2001), Loungani and Rodriguez (2008) and Lahiri et al. (2006)). The quarterly forecasts have not been used before, except in a case study for the Netherlands by de Winter (2011).

Table 2.4: Incremental forecasting accuracy model structure, average country effect

Modeltype	AR				Mixed Frequency				Factor			
Competitor	FA-MIDAS-AR	MIDAS-AR	BEQ-AR	Avrg.	MIDAS	MIDAS-AR	Avrg.	MFVAR	FA-MFVAR	FA-MIDAS	FA-MIDAS-AR	Avrg.
Base	FA-MIDAS	MIDAS	BEQ	Avrg.	BEQ	BEQ-AR	Avrg.	QVAR	MFVAR	MIDAS	MIDAS-AR	Avrg.
Evaluation period 1996Q1.2011Q3 (N=63)												
2Q ahead	0.01	0.01	0.00	**0.01**	0.00	0.01	**0.00**	-0.03	0.11	0.04	0.04	**0.06**
1Q ahead	-0.01	-0.01	0.00	**-0.01**	0.02	0.01	**0.01**	-0.05	0.06	0.04	0.04	**0.05**
nowcast	-0.01	-0.03	-0.02	**-0.02**	0.03	0.01	**0.02**	-0.07	-0.06	-0.06	-0.05	**-0.06**
backcast	0.00	-0.04	-0.03	**-0.02**	-0.02	-0.03	**-0.02**	-0.03	-0.16	-0.12	-0.08	**-0.12**
Evaluation period 1996Q1.2007Q4 (N=48)												
2Q ahead	0.00	0.01	0.00	**0.00**	-0.01	0.00	**0.00**	-0.02	0.04	0.04	0.03	**0.04**
1Q ahead	0.01	0.00	0.00	**0.00**	0.01	0.00	**0.00**	0.01	-0.01	0.02	0.03	**0.01**
nowcast	0.00	-0.02	-0.01	**-0.01**	0.01	0.01	**0.01**	0.00	-0.05	-0.03	0.00	**-0.03**
backcast	0.02	-0.01	0.01	**0.01**	0.00	-0.01	**0.00**	0.04	-0.09	-0.06	-0.03	**-0.06**
Evaluation period 2008Q1.2011Q3 (N=15)												
2Q ahead	0.02	0.01	0.00	**0.01**	0.00	0.01	**0.01**	-0.04	0.14	0.04	0.05	**0.08**
1Q ahead	-0.02	-0.01	0.00	**-0.01**	0.03	0.02	**0.02**	-0.08	0.10	0.05	0.04	**0.06**
nowcast	-0.03	-0.04	-0.02	**-0.03**	0.04	0.02	**0.03**	-0.11	-0.06	-0.09	-0.08	**-0.08**
backcast	-0.01	-0.07	-0.07	**-0.05**	-0.03	-0.03	**-0.03**	-0.07	-0.21	-0.18	-0.12	**-0.17**

Notes: Average incremental forecasting accuracy (competing model - base model) across the euro area, Germany, France, Italy and the Netherlands.

AR: Autoregressive model, BEQ: Pooled bridge equations, BEQ-AR: Pooled bridge equation with AR term, QVAR: Pooled vector autoregressive models, FA-MIDAS: Factor augmented MIDAS, FA-MIDAS-AR: Factor augmented MIDAS with AR term, MFVAR: Mixed frequency vector autoregressive model, FA-MFVAR: Factor augmented mixed frequency vector autoregressive model, MIDAS: MIDAS model, MIDAS-AR: MIDAS with AR-term.

analysis. Fresh Consensus forecasts become available only once a quarter, in the second week of the last month of each quarter. For the information set this means Consensus forecasts are not updated in the first and second month in a quarter, while monthly indicator series are updated every month. Moreover, at the time panelists form their expectations they have information on GDP growth in the preceding quarter. The backcast for quarter t is therefore equal to the non-updated Consensus forecast published in the last month of quarter t.

Table 2.5 presents the results for Consensus forecasts for six countries for the complete sample period, the pre-crisis period and the crisis period[17]. For 2 quarter ahead forecasts Consensus forecasts are better than the best statistical model in case of the euro area and Spain, while they are a competitive model in another three cases over the whole sample. When the horizon shortens, however, the relative performance versus the best model deteriorates starkly in all countries except for Germany and Spain. Consequently, purely mechanical models seem to be (much) more adept at learning when monthly information becomes available. In the relatively quiet pre-crisis period, Consensus forecasts fare very poorly, usually ranking at the bottom of the list. However, Consensus forecasts do very well in the case of Germany. By contrast, during the crisis period, when GDP displayed extreme fluctuations, Consensus forecasts perform much better. At the 2 quarter ahead horizon Consensus forecasts are the best model for five countries, and at the 1 quarter ahead horizon they consistently belong to the top-three models. For Spain and the Netherlands the difference is substantial. This suggests that analysts are better able to handle extreme observations of GDP growth once they have occurred, while the quality of estimated models in mechanical procedures is more susceptible to extreme observations in the sample, in particular when truly forecasting. This supports the outcome of Lundquist and Stekler (2012) who conclude professional forecasters are very responsive to the latest information about the state of the economy and adjust their predictions quickly. We find that despite this head start, in most cases private sector forecasts still fall behind the best model as the horizon becomes shorter and more monthly information can be used to improve forecasts. For example, the RMSE of backcasts by Consensus forecasts is between 20% and 84% larger than the RMSE associated with the best model (static or dynamic factor model).

Despite the fact that Consensus forecasts are a rather poor predictor of GDP on their own, the results for the encompassing test show that they often contain valuable extra information, which may be used to improve mechanical forecasts for all countries except Italy. The most striking results concern the backcasts by Consensus forecasts, even though

[17]Consensus forecasts are available for the euro area from March 2002 onward only, so results in Table 4 refer to the period 2003.III-2011.III for the euro area.

these actually reflect relatively dated information. Measured over the whole sample period, enriching mechanical forecasts with subjective Consensus forecasts delivers a gain in accuracy of around 10% on average. During the crisis period Consensus forecasts, unlike their statistical competitors, still offer added value for some countries. This holds in particular for Spain and to a lesser extent for the Netherlands. During the pre-crisis period forecasts for Germany and the euro area may benefit from Consensus forecasts. The outcomes of the encompassing test suggests that subjective private sector forecasts potentially contain information that cannot be picked up by sophisticated mechanical forecasting procedures. An interesting topic for future research is to investigate how the potential of judgmental forecasts may be taken on board in mechanical procedures in a real time context.

2.6 Conclusion

This chapter makes two contributions to the empirical literature on forecasting real GDP in the short run. The first contribution is a systematic comparison of eleven statistical linear models for six countries (Germany, France, Italy, Spain, the Netherlands and the euro area), utilizing the same information set across countries. Our sample period (1996.I-2011.III) allows us to compare the models' forecasting abilities in the period before the financial crisis of 2008 (Great Moderation) and the much more volatile subsequent period (financial crisis and its aftermath). The second contribution concerns the potential usefulness of (subjective) forecasts made by professional analysts. Such forecasts are very cheap and easy to use, and they may incorporate valuable information that goes beyond purely statistical data.

We summarize our findings in five points. First, monthly indicators contain valuable information that can be extracted by mechanical statistical procedures, in particular as the horizon shortens and more monthly information is processed. The largest gains in accuracy are for nowcasts and backcasts, suggesting that statistical models are especially helpful when they are able to use information that pertains to the quarter of interest. Moreover, statistical models are generally more efficient in extracting monthly information in volatile times. Their relative strength is thus to improve the assessment of the current state of the economy. By contrast, predictions by statistical models generally incorporate little information at the two quarters ahead horizon.

Second, the dynamic and static factor models consistently display the best forecasting capabilities across countries in the period 1996.I-2011.III. Their relatively strong performance in the volatile crisis episode is key to this result. The dominance of factor models

Table 2.5: Comparison Consensus Forecasts with best linear model

Eval. period	1996Q1-2011Q3 (N=63)			1996Q1-2007Q4 (N=48)			2008Q1-2011Q3 (N=15)		
Indicator	rRMSE	rank	gain	rRMSE	rank	gain	rRMSE	rank	gain
euro area									
2Q ahead	**0.99**	1	**0.98**	1.21	13	.	**0.98**	1	**0.97**
1Q ahead	**1.06**	3	**0.98**	1.42	13	.	**1.04**	2	**0.98**
nowcast	1.34	5	**0.99**	1.43	13	**0.99**	1.34	4	.
backcast	1.77	7	**0.93**	1.54	13	**0.83**	1.84	7	**0.95**
Germany									
2Q ahead	**1.04**	7	.	**1.05**	10	**0.99**	**1.04**	7	.
1Q ahead	**1.06**	3	**0.99**	**0.99**	1	**0.95**	1.12	3	.
nowcast	1.12	3	**0.98**	**0.91**	1	**0.90**	1.35	6	.
backcast	**1.00**	2	**0.91**	**0.76**	1	**0.76**	1.40	6	**0.99**
France									
2Q ahead	**1.05**	8	**0.98**	1.23	13	.	**0.98**	1	**0.97**
1Q ahead	1.18	7	.	1.30	13	.	1.11	3	.
nowcast	1.29	8	**0.99**	1.32	13	**0.99**	1.27	4	.
backcast	1.27	6	**0.95**	1.30	11	**0.97**	1.36	6	**0.95**
Italy									
2Q ahead	1.11	13	.	1.32	13	.	**0.99**	1	**0.99**
1Q ahead	1.17	12	.	1.40	13	.	**1.01**	2	**0.97**
nowcast	1.36	11	.	1.45	13	**0.99**	1.27	4	.
backcast	1.47	12	**0.99**	1.43	13	**0.97**	1.67	6	.
Spain									
2Q ahead	**0.98**	1	**0.94**	1.23	13	.	**0.88**	1	**0.87**
1Q ahead	1.10	3	**0.96**	1.27	13	.	**0.99**	1	**0.90**
nowcast	1.21	2	**0.96**	1.36	13	.	1.13	2	**0.85**
backcast	1.10	2	**0.88**	1.27	12	.	**0.85**	1	**0.57**
Netherlands									
2Q ahead	**1.06**	9	**0.98**	1.27	13	.	**0.90**	1	**0.90**
1Q ahead	1.11	3	**0.97**	1.17	13	**0.99**	1.05	2	**0.95**
nowcast	1.23	9	**0.98**	1.29	13	**0.99**	1.15	3	**0.95**
backcast	1.39	12	**0.91**	1.56	13	**0.98**	1.20	4	**0.87**

Notes: rRMSFE: relative RMSFE against best linear model, rank: rank in 13 models; 12 linear models + consensus forecasts, gain: gain in forecast accuracy from combining Consensus Forecasts with best linear model.

Bold figures indicate encompassing test indicates difference is statistically significant at 5% level.

is somewhat weaker during the more tranquil period of the Great Moderation.

Third, regarding the question which modeling features are critical to success, we find that employing factor analysis to summarize the available monthly information clearly delivers better results than the alternative of averaging indicator-based forecasts in the case of nowcasts and backcasts. Factor strategies work better than pooling strategies. Moreover, allowing for mixed frequencies and autoregressive terms (GDP's own past) in forecasting procedures leads to minor improvements in forecast reliability. All of these effects are more pronounced during the crisis period, implying that the cost of employing a suboptimal forecasting model is larger in periods of high volatility.

Fourth, statistical models significantly differ in the rate at which they are able to absorb monthly information as time goes by. However, the information content of the resulting forecasts appears to overlap to a large extent and the unique model-specific component appears to be small (in relation to the best model). The different models do not seem to have a comparative advantage of extracting a certain type of information, offering perspectives that complement each other. The scope for improving GDP forecasts by combining the 'views' of various models is rather limited in economic terms, although there are some exceptions. This is particularly true during volatile episodes when reliable assessments of the current situation and short run prospects are most needed, unfortunately.

Lastly, forecasts by professional analysts, which contain judgmental elements, appear to be a different category. Such forecasts are in many cases a rather poor predictor of GDP compared to the best statistical model. However, they tend to perform better during the crisis, when it really counts, and they often embody information that sophisticated mechanical forecasting procedures fail to pick up. Subjective private sector forecasts thus seems to offer the potential of enhancing mechanical forecasts.

The results of our large-scale comparative analysis may useful to policy makers, financial analysts and economic agents alike, as information on where the economy stands and where it is heading to in the immediate short run is particularly valuable in times of great uncertainty. The dynamic factor model and the static factor model, which is a quite simple procedure from a technical point of view, are obvious candidate models for generating short term forecasts in practice. An interesting topic for future research is to investigate how the potential of judgmental forecasts may be taken on board in mechanical procedures in a real time context.

2.7 References

Ager, P., Kappler, M., Osterloh, S., 2009. The accuracy and efficiency of the Consensus Forecasts: a further application and extension of the pooled approach. *International Journal of Forecasting* 25, 167-181.

Baffigi, A., Golinelli, R., Parigi G., 2004. Bridge models to forecast the euro area GDP. *International Journal of Forecasting* 20, 447-460.

Bai, B.J., Ng S., 2002. Determining the number of factors in approximate factor models. *Econometrica* 71, 191-221.

Bai, B.J, Ng S., 2007. Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics* 25, 52-60.

Bańbura, M., Giannone, D., Reichlin, L., 2011. Nowcasting. In *The Oxford Handbook of Economic Forecasting*, Clements M. and Hendry D (eds). Oxford University Press: Oxford, 193-224.

Bańbura, M., Rünstler, G., 2011. A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting* 27, 333-346.

Batchelor, R., 2001. How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *Applied Economics* 33, 225-235.

Bernanke B.S., Boivin J., 2003. Monetary policy in a data-rich environment. *Journal of Monetary Economics* 50, 525-546.

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

Camba-Mendez, G., Kapetianos, G., Smith, M., Weale, R., 2001. An automatic leading indicator of economic activity: forecasting GDP growth for European countries. *Econometrics Journal* 4, 56-80.

Camacho, M., Perez-Quiros, G., 2010. Introducing the euro-sting: short-term indicator of euro area growth. *Journal of Applied Econometrics* 25, 663-694.

Cattell, R.B., 1966. The scree test for the number of factors. *Multivariate behavioral research* 1, 245-276.

Clark, T.E., McCracken, M.W., 2010. Averaging forecasts from VARs with uncertain instabilities. *Journal of Applied Econometrics* 25, 5-29.

Clements, M.P., Galvão, M.P., 2008. Macroeconomic forecasting with mixed-frequency data: forecasting output growth in the United States. *Journal of Business and Economic Statistics* 26, 546-554.

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

Diebold, F.X., Mariano, R.S., 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13(3), 253-263.

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, 188-205.

Durbin, J., Koopman, S.J., 2001. *Time series analysis by state space methods.* Oxford University Press: Oxford.

ECB. 2009. *Results of a special questionnaire for participants in the ECB survey of professional forecasters.* ECB: Frankfurt.

Fair, R.C., Shiller R.J., 1990. Comparing information in forecasts from econometric models. *The American Economic Review* 80, 375-389.

Foroni, C., Marcellino, M., 2012. A comparison of mixed frequency approaches for modelling euro area macroeconomic variables. *EUI Working Papers* No. 2012/07.

Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: the real-time informational content of macroeconomic data. *Journal of Monetary Economics* 55, 665-676.

Ghysels., E., Sinko, A., Valkanov, R., 2007. MIDAS regressions: further results and new directions. *Econometric Reviews* 26, 53-90.

Kitchen, J., Monaco, R., 2003. Real-time forecasting in practice: the US treasury staffs real-time GDP forecast system. *Business Economics* 38, 10-28.

Kuzin, V., Marcellino, M, Schumacher, C., 2011. MIDAS vs. mixed-frequency VAR: nowcasting GDP in the euro area. *International Journal of Forecasting* 27, 529-542.

Lahiri, K., Isiklar, G., Loungani, P., 2006. How quickly do forecasters incorporate news? Evidence from cross-country surveys. *Journal of Applied Econometrics* 21, 703-725.

Lieberman, J., 2012. *Real-time forecasting in a data-rich environment.* MPRA Working Paper No. 39452.

Loungani, P., Rodriguez. J., 2008. Economic Forecasts. *World Economics* 9, 1-12.

Lundquist, K., Stekler, H.O., 2012. Interpreting the Performance of Business Economists During the Great Recession. *Business Economics* 47, 148-154.

Marcellino, M., Schumacher, C., 2010. Factor MIDAS for nowcasting and forecasting with ragged-edge data: a model comparison for German GDP. *Oxford Bulletin of Economics and Statistics* 72, 518-550.

Mariano, R.S., Murasawa, Y., 2010. A Coincident Index, Common Factors, and Monthly Real GDP. *Oxford Bulletin of Economics and Statistics* 71, 27-48. *Journal of Econometrics* 135(1-2), 499-526.

den Reijer, A.H.J., 2005. *Forecasting Dutch GDP using large scale factor models.* DNB

Working Paper No. 28.

Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., den Reijer, A., Jakaitiene, A., Jeloniek, P., Rua, A., Ruth, K., van Nieuwenhuyze, C. 2009. Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting* 28, 595-611.

Schumacher C, Breitung J. 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting* 24, 386-398.

Stekler, H.O., 1991. Macroeconomic forecast evaluation techniques. *International Journal of Forecasting* 7, 375-384.

Stock J.H., Watson, J., 2002. Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics* 20, 147-162.

Stock J.H., Watson J. 2004. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting* 23, 405-430.

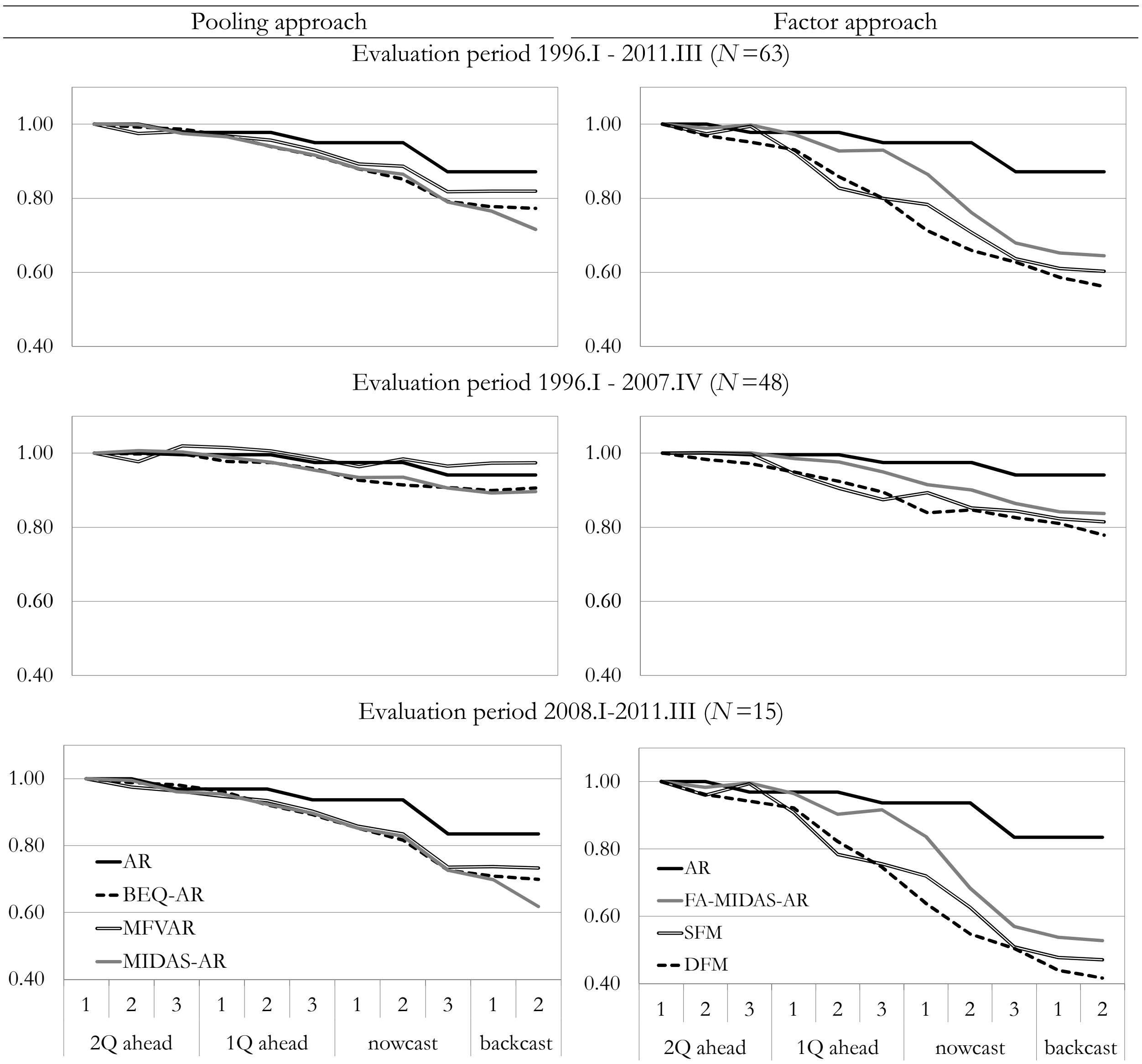
Timmerman, A., 2006. Forecast Combinations. In *Handbook of Economic Forecasting*, Elliott, G., Granger, C.W.J., Timmerman, A. (eds). Elsevier: Amsterdam, 135-196.

Wallis K. 1986. Forecasting with an econometric model: the “ragged edge” problem. *Journal of Forecasting* 5, 1-3.

de Winter JM. 2011. Forecasting GDP growth in times of crisis: private sector forecasts versus statistical models. *DNB Working Paper* No. 28.

2.8 Appendix

2.8.1 Database



Notes: RMSFE 2Q ahead forecast=100; all lines excluding Spain.

Figure 2.3: Learning curve linear models

Table 2.6: Database description

nr	description	type	transform			country					
			log	dif.	filt.	EA	DE	FR	IT	ES	NL
1	World Trade (CPB)	Sales	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
2	World Industrial Production (CPB)	Sales	1	1	3	‘91	‘91	‘91	‘91	‘91	‘91
3	Ind. production United States	Sales	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
4	Ind. production United Kingdom	Sales	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
5	Ind. production (excl. construction)	Sales	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
6	Ind. production, cons. goods ind.	Sales	1	1	3	‘80	‘80	‘77	‘77	‘77	‘90
7	Ind. production, energy	Sales	1	1	3	‘80	‘80	‘77	‘80	‘80	‘90
8	Ind. production, interm. goods ind.	Sales	1	1	3	‘77	‘80	‘77	‘77	‘77	‘95
9	Ind. production, capital goods	Sales	1	1	3	‘77	‘80	‘77	‘77	‘77	‘77
10	Ind. production, manufacturing	Sales	1	1	3	‘77	‘78	‘77	‘77	‘80	‘77
11	Ind. production, construction	Sales	1	1	3	‘85	‘78	‘85	‘95	‘88	‘85
12	New orders manufacturing	Sales	1	1	3	‘95	‘91	‘00	‘90	‘00	‘95
13	New passenger cars (reg.)	Sales	1	1	3	‘90	‘90	‘90	‘90	‘90	‘90
14	New commercial vehicles (reg.)	Sales	1	1	3	‘90	‘90	‘90	‘90	‘90	‘90
15	Retail trade volume	Sales	1	1	3	‘77	‘77	‘77	‘90	‘95	‘77
16	Unemployment rate	Sales	0	1	3	‘83	‘91	‘83	‘83	‘86	‘83
17	Unemployment rate United Kingdom	Sales	0	1	3	‘83	‘83	‘83	‘83	‘83	‘83
18	Unemployment rate United States	Sales	0	1	3	‘83	‘83	‘83	‘83	‘83	‘83
19	Exports	Sales	1	1	3	‘00	‘89	‘89	‘89	‘89	‘89

Continued on next page...

Table 2.6 – Continued

nr	description	type	transform			country					
			log	dif.	filt.	EA	DE	FR	IT	ES	NL
20	Imports	Sales	1	1	3	‘00	‘89	‘89	‘89	‘89	‘89
21	Total HICP-index	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
22	Core HICP-index	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
23	CPI, food	Prices	1	2	3	‘90	‘77	‘77	‘77	‘93	‘77
24	CPI, energy	Prices	1	2	3	‘90	‘77	‘77	‘77	‘77	‘77
25	HICP, services	Prices	1	2	3	‘90	‘85	‘90	‘87	‘92	‘87
26	Producer prices (total, excl. constr.)	Prices	1	2	3	‘81	‘77	‘77	‘77	‘77	‘77
27	World commodity prices, total	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
28	World commodity prices, raw mat.	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
29	World commodity prices, food	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
30	World commodity prices, metals	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
31	World commodity prices, energy	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
32	Oil price (1 month future Brent)	Prices	1	2	3	‘77	‘77	‘77	‘77	‘77	‘77
33	M1	Finan.	1	1	3	‘77	‘80	‘77	‘80	‘80	‘80
34	M3	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
35	Interest rate on mortgage	Finan.	0	1	3	‘03	‘82	‘80	‘95	‘84	‘80
36	3 month interest rate euro	Finan.	0	1	3	‘94	‘77	‘77	‘77	‘77	‘77
37	10 year government bond yield	Finan.	0	1	3	‘77	‘94	‘77	‘77	‘80	‘77
40	Headline stock-index	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘87	‘77
41	Basic Material-index	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘87	‘77

Continued on next page…

Table 2.6 – Continued

nr	description	type	transform			country					
			log	dif.	filt.	EA	DE	FR	IT	ES	NL
42	Industrials stock-index	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘87	‘77
43	Consumer goods stock-index	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘87	‘77
44	Consumer services stock-index	Finan.	1	1	3	‘77	‘77	‘77	‘87	‘77	‘77
45	Financial stock-index	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘87	‘77
46	Technology stock-index	Finan.	1	1	3	‘77	‘88	‘77	‘86	‘99	‘85
47	Loans to the private sector	Finan.	1	1	3	‘91	‘80	‘80	‘83	‘80	‘82
48	Exchange rate, US-Dollar per Euro	Finan.	1	1	3	‘80	‘80	‘80	‘80	‘80	‘80
49	Real effective exchange rate (CPI)	Finan.	1	1	3	‘77	‘77	‘77	‘77	‘77	‘77
50	Ind. conf. - headline	Survey	0	1	3	‘85	‘85	‘85	‘85	‘87	‘85
51	Ind. conf. - Order-book expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘87	‘85
52	Ind. conf. - Stocks expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘87	‘85
53	Ind. conf. - Production expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘87	‘85
54	Ind. conf. - Employment expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘87	‘85
55	Cons. conf. - headline	Survey	0	1	3	‘85	‘85	‘85	‘85	‘86	‘85
56	Cons. conf. - Financial sit.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘86	‘85
57	Cons. conf. - General ec. sit.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘86	‘85
58	Cons. conf. - Unemployment expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘86	‘85
59	Cons. conf. - Major purchases expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘86	‘85
60	Constr. conf. - Headline	Survey	0	1	3	‘85	‘85	‘85	‘85	‘89	‘85
61	Constr. conf. - Order book (evolution)	Survey	0	1	3	‘85	‘85	‘85	‘85	‘89	‘85

Continued on next page...

Table 2.6 – Continued

nr	description	type	transform			country					
			log	dif.	filt.	EA	DE	FR	IT	ES	NL
62	Constr. conf. - Employment expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘89	‘85
63	Retail conf. - Headline	Survey	0	1	3	‘85	‘85	‘85	‘85	‘88	‘86
64	Retail conf. - Current Stocks (volume)	Survey	0	1	3	‘85	‘85	‘85	‘85	‘88	‘86
65	Retail conf. - Orders expectations	Survey	0	1	3	‘85	‘85	‘85	‘85	‘88	‘86
66	Retail conf. - Business expect.	Survey	0	1	3	‘85	‘85	‘85	‘85	‘88	‘86
67	Retail conf. - Employment expect.	Survey	0	1	3	‘86	‘85	‘85	‘86	‘88	‘86
68	PMI United States	Survey	0	1	3	‘77	‘77	‘77	‘77	‘77	‘77
69	PMI United Kingdom	Survey	0	1	3	‘92	‘92	‘92	‘92	‘92	‘92
70	OECD Composite Leading ind. UK	Other	0	1	3	‘77	‘77	‘77	‘77	‘77	‘77
71	OECD Composite Leading ind. US	Other	0	1	3	‘77	‘77	‘77	‘77	‘77	‘77
72	OECD Composite Leading ind.	Other	0	1	3	‘77	‘77	‘77	‘77	‘77	‘77

Notes: type: sales=production and sales, finan.= monetary and financial, price= price data, survey= surveys; transform: log: 0=no logarithm, 1=logarithm, dif.: degree of differentiating 1=first difference, 2=second difference; filt.: 3= change against the same month of the previous month.

2.8.2 State space representation of mixed frequency VAR and dynamic factor model

Mixed frequency VAR

This section describes the state space representation of the mixed frequency VAR described in section 2.2.3. Let $p^* = \max(p, 3)$ and the transition equation of state vector is as follows:

$$\begin{bmatrix} z_{i,t+1} - \mu_{z_i} \\ z_{i,t} - \mu_{z_i} \\ \dots \\ z_{i,t-p^*+2} - \mu_{z_i} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \dots & A_p & 0_{2\times 2(3-p^*)} \\ I_{2(p^*-1)} & & & & 0_{2(p^*-1)\times 2} \end{bmatrix} \begin{bmatrix} z_{i,t} - \mu_{z_i} \\ z_{i,t-1} - \mu_{z_i} \\ \dots \\ z_{i,t-p^*+1} - \mu_{z_i} \end{bmatrix} + \begin{bmatrix} \Sigma_{\varepsilon}^{1/2} \\ 0_{2(p^*-1)\times 2} \end{bmatrix} v_t, \tag{2.19}$$

where $v_t \sim N(0, I_2)$. The measurement equation is:

$$z_{i,t}^{Q} - \mu_{z_i^Q} = \begin{bmatrix} 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0_{1\times(p^*-6)} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0_{1\times(p^*-6)} \end{bmatrix} \begin{bmatrix} z_{i,t} - \mu_{z_i} \\ z_{i,t-1} - \mu_{z_i} \\ \dots \\ z_{i,t-p^*+1} - \mu_{z_i} \end{bmatrix} \tag{2.20}$$

Since, y_t^Q is only available on the third month of the quarter we fill the missing observations in month 1 and 2 with a random draw from the standard normal distribution $N(0, 1)$, in accordance with Mariano and Murasawa (2010). As a consequence, we modify the measurement equation of month 1 and month 2, to accommodate the missing observation treatment. For the month y_t^Q is unavailable, the upper row of the matrix in the right hand side of equation 2.20 is all equal to zero, and white noise is added.

Dynamic factor model

The equations of the DFM, i.e equation 2.3 and 2.5 to 2.7 can be cast in state space form as illustrated below for the case of $p = 1$. The aggregation rule is implemented in a recursive way in equation 2.22 by introducing a latent cumulator variable Ξ for which: $\Xi_t = 0$ for t corresponding to the first month of the quarter and $\Xi_t = 1$ otherwise. The monthly state space representation is given by the following observation equation:

$$\begin{bmatrix} x_t \\ y_t^Q \end{bmatrix} = \begin{bmatrix} \Lambda & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_t \\ y_t \\ \hat{y}_t^Q \end{bmatrix} + \begin{bmatrix} \xi_t \\ \varepsilon_t^Q \end{bmatrix} \tag{2.21}$$

and the transition equation:

$$\begin{bmatrix} I_r & 0 & 0 \\ -\beta' & 1 & 0 \\ 0 & -\frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} f_{t+1} \\ y_{t+1} \\ \hat{y}_{t+1}^Q \end{bmatrix} = \begin{bmatrix} A_{r1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Xi_{t+1} \end{bmatrix} \begin{bmatrix} f_t \\ y_t \\ \hat{y}_t^Q \end{bmatrix} + \begin{bmatrix} \zeta_{t+1} \\ \varepsilon_t \\ 0 \end{bmatrix} \tag{2.22}$$

The application of the Kalman filter and smoother provides the minimum mean square linear estimates (MMSLE) of the state vector $\alpha_t = (f_t, y_t, \hat{y}_t^Q)$ and enables the forecasting of quarterly GDP growth y_t^Q and dealing efficiently with an unbalanced dataset of missing observations at the beginning and at the end of the series by replacing the missing data with optimal predictions. Moreover, when compared with the use of principal components technique alone, the two-step estimator enables the dynamics of the common factors and the cross-sectional heteroskedasticity of the idiosyncratic component.

2.8.3 Sensitivity Analysis

Table 2.7: Pooling schemes for **Quarterly Bridge Equation (BEQ)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,90	0,90	0,89	0,52	0,52	0,52	0,74	0,74	0,74	0,60	0,59	0,58	0,70	0,70	0,70	0,62	0,62	0,62
2Q Forecast	month 2	0,90	0,90	0,89	0,52	0,52	0,51	0,73	0,73	0,74	0,59	0,58	0,58	0,69	0,69	0,69	0,62	0,61	0,61
	month 3	0,90	0,90	0,90	0,51	0,51	0,51	0,72	0,72	0,73	0,57	0,57	0,55	0,68	0,68	0,67	0,61	0,60	0,61
	month 1	0,89	0,89	0,90	0,50	0,50	0,50	0,72	0,72	0,73	0,57	0,56	0,55	0,67	0,67	0,66	0,60	0,59	0,59
1Q Forecast	month 2	0,88	0,88	0,87	0,49	0,49	0,48	0,71	0,71	0,71	0,56	0,55	0,53	0,66	0,66	0,64	0,59	0,58	0,57
	month 3	0,87	0,87	0,86	0,48	0,47	0,45	0,70	0,69	0,69	0,55	0,54	0,51	0,65	0,64	0,63	0,57	0,56	0,55
	month 1	0,85	0,85	0,85	0,46	0,46	0,44	0,68	0,68	0,67	0,55	0,53	0,49	0,64	0,63	0,61	0,55	0,54	0,53
Nowcast	month 2	0,84	0,84	0,84	0,46	0,45	0,43	0,67	0,66	0,65	0,54	0,52	0,47	0,63	0,62	0,57	0,54	0,53	0,52
	month 3	0,81	0,80	0,82	0,44	0,43	0,40	0,65	0,64	0,62	0,52	0,50	0,44	0,61	0,60	0,58	0,51	0,49	0,48
Backcast	month 1	0,80	0,79	0,78	0,43	0,42	0,40	0,64	0,63	0,62	0,52	0,49	0,43	0,61	0,60	0,58	0,50	0,48	0,46
	month 2	0,80	0,79	0,77	0,43	0,42	0,39	0,64	0,62	0,60	0,51	0,49	0,42	0,61	0,60	0,57	0,50	0,48	0,47

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Table 2.8: Pooling schemes for **Quarterly Vector Autoregressive model (QVAR)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,92	0,92	0,92	0,53	0,53	0,53	0,76	0,76	0,76	0,60	0,60	0,60	0,71	0,71	0,71	0,65	0,65	0,65
2Q Forecast	month 2	0,92	0,92	0,92	0,53	0,53	0,53	0,76	0,76	0,76	0,60	0,60	0,60	0,71	0,71	0,71	0,65	0,65	0,65
	month 3	0,92	0,92	0,91	0,53	0,53	0,55	0,74	0,74	0,74	0,58	0,57	0,56	0,70	0,70	0,70	0,64	0,64	0,64
	month 1	0,92	0,92	0,91	0,52	0,52	0,55	0,74	0,74	0,74	0,57	0,57	0,56	0,70	0,70	0,70	0,64	0,64	0,64
1Q Forecast	month 2	0,92	0,92	0,91	0,52	0,52	0,55	0,74	0,74	0,74	0,57	0,57	0,56	0,70	0,70	0,70	0,64	0,64	0,64
	month 3	0,91	0,91	0,92	0,48	0,48	0,49	0,73	0,73	0,73	0,50	0,50	0,49	0,69	0,69	0,69	0,62	0,62	0,62
	month 1	0,91	0,91	0,91	0,48	0,48	0,49	0,73	0,73	0,73	0,50	0,49	0,48	0,69	0,69	0,69	0,62	0,62	0,62
Nowcast	month 2	0,91	0,91	0,91	0,48	0,48	0,49	0,73	0,73	0,73	0,50	0,49	0,48	0,69	0,69	0,69	0,62	0,62	0,62
	month 3	0,87	0,87	0,88	0,42	0,42	0,41	0,63	0,63	0,62	0,48	0,46	0,44	0,60	0,60	0,60	0,50	0,50	0,49
Backcast	month 1	0,87	0,87	0,88	0,42	0,42	0,41	0,63	0,63	0,62	0,48	0,47	0,44	0,59	0,59	0,59	0,50	0,50	0,49
	month 2	0,87	0,87	0,88	0,42	0,42	0,41	0,63	0,63	0,62	0,48	0,47	0,44	0,59	0,59	0,59	0,50	0,50	0,49

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Table 2.9: Pooling schemes for **Mixed Frequency VAR (MFVAR)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,91	0,91	0,92	0,51	0,51	0,52	4,31E+05	0,74	0,74	0,61	0,60	0,61	0,71	0,70	0,70	0,62	0,62	0,63
2Q Forecast	month 2	0,91	0,91	0,86	0,50	0,50	0,50	5,08E+05	0,74	0,74	0,61	0,60	0,59	0,70	0,70	0,74	0,61	0,61	0,61
	month 3	0,90	0,90	0,91	0,49	0,49	0,50	3,44E+05	0,73	1,49	0,60	0,59	0,59	0,68	0,68	0,69	0,61	0,60	0,61
	month 1	0,89	0,89	1,00	0,49	0,49	0,49	5,56E+05	0,72	0,74	0,59	0,58	0,62	0,68	0,68	0,71	0,60	0,60	0,60
1Q Forecast	month 2	0,88	0,88	0,89	0,48	0,48	0,49	1,52E+06	0,71	0,71	0,59	0,58	0,64	0,68	0,67	0,69	0,59	0,58	0,59
	month 3	0,87	0,87	0,88	0,45	0,45	0,47	1,78E+06	0,70	0,70	0,53	0,52	0,52	0,66	0,66	0,67	0,57	0,57	0,56
	month 1	0,86	0,86	0,85	0,44	0,43	0,44	1,96E+06	0,68	0,86	0,53	0,52	0,53	0,63	0,63	0,64	0,55	0,54	0,54
Nowcast	month 2	0,86	0,84	0,86	0,43	0,42	0,42	1,23E+06	0,67	0,66	0,53	0,51	0,52	0,63	0,62	0,64	0,54	0,54	0,56
	month 3	0,85	0,84	0,87	0,38	0,37	0,37	1,81E+06	0,59	0,59	0,58	0,56	0,57	0,61	0,61	0,60	0,47	0,47	0,49
Backcast	month 1	0,83	0,81	0,86	0,36	0,36	0,37	2,21E+06	0,59	0,59	0,57	0,55	0,53	0,60	0,60	0,61	0,47	0,46	0,48
	month 2	0,80	0,79	0,89	0,36	0,35	0,41	4,46E+05	0,57	0,59	0,56	0,55	0,54	0,59	0,59	0,62	0,45	0,44	0,45

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Table 2.10: Pooling schemes for **Mixed Data Sampling (MIDAS)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,91	0,91	0,91	0,52	0,52	0,51	0,73	0,73	0,73	0,62	0,60	0,60	0,70	0,70	0,69	0,62	0,63	0,62
2Q Forecast	month 2	0,91	0,91	0,90	0,52	0,52	0,51	0,73	0,73	0,72	0,60	0,58	0,58	0,70	0,69	0,70	0,62	0,62	0,61
	month 3	0,90	0,90	0,90	0,51	0,51	0,51	0,73	0,72	0,72	0,59	0,54	0,57	0,69	0,69	0,68	0,61	0,61	0,61
	month 1	0,89	0,89	0,89	0,50	0,50	0,51	0,72	0,72	0,72	0,59	0,54	0,56	0,68	0,68	0,68	0,60	0,60	0,60
1Q Forecast	month 2	0,89	0,89	0,88	0,50	0,50	0,49	0,71	0,71	0,72	0,58	0,53	0,55	0,67	0,67	0,67	0,59	0,59	0,59
	month 3	0,87	0,87	0,88	0,48	0,48	0,48	0,70	0,69	0,69	0,56	0,51	0,52	0,66	0,66	0,66	0,57	0,57	0,56
	month 1	0,86	0,86	0,85	0,47	0,47	0,46	0,68	0,67	0,67	0,54	0,49	0,48	0,65	0,64	0,64	0,56	0,55	0,54
Nowcast	month 2	0,85	0,85	0,84	0,46	0,46	0,45	0,67	0,66	0,67	0,54	0,49	0,48	0,64	0,64	0,63	0,55	0,54	0,53
	month 3	0,82	0,82	0,81	0,44	0,43	0,42	0,65	0,64	0,65	0,51	0,46	0,44	0,62	0,62	0,60	0,51	0,50	0,50
Backcast	month 1	0,81	0,80	0,77	0,43	0,42	0,40	0,64	0,63	0,62	0,51	0,46	0,44	0,61	0,61	0,58	0,50	0,49	0,47
	month 2	0,80	0,79	0,68	0,42	0,41	0,39	0,63	0,62	0,60	0,52	0,46	0,42	0,61	0,60	0,56	0,50	0,48	0,42

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Table 2.11: Pooling schemes for **Quarterly Bridge Equation with AR-term (BEQ-AR)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,90	0,90	0,90	0,52	0,52	0,52	0,74	0,74	0,74	0,60	0,59	0,60	0,70	0,70	0,70	0,62	0,62	0,62
2Q Forecast	month 2	0,90	0,90	0,89	0,51	0,51	0,51	0,74	0,74	0,74	0,59	0,59	0,58	0,69	0,69	0,69	0,62	0,61	0,61
	month 3	0,90	0,90	0,90	0,51	0,51	0,51	0,72	0,72	0,73	0,58	0,57	0,57	0,68	0,68	0,68	0,61	0,61	0,61
	month 1	0,89	0,89	0,90	0,50	0,50	0,50	0,72	0,72	0,72	0,58	0,57	0,55	0,67	0,67	0,66	0,60	0,60	0,59
1Q Forecast	month 2	0,88	0,88	0,87	0,49	0,49	0,48	0,71	0,71	0,72	0,57	0,56	0,54	0,66	0,66	0,64	0,59	0,58	0,57
	month 3	0,87	0,87	0,87	0,46	0,45	0,45	0,70	0,70	0,69	0,54	0,53	0,52	0,65	0,64	0,62	0,57	0,56	0,56
	month 1	0,86	0,85	0,84	0,44	0,44	0,43	0,68	0,68	0,67	0,54	0,52	0,48	0,64	0,63	0,61	0,55	0,54	0,53
Nowcast	month 2	0,84	0,84	0,82	0,43	0,43	0,41	0,67	0,67	0,65	0,54	0,52	0,47	0,63	0,62	0,58	0,54	0,53	0,52
	month 3	0,83	0,82	0,80	0,40	0,40	0,39	0,61	0,60	0,60	0,56	0,54	0,51	0,57	0,57	0,55	0,46	0,45	0,44
Backcast	month 1	0,81	0,81	0,78	0,39	0,39	0,38	0,60	0,60	0,59	0,56	0,54	0,49	0,57	0,56	0,55	0,45	0,44	0,43
	month 2	0,81	0,80	0,78	0,39	0,39	0,38	0,60	0,59	0,58	0,56	0,54	0,48	0,57	0,56	0,55	0,45	0,44	0,43

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Table 2.12: Pooling schemes for **Mixed Data Sampling with AR-term (MIDAS-AR)**: Equal weights, Recursive RMSFE sample, four-quarter moving moving RMSFE

		DE			FR			IT			ES			NL			EA		
	Month	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q	AV	RC	4Q
	month 1	0,95	0,95	0,95	0,52	0,52	0,52	0,74	0,74	0,73	0,62	0,60	0,60	0,70	0,70	0,70	0,62	0,63	0,62
2Q Forecast	month 2	0,94	0,94	0,95	0,51	0,51	0,51	0,74	0,73	0,74	0,61	0,59	0,59	0,70	0,70	0,70	0,62	0,62	0,62
	month 3	0,91	0,90	0,90	0,51	0,51	0,50	0,74	0,73	0,73	0,53	0,51	0,52	0,68	0,68	0,68	0,61	0,61	0,61
	month 1	0,90	0,90	0,89	0,50	0,50	0,50	0,73	0,72	0,72	0,53	0,51	0,52	0,67	0,67	0,67	0,60	0,60	0,61
1Q Forecast	month 2	0,89	0,89	0,88	0,49	0,49	0,49	0,72	0,71	0,71	0,53	0,50	0,51	0,66	0,66	0,64	0,59	0,59	0,58
	month 3	0,88	0,88	0,89	0,46	0,46	0,45	0,70	0,69	0,69	0,45	0,44	0,44	0,66	0,66	0,65	0,57	0,57	0,57
	month 1	0,86	0,86	0,84	0,44	0,44	0,43	0,68	0,67	0,67	0,43	0,43	0,41	0,64	0,64	0,64	0,55	0,54	0,53
Nowcast	month 2	0,85	0,85	0,85	0,43	0,43	0,42	0,67	0,66	0,66	0,42	0,41	0,39	0,64	0,63	0,62	0,53	0,52	0,52
	month 3	0,84	0,84	0,82	0,40	0,40	0,39	0,61	0,60	0,60	0,54	0,48	0,47	0,57	0,57	0,55	0,46	0,45	0,45
Backcast	month 1	0,82	0,82	0,79	0,40	0,39	0,38	0,60	0,60	0,60	0,55	0,48	0,48	0,57	0,57	0,54	0,45	0,44	0,42
	month 2	0,82	0,81	0,72	0,39	0,38	0,37	0,60	0,59	0,58	0,55	0,48	0,46	0,57	0,57	0,54	0,45	0,44	0,35

Notes: AV: equal weights pooling scheme, RC: Weight according to RMSFE, recursively calculated, 4Q: Weights according to RMSFE, calculated over the last four quarters. Shaded areas indicate one of the criteria has lower RMSFE than the two other methods.

Chapter 3

A Factor Augmented Mix Frequency VAR/MA Model for Monthly GDP Estimation

3.1 Introduction

The gross domestic product (GDP), which refers to the market value of all officially recognized final goods and services produced in a given period, is the most comprehensive and reliable macroeconomic indicator. However, the lack of monthly frequency data hinders its function for timely monitoring the state of economy.

The aim of this chapter is to construct a GDP based monthly coincident index, which sufficiently accommodate information from both quarterly GDP and numerous monthly indicators. Three elements are pivotal for estimating monthly GDP. Firstly, how many monthly indicator should be contained? inter alia the trade-off between more indicators and the curse of dimensionality. Secondly, how to connect the quarterly and monthly data, namely the mix frequency problem. Thirdly, which kind of model should be assumed for the latent monthly GDP data?

There are two state of art methodologies arise our attention. One is Angelini et al. (2008)'s dynamic factor model, which integrates bridge equation with factor model to estimate monthly GDP. Albeit this method alleviates the large dimension issue successfully, whether the dynamic factor model is plausible for GDP construction is indeterminate. Another one is Mariano and Murasawa (2010)'s mix frequency VAR (MFVAR) model. This structure preserves the inherent nature of time series-autoregression, and the multivariate setup dissolves the underlying endogeity issue. Nonetheless, attributing to the overparametrization issue, it constrains to limited amount of variables.

Absorbing essence from both approaches, we answer aforementioned three questions by a sophisticated factor augmented mix frequency VAR/MA (FA-MFVAR/MA) model. Instead of extracting factor and estimating the monthly series simultaneously in a state space form (see Angelini et al. 2008), we proceed the assessment in two stages. At first, a single factor is extracted from vast monthly indicators. Subsequently the foregoing extracted factor and quarterly GDP are conjuncted by a bivariate MFVAR model. Realizing the potential moving average component, the model is extended to VARMA case.

To the best of our knowledge, it is the first time that FA-MFVAR/MA model is exploited. The intriguing incorporation of VAR/MA model and factor extraction compensates the shortcoming of solely MFVAR/MA or dynamic factor model, produces a monthly indicator enable inclusion of large magnitude and mix frequency datasets.

The evaluation based on the estimation of monthly German GDP. We investigate the pros and cons of our model, in contrast with the parsimonious Chow and Lin method, single indicator MFVAR model and Angelini et al. (2008)'s dynamic factor model. The reference models represent for the benchmark interpolation technique, MFVAR model without factor extraction, and the large datasets but purely factor based methodology. The out of sample root mean square error (RMSE) and estimated monthly GDP delineation uniformly confirm that the FA-MFVAR model unilaterally beats the remaining techniques. More precisely, the substantial distinction from dynamic factor model reflects the strengthen of VAR based method over factor based one for monthly GDP estimation, which adopts the identical large information set.

This chapter is constructed as following. The literatures are summarized in Section 3.2. We sketch out the factor augmented mix frequency VAR and VARMA model respectively in Section 3.3. Section 3.4 implements the evaluation with German data. Section 3.5 concludes.

3.2 Related Literatures

3.2.1 Interpolation

Temporal disaggregation for quarterly GDP and constructing a monthly macroeconomic coincident index are two disparate but also correlated problem. Issue of interpolation, namely disaggregate high frequency data from low frequency data, received much attention in literature. The Chow & Lin (1971) approach is quite convenient and canonical for interpolation. Their approach established linear regression between low frequency data and related series firstly, afterwards assumed the same regressor coefficient for unobservable high frequency data and corresponding high frequency related series, to implement the disaggregation. The linear regression is estimated by generalized least square. Geared toward the residuals assumption, Fernández (1981) suggested the residual follows a random walk, while Litterman (1983) assumed ARIMA (1,1,0) process with unit root. Santo and Cardoso (2001) and Mitchell et al. (2005) pursued the approach on another direction, by involving divergent dynamic structure to the Chow & Lin technique.

Structural time series model, which mainly concerns components like trend, seasonal, cycle and irregular, is prevailing for disaggregation. Harvey & Chung (2000) casted a bivariate local linear trend model, to estimate the underlying change in unemployment in UK. Moauro & Savio (2005) put forward local linear trend model for the multivariable, and further developed seemingly unrelated time series equations model.

One branch of literatures merely utilized the dynamic structure of low frequency data to execute the disaggregation, without the assist of any other series. Wei & Stram (1990) and Harvey & Pierse (1984) are two distinguished examples, both assumed a ARIMA model of unobservable time series, but divergent from the approach to interpolate. The former's point of entry is the autocovariances, they used available autocovariances of an aggregated model, to access the autovariance structure of the unobservable disaggregated series. While the latter treated the disaggregated series as missing observations, introduced state space representation to obtain the unobservable high frequency data. Since Harvey & Pierse (1984) and Harvey (1989), Kalman Filter became a versatile method to solve disaggregation issue, such as Bernanke, Gertler and Watson (1997), Cuche & Hess (2000), Liu & Hall (2001).

3.2.2 Construction of Coincident Index

The construction of coincident economic index solely relies on the monthly macroeconomic indicators, precludes influence of quarterly GDP data. Factor model is the underpinning of this kind of research. Stock & Watson (1991) firstly provided a dynamic factor model for generating coincident and leading economic indicators. The standard estimation method of small scale factor models is solved by maximizing the likelihood function and means of the Kalman filter.

For high-dimension system, Stock and Watson (2002a,b) heuristically devised a nonparametric static principal components method. From then on, the tractable principal component algorithm holds powerful appeal, i.e. Boivin and Ng (2006)'s weighted version, Forni et al. (2000)'s dynamic generalized version, for consistent estimation of the factors in approximate dynamic factor models. More specifically, perceiving the factor dynamics and idiosyncratic heteroscedasticity, Doz et al. (2011) provided a two step estimator for large approximate dynamic factor model.

One disadvantage of Stock & Watson (1991) and further factor model is that the constant parameter assumption. Hamilton (1989)'s Markov switching model proposed an famed approach for regime change, via supposing the parameters of an autoregression as the outcome of a discrete-state Markov process. As a result, the variables' growth rate are allowed to correlated with the status of business cycle, compensates the drawback of factor model. Carriero & Marcellino (2007) compared the performance of constructing coincident indices for European countries by both factor and Markov switching model, and found out the performance of Markov switching is implausible, while factor model renders favorable result.

3.2.3 Estimation of Monthly Coincident Index based on GDP

A flood of papers synthesized the interpolation and coincident index technique, gauging GDP based monthly economic coincident index. The methodologies which associate with factor model are distinguished by the scale of information set. The small scale models are all motivated by the seminal contribution of Stock and Watson (1991). Mariano and Murasawa (2003) incorporated the mix frequency data, modeled the mixture of quarterly and monthly series in a static factor model as Stock and Watson (1991), casted in a linear state space framework. Proietti and Moauro (2006) extended their method by introducing a non-linear temporal aggregation constraints, while Frale et al. (2011) further exploited the decomposition sector of GDP components. The static factor model suffers from the number limitation of series. For large information set, researchers concentrate on finding the optimal connection between factor and quarterly GDP data. Angelini et al. (2008) laid out a bridge with factor framework to estimate the monthly GDP, which integrates bridge equation with factor model.

3.3 Methodology

3.3.1 Factor Model

The striking advantage of factor model for large datasets is creating a explicitly specified and statistically meticulous dimension reduction scheme. In what follows, we adopt two algorithms, viz the prototypical Stock and Watson method and the state of art two step estimator.

Stock and Watson Approach

Presumed that the N monthly time-series observations ξ_t with length T observed satisfies a factor model representation:

$$\xi_t = \Lambda F_t + e_t, \quad E(e_t e_t') = \Psi \tag{3.1}$$

where $F_t = (F_{1,t}, F_{2,t}, \dots, F_{r,t})$ is a vector of r common dynamic factors, Λ is N $\times$ r matrix of loading, and e_t comprises the idiosyncratic disturbances.

Stock and Watson (2002a,b) delivered analytical solution of factors by principal component algorithm. Considering the following least square issue:

$$min\ T^{-1} \sum_{t-1}^{T} (\xi_t - \Lambda F_t)^{'} (\xi_t - \Lambda F_t) \tag{3.2}$$

The minimization issue amounts to maximizing $trace(\Lambda \xi^{'} \xi \Lambda)$, after concentrating out F. Where $\xi = (\xi_1', \xi_2', ..., \xi_T')'$, and subject to $\Lambda\Lambda^{'} = I_r$. The solution of $\hat{\Lambda}$ is the eigenvectors of $\xi\xi^{'}$ corresponding to the r largest eigenvalues, and the estimator of F is:

$$\hat{F} = \hat{\Lambda}' \xi \tag{3.3}$$

In practice, the number of optimal factor r could be determined by eminent Bai and Ng (2002) information criterion. In virtue of our intention to diminish the model dimension, we directly choose the single factor. Generally speaking, the first factor contains the most information.

Two Step Estimation

Maintaining the factor structure as equation 3.1, Doz et al. (2011) introduced autoregressive factors and idiosyncratic heteroscedasticity of residuals, imposed a extra VAR(p) process on F_t.

$$F_t = \sum_{i=1}^{p} H_i F_{t-i} + J u_t, \quad u_t \sim N(0, I_q) \tag{3.4}$$

The estimation proceeds in four stages. Firstly, the factors loadings Λ and initial estimated factors $\hat{F}$ of the static factors are obtained by the standard principal component analysis. Secondly, the coefficient matrix H_i in equation 3.4 is estimated by Ordinary Least Squares in use of $\hat{F}$. Thirdly, the covariance matrix JJ' is decomposed into one or more principal components in equation 3.4. Ultimately, the common factor is updated in a state space form, when the Kalman smoother facilitates the treatment of the underlying ragged edge issue [1].

3.3.2 Temporal Aggregation Strategy

What need reinforce is that, the growth rate of GDP instead of stock data is estimated, avoiding the possible complication from parameter estimation and diffuse Kalman filter iteration of non-stationary series. The aggregation strategy refers to Mariano and Murasawa (2003), which takes logarithm difference transformation to the original non-stationary data. Let x_t be the quarterly GDP series observable every third period, x_t^* be the latent monthly GDP. They supposed x_t is the geometric mean of x_t^*,x_{t-1}^*,x_{t-2}^*, namely $lnx_t = \frac{1}{3}(lnx_t^* + lnx_{t-1}^* + lnx_{t-2}^*)$, which is confirmed as a good approximation by Mitchell et al. (2005).

Let $y_{t,1} = \triangle_3 lnx_t$ to be the quarterly GDP growth rate, $y_{t,1}^* = \triangle lnx_t^*$ to be the monthly GDP growth rate. The first difference form is written as:

$$y_{t,1} = \frac{1}{3}y_{t,1}^* + \frac{2}{3}y_{t-1,1}^* + y_{t-2,1}^* + \frac{2}{3}y_{t-3,1}^* + \frac{1}{3}y_{t-4,1}^* \tag{3.5}$$

3.3.3 Mix Frequency Model

The combination of unobservable monthly GDP growth $y_{t,1}^*$ and the extracted single factor $y_{t,2}$ is presumed a VAR/MA process. The involvement of moving average component is always controversial for macroeconomic estimation and forecast. In this chapter, we exploit both VAR and VARMA models, inspect their performance by out of sample RMSE. The MFVARMA model can be written as:

$$\Phi(L)(y_t^* - \mu_{y^*}) = \Theta(L)\varpi_t, \quad \varpi_t \sim IN(0, \Sigma_2) \tag{3.6}$$

where ϖ_t is the disturbance, $\Phi(L)$ indicates all coefficient matrices, μ denotes the expectation of corresponding variable. $\mu_{y_{t,1}} = 3\mu_{y_{t,1}^*}$, due to the association between monthly

[1] The operation detail see Marcellino and Schumacher (2010) and Giannone et al. (2005)

and quarterly GDP[2]. The VARMA Model could be reduced to VAR model by setting $\Theta(L) = I_2$.

Let $y_t = [y'_{t,1}, y'_{t,2}]'$ to indicate the mixture of quarterly growth rate of GDP and extracted single factor. y^*_t and y_t is associated by the equation:

$$y_t - \mu_{y_t} = \left(\begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} L + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L^2 + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} L^3 + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L^4 + \begin{bmatrix} 1/3 & 0 \\ 0 & 0 \end{bmatrix} L^5 \right) (y^*_t - \mu_{y^*_t}) \tag{3.7}$$

3.3.4 Algorithm

In this section, we elaborate the algorithm. The estimation of the parameters and monthly GDP growth rate $y^*_{t,1}$ is realized by Kalman filter and maximum likelihood estimation [3]. Beforehand, the mix frequency model should be casted in a state space form.

State space representation

For MFVAR model, the state space representation is in line with Mariano and Murasawa (2010). Let p*=max(p, 5), z_t is the disturbance, $z_t \sim IN(0, I_2)$, we define the state vector as:

$$s_t := \begin{bmatrix} y^*_t - \mu_{y^*} \\ y^*_{t-1} - \mu_{y^*} \\ \dots \\ y^*_{t-p*+1} - \mu_{y^*} \end{bmatrix} \tag{3.8}$$

The state space representation is:

$$s_{t+1} = As_t + Bz_t, \tag{3.9}$$

$$y_t - \mu_{y_t} = C_t s_t, \tag{3.10}$$

[2]For simplifying estimation, we initially use the demeaned series, and afterwards add mean back to get final result.

[3]For VAR model expectation maximization could be easily applied. However, for VARMA model, it is hard to impose restriction on parameters. To make sure the comparability of disparate models, we apply MLE for both model.

where

$$A := \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_p & O_{2\times 2(p*-5)} \\ I_{2(p*-1)} & & & O_{2(p*-1)\times 2} & O_{2\times 2(p*-5)} \end{bmatrix}, \quad B := \begin{bmatrix} \Sigma^{1/2} \\ O_{2(p*-1)\times 2} \end{bmatrix} \tag{3.11}$$

$$C_t := \begin{pmatrix} 1/3 & 0 & 2/3 & 0 & 1 & 0 & 2/3 & 0 & 1/3 & 0 & 0_{1\times(p-10)} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0_{1\times(p-10)} \end{pmatrix} \tag{3.12}$$

With respect to MFVARMA model, the structure is more strenuous, see Harvey and Pierse (1984). The state vector and the matrices in state space form are modified as follows:

$$s_t := \begin{bmatrix} y_t^* - \mu_{y^*} \\ \Sigma_{i=2}^m \Phi_i(y_{t-i+1}^* - \mu_{y^*}) + \Sigma_{i=2}^m \Theta_{i-1} z_{t-i+2} \\ \\ \dots \\ \Phi_m(y_{t-1}^* - \mu_{y^*}) + \Theta_{m-1} z_t \\ y_{t-1}^* - \mu_{y^*} \\ \dots \\ y_{t-4}^* - \mu_{y^*} \end{bmatrix}, A := \begin{bmatrix} \Phi_1 & I_2 & O_2 & ... & O_2 & O_2 & O_2 & O_2 & O_2 \\ \Phi_2 & O_2 & I_2 & ... & O_2 & O_2 & O_2 & O_2 & O_2 \\ . & . & . & . & . & . & . & . & . \\ \Phi_m & O_2 & ... & & O_2 & O_2 & O_2 & O_2 & O_2 \\ I_2 & O_2 & ... & & O_2 & O_2 & O_2 & O_2 & O_2 \\ O_2 & O_2 & ... & & O_2 & I_2 & O_2 & O_2 & O_2 \\ O_2 & O_2 & ... & & O_2 & O_2 & I_2 & O_2 & O_2 \\ O_2 & O_2 & ... & & O_2 & O_2 & O_2 & I_2 & O_2 \end{bmatrix}, \tag{3.13}$$

$$B := \begin{bmatrix} I_2 \\ \Theta \\ O_{8\times 2} \end{bmatrix} \Sigma_2^{1/2}, \quad C_t := \begin{pmatrix} 1/3 & 0 & 0 & \dots & 0 & 2/3 & 0 & 1 & 0 & 2/3 & 0 & 1/3 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.14}$$

where $z_t \sim (0, I_2), m = max(p,\ q+1), \Theta = (\Theta_1, \dots, \Theta_{m-1})'$.

Missing observation

Aware of the fact that $y_{t,1}$ is only available on the third month of the quarter, we fill the missing observations in month 1 and 2 with 0. Referring to Durbin and Koopman (2001) Ch 4.8, we modify the measurement equation to accommodate the missing observation treatment. For the month when $y_{t,1}$ is unavailable, the upper row of the matrix in C_t is all equal to zero. Subsequently, the Kalman filter and smoother would simply skip the missing observation.

Identification

The parameters for estimation includes the coefficients for autoregressive component $\Phi(L)$, covariance component Σ_2 and coefficients for moving average component $\Theta(L)$ in the case of VARMA model. We execute maximum likelihood estimation accompanied by Quasi-Newton algorithm, and update the Hessian Matrix by Broyden, Fletcher, Goldfarb, and Shanno (bfgs) approach.

Specific constraint for parameters need be emphasized. First and foremost, either the VAR or VARMA model should be stationary according to our methodology. For this reason, the parameters in $\Phi(L)$ should satisfy that the eigenvalues are all less than 1. Instead of imposing this constraint on MLE process, we estimate the parameters without constraint and investigate whether $\Phi(L)$ satisfies the stationary condition in the end. Besides, we assume the covariance matrix Σ_2 is positive definitive, which rules out the identities in data. This condition is satisfied if Σ_2 could be Cholesky decomposed, namely $\Sigma_2 = L * L'$, L is a lower triangular matrix, and the diagonal element of L is strictly positive.

3.4 Empirical Result

3.4.1 Data Set

The performance of our FA-MFVAR/MA model is judged by estimating monthly real GDP of Germany. The selection of monthly variables refers to Stock and Watson (2002b) for the US data, Proietti (2008) and Banbura et al. (2010) for the euro area. Only balanced data is taken account for, consequently there are 91 monthly series available[4]. They are classified into several categories: industrial production (index of industrial, turnover and capacity utilization rate), labor market (employment,unemployment and wage), construction, trade, orders, monetary and financial indicators, price index, external trade, survey and cyclical index. Data spans from 1991Q1 to 2010Q4 [5].

We take priority to seasonal adjusted series, and apply ARIMA-X12 seasonal adjustment (Eviews) for series without seasonal adjustment [6]. To reach the requirement of factor model, we treat monthly indicators as follows. First of all, ξ_t should be stationary series. Therefore we examine the stationarity by augmented Dicky-Fuller test. The first difference

[4]The two step method is fairly convenient for ragged-edge data. Nevertheless, we believe balanced data is better for estimating monthly GDP.

[5]Detailed in the appendix 3.7.2.

[6]This seasonal adjustment depends on the trend of each series, we make this adjustment only for the series having discernible seasonal element.

or log difference transformation is adopted for non-stationary data. Subsequently, all series are demeaned, standardized and screened for outliers [7].

Figure 3.1 manifests that the distinction of single factor series estimated from Stock and Watson (short for PCA) and two step (short for TS) method is tiny. Due to the stagnation in 2008, there is an abrupt incline in the single factor, akin to the sudden decline in the quarterly GDP growth several months later. The extracted factor appears to a quick signalling of growth slowdown, in contrast with the quarterly GDP growth. It is not hard to image, since the single factor is extracted from monthly indicators, which are regarded as more timely data. Moreover, the volatility bands of the single factor and quarterly GDP growth are essentially identical, mainly between ± 0.05, except the break due to financial crisis.

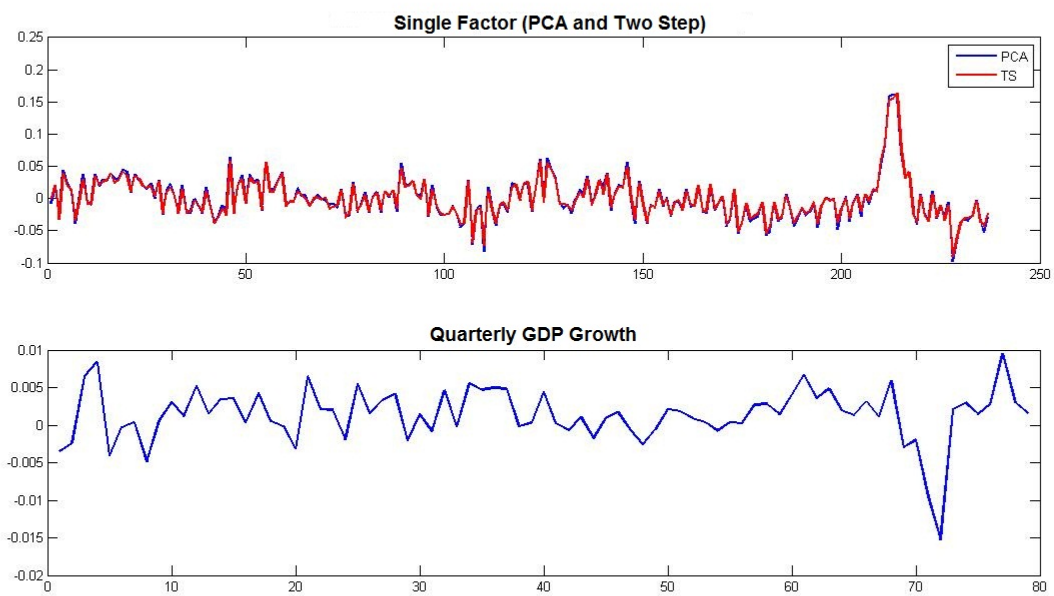


Figure 3.1: Monthly Single Factor VS Quarterly GDP growth

The PCA single factor could explain almost 20% variance of full data set. More concretely, the constitute of single factor is under scrutiny. Figure 3.2 depicts the association between the single factor and individual monthly indicator. The first two bars describe the R2 and correlation coefficient from OLS regression between factor and individual series, while the third bar plots the loading of each series. We can perceive that none of series

[7] Outliers are defined as the observation exceeding 10 times the interquartile range from the median, and they were replaced by the previous one period's value. See Stock and Watson (2002a) for detail.

extremely dominates the single factor, whereas the importance of industrial production and orders series is notable, afterwards is the price and survey data[8].

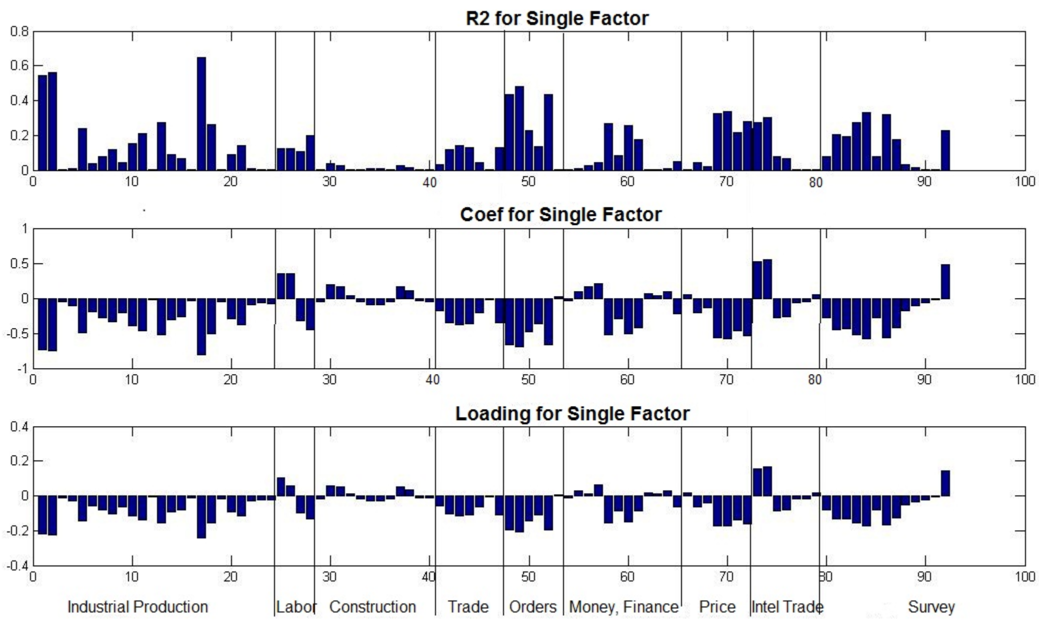


Figure 3.2: PCA Single Factor's Correlation with Individual Series

3.4.2 Model Specification and Evaluation

For the sake of performance evaluation, we consider three methodologies as counterpart. Firstly, Chow and Lin method is regarded as a benchmark model, they introduced how related series can be used to interpolate with a GLS estimator. This approach is based on the assumption that high frequency observations of the series could be expressed as a multiple regression relationship with several high frequency related series. For this method, stationary series is not necessary. Hereby, we measure the level of monthly GDP, and then transform to the growth rate for comparison. The four related series we employ are industrial production, total retail, number of employed person and ifo business climate index.

Secondly, to investigate the implication of factor extraction, we parallel the result of our methodology with single indicator MFVAR model, by exploiting industrial production and ifo business climate index respectively.

[8]For the two step method, only the R2 and correlation coefficients can be plotted, result is quite closed to PCA's, hence the graph is skipped.

Thirdly, Angelini et al. (2008)'s dynamic factor model need be emphasized, as it is enable inclusion of the same large datasets, but distinct from the structure of factor and underlying monthly GDP. They supposed that the monthly growth rate y_t^* is conjectured with the common factor by static equation 3.15, where β is the coefficient matrix for factor. Combined with equation 3.5, the estimation and forecasting can be reckoned in state space form as well[9].

$$y_t^* = \beta f_t + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma_\epsilon) \tag{3.15}$$

Predictive ability exerts prominent function in evaluation of econometric models, except the Chow and Lin approach[10]. The out of sample RMSE is acknowledged as our principal criterion. We split the full sample into two parts, leaving 2 quarterly GDP data for out of sample test[11], the remainder is for in sample estimation. Initially, either the VAR or VARMA model is applied for in sample data to measure parameters. The lag of order is determined by information criteria (AIC, BIC) and the specific performance of estimation. Subsequently, these parameters estimated are used for forecasting. The first period of Kalman filter forecasting is:

$$E(s_{T+1} \mid \Omega_T, s_T) = AE(s_T \mid s_T) \tag{3.16}$$

$$E(y_{T+1} - \mu_{y_t} \mid \Omega_T) = C_{t+1} E(s_{T+1} \mid \Omega_T) \tag{3.17}$$

where $\Omega_T = (y_1, y_2, , , , y_T)$, T is the end period of in sample data[12].

The rest of forecasting could be implemented by recursively applying Kalman filter forecasting. Posterior to obtaining a quarterly forecasts of GDP growth, the out of sample RMSE between quarterly forecasts and real GDP growth data are computed to assess the relative performance of diverse models.

$$RMSE = \sqrt{\frac{1}{n}[\sum_{i=1}^{n}(y_{T+i} - E(y_{T+i} \mid \Omega_T))^2]} \tag{3.18}$$

[9] Model detail see Angelini et al. (2008).

[10] Chow and Lin approach could not extend to forecasting so convenient as the rest models, the comparison with Chow and Lin approach mainly relies on graphical analysis.

[11] We only consider two period for forecast due to two reason. On one side, Kuzin et al. (2011) stated this sophisticated model benefit for short run forecasting. On the other side, the performance of interpolation should be more associated with short-run forecasting.

[12] Notice μ_{y_t} is the mean for in sample part's quarterly observation, not full sample.

Table 3.1: Model Evaluation by RMSE

Model	Monthly Series	RMSE	Model	Monthly Series	RMSE
MFVAR(1)	PCA	1.492e-4	MFVARMA(1,1)	PCA	3.615e-4
	TS	1.824e-4		TS	9.537e-4
	IP	6.900e-4		IP	1.157e-3
	IFO	2.524e-3		IFO	1.461e-3
Dynamic Factor Model		2.400e-3			

Note: PCA represents single factor estimated by Stock and Watson's method, TS represents single factor estimated by two step approach, IP means industrial production as the monthly indicator for interpolation, IFO is the IFO business climate index. Full sample period is from 1991Q2 to 2010Q4.

where y_{T+i} is the realization of quarterly GDP growth during out of sample period, $E(y_{T+i} \mid \Omega_T)$ is predicted GDP growth based on information until time T. n is the length of quarterly data for out of sample evaluation.

3.4.3 Result

For the VAR model, we select the lag of order from 1 to 4, and chose VAR(1) ultimately. And for VARMA model, due to the difficulty of identification, we only consider VARMA(1,1) model. Table 3.1 reports the one step ahead RMSE for MFVAR(1) and MFVARMA(1,1), with the monthly indicator choose among factor analysis (PCA and TS) and single series (industrial production, ifo business climate index), and dynamic factor model. The first four lines on the left side distinguish among the selection of monthly indicator of MFVAR model. Parallelling the factor estimated by principal component and two step approach, PCA factor outperforms the TS factor slightly. This is in conformity with recently research finding of factor extraction, the parsimonious and standard principal component method generates the optimal factors. To benchmark MFVAR model between single factor and monthly indicator, it is conspicuous that single factor's RMSE is smaller than monthly indicator's. The statement stands also for FA-MFVARMA's result. More specifically, when we benchmark among disparate approaches, the FA-MFVAR model yields the lowest RMSE, FA-MFVARMA model ranks second, while the dynamic factor model produce much worse result.

We presents the estimation results of FA-MFVAR and FA-MFVARMA model for PCA factor in Table 3.2. Majority of parameters are significant at the level of 1% and coefficient matrix and residual covariance matrix are all plausible for FA-MFVAR model. Whilst the

Table 3.2: FA-MFVAR/MA Estimation Result for PCA Factor

		FA-MFVAR		FA-MFVARMA	
Parameter	Φ	-0,7398*	-0,8027*	0,4822*	0,0009
Matrices		-0,0343*	0,4742*	-6,6518*	0,5784*
	Σ	9,33e-06	-3,89e-05	3,05e-06	7,43e-06
		-3,89e-05	8,22e-04	7,43e-06	6,80e-04
	Θ	.	.	-0,3881*	-0,0464*
				3,4864*	-0,3078*
AR Roots		-0,7621	0,4965	0,5303± 0,0580i	
MA Roots		.	.	0,3480± 0,4003i	
Log Likelihood		700,858		718,193	

Note: * denotes the corresponding coefficient is significant at the level of 1%. Sample Period:1991Q2-2010Q4.

moving average component is included, the parameters estimation seems problematic, and causes unreasonable large coefficient for VARMA(1,1). The AR characteristic roots are all inside the unit root, implies that the estimated monthly GDP growth series is stationary.

After carefully description of model estimation and RMSE, we plot the quarterly and estimated monthly GDP growth from PCA factor in Figure 3.3. The trend of monthly and quarterly GDP growth is consistent, except Chow and Lin approach. One explanation is that Chow and Lin approach is based on the level data, while the rest models consider growth rate data. Furthermore, the growth rate estimation from FA-MFVAR is much more volatile than the remainder models. Whereas, it is hard to recognize the optimal model only from the growth plot. To provide more intuitive evidence of graphical analysis, we transform the monthly GDP growth back to level series. The monthly GDP index is achieved by making the cumulative sum of the growth rate data and taking the exponential transformation afterwards. Figure 3.4 delineates the monthly GDP index estimated by four models, which displays the huge distinction among various models. From the aspect of tracking the movement of quarterly GDP, FA-MFVAR model exhibits an unbelievable excellent performance, unilaterally beats all other models. FA-MFVARMA model still ranks the second. Although the growth rate plot of Chow and Lin method is not ideal, it hits the dynamic factor model from Figure 3.4. The latter completely fails to trace the trend of quarterly GDP.

To summarize, both RMSE comparison and graphical analysis verify the goodness of FA-MFVAR model. FA-MFVAR outperforms the single indicator MFVAR and Chow and Lin method, implies that the construction of GDP based monthly indicator benefits

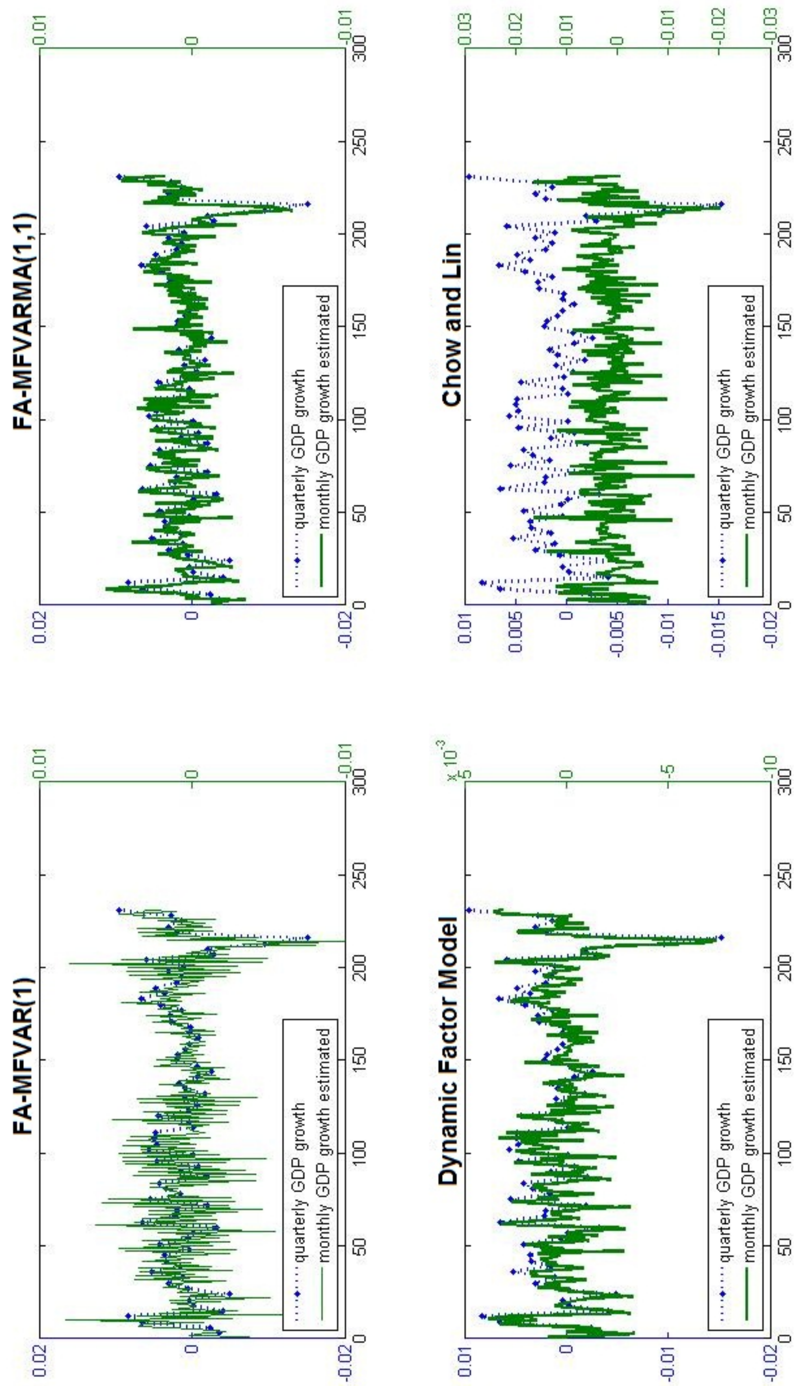


Figure 3.3: GDP Growth Comparison

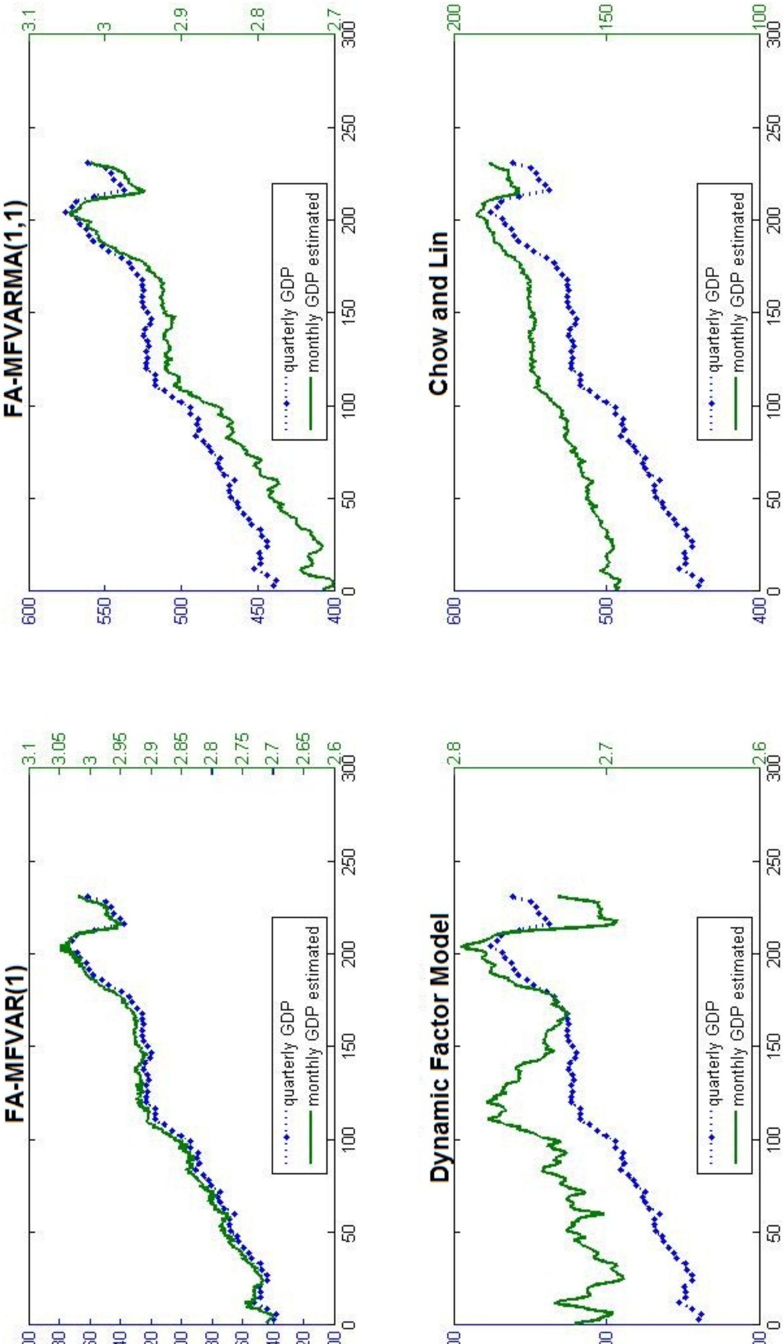


Figure 3.4: GDP Comparison

from large information set. Whereas, the unfavorable dynamic factor estimate reveals the linchpin of appropriate model setup. Large dynamic factor model earns good reputation of short run forecasting, but it does not guarantee its fitness for GDP interpolation. Last but not least, adding the moving average component fuels concern about strenuous parameter estimation, and does not improve the performance. It is better to constraint to VAR model.

3.5 Conclusion

In closing, we design a sophisticated mechanism of generating GDP based monthly coincident index in this chapter, which successfully combines a large scale monthly indicators with quarterly GDP data, incorporates factor analysis with traditional vector autoregressive model. We could not promise our method is the optimal for estimating monthly GDP, as the performance indeed relies on the specific data set. Whereas, our estimator contains large information set and could be deemed to a favorable GDP based monthly coincident index.

Furthermore, the application of German data exhibits substantial advantage over the remaining methodologies, i.e. single indicator MFVAR model, Chow and Lin approach, and dynamic factor model, through the diagram analysis and out of sample RMSE comparison. It implies that large dataset might improve the performance of index construction, but it relies on an appropriate model setup. In addition, adding the moving average term may jeopardize the accuracy of parameter estimation, leads to a deterioration of performance.

One crucial issue for estimating monthly GDP is the difficulty to evaluate the performance of diverse method. Up to now, the evaluation relies on out of sample RMSE. Nonetheless, it virtually is not one criterion for estimation, but for forecasting. In future, it could be splendid if an adequate approach is devised for evaluation performance of disaggregation series.

3.6 References

Angelini, E., Canba-Mendez, G., Giannone, D., Reichlin L. and Rünstler, G., 2011. Short-term Forecasts of Euro Area GDP Growth. *The Econometrics Journal* 14, 25-44.

Angelini, E., Bańbura, M., Rünstler, G., 2008. Estimating and Forecasting Euro Area Monthly National Accounts from a Dynamic Factor Model. *ECB Working Paper Series*, No 953.

Bańbura, M., Giannone, D., Reichlin, L., 2011. Nowcasting. In *The Oxford Handbook* *of Economic Forecasting*, Clements M. and Hendry D (eds). Oxford University Press: Oxford, 193-224.

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 and Economic Statistics* 25(1), 52-60.

Bernanke, B., Gertler, M., Watson, M., 1997. Systematic Monetary Policy and the Effects of Oil Price Shocks. *Brookings Papers on Economic Activity*, 91-142.

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

Carriero, A., Marcellino, M., 2007. Monitoring the Economy of the Euro Area: A Comparison of Composite Coincident Indexes. *Bocconi University Working Papers Series*, No 319.

Chow, G.C., Lin, A., 1971. Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series. *Review of Economics and Statistics* 53, 372-375.

Cuche, N., Hess, M., 2000. Estimating Monthly GDP in a General Kalman Filter Framework : Evidence from Switzerland. *Economics and Financial Modeling*, Winter 2000.

De Jong, P., 1989. Smoothing and Interpolation with the State-Space Model. *Journal of the American Statistical Association* 84(408), 1085-1088.

Doz, C., Giannone, D., Reichlin, L., 2011. A two step estimator for large dynamic factor models based on Kalman Filtering. *Journal of Econometrics* 164, 188-205.

Fernández, R.B., 1981. A methodological note on the estimation of time series. *Review of Economics and Statistics* 63, 471-476.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2000. The Generalized Factor Model: Identification and Estimation. *Review of Economics and Statistics* 82, 540-554.

Frale, C., Marcellino, M., Mazzi, G., Proietti, T., 2011. EUROMIND: a monthly indicator of euro area economic conditions. *Journal of the Royal Statistical Society A* 74(2), 439-470.

Giannone, D., Reichlin, L., Sala, L., 2005. Monetary policy in real time. *NBER Macroeconomics Annual* 19, 161-224.

Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: the Real Time Informational Content of Macroeconomic Data Releases. *Journal of Monetary Economics* 55, 665-676.

Hamilton, D.J., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57(2), 357-384.

Harvey, A.C., Pierse, R.D., 1984. Estimating Missing Observations in Economic Time Series. *Journal of the American Statistical Association* 79, 125-131.

Harvey, A.C., 1989. Forecasting, Structural Time Series Models and the Kalman Filter. *Cambridge University Press*.

Harvey, A.C., Chung, C.H., 2000. Estimating the underlying change in unemployment in the UK. *Journal of Royal Statistic Society A* 163,303-339.

Litterman, R.B., 1983. A random walk, Markov model for the distribution of time series. *Journal of Business and Economic Statistics* 1(2), 169-173.

Liu.H., Hall., S.G., 2001. Creating high-frequency national accounts with state-space modeling: a Monte Carlo experiment. *Journal of Forecasting* 20, 441-449.

Kuzin, V., Marcellino, M., Schumacher C., 2011. MIDAS vs. mixed-frequency VAR: nowcasting GDP in the euro area. *International Journal of Forecasting* 27, 529-542.

Marcellino, M., Schumacher C, 2010. Factor MIDAS for nowcasting and forecasting with ragged-edge data: a model comparison for German GDP. *Oxford Bulletin of Economics and Statistics* 72(4), 518-550.

Mariano, R.S., Murasawa,Y., 2003. A New coincident Index of Business Cycles Based on Monthly and Quarterly Series. *Journal of Applied Econometrics* 18, 427-443.

Mariano, R.S., Murasawa, Y., 2010. A Coincident Index, Common Factors, and Monthly Real GDP. *Oxford Bulletin of Economics and Statistics.* 71(1), 27-48.

Mitchell, J., Weale, M., 2005. Estimates of monthly GDP for Euro Area: an Application of Regression Approach to Interpolation. *European Communities Working Paper*.

Proietti, M., 2008. Estimation of Common Factors under Cross-Sectional and Temporal Aggregation Constraints: Nowcasting Monthly GDP and its Main Components. *Working Paper*.

Proietti, M., Moauro, F., 2006. Dynamic factor analysis with non-linear temporal aggregation constraints. *Journal of the Royal Statistical Society C* 55(2), 281-300.

Santo, S., and Cardoso, F.N., 2001. The Chow-Lin method using dynamic models. *Economic Modeling* 18(2), 269-280.

Schumacher, C., Breitung, J., 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting* 24, 368-398.

Wei, W.S., Stram, D.O., 1990. Disaggregation of Time Series Models. *Journal od the Royal Statistical Society.Series B* 52(3), 453-467.

Stock,J., Watson, M., 1991. A Probability Model of the Coincident Economic Indicators In Leading Economic Indicators (eds K. Lahiri and G. H. Moore). *Cambridge University Press*.

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

Stock,J., Watson,M., 2002a. Forecasting Using Principal Components From a Large Number of Predictors.*Journal of the American Statistical Association* 97, 1167-1179.

3.7 Appendix

3.7.1 Kalman Filter and Smooth

Kalman Filter Iteration

Initial state:

$$s_{1|0} = 0$$

$$P_{1|0} = B * B'$$

where $P_{t|t-1} = E((s_{t|t-1} - E(s_{t|t-1}))(s_{t|t-1} - E(s_{t|t-1}))', s_{t|t-1} = E(s_t|s_{t-1})$

Update equation:

$$s_{t|t-1} = A * s_{t-1|t-1}$$

$$P_{t|t-1} = A * P_{t-1|t-1} * A' + BB'$$

Prediction equation:

$$y^+_{t|t-1} = C_t s_{t|t-1}$$

Prediction error:

$$v_t = y^+_t - y^+_{t|t-1}$$

Variance of y^+_t:

$$D_t = C_t P_{t|t-1} C'_t$$

Kalman filter:

$$K_t = P_{t|t-1} C'_t D_t^{-1}$$

Adjust prediction by Kalman filter:

$$s_{t|t} = s_{t|t-1} + K_t v_t$$

$$P_{t|t} = P_{t|t-1}(I - K_t C_t P_{t|t-1})$$

Log likelihood function:

$$logL = -T * log(2\pi) - \frac{1}{2}log(|D_t|) - \frac{1}{2}v_t D_t^{-1} v_t'$$

One problem is that we need the inverse of D_t for Kalman filter calculation. As D_t is a 2*2 matrix, there will be situation that the inverse does not exist. In that case, we take the Moore-Penrose pseudoinverse if D_t is not nonsingular.

Fixed-iterative Smoothing

In order to obtain expectation and variance of s_t based on whole information set from time period 1 to T, instead of only from 1 to t-1, fixed iterative smoothing method suggested by de Jong(1989) is utilized. Assume $r_{T+1} = 0, R_{T+1} = O$, and for t=T,,,1,

$$r_t := C_t' D_t^{-1} v_t + L_t' r_{t+1}$$

$$R_t := C_t' D_t^{-1} C_t + L_t' R_{t+1} L_t$$

where

$$L_t := A(I - K_t C_t)$$

The smoothing equations are for t=1,...,T,

$$s_{t|T} = A s_{t|t-1} + P_{t|t-1} r_t$$

$$P_{t|T} = P_{t|t-1} - P_{t|t-1} R_t P_{t|t-1}$$

3.7.2 Monthly Indicator Selection

Table 3.3: Database description

Series	Description	Transform type
Industrial Production		
IP including construction	Volume index,sa	2
IP manufacturing	Volume index,sa	2
IP: mining and quarrying	Volume index,sa	2
IP: manufacture of food,beverages and tobacco	Volume index,sa	2
IP: textile and apparel	Volume index,sa	2
IP: leather goods	Volume index,sa	2
IP: wood and wood prod	Volume index,sa	0
IP: pulp,paper and paper prods,publishing and printing	Volume index,sa	2
IP: coke,refined petroleum prod and nuclear fuel	Volume index,sa	0
IP: chemical prod	Volume index,sa	2
IP: rubber and plastic prod	Volume index,sa	2
IP: other non-metallic mineral prod	Volume index,sa	0
IP: computer,electronic and optical prod	Volume index,sa	2
IP: furniture, jewelery, musical instrument N.E.C.	Volume index,sa	2
IP: electricity, gas ,steam and air conditioning supply	Volume index,sa	0
IP: extraction of crude petroleum and natural gas	Volume index,sa	0
IND. Total Turnover: intermediate goods	Price index,sa	2
IND. Total Turnover: capital goods	Price index,sa	2
IND. Total Turnover: durable cons.goods	Price index,sa	0
IND. Total Turnover: non-durable cons.goods	Price index,sa	2
IND. Total Turnover: cons.goods	Price index,sa	2
Civil Engineering : capacity utilization rate	%,sa	2
CNSTR.IND.: capacity utilization rate	%, sa	0
BLDG.CNSTR.: capacity utilization rate	%, sa	2
Labor Market		
Unemployment rate(pan from sept 1990)	Volume,na	2,X12
Unemployment : % civilian labor	sa	2
Employed persons(work place concept)	na	2
Vacancies(pan from M0790)	Volume,na,	2,X12
Construction		
Housing permits issued for BLDG.CNSTR, new	Volume,na	2,X12
Building permits granted: non-resid,indl. CNSTR	Current prices,sa	0
Construction permits granted-non-resid	Current prices, sa	0
Construction permits granted-new homes	sa	2

Continued on next page. . .

Table 3.3 – Continued

Series	Description	Transform type
Construction orders received	sa	2
Construction orders received: buildings	sa	2
Construction orders received: residential buildings	sa	2
Construction orders received: industrial buildings	sa	2
Construction orders received: public buildings	sa	0
Construction orders received: civil engineering	sa	0
Construction orders received: industrial clients	sa	2
Employment in construction	sa	0
Trade		
Total retail trade(value)	Price index,sa	0
Total wholesale trade(value)	Price index,na	2,X12
Sales of manufactured durable consumer goods	Volume index,na	2,X12
Sales of manufactured intermediate goods	Volume index,na	2,X12
Sales of manufactured investment goods	Volume index,na	0,X12
Sales of manufactured non-durable consumer goods	Volume index,na	0,X12
Sales of total manufactured goods	Volume index,na	2,X12
Orders		
New orders to manufacturing	price index,sa	2
New orders to manufacturing-intermediate goods	Price index,sa	2
New orders to manufacturing-capital goods	Price index,sa	2
New orders to manufacturing-consumer goods	Price index,sa	2
Manufacturing orders	Price index,sa	2
New registrations-cars	Volume,na	0,X12
Money and Finance		
Money supply M0	Current prices,na	2,X12
Money supply,germen contribution to euro M1	Current prices,na	2
Money supply M2	Current prices,na	2
Money supply M3	Current prices,na	2
FIBOR - 3 month(monthly avg)	%	2
Long term government bond yield(9-10 years)	%	2
Three-month money market rate-frankfurt	%	2
Day -to -day money market rate -frankfurt	%	2
Bank lending to entps. and indiv-median/long	Current prices,na	2
Bank lending to entps. and indiv-long/advanced	Current prices,na	2
Bank lending to government-long term loan	Current prices,na	2
DAX share price index	Price index,na	2

Continued on next page...

Table 3.3 – Continued

Series	Description	Transform type
Price Index		
CPI (excluding energy)	sa	2
CPI - food	sa	2
CPI	sa	2
export price index	sa	2
import price index	sa	2
PPI:indl. prod,total,sold on the domestic market	na	2
WPI	na	2,X12
International Trade		
terms of trade-on the basis of price indices	Price index,sa	2
terms of trade(pan frm 1991)	Price index,na	2
exports FOB(pan M0790)	Current prices,sa	2
imports CIF(pan M0790)	Current prices,sa	2
visible trade balance	Current prices,sa	2
German Mark to US Dollar	MTH.AVG.	2
US Dollar to 1 EURO	MTH.AVG	2
Survey and Cyclical Index		
Business expectations	Index, sa	2
Consumer confidence indicator	Index,sa	1
Consumer confidence index	sa	2
IFO business climate index	Index, sa	2
Assessment of business situation	Index, sa	2
Productivity: output per man-hour worked in industry	Index, sa	2
GFK consumer survey: economic situation next 12 mth	Index, sa	1
GFK consumer survey: financial situation next 12 mth	Index, sa	1
GFK consumer survey: major purchases over next 12 mth	Index, sa	0
GFK consumer survey: prices next 12 mth	Index, sa	1
GFK consumer survey: saving next 12 mth	Index, sa	0
GFK consumer survey: unemployment next 12 mth	Index, sa	1

Notes: 0 denotes the level data, 1 denote differentiated data, 2 denoted the log differentiated data, X12 denotes X12-ARIMA seasonal adjustment for corresponding data series. Data period from 1991Q1 to 2010Q4.

Curriculum Vitae

2008 - 2013	Ph.D. program in economics,
	Munich Graduate School of Economics,
	University of Munich, Germany
2003 - 2007	Bachelor of Economics,
	Major in Mathematical Finance,
	Zhejiang University, China
2001 - 2003	Hangzhou No 1 High School,
	Zhejiang Province, China
Nov 15, 1985	Born in Hangzhou, China