Three Essays on Econometric Analysis of the Chinese Macroeconomy

by **Qin Zhang**

Principal Supervisor: Dr. Natalia Ponomareva

Associate Supervisor: Dr. Christopher Heaton

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Economics

Macquarie University

September 14, 2020

Contents

List of Figures

List of Tables

List of Abbreviations

ADF	Augmented Dickey–Fuller
AE	Aggregate Expenditure
AIC	Akaike Information Criterion
APP	Agricultural Producer Index
AR	Autoregressive
ARMA	Autoregressive Moving Average
CNY	Chinese Yuan
CPC	Communist Party of China
CPI	Consumer Price Index
DFM	Dynamic Factor Model
DI	Diffusion Index
DRS	Direct Reporting System
EM	Expectation Maximisation
FA-AR	Factor Augmented Auto-regressive
FA-VAR	Factor Augmented Vector Auto-regressive
FFNN	Feed Forward Neutral Network
GDP	Gross Domestic Product
K-NNR	K-Nearest Neighbour Regression

LASSO	least absolute shrinkage and selection operator
ML	Machine Learning
MLD	Multivariate Leading Indicator
MSE	Mean squared error
NBSC	National Bureau Statistics of China
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PPI	Producer Price Index
RPI	Retail Price Index
SMPS	Soviet Material Product System
SOE	State-owned Enterprise
SVR	Support Vector Regression
VA	Value-added
VAR	Vector Auto-regressive

Summary

China has consistently achieved high and stable rates of economic growth since 1970s. China is now the second biggest economy in the world, and its goods and services are highly competitive in international markets. These renewed academic interest in understanding the country's macroeconomic policy-making. Macroeconomic forecasting and monitoring are established practices that are vitally important for real-world macroeconomic policy-making. Having this in mind, the literature on macroeconomic forecasting and monitoring in China is limited compared to the research in Western economies. As such, this thesis consists of three individual chapters that are aimed towards contributing to the empirical investigation of macroeconomic forecasting in China.

Chapter 2 aims to evaluate the quality of China's official macroeconomic statistics by examining three factors. The first is how political interference may affect the statistical reporting system. At local government level, the career incentives and the Chinese cadre evaluation system, alongside the geography-based governing logic, have motivated local officials to compete to influence the reported growth rate. At central government, the notion of independence of the National Bureau of Statistics of China is criticised because it has limited authority over the statistical divisions of other government institutions and provincial bureau of statistics. Second, it reviews the ongoing challenges of gathering, measuring and presenting economics, with focuses on incomplete survey data, issues with direct reporting systems and revisions of economic data. Third, it investigates the internal inconsistency by using quantitative methods to explain where discrepancies come from.

Chapter 3 studies forecasting Chinese macroeconomic variables using large-scale factor models with mixed-frequency data and missing observations component. The factor models are particular compatible with potential data contamination, rapid institutional

and structural change, which are prevalent in China. Using 251 monthly variables and 34 quarterly variables over the December 2001 to June 2018 period, we find statistical evidence that mixed-frequency factor models, especially mixed-frequency factor-augmented vector autoregressive models, generated superior forecasts to the univariate and multivariate models for price series, nominal investment and nominal consumption, except for the CPI inflation rate and nominal consumption at one month ahead. Therefore, the results of this chapter provide clear guidance and important implications for academics, practitioners and the public who are interested in macroeconomic forecasting in China.

Chapter 4 undertakes the task of nowcasting GDP for mainland China using machine learning algorithms. Using a large set of quarterly macroeconomic indicators and monthly indicators, we train eight popular machine learning algorithms and nowcast GDP growth for each quarter over the 1993Q1-2018Q2 period. We compare the predictive accuracy of these nowcasts with those of AR model and dynamic factor model in the state-space representation. We use the model confidence set to obtain a set of best model(s) with 10% level of confidence. Our results show that shrinkage methods are covered by the model confidence set and therefore are in the set of best models. As such, ML algorithms proved useful for improving the accuracy of nowcasting the Chinese GDP growth rate.

Overall, this thesis enriches our understanding of the quality of Chinese official macroeconomic data and guides practitioners toward selecting the appropriate forecasting and nowcasting models for China's economy in a data-rich environment and provides considerable scope for future research.

Declaration

I, Qin Zhang, declare that this thesis titled: 'Three Essays on Econometric Analysis of the Chinese Macroeconomy' and the work presented in it are my own. I confirm that:

- This work was done wholly while in candidature for a research degree at Macquarie University.
- This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.

Signed:

Date: September 14, 2020

Acknowledgements

This dissertation would not have been completed without the help and the support of many people. First and foremost, I would like to express my gratitude to my supervisors Dr. Natalia Ponomareva and Dr. Christopher Heaton who have stood by me since my first day. Their sincere comments and endless help enabled me to come this far. I truly appreciate what they taught me.

I am thankful to be given the opportunity to be a part of Department of Economics in Macquarie University. I am also grateful to my peers in the department. I could not have picked a better group to go through four years with. Their friendship, encouragement and passion for economics truly brought out the best in me.

I would also like to thank my thesis examiners who have offered valuable feedback and comments to produce a more-developed thesis – Associate Professor Nektarios Aslanidis, Associate Professor Dalibor Stevanovic and Dr Wenying Yao.

Lastly, I would like to thank my parents and wife. Even through my parents are in China, they have mentally encouraged me to do something that was truly challenging. I could not ask for more in a family and I am truly grateful to have such as fantastic and loving family. My wife Ya Wen is always by my side supporting me through the depressed and amused time.

Chapter 1

Introduction

China has consistently achieved high and stable rates of economic growth since the country reformed its economy and accepted foreign trade in the late 1970s. China has attained an annual gross domestic product (GDP) growth rate of approximately 10% in real terms, positioning the economy as the second largest in the world. Moreover, the country raised more than 850 million people out of poverty.[1] Other countries are becoming more exposed to China's economy. International investors flock to Chinese markets to invest in stocks, bonds and real estate. China is a large importer of foreign goods and services, and its goods and services are highly competitive in international markets. During the aftermath of the 2008 Global Financial Crisis, China's GDP growth slowed to below double-digits and the country's focus moved to countercyclical government policy. The government injected four trillion Chinese yuan into investment to combat the decrease in output growth. Even if growth does stabilise, China is likely to become the world's largest economy by 2030. However, the country's per capita income would still be a fraction of the average in advanced economies.

China's incredible economic performance renewed academic interest in understanding the country's macroeconomic policy-making. Macroeconomic forecasting and monitoring are established practices that are vitally important for real-world macroeconomic policy-making. Policy-makers, central banks, the public and academics are interested in producing accurate forecasts for various reasons. For example, forecasting inflation is fundamen-

[1]Sources: Development Research Center of the State Council & The World Bank. (2013). China 2030: Building a Modern, Harmonious, and Creative Society. World Bank Publications.

tal in monetary policy because inflation expectations affect the central banks' decisions regarding future monetary policy and, consequently, the private sector's consumption and investment decisions.[2] Considering this, the literature on macroeconomic forecasting and monitoring in China is limited compared to the research in Western economies. As such, this thesis devotes three chapters to the empirical investigation of macroeconomic forecasting in China.

Macroeconomic forecasting in emerging countries must manage problems regarding the quality of official economic data. There is a longstanding debate among academics over the reliability of China's official GDP figures. Rawski (2001), Holz (2014, 2008, 2004, 2003), Hsien & Song (2019), Xiong (2019), Ma et al. (2014), Koch-Weser (2013) believed that the official GDP data are not reliable, whereas Fernald et al. (2015), Kerola (2019), Fernald et al. (2013) asserted that the official GDP data can approximately reflect actual economic growth. This debate has revealed a range of alternative indicators for assessing the quality of GDP data, including the Li Keqiang index, Fernald et al. (2015)'s trade partner index, Barclay's index, Bloomberg economic indices and Capital Economic indices. Based on these indicators, Chinese GDP annual growth from 2012 to 2017 is estimated to be between 3% and just over 10%. These large data discrepancies can distort assessments of the economic situation, prompting inappropriate economic policies and business decisions. Moreover, understanding the nature of data quality is vital for producing accurate macroeconomic forecasts. If data contamination is a severe problem, forecasting models may be misspecified, and forecasting results may be biased.

In response to this issue, Chapter 2 aims to evaluate the quality of China's official macroeconomic statistics by examining three factors. The first is how political interference may affect the statistical reporting system. At the central government level, the

[2]By informing the public about the likely inflation trends, forecasters can influence expectations, therefore serving as a nominal anchor (e.g., in the wage bargaining process or for other nominally fixed contracts like housing rents or interest rates).

independence of the National Bureau of Statistics of China (NBSC) is criticised because commissioners and broad members are appointed by the State Council. At the local government level, the evaluations and promotions of officials are strongly linked to local economic performance, which incites the officials to manipulate output data. Second, the ongoing challenges of gathering, measuring and presenting economic data are reviewed, focusing on incomplete survey data, issues with direct reporting systems (DRS) and revisions of economic data. Third, several internal inconsistencies are examined. The difference between results calculated using the value-added (VA) approach and those that used the aggregate expenditure (AE) approach became severe in recent years. Provincial governments tend to over-report output level, especially industrial production and gross capital formation, and the NBSC can influence the real GDP growth rate through its choice of implicit deflator for each sector and through revising the nominal values of previous years. This chapter contributes to literature by providing a comprehensive review for scholars, analysts and politicians who are interested in the quality of China's official economic data and using quantitative methods to find the origin of these discrepancies.

Recent advances in information technology have allowed economic datasets to reach a tremendous size in terms of the number of variables and the number of observations. However, the performance of standard forecasting models such as vector autoregressive (VAR) models tends to deteriorate as the number of time observations and the number of variables increases, which is the well-known curse of dimensionality. Another issue associated with forecasting in a data-rich environment is incomplete statistical information. Several key economic variables are released at different frequencies with considerable lags. Therefore, macroeconomic forecasters must design computationally efficient forecasting models that can transform large datasets into concise information without needing to discard predictors with missing observations and different sampling frequencies. In the past twenty years, the literature proposed two main methods for overcoming the curse of dimensionality. In the Bayesian domain, Bańbura et al. (2010), Koop (2013), Carriero et al. (2015), Giannone et al. (2015) have contributed to theoretical developments in large Bayesian VAR models that handle larger number of predictors than usual VAR models. In the frequentist domain, factor models by Stock & Watson (2002*b*) have been used to include an arbitrary amount of series in the forecasting system, which led to extensive improvements in macroeconomics forecasting for Western economies (e.g., Gupta & Kabundi (2011) for forecasting macroeconomic variables in South Africa, Artis et al. (2005) for inflation forecasting in the United Kingdom, Schumacher (2007), Schumacher (2007) for forecasting German GDP and Stock & Watson (2002*b*) for macroeconomic forecasting in the United States).

Compared to Western countries, there is little research on the challenge of macroeconomic forecasting in China in a data-rich environment. Chapter 3 considers forecasting Chinese macroeconomic variables using large-scale factor models with mixed-frequency data and missing observations. Stock & Watson (2011) showed that the factor models are compatible with potential data contamination and rapid institutional and structural change, which are prevalent in China. To investigate whether the factor models for forecasting Western macroeconomics could be reliable for China, the speed of the large set of traditional models is compared, as is the speed of the large dimensional approximate factor models. The set of factor models includes the diffusion index (DI) of Stock & Watson (2002*b*), factor-augmented autoregression (FA-AR) models and factor-augmented vector autoregression (FA-VAR) models. In contrast, the set of univariate and multivariate models contains the mean model, autoregression (AR) model, autoregression moving average (ARMA) model, vector autoregression (VAR) model and multivariate leading indicator (MLD). The principal component analysis for the balanced panel data and the expectation maximisation (EM) algorithm for the mixed-frequency data with missing observations are used to estimate the factors. We forecast two measures of price inflation and three measures of real economic activity. The measures for price inflation are consumer price index (CPI) inflation rate and retail price index (RPI) inflation rate, while the measures of real economic activity are nominal investment, nominal consumption and railway cargo, with the forecast horizons ranging from one to twelve months from December 2001 to June 2018. Forecasts are produced for one-month-, three-months-, six-months-, nine-months- and twelve-months-ahead and the forecasting performances are compared by the relative mean squared forecasting errors, with the benchmark model being the AR(p) model.[3]

We find evidence that mixed-frequency factor models, especially mixed-frequency FA-VAR models, provide more accurate forecasts than the benchmark model for inflation price series, nominal investment and nominal consumption, but not for railway cargo. The forecasting results for the global financial crisis period are markedly different. In most cases, the AR(p) model provides the most accurate forecasts for real economic activity. For inflation, only the mixed-frequency FA-VAR(p) model provides accurate forecasts for the CPI at one-month-ahead and for the RPI at one-month, three-months and six-months ahead. The main contributions to the literature are twofold. First, we include the factor models based on a very large number of predictors with mixed frequency and missing observation components, as well as factor models based on preselected targeted variables. The number of predictors used in other studies for China is smaller-usually around 40 variables or less. Second, we run a sufficiently large number of training sets and validation sets, whereas other studies normally have 5 to 60 observations for the validation set.

Chapter 4 considers macroeconomic monitoring in China. Due to the economic turmoil affecting the world economy in 2008 financial crisis, there has been an urgent need for the early and accurate assessment of the economic activity in real time for the benefit of academics and practitioners. Macroeconomic monitoring models have been used

[3]The AR(p) model is the autoregression model with lag order that is selected according to Akaike information criteria.

to mitigate uncertainties by assessing the current or most recent aggregate state of an economy using a range of partial indicators before the official economic statistics are released. Macroeconomic monitoring, or ‘nowcasting’, plays a crucial role in economic policy-making because the availability of crucial statistics concerning the current state of the economy is significantly delayed. In the past decades, economists used indices of economic indicators and regression models for updating expectations of new data releases, combined with professional judgement to monitor the economy in real-time. This method is not scientific in the sense of being replicable, using well-understood methods, quantifying uncertainty, and being amenable to later evaluation. Moreover, this method runs risk of assigning incorrect weights for each new data release, which causes internal inconsistencies because each series is handled separately (Bańbura et al. 2013).

These weaknesses have prompted central banks and academics to apply considerable effort to improving the foundations and reliability of real-time nowcasting. The famous paper of Giannone et al. (2008) showed that the dynamic factor model in the state-space representations can incorporate real-time data releases to provide an internally consistent framework for estimating current economic conditions. Their framework has three important features: the model can be automatically updated when new data are released in a nonsynchronous manner because the Kalman filter generates projections for all variables, the model bridges monthly indicators with quarterly GDP to manage the issue of mixed frequencies, and the model is economically dynamic because it accounts for the dynamics of the indicators used in the analysis. The dynamic factor model is now a standard macroeconomic monitoring model used by several international institutions and central banks for current GDP nowcasting, including the Federal Reserve Bank of Atlanta (Higgins 2014), Federal Reserve Bank of New York (Aarons et al. 2016), the European Central Bank (Bańbura et al. 2013) and the International Monetary Fund (Matheson 2011). However, the literature shows a growing trend that the dynamic factor model can discard potentially important information because of parametric restrictions (Stock & Watson 2017, Athey 2018). Alternatively, machine learning (ML) algorithms can improve macroeconomic forecasting and nowcasting by exploring the nonlinear and non-parametric structure in the macroeconomic dataset. Therefore, ML algorithms can solve the the curse of dimensionality in standard regression analysis without transforming variables into latent and unobservable factors (Athey 2018).

Despite the importance of macroeconomic nowcasting, only Yiu & Chow (2010), Barnett & Tang (2016), and Zhang et al. (2018) addressed the problem of nowcasting China's GDP. These studies are inadequately designed because the chosen benchmark models are not appropriate and the researchers only consider one model at a time. Chapter 4 addresses the issue of nowcasting China's GDP in a data-rich environment by exploring the use of ML algorithms. Using a large set of quarterly macroeconomic indicators, we train eight machine learning algorithms and nowcast GDP growth for each quarter from the first quarter in 1993 to the second quarter in 2018. We compare the predictive accuracy of these nowcasts with those of AR model and dynamic factor model in the state-space representation. The model confidence set of Hansen et al. (2011) is used to obtain a set of best models with the 10% level of significance. The results indicated that shrinkage methods are covered by the model confidence set, positioning them in the set of best models. As such, ML algorithms proved useful for improving the accuracy of nowcasting the Chinese GDP growth rate. Overall, this chapter joins the growing literature that examines the usefulness of ML learning for nowcasting GDP and guides practitioners toward selecting the appropriate nowcasting models for China's economy.

Chapter 2

How Trustworthy are Chinese Official Statistics

2.1 Introduction

China is the world's second-largest economy and is a prime destination for international investors, as the country has sustained strong growth rates for an extended period. Despite China's great economic achievements in recent decades, economists and data analysts have questioned the accuracy and reliability of official statistics, especially considering the rapid changes in China's economic structure (Rawski 2001, Holz 2014, 2008, 2004, 2003, Hsien & Song 2019, Xiong 2019, Ma et al. 2014, Koch-Weser 2013). The rapid development of China's economy requires official statistics to reflect economic activities accurately. Accurate and timely economic data assists the policy-making process, not only for economic researchers, financial analysts and policy-makers, but also for other scientific fields, such as geography and medicine. Economists and other stakeholders need to consider whether the Chinese government is willing to and capable of providing accurate and timely economic data. The answer is vitally important for empirical research and the analysis of Chinese macroeconomics.

Several examples indicated that Chinese data may be unreliable and may cause conflict with other parties' data. Over 60,000 cases of illegal practice were found in the national inspection of statistics in 1997 (Wang & Meng 2001, Wang 1998)[1]. In October 2011,

[1]Volume 194 of the China Statistics Journal reported two severe cases of illegal practice. In the first case, Rongzhi Liu, the head of government of the Xichong county in the Sichuan province, abused his authority by falsifying local statistics and exaggerating the county's gross output by 61.75%. This led

the Chinese reading of the air quality in Beijing was ‘slightly polluted’, whereas the US Embassy in Beijing reported ‘crazy bad’ air quality. The Chinese Foreign Ministry complained that the conflict between the US air quality data and the China data caused ‘confusion’ and undesirable social consequences. In 2007, the now Chinese Premier, Li Keqiang, described the Chinese GDP figure as ‘man-made’ and unreliable.[2] Rather than examining GDP figures, Keqiang used electricity consumption, rail cargo volume and bank lending to evaluate the speed of economic growth in Liaoning province, where he was the head of the Communist Party. All other figures, especially GDP, were for ‘reference only’. During the local congress, the current Head Governor of Liaoning Province, Qiufa Chen, publicly admitted falsifying economic data in the past five years.[3] Similarly, Binhai District in Tianjin city revised their GDP from above 1000 billion Chinese Yuan to 665.4 billion Chinese Yuan, indicating that approximately a third of GDP was ‘water content’.

The suspicion of China’s economic data is a recent phenomenon. Most researchers held a positive view of China’s economic data until the late 1980s. Rawski (2001) concluded that most foreign specialists agree the official statistics were generally accurate and reliable until the 1980s. Chow (1986) wrote: ‘by and large Chinese statistics are honest although measuring errors or statistical discrepancies do exist’. In the late 1990s, concerns regarding the quality of China’s official GDP data were common, as the Asian financial crisis and natural disasters were expected to significantly weaken the economy. However, the Chinese government was still able to ‘guarantee’ a double-digit real GDP growth rate (Rawski 2001, Holz 2004). Since this implausible guarantee, many economists have expressed concerns and doubts about the accuracy and the quality the official Chi-

to the dismissal of the county head. In the second case, Xicheng Li, the first secretary of the ruling party of the Chixiang county in the Jiangsu province, was dismissed from his position due to severe data manipulation Holz (2014).

[2]see New York Times (22 June 2012) and the Wall Street Journal (6 December 2010) and Reuters (28 January 2010). The source is an American diplomatic cable released by WikiLeaks.

[3]See Financial Times, China fake data mask economic rebound 16 January 2018.

nese GDP growth rate, especially the possible government manipulation of the growth rate (Holz 2004, 2008, 2014, Koch-Weser 2013, Kerola 2019, Xiong 2019, 2018, Hsien & Song 2019, Bailliu et al. 2019). Economists keyed the phrase 'wind of falsification and embellishment' to describe their suspicion that the NBSC engaged in intentional data falsification and manipulation, which may have been ordered by the Communist Party of China (CPC) and the State Council. Additionally, the defective statistical system has contributed to the scepticism surrounding official data (Holz 2014, 2008).

This paper investigates the quality of Chinese official economic statistics and offers several contributions to the existing literature. First, it provides a comprehensive review of the existing literature for scholars, financial analysts and politicians interested in the quality of China's official statistics. Second, it uses eight different price indices to evaluate the quality of implicit GDP deflators. Third, it resolves the discrepancies of GDP using multiple methods and the discrepancies for individual sectors, producing constructive solutions based on quantitative methods. Fourth, it presents an examination of the institutional framework as it relates to statistical work in China, which is useful for explaining the discrepancies within the country.

The rest of the paper is organised in sections. Section 2.2 discusses the impact of political interference. Section 2.3 reviews and examines the quality of the statistical frameworks. Section 2.4 examines internal inconsistency and Section 2.5 concludes. Regarding sources, academic articles from Chinese journals proved to be useful in the preparation of this paper and the data used were collected from China Statistical Yearbook 2017.

2.2 Political Interference

The political economic literature raised concerns about the possible effects of politics on the quality of China's official economic data. A popular opinion is that the publication of economic data acts is a crude political farce; the head of the CPC and the State Council intervene arbitrarily to direct the level of some macroeconomic indicators. This view is not wholly accurate, but it is not far from reality (Holz 2014, Koch-Weser 2013, Xu et al. 2000, Xianchun 2002, Maddison 2007, Rawski 2001, Cai 2000, Xiong 2019). Political interference plays a crucial role in the statistical reporting system; it is widely believed to be one of the most influential causes of inaccuracy in China's official statistics (Holz 2014, Cai 2000, Ma et al. 2014, Xiong 2019). Political interference exists in various forms, including new policies, new growth targets and new orders. This applies the pressure of superlative economic performance along the reporting hierarchy, from the State Council to the provincial level and then to the municipal and county levels. This section discusses the effect of political interference on the central government level, as well as the local government level.

2.2.1 The Central Government and the NBSC

The NBSC is under the direct leadership of the State Council of the People's Republic of China. The State Council appoints major personnel and provides funding to the NBSC. The Statistics Law of 1996 and its revision in 2009 assigned rights and responsibilities to the NBSC to organise, direct and coordinate statistical work throughout the country.[4] The NBSC largely collects data through the DRS, surveys and censuses. The NBSC, approved by the Statistics Law and National People's Congress, also received data from

[4] National People's Congress 1996 Art. 4 and National People's Congress 2009 Art. 27.

approximately one hundred other state-owned institutions (Holz 2014, Ma et al. 2014).[5] However, the NBSC has limited authority over the statistical divisions of other government institutions – the bureau only has direct control over its own survey teams. Therefore, data collection at the NBSC is hindered by other government departments because the quality and coverage of the collected data are determined by each government departments' needs and reach (Holz 2014). Moreover, the NBSC has little control over the statistical bureaus in provincial, municipal and county areas. Key data, such as that of the GDP, compiled by local statistical bureaus must be approved by local government leaders before it reaches the next highest statistics department (Holz 2014, Koch-Weser 2013, Landry 2008, Jia et al. 2015).

The notion that the NBSC is independent has been criticised (Koch-Weser 2013, Holz 2014, Mei & Wang 2017, Landry 2008, Jia et al. 2015). The commissioner, the deputy commissioners and the CPC's cell members of the NBSC are appointed by the State Council. The current NBSC commissioner, Jizhe Ning, is also the CPC's first secretary of the NBSC. Similarly, the members of the CPC cells are also deputy commissioners. Consequently, the NBSC primarily serves the interests of the State Council and the CPC, rather than the interests of the public.[6] The work regulations described in Act 13 of the NBSC explicitly state that the NBSC implements the 'important decisions and instructions of the CPC and the State Council'. The former NBSC commissioner and first secretary of the CPC from 1984 to 1997, Sai Zhang, clarified the tasks and the responsibilities of the statistics system: 'the NBSC and relevant statistical organisation primarily serve the needs of macroeconomic decision-making of Party and government leaders at

[5] including other subordinate State Council departments, such as financial sector data from the People' Bank of China, trade data from China Customs and fiscal data from the Chinese State Administration of Taxation and Finance Ministry.

[6] Jizhe Ning had a PhD in economics from Renming University in Beijing, China. Before he became the commissioner and the CPC's first secretary of the NBSC in 2016, Ning served in various positions within the State Council, including the CPC's secretary of the policy analysis centre.

each administrative level and is responsible to the Party and government leaders at each administrative level'. (Sai, 2001, p.319). The commissioners and the CPC cell members do not seem to prioritise serving the interest of the general public—they give priority to the CPC leaders rather than to the congress and the parliament of China. Additionally, the Statistic Law states that the National Development and Reform Commission can access the NBSC data before they are published.[7]

The Statistics Law implements articles against data falsification and fabrication, but only in an ethical and professional manner (daode) (Article 29 in The Statistics Law 2009) (Holz 2014, Jia et al. 2015, Shih et al. 2012). Article 6 warns that neither the leaders, the CPC secretaries at local government level, the statistical institutions nor the relevant departments of the people's government will, without authorisation from the senior government level, revise, fabricate or tamper with statistical data. Moreover, the CPC will not retaliate against statisticians who perform their duties following the law or who refuse to violate the Statistics Law. However, the Statistics Law does not specify how information would be reported to a higher level of government if a violation of the Statistics Law occurs or if the supervisory government is the originator of the violation.

2.2.2 Manipulation along the Hierarchy: Local Government

Clarifying data manipulation and falsification throughout the reporting hierarchy and among different reporting units in the statistical chain assists understanding of the underlying mechanisms that regularly cause manipulation. Substantial inconsistencies between China's national and provincial GDP were observed in recent years, mainly in the reported industrial output. The origin of these inconsistencies may be deeply rooted—local cadres

[7]The Statistics Law 1996 Article 2: 'the fundamental task of statistical work is to conduct statistical examination of the implementation of the national economic and social development plan, to analyse the statistics, to provide statistical advice and suggestions, and to supervise through the use of statistics'.

are incentivised to over-report regional output levels and the provincial governments have strong influence over the provincial bureaus of statistics, which report the provincial economic statistics (Koch-Weser 2013, Jia et al. 2015, Chen & Kung 2016, Ma et al. 2014). The cases of Huanzhong Wang and Tianjin illustrate the reported data manipulation by local governments. The former deputy governor of the Anhui Province, Huaizhong Wang, led the corruption and data manipulation by over-reporting township enterprises' revenue as approximately five times the actual amount when he was the CPC secretary of an Anhui county. Wang 'ordered' the local statistical bureau to fake the GDP growth target rate, increasing the rate from 4.8% to 22% per annum during the 9th Five-Year plan period when Wang was the CPC's first secretary of Fuyang. In 2016, the Binhai district in Tianjin revised down 334.8 billion Chinese Yuan which is about one-third of its annual GDP.[8]

A large portion of the political economy literature emphasises that the career incentives and the Chinese cadre evaluation system, alongside the geography-based governing logic, have motivated local officials to compete to generate high growth (Mei & Wang 2017, Landry 2008, Li & Zhou 2005, Holz 2014, Koch-Weser 2013, Holz 2008, 2004, Xiong 2019, 2018, Xiaolu & Lian 2001, Xi et al. 2018). The Chinese government has been known to identify and credibly reward officials who can nurture economic growth and to sideline those who cannot. Although China has embraced the marketisation of its economy since the 1970s, it remains a politically centralised state and local officials are appointed by the central government that is tightly controlled by the Central Committee Party (Landry et al. 2017, Naughton & Yang 2004, Holz 2008). The centralised personnel appointment system is a critical feature that allows members of the Central Committee Party to set rules at all levels.[9] When political selection is centralised and economic growth

[8]Source: Tianjin has revised the 2016 GDP for the Binhai district, Financial Times Chinese January 12, 2018.

[9]It is commonly argued that falsification and manipulation occur more often at the local government

is prioritised, local Chinese officials engage in a promotion tournament in which they compete against peers to survive politically and gain promotion to a higher government level. Local officials know that they are very unlikely to be promoted should they fail to ensure adequate economic performance and fiscal income. Xi et al. (2018), Xiong (2018) and Xu (2011) showed that local politicians' competence clarifies much of the variation in economic growth rate since 1978.

State-owned enterprises (SOEs) are likely to inflate output and revenue figures (Holz 2014, Ma et al. 2014). Under the leadership of the CPC, the first secretary and deputy secretaries of SOEs cannot jeopardise their careers by disappointing the interest of the general public or the shareholders. Instead, the SOE personnel are appointed by the provincial government and the State Council, and promotions are primarily given based on competence, and possibly on sales and profitability figures (Holz 2004, Koch-Weser 2013, Maddison 2007). SOE personnel may even be promoted to leaders of the provincial government or of a ministry, which enhances the incentive to fake output figures. Local government leaders can also pressure SOEs, such as utility and electricity companies, to force them to report false output figures that match the inflated GDP growth.

Overall, data manipulation can occur at the central government level and the local government level. The Political Bureau of the CPC appoints the personnel of the NBSC, forcing the NBSC to report output figures that consider the expectations of the central level of government and the CPC leaders. Moreover, leaders in Beijing encourage competition between the local governors to generate better economic performance. The NBSC is more likely to report false figures and revise them in later years, than not to publish at all (Koch-Weser 2013, Holz 2004, 2008). In fact, the NBSC has recognised

level than at the central government level. At the central level, the promotion of cadres is arguably related more to the factional recruitment strategies or work college of the Central Committee members (Shih et al. 2012, Meyer et al. 2016, Landry et al. 2017, Landry 2008, Naughton & Yang 2004, Bulman 2016).

over-reporting behaviour by local governments and has adjusted nationwide GDP figures for several years (Sinclair 2019). Xiong (2018) developed the Mandarin model to systematically examine the agency problems between central and local governments in affecting the Chinese economy and Hsien & Song (2019) provided a study that routinely corrects China's GDP statistics.

2.3 Quality of Statistical Framework

After the economic reforms and opening to foreign trade in the late 1970s, the Chinese government implemented a series of statistical reforms to improve the quality of the statistical work. These changes focused on introducing new statistical methods that suit the market-oriented economy (see Appendix A) (Koch-Weser 2013, Holz 2014). These steps were necessary during the rapid transformation of the economic structure. This section reviews the ongoing challenges of gathering, measuring and presenting economic data.

2.3.1 Questions Related to Sample Surveys

With GDP measured by the VA approach, the NBSC classifies the industry structure according to the historical sequence of development. The primary sector of the economy includes any industry involved in the extraction and collection of natural resources—farming, forestry, mining and fishing. The secondary sector of the economy includes industries that produce a finished, physical product or are involved in construction. The tertiary sector of the economy involves the production of services. The NBSC uses four major surveys to calculate GDP using the VA approach. Each survey covers one sector—the large industrial sector firms, large service sector firms, qualified construction firms and smaller industrial firms.[10] Data manipulation can occur when official data are based on published and unpublished data. For the industrial sector, output data from the SOEs and listed firms are publicly available while output data from small and non–state owned enterprises are not.[11] This pattern is very similar in the construction and service sectors. The NBSC can easily manipulate unpublished data, including sample survey data and the dataset of units that do not report directly to the statistics system. From 2008, it is possible the published nominal VA data was also manipulated because the NBSC limited publication to the aggregate number for each industry.

Economic units outside of the traditional reporting hierarchy are more likely to be manipulated, such as the 'shadow' or 'underground' economy. The shadow economy comprises economic activities that are illegal, unrecorded or that are not captured in the statistics system (e.g., because they are too small or dispersed). Moreover, it is difficult to estimate the rate of unpaid home production, including household cleaning, childcare, house repair and maintenance. The proportion of unpaid home production may be large compared to Western economies; however, unpaid home production, is not included when calculating the Chinese GDP. The NBSC uses benchmark revision to revise the underground economic activities and unpaid home production. The effects of benchmark revisions can be significant. For example, tertiary census [12] in 1993 revised the tertiary VA upward by 32% and the GDP by 10%. Following the 2004 economic census, the 2004 tertiary sector VA was revised upward by 48.71% and the GDP was revised upward by 16.81%. However, the Chinese macroeconomic data revisions are not well-behaved, which leads to a biased estimate of the tertiary sector VA (Sinclair 2019).

The issue of incomplete sample surveys extends to the enterprise sector and the bank-

[10] As previously mentioned, the NBSC also works with administrative data from other government departments.

[11] Direct-reporting industrial enterprises are required to regularly report variables (e.g., output, employment, balance sheet and income statement) to the statistics system.

[12] Tertiary census refers to national economics census for the tertiary sector.

ing sector. The NBSC uses the DRS and sample survey to measure and gather enterprises' output and revenue data. Firms with revenue greater than the threshold amount are eligible to be included in the DRS; firms with revenue less than the threshold amount are covered by sample surveys. Private firms are tempted to avoid tax by under-reporting output and over-reporting production costs, regardless of whether the firms are covered by the DRS or the sample survey. Privates businesses often keep three sets of books: one for tax officials, one for investors and employees and one for themselves. Detecting tax dodging in these businesses is complicated and the punishment is often less expensive than the money saved through dodging tax (Koch-Weser 2013). Several corporate entities, such as fund management and real estate management, are often not included in the DRS or the sample survey. Banks heavily favour lending to SOEs, forcing small businesses to approach the underground banking network to access credit. This underground system is hard to track and capture using surveys. Additionally, there are unaccounted economic activities in the service sector, in which cash transactions that are difficult to measure by sample surveys are widespread (Koch-Weser 2013).

2.3.2 Data Measurement Issues for Consumption and Investment

On the aggregate expenditure side, local statistical authorities supply estimates of local consumption, investment, government spending and net exports. The two main sources are the survey of household income and expenditures, which provides an estimate of local consumption, and the survey of investment projects, which provides an estimate of local investment. The main issue with the measurement of Chinese household consumption is that it primarily relies on retail sales data. Other sources (e.g., household surveys) are not sufficiently used by the NBSC to measure household consumption. Some economists argued that local statistical authorities measure retail sales by measuring the goods that are shipped to the warehouses, rather than when the goods are sold (Koch-Weser 2013, Holz 2004, Rawski 2001). As such, the retail sales figures can be over-estimated when the goods are taken to the warehouse but not sold to the public. Consequently, local household consumption and GDP can be inflated by local statistical authorities. Addressing this issue is difficult because of the limited inventory data available to confirm the accuracy of the retail sales data. Although the NBSC sets the strict statistical requirement that goods and services must be counted based on the sales made by wholesalers and retailers, there is little evidence of whether this requirement is met. Additionally, surveys by the NBSC did not capture household income and personal tax accurately. Personal income tax accounts for approximately 40% of tax revenue in the US, while it only accounts for 6% of tax revenue in China. As the high-income earners in China are only taxed directly on their salary, they often refuse to disclose other income, such as gains in real estate, stocks, fixed income security and gifts, in the sample surveys. This phenomenon is partially caused by the Chinese custom in which households are unwilling to disclose information on their financial conditions. Many households either significantly under-report their income or refuse to take sample surveys to avoid the personal income tax (Koch-Weser 2013).

The relationship between household income, household savings and retail sales reveals several contradictions. Theoretically, an increase in retail sales should entail a proportionate increase in household income or a decrease in household saving. Rawski (2001), Koch-Weser (2013), and Holz (2014) suggested that the Chinese retail sales figures continue to outpace household income, while household saving has not decreased commensurately. As such, retail sales seem exaggerated and consumption data are inaccurate.[13]

[13]Collecting household income data is not merely about asking survey respondents how much they earn but is a very complicated process. Working closely with the National Bureau of Statistics of Canada, the National Bureau of Statistics of China (NBSC) has been instituting a series of reforms since 2008 to ensure that the collection of household income data meets international standards. The NBSC scientifically selects 160,000 households from 1650 counties, districts and cities. Then, sampling teams (directly led by

Additionally, the housing sector can affect the measurement of household consumption. Housing is a durable good that generates a stream of welfare to owners. In the Chinese national account setting, house owners pretended they were tenants and paid 'inputted rents' (artificial rent) based on the operational costs of household consumption. Conversely, house owners in Western countries use the market rental price to calculate prices. Chinese economists Zhang Jun and Zhu Tian argued that household consumption was over-estimated by approximately 2% p.a. over the past two decades because of the use of operational costs when calculating 'inputted rents'.[14] NBSC researchers conducted a study in 2012 to understand how much housing rent affects consumption by surveying housing rent in China's four largest cities and found similar results.[15]

There are similar issues with fixed asset investment—over-estimation of fixed asset investment occurs when funds are disbursed rather than using the working capital (Koch-Weser 2013). This is a bias towards the supply measurement of China's GDP—the VA approach. Using the VA approach, fixed investment could be inflated by releasing the funds into the market but not using them as working capital to stimulate the economy. To overcome this bias, the NBSC set regulations that only investments using working capital should be counted; however, the execution of such regulations is difficult to observe because there is no adequate public data on how investment funds are utilised. A prominent issue of over-reported fixed asset investment originates from infrastructure projects—a core component of investment statistics. Property developers benefit signif-

the NBSC) survey and record the salary, government subsidy, financial income, business activity, rental income, superannuation, tax and insurance expenses of these households. At every survey collection spot, there are professional surveyors to advise on and supervise the data recording process. Using standard guidelines set by the NBSC, local bureaus of statistics check, review and report household income data to the NBSC. At the end of each quarter, the NBSC randomly calls 6000 survey respondents to verify the validity of the recorded data. Finally, the NBSC reviews and releases household income data. Therefore, the act of households under-reporting their true income is unlikely to lead to such discrepancies.

[14] Sources: Yao Yang, 'Quality of GDP statistics improving', China Daily.

[15] Xu Xianchun et al, 'Jumin Zhufang Zuping Hesuan Ji Dui Xiaofulu de Yingxiang [On Residents House Rents and Its Impact on the Consumption Rate],' Kaida Daobao [China Opening Journal] 2 (April 2012): 12–15.

icantly from rising land values, giving them an incentive to keep undeveloped land and resell it in the future. By law, property developers must build on the land within two years or it will be confiscated. They may tell the government that a project is 'under construction' and make minor investments to avoid land confiscation. This phenomenon is likely to cause an overstatement of fixed asset investment.

2.3.3 Revision and Publication Lag

China is always one of the first countries to announce GDP statistics, usually two to three weeks after the end of each quarter. Comparatively, most Western economies, which generally collect a smaller amount of data more efficiently, typically take over six weeks to publish their quarterly GDP figures. The quick release and premature publication of Chinese GDP statistics are essential reasons to criticise the reliability of output data (Holz 2014, Koch-Weser 2013). Conversely, statistics for the other components of GDP, such as international trade and investment, are released infrequently and with a significant publication lag.

In addition to the publication lag and release timing issue, the NBSC has conducted revisions of the published GDP data to capture the unaccounted economic activities in previous years. Revisions of GDP statistics are common across Western countries. For GDP data to be published in a timely manner, statistical authorities often release a portion of the data and later revise it if necessary. However, for China, benchmark revisions are frequent, large and not always clearly justified (Koch-Weser 2013). In the Chinese Statistical Yearbook, the NBSC revises the previously published nominal GDP and sectoral VA output data for the last several years. However, the real GDP growth rates are typically not revised as often as the nominal GDP growth rates and the GDP deflators generally are not disclosed. For some data series (e.g., fixed asset investment and industrial production), the sum of monthly data differs from the annual data, indicating a substantial revision of the monthly figures. Sinclair (2019) found the Chinese macroeconomic data revisions were not well-behaved and generally failed Aruoba (2008)'s test, which indicates that further research is needed to improve China's data gathering processes.[16]

Revising the GDP growth rate is a serious matter, especially after the national economic census. The NBSC partnered with Japan's Hitotsubashi University to conduct the earliest revision in 1997. Together, they re-estimated and revised China's historical GDP growth rate from 1952 to 1997 (Koch-Weser 2013, Holz 2014). Maddison (2007) and Koch-Weser (2013) suggested that this revision over-estimated the real growth rate by an average of 3% to 4% because of the miscalculation of the price level and productivity. Their revised rate was well below the revised official annual growth rate of 7.49%. The issue became severe after the state government passed a law in 2003 that allowed the NBSC to revise long-run GDP data based on findings from the national economic census. A common criticism is that it is unclear which price indices are used by the NBSC to calculate the GDP deflator, as the NBSC does not disclose how the GDP deflators are calculated and adjusted (Holz 2014). Moreover, sample surveys conducted by the NBSC do not cover many self-employed workers and labour migrants in urban areas, which distorts the results of the census. The financing ability of the local governments can also affect the results of the census because local governments cover a large portion of costs and their fiscal capacity varies between different provinces and cities. Further, the quality of the census work is not consistent across the different regions.

The benchmark revisions following the 1995 industrial census and the 2004 national economic census showed anomalies. In the benchmark revisions that followed the 1995

[16] From a statistical point of view, Aruoba (2008) proposed three properties to satisfy a 'well behaved' revision. First, the revision is expected to have mean zero, indicating that the initial announcement of the statistical agency is an unbiased estimate of the final value. Second, the expected variance of the final revision is small compared to the variance of the final value. Third, the final revision is unpredictable, given the information available at the time of the initial announcement.

industrial census, the gross output value of the industrial sector showed a decrease between 1991 and 1994, which should have triggered corresponding revisions to the sectoral VA. However, the industrial VA was not revised, implying that intermediate inputs were revised by an equal amount, which is not plausible (Koch-Weser 2013, Holz 2014). The NBSC did not explain why industrial VA data were unchanged following the industrial census. Holz (2014) argued that intentional falsification of industrial VA and problematic gross output value data from enterprises outside the DRS are two possible explanations. During the 2006 benchmark revision that followed the 2004 national economic census, the nominal VA of all sectors between 1993 and 2004 and the real growth rates of tertiary sector VA were revised. However, the real growth rates of the primary and secondary sectors were not revised. A possible implication is that the NBSC raised the implicit deflator for the industrial sector and lowered the implicit deflator for the construction sector. However, the 2004 economic census collected no price data and the NBSC offered no explanation of why and how it revised the implicit sectoral deflators. Moreover, the secondary sector's real growth rate is based on a weighted average of the real growth rate of the industrial and construction sectors. Retention of the real growth rate for the secondary sector implies that the NBSC did not change the weights of industry and construction in the calculation of the secondary sector's real growth rates. This is inconsistent as the VA of the industry was revised upward by 3.8%, whereas the VA of construction was revised downwards by 9.2% between 1993 and 2004.

2.4 Internal Inconsistencies

The reliability and credibility of China's national output statistics can also be accessed and tested by examining the internal inconsistencies within the country's data. Inconsistencies within the data imply a lack of accuracy. In this section, four types of inconsistencies are examined: 1) internal discrepancies between the GDP calculated using the AE approach and the GDP calculated using the VA approach;[17] 2) internal inconsistency between nationwide GDP data and the sum of provincial output data; 3) sectoral discrepancies to identify the sector accounts with the largest discrepancies; and 4) discrepancies between implicit GDP deflators and various price indexes.

2.4.1 Discrepancies between the VA Approach and the AE Approach

Let $Y_{t,VA}$ denote the nominal GDP by the VA approach in year t and denote $Y_{t,AE}$ to be the nominal GDP by the AE in year t. The percentage difference between GDP by the VA approach and the AE approach D_t in year t is then:

$$D_t = \frac{Y_{t,VA} - Y_{t,AE}}{Y_{t,AE}}. \tag{2.1}$$

Let $g_{t,VA}$ and $g_{t,AE}$ denote the growth rates of nominal GDP by the VA approach and the AE approach in year t, respectively. The absolute difference between the growth rate of nominal GDP by the two approaches in year t is

$$d_t = |g_{t,VA} - g_{t,AE}|. \tag{2.2}$$

Table 2.1 presents the percentage differences between the nominal GDP by the VA and the AE methods (D_t in (2.1)) and the absolute difference between the nominal growth rate of the two approaches from 1979 to 2017 (d_t in (2.2)). Overall, the reported percentage

[17]Theoretically, different approaches of calculating GDP should produce the same figure if they are compiled independently and accurately. Analysis of the gap between the two approaches can provide an insight into the quality of the GDP figure.

differences by which the nominal GDP calculated using the VA approach exceeded the GDP calculated using the AE approach were between –1.44% and 1.86%, with an average of –0.43% during the sample period. Most years, the percentages were generally negative; therefore, GDP measured by the VA approach was usually smaller than GDP measured by the AE approach. The reported absolute difference between the growth rates of the nominal GDP using the two approaches ranged from 0.01% to 2.33%, with an average of 0.34% during the sample periods.

There are two noticeable results and interpretations in Table 2.1. First, the percentage differences for the nominal GDP of the two approaches are significant in some years when compared to the US. For example, in the years 1989, 1990, 2015 and 2017, the percentage differences in China exceeded 1%. These percentages appear to be small but the discrepancies are significant because of the size of the Chinese economy.[18] Second, reported absolute discrepancies of the growth rates became large after 2007, averaging 0.75% between 2008 and 2017. Absolute discrepancies were larger at the end of the sample period—the reported numbers exceed 1% in 2015 and 2016 and 2% in 2017. Comparatively, the average difference in the annual growth rate between the two approaches was smaller from 1979 to 2007—only 0.2%. The unusual range in the size of the discrepancies indicates that the NBSC has not had a consistent standard for nominal GDP calculations during the past four decades. In contrast, the discrepancies between two approaches for United States and United Kingdom were negligible from 1979 to 2017 as shown in Figure 2.1, with maximum percentage differences being -0.17% for United States in 2017.[19]

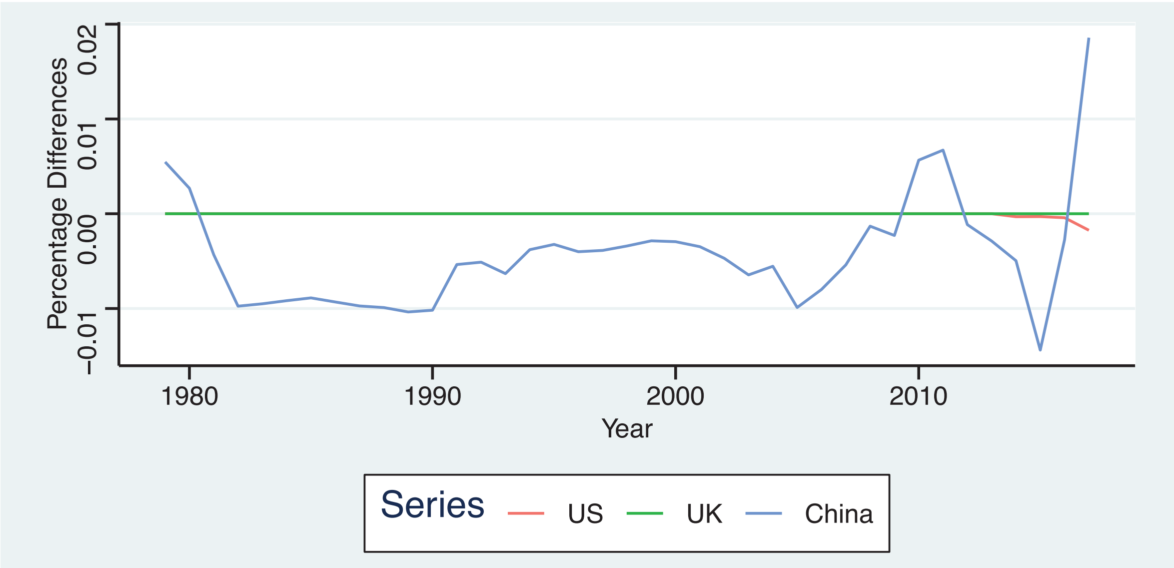
The difference between using the VA approach and using the AE approach to calculate the Chinese GDP has become particularly pronounced in recent years. The mirror image

[18]For example, the reported difference between the two approaches was 1.86% in 2017, which is about 1.5 trillion Chinese Yuan.

[19]Chinese GDP data from the Federal Reserve Economic Data (FRED) are sourced from the World Bank. They are identical to the NBSC data.

comes from growing discrepancies between aggregate net export and gross capital formation. The calculation of local net exports by provincial bureaus of statistics is not based on actual trade data; instead, the provincial governments report local net exports as a residual to balance local production and GDP expenditure. The gross capital formation figure is primarily based on reports of fixed asset investment by local governments. There is no audit for the gross capital formation and local net exports data, nor are there any consequences for misreporting behaviour. Conversely, household consumption and government expenditure data by provincial governments are relatively consistent with the NBSC figures, provided that the NBSC directly sends survey teams to collect household consumption data through the Household Survey.

Figure 2.1: Discrepancies Between GDP by the VA Approach and the AE Approach for Three Countries



Source: China Statistical Yearbook 2017 and OECD Statistics

Table 2.1: Discrepancies Between GDP by the VA Approach and the AE Approach

	Nominal Value (in Billion Yuan)			Annual Growth		
	GDP Value-added	GDP Expenditure	Percentage Difference	GDP Value-added	GDP Expenditure	Absolute Difference
1979	410.0	407.8	0.55%	11.5%	12.2%	0.75%
1980	458.8	457.5	0.27%	11.9%	12.2%	0.31%
1981	493.6	495.7	-0.43%	7.6%	8.3%	0.76%
1982	537.3	542.6	-0.98%	8.9%	9.5%	0.60%
1983	602.1	607.9	-0.95%	12.1%	12.0%	0.03%
1984	727.9	734.6	-0.92%	20.9%	20.8%	0.04%
1985	909.9	918.0	-0.89%	25.0%	25.0%	0.04%
1986	1037.6	1047.4	-0.93%	14.0%	14.1%	0.05%
1987	1217.5	1229.4	-0.97%	17.3%	17.4%	0.05%
1988	1518.0	1533.2	-0.99%	24.7%	24.7%	0.02%
1989	1718.0	1736.0	-1.04%	13.2%	13.2%	0.05%
1990	1887.3	1906.7	-1.02%	9.9%	9.8%	0.02%
1991	2200.6	2212.4	-0.54%	16.6%	16.0%	0.57%
1992	2719.5	2733.4	-0.51%	23.6%	23.5%	0.03%
1993	3567.3	3590.0	-0.63%	31.2%	31.3%	0.16%
1994	4863.7	4882.3	-0.38%	36.3%	36.0%	0.35%
1995	6134.0	6153.9	-0.32%	26.1%	26.0%	0.07%
1996	7181.4	7210.2	-0.40%	17.1%	17.2%	0.09%
1997	7971.5	8002.5	-0.39%	11.0%	11.0%	0.02%
1998	8519.6	8548.6	-0.34%	6.9%	6.8%	0.05%
1999	9056.4	9082.4	-0.29%	6.3%	6.2%	0.06%

Table 2.1: Discrepancies between GDP VA method and AE method (continued)

	Nominal Value (in Billion Yuan)			Annual Growth		
	GDP Value-added	GDP Expenditure	Percentage Difference	GDP Value-added	GDP Expenditure	Absolute Difference
2000	10028.0	10057.7	-0.29%	10.7%	10.7%	0.01%
2001	11086.3	11125.0	-0.35%	10.6%	10.6%	0.06%
2002	12171.7	12229.2	-0.47%	9.8%	9.9%	0.13%
2003	13742.2	13831.5	-0.65%	12.9%	13.1%	0.20%
2004	16184.0	16274.2	-0.55%	17.8%	17.7%	0.11%
2005	18731.9	18919.0	-0.99%	15.7%	16.3%	0.51%
2006	21943.8	22120.7	-0.80%	17.1%	16.9%	0.22%
2007	27023.2	27169.9	-0.54%	23.1%	22.8%	0.32%
2008	31951.6	31993.6	-0.13%	18.2%	17.8%	0.48%
2009	34908.1	34988.3	-0.23%	9.3%	9.4%	0.11%
2010	41303.0	41070.8	0.57%	18.3%	17.4%	0.93%
2011	48930.1	48603.8	0.67%	18.5%	18.3%	0.12%
2012	54036.7	54098.9	-0.11%	10.4%	11.3%	0.87%
2013	59524.4	59696.3	-0.29%	10.2%	10.3%	0.19%
2014	64397.4	64718.2	-0.50%	8.2%	8.4%	0.23%
2015	68905.2	69910.9	-1.44%	7.0%	8.0%	1.02%
2016	74358.6	74563.2	-0.27%	7.9%	6.7%	1.26%
2017	82712.2	81203.8	1.86%	11.2%	8.9%	2.33%

Source: China Statistical Yearbook 2017. Nominal GDP is in current prices and is not seasonally adjusted.

2.4.2 Sum of Provincial Data Versus Nationwide Data

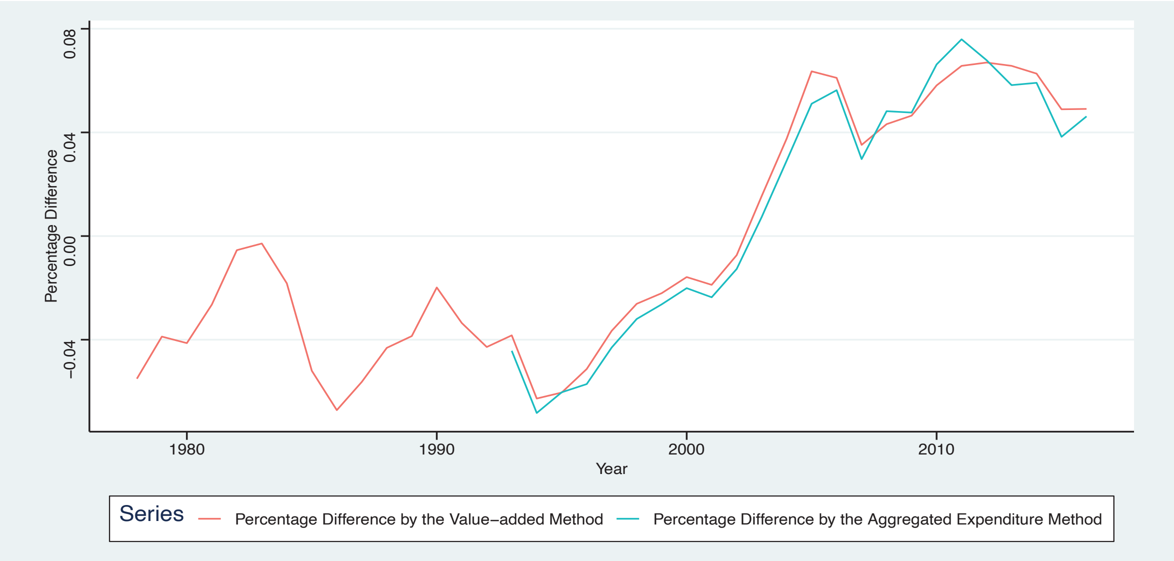
A key institutional feature of the Chinese National Accounts is that the underlying data are compiled by the statistical bureaus of local governments. Local statistical bureaus supply estimates of local GDP, including the production and expenditure components, to the higher levels of government. For the VA approach, provincial statistical departments combine the four major surveys' data and the administrative data from other government departments to calculate local GDP.[20] For the expenditure approach, provincial statistical bureaus estimate the local consumption, investment, government spending and net exports using surveys of household income and expenditures and local investment projects. There have been longstanding inconsistencies between the nationwide GDP statistics and the sum of the corresponding provincial outputs in the past decades. These inconsistencies have become a well-known issue in academia and in the mass media, which undermines the public's trust in China's statistical work (e.g., Ma et al. (2014), Hsien & Song (2019), Koch-Weser (2013) and Xiong (2019)).

Figure 2.2 demonstrates the time series plot of percentage differences between the sum of provincial output numbers and the nominal GDP calculated using the VA approach from 1978 to 2016 and using the AE approach from 1993 to 2016. For both approaches, the percentage differences were relatively small until 2001. From 2002 onwards, the percentage differences increased sharply, peaking around 2011 at 6.2%-approximately 4 trillion Chinese Yuan.

In addition to simply discussing and criticising the inconsistencies between the national GDP and the sum of provincial numbers, this section considers using index decomposition to identify the sectoral discrepancies in the VA approach and the AE approach. This

[20]The four surveys cover the large industrial sector firms, large service sector firms, qualified construction firms and smaller industrial firms.

Figure 2.2: Percentage Differences Between the Sum of Provincial Output and Nominal GDP by the VA Approach and the AE Approach



Source: China Statistical Yearbook 2017. Nominal GDP is in current prices and is not seasonally adjusted.

would identify the sectors or sub-sectors that produce the largest contributions to the discrepancies between the national GDP and the sum of the provincial output.

Given that the sum of provincial output typically exceeds the national GDP number, sectoral discrepancies are defined here as the percentage difference between the sum of provincial output and the national GDP in a given year. The following expression was derived:

$$D_{y,t} = \frac{\sum_{i=1}^{m} y_{i,t} - Y_t}{Y_t} = \sum_{j=1}^{n} \frac{Y_{j,t}}{Y_t} \cdot \frac{\sum_{i=1}^{m} y_{i,j,t} - Y_{j,t}}{Y_{j,t}} = \sum_{j=1}^{n} w_{j,t} \cdot d_{j,t}, \tag{2.3}$$

where in (2.3) $D_{y,t}$ represents discrepancies in year t, Y_t denotes the nationwide GDP sourced from the NBSC and $y_{i,t}$ refers to the provincial output sources from the provincial bureaus of statistics. The subscripts i,j and t represent provinces, economic sectors and years, respectively. Based on the sectoral classification of China's National Account, the total discrepancy can be decomposed into sectoral discrepancies, which are denoted by $d_{j,t}$ in (2.3) (where $w_{j,t}$ is sectoral weight in year t). According to Ma et al. (2014), the contribution rate for sector j at time t is defined as

$$c_{j,t} = \frac{w_{j,t} \cdot d_{j,t}}{D_{y,t}}. \tag{2.4}$$

Table 2.2 and Figure C.1.1 present total discrepancies ($D_{y,t}$) and sectoral discrepancies ($d_{j,t}$) using the VA method from 1978 to 2016. Overall, provincial governments under-reported output statistics between 0.3% and 6.7% from 1978 to 2002, with an average discrepancy of -3.5% during that period. From 2003 onwards, provincial governments collectively over-reported the local output level by between 1.5% and 6.7%, with an average discrepancy of 5.1%. A number of reasons have been advanced to explain under-reporting of provincial GDP before 2002. First, China's sectoral classification system has changed over time. The national sectoral classification criterion was updated for the second time in 2002, which led to a wide range of re-classifications. Most of China's provinces did not begin to restructure the sectoral GDP until 2005. Second, the NBSC tended to perform downward revisions to reconcile over-reporting behaviour by provincial governments. For instance, the NBSC performed large downward adjustments to provincial production in its approach to calculating GDP in 1993 (because the 1993 provincial-level data, published a year late, already incorporated the retrospective upward revisions to GDP following the 1993 tertiary sector census, but the nationwide data did not). Third, the NBSC explanations of how provincial output and national GDP are calculated are not consistent across different NBSC sources. The NBSC tended to publish incomplete statistics first and adjust them later through the National Economics Census rather than not publish them at all. Holz (2014) found that local governments over-reported provincial GDP based on first published values from 1993–2003 and the NBSC used the census to adjust provincial output data downwards.

Regarding the sectoral discrepancy produced by the VA approach, one notable irregularity is that local governments tended to significantly under-report tertiary sector output from 1978 to 2002, with an average discrepancy of –10.2% during this period. After 2003, the percentage of discrepancies dropped to a relatively low level. This trend in the inaccuracy of the statistics could be traced back to the 1993 tertiary sector census when tertiary sector value-added output for the years 1978–1993 was largely revised upward. Tertiary sector value-added output for 1993 was revised upward by 32.04% to reconcile the figures for over-reporting by provincial governments.

The sectoral discrepancies in the construction sub-sector were significant from 1978 to 1985 and from 2002 to 2005, exceeding 20% in 1978, 1979, 1982 and 1984. The percentage differences for the industrial sub-sector were relatively small until 2004. From this time onwards, the differences increased significantly from 1.5% in 2003 to 15% in 2016. The discrepancies in the primary sector figures, averaging –0.4%, were relatively small compared to the tertiary sector and the secondary sector. The increased discrepancies in the secondary sector since 2004 can be rooted in the career incentives of local leaders. Compared to the tertiary sector, the secondary value-added output is relatively easy to manipulate. Provinces often double-count cross-provincial economic activities and still use (presumably questionable) report forms for industrial enterprises with annual sales revenues below 5 million yuan.

The sectoral contribution rates to the overall discrepancies when using the VA approach are presented in Table 2.3 and Figure C.1.2. The primary sector has contributed very little to the overall discrepancies, except for during 1982 and 1983. The tertiary sector was the greatest contributor to nationwide discrepancies in the years leading to

1994, with an average contribution rate of 140.7% from 1978 to 1994. With the exception of 2003, from 1995 to 2007, the tertiary sector contributed positively to the overall discrepancies and the level of contribution was smaller than that of the secondary sector. From 2007 onwards, the tertiary sector contributed negatively, indicating that provincial governments collaboratively under-reported output statistics in the tertiary sector. The secondary sector was the first to contribute negatively to the overall discrepancies between 1978 and 1985, with an average contribution rate of 57.6%. Contribution rates dramatically increased from 18.3% in 1986 to 102% in 2016, averaging 69% during the same period. The contribution rates of the industrial sub-sectors have historically followed similar patterns to the contribution rates of the secondary sector, especially after 1986. Contribution rates for the construction sub-sector are relatively low compared to the industry sub-sector during the sample period. Notably, from 1978 to 1985 and 1999 to 2002, the construction sub-sector contributed an average of –70.5% and –37.3% to the nationwide discrepancies, respectively.

Similarly, discrepancies can also be found in the AE method. Table 2.4 and Figure C.1.3 present the discrepancies as the percentage difference by which the sum of provincial data exceeds the national GDP data for total GDP, final consumption expenditure and its two sub-sectors (household and government), and gross capital formation from 1993 to 2016.[21] Overall, the nationwide discrepancies by the AE were significant, fluctuating between –1.3% and –6.8% from 1993 to 2002, with an average of –3.9%. The percentage differences became positive in 2003 and were particularly large after 2005, reaching more than 5% most years. The largest over-reporting of provincial output data was found in the gross capital formation component. The discrepancies increased sharply from only –2.5% in 1993 to 29.7% in 2016, averaging 20.4% during this period. The discrepancies are

[21] Only percentages from 1993 were reported because of the availability of provincial data. Data from some provinces (e.g., Tibet) were not available until 1993. The export and the import sector were not reported, as proportions of the net export were small over the sample period—often less than 3%.

particularly large after 2005, with the percentage differences reaching more than 22.7% in all years.

Overall, provincial governments tended to under-report their output figure in the final consumption expenditure component. The discrepancies for total final consumption expenditure exhibited an upward trend, rising from –15.7% in 2000 to –0.7% in 2016. The household expenditure sub-component followed similar patterns to the discrepancies of total final consumption expenditure. The percentage differences first decreased from –5.6% in 1993 to –13.1% in 2002, then increased to 0.3% at the end of the study period. Provincial governments heavily under-reported government expenditure from 1993 to 2002. On average, they reported 20.7% less than the nationwide GDP figure. Since 2003, provincial governments began over-reporting the provincial output number. Still, the level of over-reporting was not nearly as substantial as the level of under-reporting from 1993 to 2002.

The sectoral contribution rates calculated using the AE approach are presented in Table 2.5 and Figure C.1.4. The greatest contributor to the discrepancies between nationwide GDP and the sum of the provincial outputs was in the gross capital formation component, with discrepancy percentages ranging from –2.3% to 609% in the considered period. The gross capital formation first contributed negatively to the overall discrepancies, decreasing from –3% in 1994 to –378.9% in 2002. However, from 2003, the gross capital formation contributed positively, with an average of 253.8%. The contribution rate for total final consumption expenditure first increased from 133.4% in 1993 to its peak of 522.8% in 2002 and then sharply decreased to –41.1% in 2007. The contribution rate for total final consumption expenditure became relatively small after 2008, with the figures often approximately 10%. This trend is followed by the significant drop of discrepancies for final consumption expenditure, as well as increased discrepancies for gross capital formation. The household and government consumption expenditure presented very similar patterns to the total consumption expenditure, increasing from 54.8% and 83.1% in 1993 to 182.7% in 2000 and 463% in 2002 and then decreasing to 2.4% and –11% in 2016 respectively.

Internal inconsistency between the VA method and the AE method and discrepancies between the sum of provincial output and nationwide data reveal abnormalities. Ma et al. (2014) indicated that the data-generating process cannot explain the discrepancy. Instead, the problem of over-reporting in provincial governments may be rooted in the career incentives of provincial officials. Provincial governments play vital roles in China's economic development, carrying out more than 70% of fiscal spending to develop economic institutions and infrastructure. The provincial government leaders are appointed by the central government rather than being elected by the local electorate. Moreover, the central government has established competition among officials across the regions by using economic performance to determine their career advancement. Consequently, provincial governments pressure the provincial bureaus of statistics to inflate output levels to please the highest ranking leaders in Beijing (Xiong 2019, Hsien & Song 2019).

The NBSC may be aware of the incentives of provincial leaders to over-report output level; however, how the bureau can respond to over-reporting behaviour is unclear (Xiong 2019, Holz 2014, Ma et al. 2014). The NBSC does not indicate its awareness of the over-reporting by provincial governments, possibly because the leaders in Beijing do not want to publicly embarrass the provincial leaders, some of whom are already members of the Politburo.

Table 2.2: Sectoral Discrepancies by the VA Approach

	Nationwide GDP	Primary Sector	Secondary Sector			Tertiary Sector
			Total	Industry	Construction	
1978	-5.5%	-0.6%	2.9%	0.4%	28.3%	-27.2%
1979	-3.9%	-1.2%	2.4%	-0.9%	39.6%	-20.7%
1980	-4.1%	-2.4%	0.4%	-1.5%	16.9%	-16.3%
1981	-2.6%	-0.1%	0.6%	-1.3%	15.9%	-12.6%
1982	-0.5%	1.3%	1.5%	-1.0%	22.7%	-7.3%
1983	-0.3%	2.2%	1.1%	-0.8%	14.6%	-6.4%
1984	-1.8%	0.6%	3.3%	1.0%	20.6%	-13.3%
1985	-5.2%	0.7%	2.8%	0.6%	18.2%	-22.4%
1986	-6.7%	1.3%	-2.8%	-4.0%	3.6%	-19.5%
1987	-5.6%	0.4%	-1.2%	-1.3%	-2.3%	-17.2%
1988	-4.3%	1.8%	-1.7%	-2.0%	-1.8%	-12.9%
1989	-3.9%	1.3%	0.2%	-0.4%	2.5%	-12.9%
1990	-2.0%	1.0%	0.5%	0.0%	1.9%	-7.6%
1991	-3.4%	0.1%	-1.2%	-1.8%	1.0%	-9.3%
1992	-4.3%	-0.3%	-1.5%	-2.2%	1.5%	-10.9%
1993	-3.8%	-0.9%	-1.7%	-1.0%	-8.2%	-8.3%
1994	-6.3%	-2.7%	-4.7%	-4.1%	-10.8%	-10.4%
1995	-6.0%	-1.8%	-8.2%	-8.0%	-11.5%	-5.6%
1996	-5.1%	-1.4%	-8.2%	-7.8%	-13.2%	-2.9%
1997	-3.7%	1.4%	-6.8%	-7.2%	-6.3%	-2.0%
1998	-2.6%	1.3%	-4.0%	-4.9%	0.0%	-2.7%
1999	-2.2%	0.2%	-3.1%	-4.4%	3.8%	-2.2%

Table 2.2: Sectoral discrepancies by the VA method (continued)

	Nationwide	Primary	Secondary Sector			Tertiary
	GDP	sector	Total	Industry	Construction	Industry
2000	-1.6%	0.7%	-2.7%	-4.2%	6.2%	-1.3%
2001	-1.9%	0.0%	-2.2%	-4.1%	9.5%	-2.3%
2002	-0.7%	0.0%	-0.2%	-2.3%	12.6%	-1.7%
2003	1.5%	1.0%	3.5%	1.5%	15.7%	-0.6%
2004	3.8%	-0.7%	7.8%	6.2%	17.6%	0.5%
2005	6.4%	4.1%	10.6%	9.9%	12.5%	2.2%
2006	6.1%	3.5%	11.0%	11.2%	6.9%	1.2%
2007	3.5%	2.9%	9.7%	10.6%	0.4%	-3.1%
2008	4.3%	2.2%	11.5%	12.7%	-0.3%	-3.1%
2009	4.6%	3.1%	12.3%	14.0%	-1.2%	-3.0%
2010	5.8%	3.0%	14.8%	17.1%	-1.9%	-3.1%
2011	6.6%	2.8%	16.2%	18.8%	-3.0%	-2.7%
2012	6.7%	2.9%	16.8%	19.6%	-3.3%	-2.6%
2013	6.6%	2.8%	16.8%	20.3%	-4.0%	-3.6%
2014	6.3%	0.0%	15.5%	18.6%	-1.2%	-0.8%
2015	4.9%	0.0%	13.7%	16.3%	0.5%	-1.5%
2016	4.9%	0.0%	12.6%	15.0%	0.0%	-0.2%

Source: China Statistical Yearbook 2017.

Table 2.3: Sectoral Contribution Rate by the VA Approach

	Primary	Secondary Sector			Tertiary
	Sector	Total	Industry	Construction	Sector
1978	3.3%	-24.9%	-3.2%	-19.4%	121.6%
1979	9.9%	-29.1%	10.8%	-36.0%	119.1%
1980	17.1%	-4.8%	16.8%	-17.5%	87.7%
1981	1.6%	-10.1%	20.8%	-25.3%	108.6%
1982	-78.9%	-124.0%	71.6%	-171.7%	302.2%
1983	-246.4%	-168.4%	105.2%	-228.4%	514.3%
1984	-9.9%	-76.6%	-20.8%	-49.4%	186.4%
1985	-3.6%	-23.1%	-4.9%	-16.1%	126.7%
1986	-5.0%	18.3%	24.5%	-2.7%	86.7%
1987	-1.7%	8.9%	9.3%	2.2%	92.8%
1988	-10.7%	17.3%	18.8%	2.2%	93.4%
1989	-8.2%	-2.1%	4.2%	-3.0%	110.3%
1990	-13.8%	-10.7%	-0.4%	-4.4%	124.5%
1991	-0.5%	14.6%	20.3%	-1.4%	95.3%
1992	1.7%	14.7%	20.3%	-1.9%	90.4%
1993	4.6%	20.9%	11.1%	13.6%	74.5%
1994	8.4%	34.5%	27.8%	10.5%	57.1%
1995	5.7%	63.3%	57.3%	11.6%	31.0%
1996	5.3%	75.6%	65.8%	15.8%	19.0%
1997	-6.7%	88.0%	84.8%	9.9%	18.7%
1998	-8.5%	70.1%	77.0%	0.1%	38.4%
1999	-1.8%	63.3%	80.8%	-9.9%	38.4%

Table 2.3: Sectoral contribution rate by the VA method (continued)

	Primary	Secondary Sector			Tertiary
	Sector	Total	Industry	Construction	Sector
2000	-6.9%	77.2%	108.8%	-21.7%	32.0%
2001	0.0%	52.0%	88.4%	-26.9%	50.3%
2002	-0.2%	9.4%	120.6%	-90.6%	97.2%
2003	8.3%	104.8%	39.0%	55.6%	-16.3%
2004	-2.5%	95.8%	64.7%	25.2%	5.3%
2005	7.5%	78.3%	61.2%	10.9%	14.2%
2006	6.2%	85.9%	72.6%	6.4%	7.9%
2007	8.6%	129.0%	119.9%	0.6%	-37.6%
2008	5.3%	125.3%	116.4%	-0.3%	-30.6%
2009	6.6%	121.6%	114.3%	-1.7%	-28.2%
2010	4.9%	118.4%	111.0%	-2.1%	-23.3%
2011	4.0%	114.4%	107.2%	-3.1%	-18.4%
2012	4.1%	113.3%	106.3%	-3.3%	-17.4%
2013	4.0%	112.4%	108.2%	-4.2%	-25.5%
2014	0.0%	106.4%	101.3%	-1.3%	-6.3%
2015	0.0%	114.9%	109.2%	0.6%	-14.9%
2016	0.0%	102.0%	97.5%	-0.1%	-2.0%

The NBSC classify the secondary sector into industry, construction and other. Therefore, the sum of sectional contribution rate by industry and construction does not necessarily equal 100%. We source our GDP data from the China Statistical Yearbook 2017. GDP is in in current price and not seasonally adjusted.

Table 2.4: Sectoral Discrepancies by the AE Approach

	GDP	Final Consumption Expenditure			Gross Capital
		Total	Household	Government	Formation
1993	-4.4%	-10.2%	-5.6%	-25.9%	-2.5%
1994	-6.8%	-11.9%	-7.4%	-27.2%	0.4%
1995	-6.0%	-12.5%	-9.6%	-23.5%	4.4%
1996	-5.7%	-12.5%	-10.7%	-20.7%	7.2%
1997	-4.3%	-11.9%	-9.4%	-20.3%	13.5%
1998	-3.2%	-13.3%	-10.5%	-22.8%	19.7%
1999	-2.6%	-14.9%	-12.1%	-24.1%	19.5%
2000	-2.0%	-15.7%	-14.1%	-22.2%	20.1%
2001	-2.4%	-13.7%	-14.0%	-13.3%	13.5%
2002	-1.3%	-11.0%	-13.1%	-6.6%	13.1%
2003	0.7%	-7.0%	-9.5%	0.6%	11.0%
2004	2.9%	-4.2%	-7.5%	3.3%	13.1%
2005	5.1%	-2.6%	-4.0%	1.7%	22.7%
2006	5.6%	-1.3%	-2.6%	1.9%	25.7%
2007	3.0%	-2.4%	-4.3%	2.7%	22.8%
2008	4.8%	-1.0%	-3.3%	4.8%	22.8%
2009	4.8%	-0.2%	-2.7%	5.8%	23.0%
2010	6.6%	1.6%	-1.2%	8.5%	23.8%
2011	7.6%	-0.6%	-2.5%	4.5%	25.5%
2012	6.8%	-0.6%	-2.4%	4.3%	29.2%
2013	5.8%	-0.3%	-1.4%	2.7%	29.8%
2014	5.9%	-0.2%	-0.8%	1.5%	29.9%

Table 2.4: Sectoral Discrepancies by the AE approach (continued)

	GDP	Final Consumption Expenditure			Gross Capital Formation
		Total	Household	Government	
2015	3.8%	-0.7%	-0.6%	-0.9%	29.3%
2016	4.6%	-0.7%	0.3%	-3.6%	29.7%

Source: China Statistical Yearbook 2017. Nominal GDP is in current prices and is not seasonally adjusted.

Table 2.5: Sectoral Contribution Rate by the AE Approach

	Final Consumption Expenditure			Gross Capital
	Total	Household	Government	Formation
1993	133.4%	54.8%	83.1%	25.2%
1994	101.3%	47.8%	55.7%	-2.3%
1995	122.1%	72.3%	51.6%	-28.9%
1996	131.1%	87.4%	47.5%	-48.0%
1997	164.0%	100.5%	64.2%	-113.5%
1998	249.5%	148.7%	105.5%	-218.5%
1999	353.1%	212.1%	148.1%	-257.6%
2000	493.5%	328.4%	182.7%	-342.7%
2001	356.6%	270.6%	90.5%	-207.1%
2002	522.8%	463.2%	80.6%	-378.9%
2003	-552.1%	-557.1%	11.6%	609.0%
2004	-78.6%	-106.0%	15.6%	191.7%
2005	-27.3%	-31.3%	4.7%	182.1%
2006	-12.2%	-17.7%	4.6%	185.5%
2007	-41.1%	-53.6%	12.3%	316.4%
2008	-9.8%	-24.8%	13.2%	204.1%
2009	-1.9%	-20.2%	16.1%	223.8%
2010	11.5%	-6.3%	16.5%	172.1%
2011	-4.0%	-11.8%	7.8%	160.9%
2012	-4.6%	-13.2%	8.6%	202.8%
2013	-2.7%	-8.9%	6.2%	241.9%
2014	-1.6%	-4.9%	3.4%	236.3%

Table 2.5: Sectoral Contribution Rate by the AE approach (continued)

	Final Consumption Expenditure			Gross Capital
	Total	Household	Government	Formation
2015	-9.7%	-6.3%	-3.4%	341.7%
2016	-8.6%	2.4%	-11.0%	284.1%

The NBSC classifies the AE approach GDP in three components: final consumption expenditure, gross capital formation and net export. The contribution rate for net exports is not reported because local net exports are calculated as a residual by the provincial bureaus of statistics and are not based on trade data. We source our GDP data from the China Statistical Yearbook 2017. GDP is in in current price and not seasonally adjusted.

2.4.3 Issues with the GDP Deflator and the Real GDP Growth Rate

While the statistical authorities of most Western countries estimate the real GDP by deflating the nominal GDP with a separate, independently constructed price index, this is not the case in China. Compared to Western countries, the GDP deflator in China is inconsistently calculated[22], and it is unclear which price indices are used to calculate it (Holz 2014, Koch-Weser 2013, Kerola 2019).[23] Another anomaly is associated with the benchmark revision and economic census of the nominal GDP. The NBSC tends to perform major revisions of the nominal VA output data while leaving the real VA output

[22]Prior to 2004, the NBSC was heavily dependent on enterprise-provided, output-based implicit deflators to deflate the GDP. Since 2004, the NBSC has adopted predominantly relevant price indices to deflate its nominal VA series. The NBSC changes its real GDP deflator calculation method over time without a clear explanation of which method is used during each period.

[23]In an interview with the press in 2015, the NBSC cell member of the Community Party and the chief economist, LaiYun Shen, explained there would be different price indices for the primary, secondary and tertiary sectors and for the intermediate input and final output. For example, when considering industrial VA, final outputs were deflated using the producer price index and intermediate input was deflated using the purchase price index.

unchanged, therefore, letting the implicit deflator bear the burden. Since the NBSC has not attempted to revise price indices in the past.[24], this implies that, if the benchmark revision of the nominal output data is correct, the real GDP growth rate is systematically incorrect. Several studies claimed that China's GDP deflator is underestimated compared to the true price level. Therefore, the real GDP growth rate is exaggerated (Movshuk 2002, Young 2003, Holz 2014).[25]

This section compares the implicit GDP deflator with various price indices, affording some insight on how often the NBSC incorrectly estimates the real growth rate. Figures 2.3 to 2.6 show comparisons between a selection of alternative price indices and the GDP deflators for the primary, secondary and tertiary sectors. The selection of alternative price indices is based on how the NBSC obtains sectoral VA and justified by the availability of data.[26] The official GDP deflators for the three sectors were calculated by dividing the nominal GDP growth rate of each sector by the real GDP growth rate. The agricultural producer price index (APP), producer price index: food (PPI: food), rural retail price index (rural RPI) and rural consumer price index (rural CPI) were selected as the measures for the primary sector GDP deflator, producer price index: industrial product (PPI:industrial product) and purchasing price index were selected as the measures for the secondary sector, and self-constructed Tornqvist index and weighted average index were selected as the measure for the tertiary sector. The weighted average index for each year is defined as:

$$I_t = \sum_{j=1}^{n} w_{j,t} \cdot i_{j,t}, \tag{2.5}$$

[24]The NBSC has only made changes to how price indices are calculated (i.e., which goods are included in the CPI).

[25]Using proxies of sectoral price indices, Movshuk (2002) found that the official implicit deflator is underestimated. Therefore, the real GDP growth rate was exaggerated by approximately 2% from 1991 to 1999. Young (2003) found an approximately 1.7% downward adjustment in the real growth rate by replacing implicit sectoral deflators with proxies of sectoral price indices from 1978 to 1998.

[26]Holz (2014)'s appendix provides a detailed explanation for how the NBSC calculates real GDP deflators.

where I_t stands for the weighted average index in year t, $w_{j,t}$ is a weight of sector j and $i_{j,t}$ is the implicit deflator of sector j in year t. The Tornqvist index is defined as:

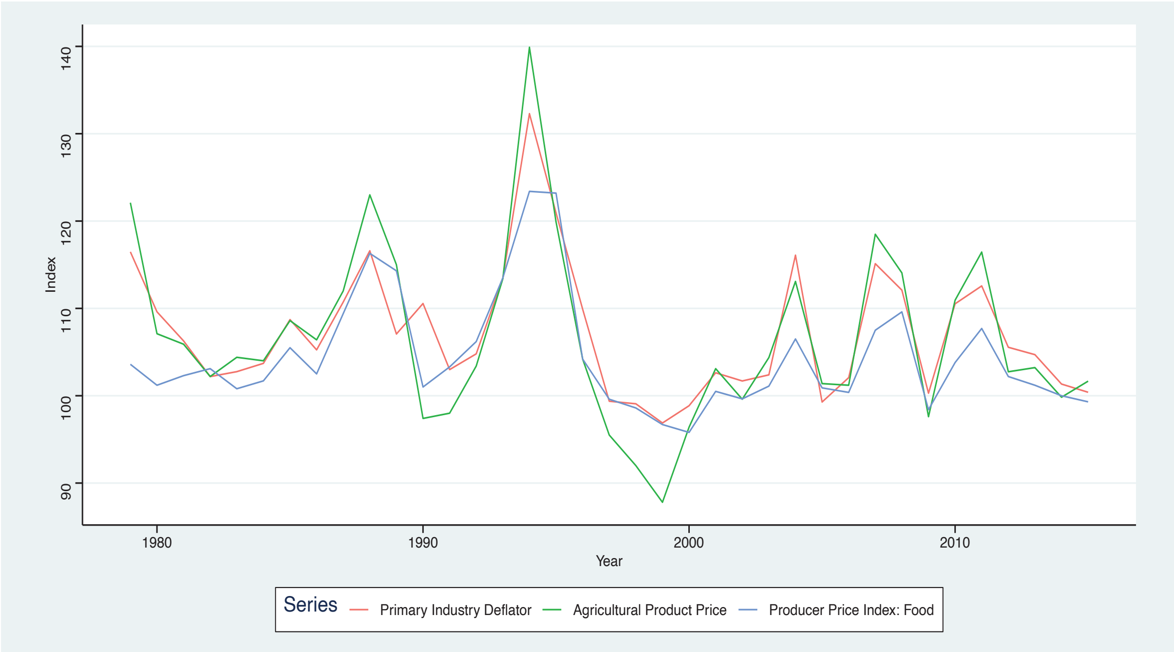
$$\frac{p_t}{p_{t-1}} = \prod_{j=1}^{n} \frac{p_{j,t}}{p_{j,t-1}}^{\frac{1}{2}[\frac{\text{value}_{j,t-1}}{\text{value}_{t-1}} + \frac{\text{value}_{j,t}}{\text{value}_t}]}, \tag{2.6}$$

where p_t is the Tornqvist index for the tertiary sector, $p_{j,t}$ is the price index for the sub-sector j in year t and value_t and $\text{value}_{j,t}$ refer to the nominal value of the tertiary sector and its sub-sectors in year t. The construction of the Tornqvist index was based on five sub-sectors in the tertiary sector, using (CPI: renting) for the real estate sub-sector, (CPI: service) for the financial intermediation sub-sector, (CPI: service) for the accommodation and catering trade sub-sector, RPI for the wholesale and retail trade sub-sector, and (CPI: transport) for the transport, storage and post sub-sector.

As in Figures 2.3 and 2.4, alternative price indices for the primary sector diverged significantly during several years, especially 1990, 1995, 2008 and 2012. The (PPI: food) index was generally smaller than the implicit deflator for the primary sector, whereas the APP index showed similar historical movements to the primary sector deflator, except between 1988 and 1990, and 1994 and 1999. The rural RPI and rural CPI were mostly smaller than the primary sector deflator in all years except 1989. Both price indices were considerably smaller than the primary sector deflator from 2003 onward. Based on the analysis above, the real growth rates of the primary sector, especially after the year 2003, are not exaggerated.

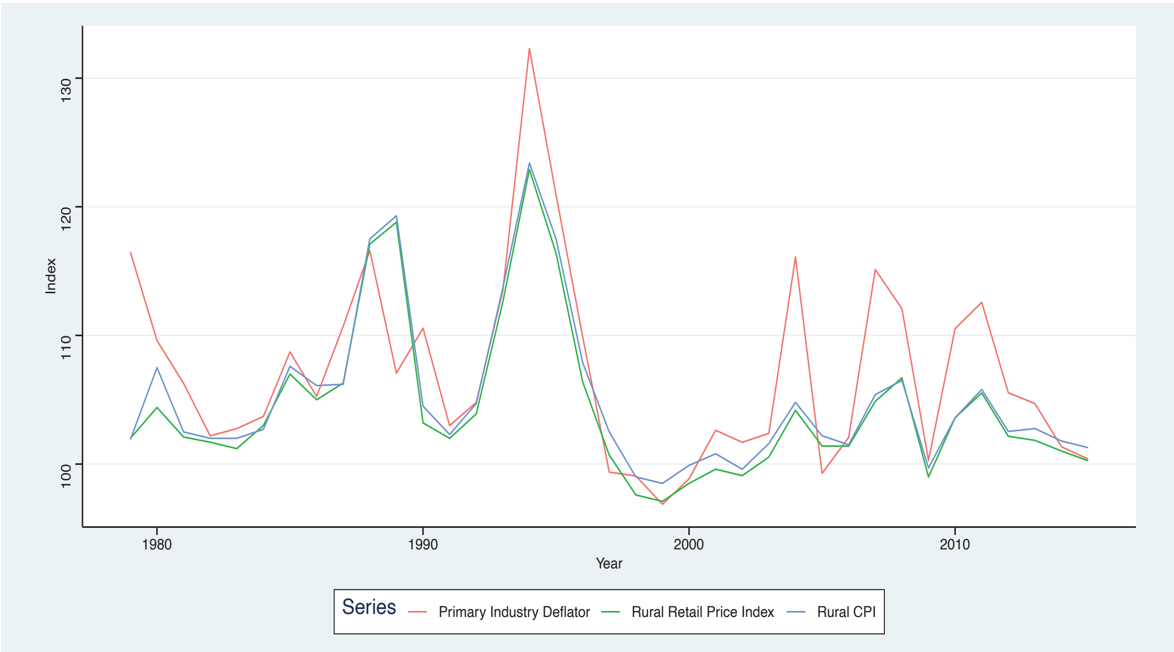
Figures 2.5 and 2.6 show the time series plots of alternative price indices for the secondary and tertiary sectors. For the secondary sector (see Figure 2.5), both price indices were higher than the official implicit deflator, especially in the years 1985, 1989, 1993, 1994 and 2004. The gap between the secondary sector deflator and the alternative price indices decreased after 2003, indicating that the exaggeration of the real growth rate

Figure 2.3: Primary Sector Deflator vs. Agriculture Product Price and PPI: food



Source: China Statistical Yearbook 2017.

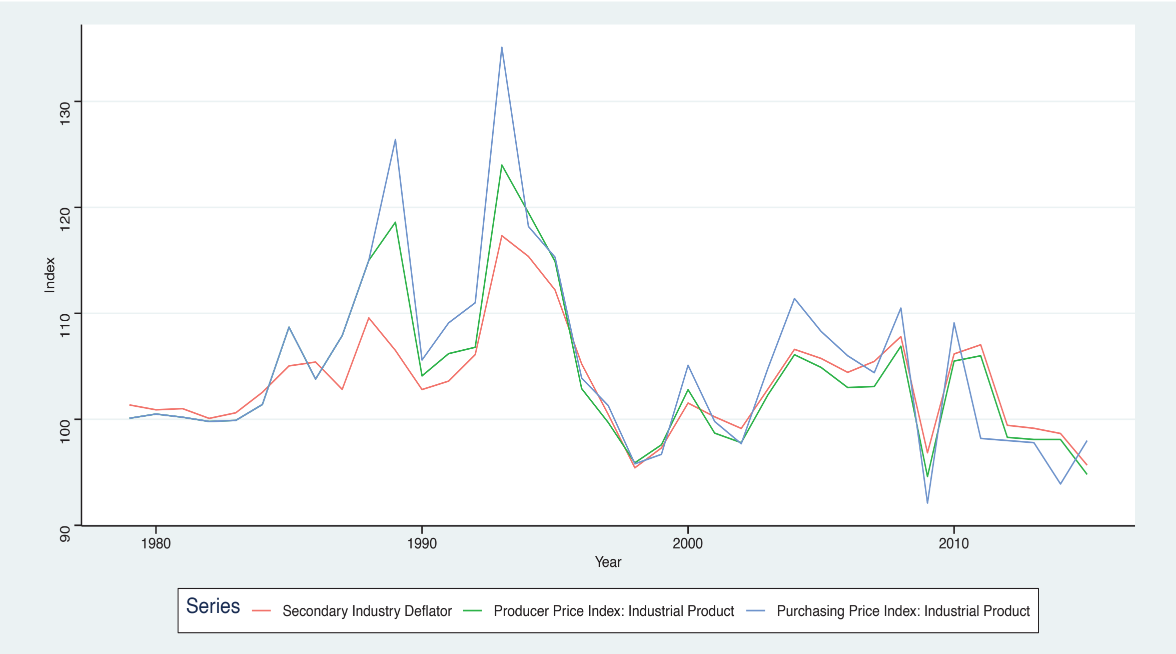
Figure 2.4: Primary Sector Deflator vs. RPI: rural and CPI: rural



Source: China Statistical Yearbook 2017.

in the secondary sector had become less severe. For the tertiary sector (see Figure 2.6), the weighted average index showed a close movement with the tertiary sector deflator, except in years 1984, 1985, 1989 and 1993. Conversely, the Tronqvist index was generally smaller than the official implicit deflator, especially after 2000.

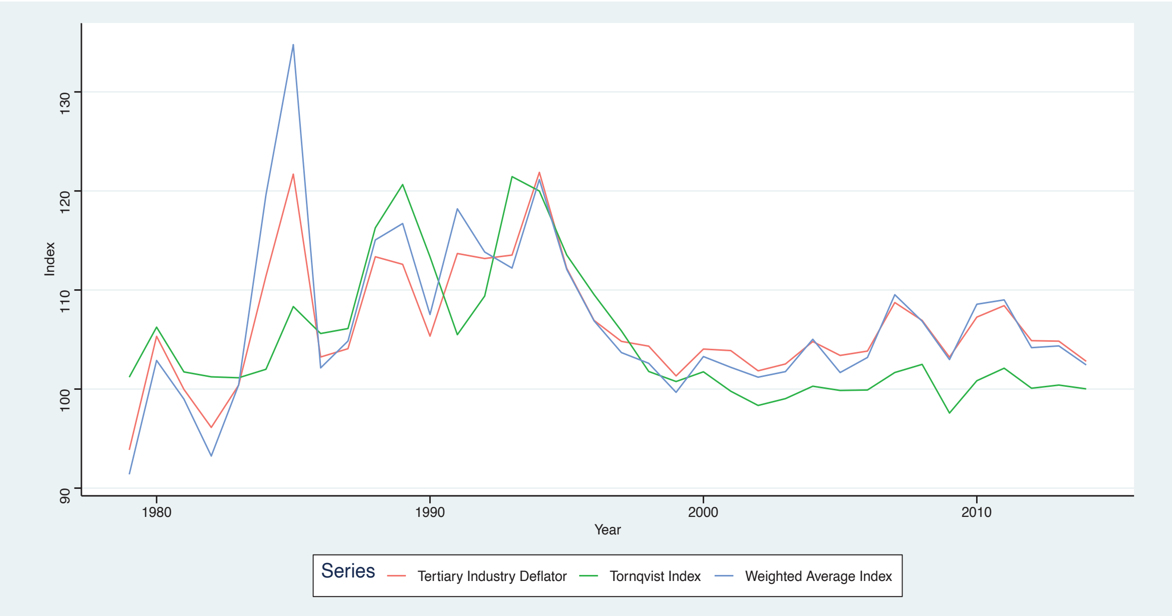
Figure 2.5: Secondary Sector Deflator vs. Producer Price Index: industrial products and Purchasing Price Index: industrial products



Source: China Statistical Yearbook 2017.

Table 2.6 presents the average annual real GDP growth rate from 1979 to 2015 in three different scenarios: 1) actual scenario, which is calculated from the Statistical Year Book 2017; 2) high real growth scenario, in which each sector's nominal VA is deflated using one of the lowest alternative price indexes; and 3) low real growth scenario, in which each sector's nominal VA is deflated using one of the highest alternative price indexes. While the official annual real growth rate is 9.7%, it could reasonably be anywhere between 8.4% and 10.6% (see Table 2.6). Based on the above analysis, it was concluded that the NBSC can obtain a relatively wide range for the real GDP growth rate by selecting different

Figure 2.6: Tertiary Sector vs. Weight Average Index and Tornqvist Index: 6 sub-sectors



Source: China Statistical Yearbook 2017.

price indices in the calculation process, which allows the NBSC to vary its final reported figure and meet the pre-announced target.

Table 2.6: Average Annual Real GDP Growth Rates for the Period 1979 to 2015 in three Different Scenarios

	Official Rate from Statistical Year Book 2017	Reasonable deflator scenario for	
		High real growth	Low real growth
Primary Sector		Rural RPI	Agricultural Product Index
Secondary Sector		Secondary Sector Deflator	Purchasing PI: Industrial goods
Tertiary Sector		Tornqvist Index	Weight Average Index
Average Annual real GDP growth rate 1979-2015	9.7%	10.6%	8.4%

Source: China Statistical Yearbook 2017.

2.5 Concluding Remarks

Understanding the reliability of China's economic statistics is a crucial challenge to overcome for the world economy. China remains open to international trade and it is the primary destination for foreign investment. Countless investors, firm managers, financial analysts and economic policy-makers depend heavily on accurate statistics to make beneficial decisions. The recent complaints followed an extended period of questioning whether China has the institutional capacity to publish accurate statistics—critics argued that official statistics overstated the economy's growth and understated trend inflation. The broad argument is that the collection of economic data behind the national accounts is under the control of local government. In Tables 2.2 and 2.4, evidence shows that local governments chose to use their power to over-report local statistics on GDP, particularly by overstating industrial output and investment since 2003. This over-reporting problem may be rooted in the career incentives of provincial officials. Also, the NBSC can obtain a relatively wide range of the GDP growth rates by selecting different price indices to calculate GDP deflators.

This paper evaluates the quality of China's official economic data and contributes to several aspects of the existing literature. First, a comprehensive review of the existing literature has been provided, focusing on the effects of political interference on the statistical reporting system and the statistical framework and data compilation methods in China. Second, a number of different price indices for evaluating the implicit GDP deflators have been constructed. Third, the discrepancies of GDP by different methods, as well as for individual sectors have been decomposed, thus making constructive solutions based on quantitative methods.

There is a considerable scope for future research. Further investigation into the issue of the quality of China's GDP statistics might proceed by studying the agency problems between China's central and local governments in affecting the economy. Using alternative indicators, such as the VA tax revenue of the local government and industrial energy consumption, to estimate the extent of over-reporting could be another future research topic. Less prone indicators, such as labour market indicators and trade data from developed countries, can be fruitful to call attention to the manipulation of official macroeconomic statistics. These alternative indicators are useful to correct ex-post GDP measures. The difference between the Chinese GDP calculated using the VA approach and the AE approach has become more pronounced in recent years; therefore, reconciling different measures of the Chinese GDP is vitally important for Chinese macroeconomics. There could also be a link between Chinese monetary policy and the quality of official statistics. The OECD suggests several tests to examine whether macroeconomic data revisions are well behaved or not. These tests consider examinations of the mean revision, median revision, adjusted t-statistic (for the significance of mean and median revisions), standard deviation of revision, root mean square revision, quartile deviation and skewness. These tests could be applied to Chinese macroeconomic data in future research.

Chapter 3

Forecasting the Chinese Macroeconomy Based on a Large Factor Model with Monthly and Quarterly Data

3.1 Introduction

As China has become the world's second-largest economy, rigorous and systematic research in the evaluation of out-of-sample forecasts of China's macroeconomy is urgently required. In this paper, we undertake the task of forecasting Chinese macroeconomic variables in a data-rich environment using a large-scale factor model with mixed frequency and missing observations components. Macroeconomic policymaking often faces the problem of forecasting the state of an economy with incomplete statistical information. Important economic variables are released at different frequencies with considerable time lags. In addition, due to the significant effect of Chinese New Year on the seasonal pattern of the data, the NBSC does not release some variables (such as industrial production) for both January and February. This break in the data dissemination pattern also affects other variables and, therefore, complicates the task of producing an accurate forecast due to the scarcity of information available over that specific time window. Another challenging issue is the quality of Chinese official economic data. The degree of data contamination has made traditional forecasting models less feasible; a high level of care must be taken regarding the selection of inputs of a specific model (Fernald et al. 2014, Higgins et al. 2016, Holz 2014). Scepticism regarding the contamination of China's official data stems from several factors, among which technical difficulties, outdated reporting systems, incomplete sample surveys and political interference are common concerns (Holz 2008, 2014, Koch-Weser 2013).[1] As suggested by Fernald et al. (2014), Stock & Watson (2002*b*), Bernanke & Boivin (2003) and Bernanke et al. (2005), the large-dimensional approximate factor model can overcome the concerns regarding such potential contamination of Chinese official economic data and the issue of mixed-frequency sampling with missing observations during Chinese New Year.[2] Moreover, as the number of time observations is small relatively to the number of economic indicators in China, pooling information from a large set of predictors can provide substantial benefits compared to standard time series forecasting models.

Factor models are a sensible way to exploit the information from a very large number of predictors. The underlying assumption is that a small number of unobservable and latent factors are the driving forces behind the state of an economy. The use of a few common factors driving all economic variables is in line with the real business cycle and dynamic stochastic general equilibrium models. This is an appealing feature for forecasting purposes since it allows researchers to concentrate on a few common factors instead of overfitting models with a potentially large number of explanatory variables. In addition, factor models ameliorate omitted variable bias, require minimal conditions on the idiosyncratic disturbances, deal with measurement errors and are relatively easy to

[1]Political economy literature often argues that data manipulation can occur at both the central government and local government levels. At the central government level, it is likely that only the cell members of Community Party of the NBSC know the choices leading to the final economic data that are presented, possibly with influences from the leaders of the state council and the Community Party (Holz 2014, Koch-Weser 2013, Xiong 2019). At the local government level, officials compete with each other for career advancement. Since economic performance is an important measure in cadres' evaluation, local officials have strong incentive to over-report key output statistics (Holz 2008, Xiong 2019).

[2]The large dimensional approximate factor model provides robustness in the presence of structural breaks and a small amount of data contamination, see Stock & Watson (2002*a*,*b*, 2016), Bai & Ng (2008*b*) for references. Not all Chinese official statistics are unreliable—only a few important variables such as GDP and industrial production are considered unreliable. Therefore, the large dimensional approximate factor model is particularly suitable in the context of macroeconomic forecasting for China.

implement.

The initial works on factor models were presented by Sargent et al. (1977) and Geweke (1977). Sargent et al. (1977) examined a large-scale macroeconometric model and concluded that two dynamic factors could explain 80% or more of the variance of major economic variables, including unemployment rate, industrial production growth, employment growth and wholesale price inflation. Geweke (1977) applied dynamic factor models to macroeconomic data and analysed these models in the frequency domain for a small number of variables. However, they imposed orthogonality on the idiosyncratic components, and such assumptions were too restrictive for the economic data.

Factor models have been improved through advances in estimation techniques proposed by Reichlin (2002), Forni et al. (2000), Kapetanios & Marcellino (2009), Stock & Watson (2002*b*), Stock & Watson (2002*a*) and Bai & Ng (2006).[3] To select the number of common factors, Bai & Ng (2002) and Hallin & Liška (2007) developed information criteria for the static factor model and the generalised dynamic factor model. Following these statistical advances, factor models have become popular in the economic literature under the name of the large-dimensional approximate factor model. In this case, '*large*' means that the sample size in both dimensions T and N tends to infinity in the asymptotic theory. '*Approximate*' means that the idiosyncratic errors are allowed to be weakly correlated across i and t. The estimated factors are useful to perform a structural analysis. They can be modelled as a dynamic structural factor model to identify the structural shocks and their dynamic impact on a large set of economic and financial indicators (Lütkepohl 2014, Van Nieuwenhuyze 2005, Bernanke et al. 2005, Bernanke & Boivin 2003, He et al. 2013). They can also be used to improve nowcasting and forecasting accuracy, possibly

[3]Stock & Watson (2002*a*) proved consistency of the principal component estimator of the static factors under the condition which the estimated factors can be treated as observed in subsequent forecasting regressions. Bai & Ng (2006) further provided improved rates for consistency of the estimator for which the factor estimated by principal component can be treated as data (that is, the error in estimation of the factors can be ignored when they are used as regressors).

in combination with autoregressive (AR) terms and/or other selected variables (Giannone et al. 2008, Stock & Watson 2002*b*, Schumacher 2007, Schumacher & Breitung 2008, Moser et al. 2007, Artis et al. 2005, Forni et al. 2003, Graff et al. 2004, Gupta & Kabundi 2011, Cristadoro et al. 2005, Giannone & Matheson 2007, Banerjee et al. 2005).

To motivate this paper, we consider the following two questions: (1) do large-dimensional factor models that are well established for forecasting Western macroeconomies also work well for China? and (2) do factor models based on mixed-frequency data with missing observations provide more accurate forecasts than factor models based on targeted predictors?[4] While evaluation of new forecasting models that have recently been proposed by other research is of considerable interest for governments, academic research, business and many other parties, the large-dimensional factor model with monthly and quarterly data is equally important for a broad program of research on Chinese macroeconomic forecasting. This is particularly relevant in light of the limited range of existing academic literature in the field.

Scholarly journals have published a small number of papers on forecasting Chinese macroeconomic variables, among which mixed-frequency datasets with missing observations are infrequently used. Several studies have used traditional models such as univariate time series and multivariate vector autoregression (VAR) models to forecast the Chinese real economic activity and inflation and have found that the growth rate of industrial value-added goods, the term structure of credit spread and the money supply are useful indicators for predicting China's macroeconomy (Zhou et al. 2013, Higgins et al. 2016, He & Fan 2015, Kamal 2013). However, these studies typically only consider a handful of predictors. Zhou et al. (2013) only incorporated the term structure of credit spreads into VAR models. Higgins et al. (2016) used only M2, inter-bank repurchase rate and

[4]We preselect variables from our mixed-frequency dataset according to Bai & Ng (2008*a*) and name them targeted predictors.

deposit rate. Kamal (2013) considered only money supply and output indicators, and He & Fan (2015) selected 14 money and credit indicators to forecast China's inflation. The literature has extended to areas of big data, but the lack of economic theory and the quality of data are two major concerns. Li et al. (2015) used the mixed-data sampling (MIDAS) method with Google search data to forecast China's inflation. They found that Google search data are strongly correlated with CPI and that the MIDAS model can outperform the benchmark autoregressive integrated moving average (ARIMA) model, with an average reduction of root mean square error of 32.9%. However, there is currently underutilisation of economic theory when using data from internet companies to predict the macroeconomy, and the quality and accessibility of Google data are questionable. Jiang et al. (2017) extracted dynamic factors from 44 monthly variables and 54 daily variables and used them in a mixed-frequency data sampling framework to predict China's GDP growth rate. However, they only covered a forecast horizon of one quarter ahead. In this case, forecasting the real GDP growth rate at one quarter ahead makes limited economic sense, as China is the only major economy that sets a rigid target for annual GDP growth.

This paper also considers a large number of forecasts than other existing studies forecasting China's macroeconomy. Mehrotra & Sánchez-Fung (2008), Lin & Wang (2013), Kamal (2013), Zhou et al. (2013) and He & Fan (2015) all performed comparisons of alternative forecasting models for China's economic indicators, but these studies only considered a small number of forecasts. Mehrotra & Sánchez-Fung (2008) and Lin & Wang (2013) only conducted forecasts up to 12 months ahead. Kamal (2013) produced nine annual out-of-sample forecasts, whereas Zhou et al. (2013) only considered six out-of-sample forecasts. He & Fan (2015) produced forecasts for five quarters. In comparison, our paper constructs at least 88 forecasts.

This study conducts a horse race among a large set of traditional forecasting models and the large-dimensional approximate factor models. We predict two measures of inflation and three measures of real economic activity. The measures of inflation are the consumer price index (CPI) and the retail price index (RPI), and the measures of real economic activity are nominal investment, nominal consumption and railway freight traffic. We use seven different forecasting models, each of which has proved useful for forecasting macroeconomic variables in Western economies. The forecasting results are presented in an out-of-sample forecasting simulation framework. In addition, we assess the performance of comparable models during the Global Financial Crisis period.

The set of univariate models covers the mean forecast, the AR(p) forecast, the AR(2) forecast, the ARMA(p,q) forecast and the ARMA(2,2) forecast, while the set of multivariate forecasts consists of specifications of the VAR model and multivariate leading indicators. Our set of factor models includes the diffusion index (DI) by Stock & Watson (2002*b*), factor-augmented autoregressive(FA-AR) models and factor-augmented vector autoregressive (FA-VAR) models. We collect our data over the December 2001 to June 2018 period and construct a 50:50 per cent split between in-sample and out-of-sample periods. To estimate factors, we use the principal component analysis for balanced panel data and the expectation–maximisation (EM) algorithm for the mixed-frequency data with missing observations. The parameter estimation for each forecasting model is based on a rolling window of 110 observations to minimise the effect of changing sample size. Then we construct one-month, three-months, six-months, nine-months and twelve-months-ahead forecasts and compare the forecasting performances by the relative mean squared forecasting errors.

The rest of this paper is organised as follows. Section 3.2 presents the forecasting methodologies of each model. Section 3.3 describes the data. Section 3.4 discusses empirical results and Section 3.4 makes concluding comments.

3.2 Factor Estimation and Forecasting Methodology

This section first outlines the estimation procedure for the common factor. The estimation procedure is presented for two cases: the case without data irregularities (in Section 3.2.1) and the case with mixed-frequency data and missing observations (in Section 3.2.2). The methodology of the forecasting models is presented in Section 3.2.3.

3.2.1 The Large-Dimensional Approximate Factor Model and Its Estimation

To motivate the representation of factor models, we can assume that the state of an economy can be represented as the sum of two mutually orthogonal components: the common component and the idiosyncratic component. The common component of each variable is a linear combination of a small number of common factors. The idiosyncratic components $e_{i,t}$ are variable specific. Theoretically, the premise of factor models is that there exist a small number of latent and unobservable factors driving the co-movements of the high-dimensional vector of macroeconomic variables.[5] We now address the representation of factor models.

Factor Models Representation

The classical factor model has been widely used in psychology and other disciplines of social science but less so in economics and finance, perhaps because the assumption that factors and idiosyncratic errors are serially and cross-sectionally independent does not align with economic data. The dynamic factor model explicitly recognises serial dependence across T and N, where N is the number of variables and T is the number of time

[5] This is also affected by a vector of zero mean uncorrelated idiosyncratic disturbances.

series observations. Let y_{t+1} denote the scalar series to be forecast and let x_{it} be a predictor variable observed for $t = 1, ..., T$ and $i = 1, ...N$. Following the work of Stock & Watson (2011), we suppose that (x_{it}, y_{t+1}) admits a dynamic factor model representation with $\bar{r}$ common dynamic factors f_t:

$$y_{t+1} = \beta(L)f_t + \gamma(L)y_t + \epsilon_{t+1}, \tag{3.1a}$$

$$x_{it} = \lambda_i(L)f_t + e_{it}, \tag{3.1b}$$

where in (3.1a) and (3.1b), $\lambda_i(L)$, $\gamma(L)$ and $\beta(L)$ are lag polynomials in non-negative powers of L.[6] Although knowledge of the dynamic factor model is useful in structural identification of the number of primitive shocks in the economy, the estimation of factors and factor loading requires the use of tools of frequency domain analysis and the proper choice of the auxiliary parameters is relatively difficult. For the purpose of multi-step forecasts, Stock & Watson (2002*b*) imposed two important modifications to (3.1a) and (3.1b). First, the lag polynomials $\lambda_i(L)$, $\gamma(L)$ and $\beta(L)$ are restricted to have finite orders of maximum q so that $\lambda_i(L)=\sum_{j=0}^{q} \lambda_{ij} L^j$ and $\beta(L) = \sum_{j=0}^{q} \beta_j L^j$. The finite lag assumption allows the dynamic factor model (3.1a) and (3.1b) to be written as a static factor model:

$$y_{t+1} = \beta' F_t + \gamma(L)y_t + \epsilon_{t+1}, \tag{3.2a}$$

$$X_t = \Lambda F_t + e_t, \tag{3.2b}$$

where in (3.2a) and (3.2b), $X_t = (x_{1t}, ..., x_{Nt})'$ is the $N \times 1$ vector of stationary variables, $F_t = (f_{1t}, ..., f_{rt})$ is the $r \times 1$ vector of unobservable factors (with $r \leq (q+1)\bar{r}$), $\Lambda = (\lambda_1', ..., \lambda_N')'$ is known as $N \times r$ matrix of factor loadings and $e_t = (e_{1t}, ..., e_{Nt})'$ is the

[6] It is assumed that $E(\epsilon_{t+1}|f_t, y_t, X_t, f_{t-1}, X_{t-1}, ...) = 0$. Therefore, in the case that f_t, $\beta(L)$ and $\gamma(L)$ were known, the minimum mean squared error forecast of y_{t+1} would be $\beta(L)f_t + \gamma(L)y_t$.

$N \times 1$ vector of idiosyncratic disturbances.[7] Second, instead of developing a vector time series model for F_t and rolling the forecasts $y_{t|F_t}$ forward, which entails estimating a large number of parameters that could erode forecasting results, the construction of the h-step-ahead forecast can be achieved by linear combination of F_t and y_t (with or without its lags) (Marcellino et al. 2006). The resulting adopted multi-step-ahead forecasting equation is:

$$y^h_{t+h} = \alpha_h + \beta_h(L)F_t + \gamma_h(L)y_t, \tag{3.3}$$

where in (3.3), y^h_{t+h} is the h-step-ahead variable to be forecast, the constant term is introduced explicitly and the subscripts reflect the dependence of the projection on the horizon.

Rewriting the dynamic factor as static format yields two advantages. First, the properties of the estimated static factors are easier to understand from a theoretical standpoint. Second, the static factor model can be easily estimated using time domain methods such as principal component analysis and involves fewer and easier choices regarding auxiliary parameters.

Restrictions and Identification

Imposing normalisation conditions on factors and factor loadings is important to pin down the rotational indeterminacy. When the number of time observations T and the number of variables N is large, we can treat Λ and F as parameters and estimate Λ and F simultaneously.[8] Considering estimating a compact formulation of (3.2b):

$$X = F\Lambda' + e, \tag{3.4}$$

[7] In factor model literature, r is known as the true number of factors whereas k refers to the estimated number of factors. Typically, r is much smaller than N and r does not necessarily equal k.

[8] After obtaining F and Λ, the residual matrix can be obtained from $e = X - F\Lambda'$.

where X is the $T \times N$ matrix of stationary variables, $F = (F_1, ...F_T)'$ is the $T \times r$ matrix of unobservable factors, Λ is the $N \times r$ matrix of the factor loadings and e is $T \times N$ matrix of the idiosyncratic error. For an arbitrary non-singular $r \times r$ matrix A, we have $F\Lambda' = FAA^{-1}\Lambda' = F^*\Lambda^{*\prime}$, where $F^* = FA$ and $\Lambda^* = \Lambda A^{-1\prime}$. (3.4) is observationally equivalent to $X = F^*\Lambda^{*\prime} + e$, and F and Λ are not separately identifiable; therefore, restrictions are required to uniquely determine F and Λ. Since the arbitrary matrix A has r^2 degree of freedom, at least r^2 restriction is required to remove the indeterminacy. Bai & Ng (2008*b*) suggested normalisation of $T^{-1}FF' = I_r$ and $\Lambda\Lambda'$ being diagonal or, alternatively, $N^{-1}\Lambda\Lambda' = I_r$ and FF' being diagonal. Since the method of principal components estimates space spanned by latent factors instead of factors themselves, further restrictions are required for F and Λ. Following the methods of Bai & Ng (2013), $\Lambda\Lambda'$ is restricted being a diagonal matrix whose diagonal elements are distinct entries, positive, and arranged in decreasing order. Such a restriction is referred to as a PC1 restriction by Bai & Ng (2013). Under PC1 conditions, the normalisation of factors provides $r \times (r+1)/2$ restrictions, whereas the diagonal matrix Λ' gives $r \times (r-1)/2$. Together, the two normalisations lead to exactly r^2 restrictions.

Uniqueness (or identification) can be ensured by obtaining PC1 restrictions. The requirements that the factor variance must be distinct, positive and arranged in descending order ensure that columns of factor loadings cannot be simply reordered. The first factor has the largest variance, explaining the largest part of the variance of X_{it}, and the second factor has the second-largest variance, and so on. If factor loadings (or factors) were known and normalised according to the work of Bai & Ng (2013), a natural estimator for factors (or factor loadings) could be obtained by left-multiplying with Λ' (or F') and dropping out the idiosyncratic term, as below:

$$F_t = \Lambda' X_t. \tag{3.5}$$

In reality, both factors and factor loadings are often unknown. Following the work of Stock & Watson (2002*b*), Bai & Ng (2002), Bai & Ng (2008*b*) and Schumacher & Breitung (2008), we use the method of asymptotic principal components analysis to estimate the factors and factor loadings simultaneously.

Asymptotic Principal Components Analysis

Asymptotic principal components analysis is one of nonparametric methods that estimate the static factors in (3.2b) without specifying a model for the factors or assuming specific distributions for the disturbances. These approaches use weighted cross-sectional averaging to remove the influence of the idiosyncratic disturbances. Such weights are the resulting factor estimators explaining as much data variance as possible. The estimator of the asymptotic principal component analysis for F and Λ can be treated as a direct outcome of the least squares problem under normalisation conditions of PC1.

For any given k not necessarily equal to the true number of factors r, the method of principal component analysis constructs a $T \times k$ matrix of estimated factors and a corresponding $N \times k$ matrix of estimated loadings by solving the following minimisation problem:

$$\min_{\lambda^{k'} F^k} S(k), \quad \text{where} \quad S(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i^{k'} F_t^k)^2, \tag{3.6}$$

subject to either the normalisation that $\Lambda^{k'}\Lambda^k/N = I_k$ and $F^{k'}F^k$ is diagonal, or the normalisation that $F^{k'}F^k/T = I_k$ and $\Lambda^{k'}\Lambda^k$ is diagonal. The estimates can be obtained by concentrating out F^k (Stock & Watson 2002*a*, Bai & Ng 2008*b*). Minimising (3.6) is equivalent to maximising:

$$V(\Lambda^k) = \text{tr}(\Lambda^{k'} X' X \Lambda^k), \tag{3.7}$$

subject to normalisation that $\Lambda^{k'}\Lambda^k/N = I_k$, where $tr(\cdot)$ denotes the matrix trace. This is similar to solving the classical principal component problem, where λ^k is equal to $\sqrt{N}$

times the eigenvectors corresponding to the k largest eigenvalues of the matrix $X'X$. Using a normalisation condition of $\Lambda^{k'}\Lambda^{k}/N = I_k$ yields $F^k = X\Lambda/N$[9].

Determining the Number of Factors in Approximate Factor Models

A critical issue to the validity of the large-dimensional approximate factor model is the correct specification of the number of static factors k. Traditional information criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), which are functions of N or T alone, provide inconsistent estimations of the number of static factors when both N and T diverge.[10] Bai & Ng (2002) argued that the penalty function for overfitting must include both N and T. They pioneered the literature by proposing information criteria to take into account both dimensions of the dataset as arguments of the function penalising over-parametrisation. Their information criteria can consistently estimate the true number of statics factors; hence, they should be given priority over the AIC or BIC (Stock & Watson 2016).

Let g(N,T) be a penalty function. Bai & Ng (2002) defined the information criteria as:

$$PCP(k) = S(k) + k\sigma^2 g(N,T), \tag{3.8}$$

$$IC(k) = \ln(S(k)) + kg(N,T), \tag{3.9}$$

where $S(k)$ is the sum of squared residuals (divided by $N \times T$) in (3.6) when k factors are estimated. σ^2 is equal to S($kmax$) for a pre-specified value $kmax$, then the estimator

[9]Another solution to obtain F^k and Λ^k is to concentrate out Λ^k and use the normalisation condition of $F^{k'}F^k/N = I_k$. This approach is less intensive when $T < N$, which is useful for the targeted factor models when T=199 and N=41.

[10]Stock & Watson (2002*b*) used a modified version of the BIC to select the optimal number of factors for forecasting a single series, but this requires $N >> T$. Additionally, the selected factors can have no predictive ability for an individual data series.

for the number of factors is defined as:

$$\begin{aligned}\widehat{k}_{PCP} &= \underset{0\leq k\leq kmax}{\text{argmin}}\ PCP(k),\\ \widehat{k}_{IC} &= \underset{0\leq k\leq kmax}{\text{argmin}}\ IC(k).\end{aligned} \tag{3.10}$$

The condition $C_{NT}^2 g(N,T) \to 0$ as $N,T \to 0$ would require divergence of $g(N,T)$. That is, the convergence of estimated factor space requires that $\min\limits_{N,T} g(N,T)$ must diverge. In a practical panel setting, examples of $g(N,T)$ that satisfy the required conditions are:

$$\begin{aligned}g_1(N,T) &= \frac{N+T}{NT}\ln\left(\frac{NT}{N+T}\right),\\ g_2(N,T) &= \frac{N+T}{NT}\ln C_{NT}^2,\\ g_3(N,T) &= \frac{\ln C_{NT}^2}{C_{NT}^2},\\ g_4(N,T) &= \frac{(N+T-k)\ln(NT)}{NT},\end{aligned} \tag{3.11}$$

where $C_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$. The literature has frequently used $g_2(N,T)$ in empirical work as it tends to be more stable in practice (Bai & Ng 2002, 2008*b*, Stock & Watson 2011, 2006). According to Bai & Ng (2002)'s work, $g_4(N,T)$ possesses good properties, especially when errors are cross-sectionally correlated. From a statistical standpoint, $g_4(N,T)$ fails when $T = \exp(N)$ or $N = \exp(T)$, but these configurations of N and T do not seem empirically relevant. Therefore, $g_4(N,T)$ should not be ruled out in practice, particularly when the errors are cross-sectionally correlated. We use $g_2(N,T)$ to determine the number of factors in our factor models.

3.2.2 The Expectation–Maximisation Algorithm for Unbalanced Data

Macroeconomic forecasting often encounters a situation in which a dataset may be highly unbalanced, possibly due to missing observations and different sampling frequencies. In practice, policymakers and economic forecasters often want to consider the most recent economic information for forecasting; however, as economic information is released with different sampling frequencies and publication lags, the standard principal component estimator cannot be applied in such circumstances. The EM algorithm is an iterative method for efficient estimation and can consistently estimate Λ^k and F^k when data irregularities are present.[11]

When data are unbalanced, minimisation of (3.6) through standard eigenvector decomposition does not apply; instead, the least squares estimators of F_t and Λ are achieved by solving the following minimisation problem:

$$V(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} I_{it}(x_{it} - \lambda_i^{k'} F_t^k)^2, \tag{3.12}$$

where $I_{it} = 1$ if X_{it} is observed and $I_{it} = 0$ if X_{it} is unobserved. Minimising (3.12) requires use of iterations and sometimes may be difficult to solve analytically (Stock & Watson 2002*a*). Stock & Watson (2002*b*) proposed a simple EM algorithm to estimate the space spanned by factors when a dataset is unbalanced. Let $\widehat{F}^{(j-1)}$ and $\widehat{\Lambda}^{(j-1)}$ denote the estimated factors and factor loadings at $(j-1)^{th}$ iteration of the EM algorithm. Then, the estimates of factors and factor loadings at j^{th} iteration solve the following minimisation

[11]The EM algorithm is also used by McCracken & Ng (2016) to build the balanced FRED-MD dataset. The FRED-MD is a large macroeconomic database designed for the empirical analysis of 'big data' and has been widely used by macroeconomists.

problem:

$$Q(X^{obs}, \widehat{\Lambda}^{(j-1)}, \widehat{F}^{(j-1)}, F, \Lambda) = \sum_{i=1}^{N}\sum_{t=1}^{T}\left\{E_{\widehat{F}^{(j-1)},\widehat{\Lambda}^{(j-1)}}(x_{it}^2|X^{obs}) + (\lambda_i^{k'}F_t^k)^2 - 2\widehat{x}_{it}(\lambda_i^{k'}F_t^k)\right\}, \tag{3.13}$$

where X^{obs} is the full set of observed data and $\widehat{x}_{it} = E_{\widehat{F},\widehat{\Lambda}}(x_{it}|X^{obs})$. The first term on the right side of (3.13) does not depend on either factors or factor loadings. This implies that, for the minimiser of (3.13) is proportional to the minimiser of:

$$\widehat{V}(F, \Lambda) = \sum_{i=1}^{N}\sum_{t=1}^{T}(\widehat{x}_{it} - \lambda_i' F_t)^2. \tag{3.14}$$

At the j^{th} step of convergence, this reduces to the usual principal component analysis wherein the missing data are replaced by their expectation, conditional on the observed data and using the parameter values from the previous iteration. Following the work of Stock & Watson (2002*b*), we proceed with the EM algorithm as follows:

1. Initial step: provide an initial estimate of missing observations. In our application, observations that are missing are initialised to the unconditional mean based on the non-missing values (which is zero, since the data are demeaned and standardised). Then, we use standard principal component analysis to estimate factors and factor loadings.

2. E-step: For j^{th} iteration, given the estimated factors and loadings from the previous $(j-1)^{th}$ iteration, compute the updated estimate of missing monthly series and monthly estimate of quarterly series by the expectation of X_i ($T \times 1$ vector), conditional on the observed data X_i^{obs} and the previous iteration factors and loadings for

variable i according to:

$$\widehat{X}_i^j = E(X_i | X_i^{obs}, \widehat{F}^{(j-1)}, \widehat{\lambda}_i^{(j-1)}) \quad (3.15)$$
$$= \widehat{F}^{(j-1)} \widehat{\Lambda}_i^{(j-1)} + A_i'(A_i A_i')^{-1}(X_i^{obs} - A_i \widehat{F}^{(j-1)} \widehat{\Lambda}^{(j-1)}),$$

where $\widehat{F}^{(j-1)}$ is the $T \times r$ matrix of factors estimated from the previous $(j-1)^{th}$ iteration, $\widehat{\lambda}_i^{(j-1)}$ is the i^{th} row of $\widehat{\Lambda}^{(j-1)}$ and A_i is a time aggregation matrix. In (3.15), part $\widehat{F}^{(j-1)} \widehat{\Lambda}_i^{(j-1)}$ is the common component from the previous iteration and the term $(X_i^{obs} - A_i \widehat{F}^{(j-1)} \widehat{\lambda}^{(j-1)})$ is the low-frequency idiosyncratic component, distributed by the projection coefficient $A_i'(A_i A_i')^{-1}$. If X_i contains no missing values, A_i is the identity matrix and (3.15) simply becomes $\widehat{X}_i = \widehat{F}\widehat{\Lambda}_i + (X_i^{obs} - \widehat{F}\widehat{\Lambda}_i) = X_i^{obs}$. Therefore, for time series without data irregularities, no EM iteration is required.

3. M-step: Repeat the E-step for each variable i. The estimated monthly observations for quarterly series, the estimated monthly missing observations and the monthly series without data irregularities are combined into a new $T \times N$ dataset. Re-estimate factors and factor loadings by principal component analysis. Go back to the E-step until convergence. In our application, we stop the algorithm if the change in the objective function (3.13) is smaller than 10^{-5}.

Time Aggregation Matrices

Correctly specifying A_i in (3.15) is key to the EM algorithm. When dealing with mixed-frequency datasets, transformation of quarterly observations into monthly estimates must be performed properly. If missing observations occur (e.g., at the end of the sample period), further transformation is required. This transformation procedure is called *mapping* by Schumacher & Breitung (2008). We distinguish X^{obs} to be the $T^{obx} \times 1$ vector of available observations and X_i to be the $T \times 1$ vector of all observations (including available and missing) for the variables i_{th}. Then, the function of mapping can be written as a linear relationship as below:

$$X_i^{obs} = A_i X_i, \tag{3.16}$$

where A_i is a $T^{obs} \times T$ matrix that tackles missing values or different sampling frequencies. Next, we provide three examples to demonstrate specifications of A_i for different variables.

Example 1: Missing Observations at the End of the Sample

If missing observations occur, rows of the identify matrix corresponding to the missing values of X_i should be removed. For instance, if the last observation in X_i is not available, the last row of the identity matrix A_i corresponding to the missing value of X_i should be removed and the mapping matrix A_i becomes:

$$A_i = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \tag{3.17}$$

Example 2: Gross Domestic Product as a Monthly Flow Variable

The key output indicator GDP is usually a quarterly flow variable, and we let the natural log of GDP be denoted as y_t^q. If we assume that the observable quarterly data are averages of unobservable monthly series, the underlying relationship between the observed natural log of quarterly GDP y_t^q and the unobserved monthly GDP in the natural log can be written as:

$$y_t^q = (1/3)(y_t^m + y_{t-1}^m + y_{t-2}^m), \tag{3.18}$$

where t is the monthly time index and (3.18) holds for t = 3,6,9,...T, assuming that quarterly observations are available in the last month of the quarter. If time index T is not a multiple of three, then the last row (or last two rows) is missing observations. The theory outlined in the section ‘Factor Models Representation’ requires stationary time series; therefore, it may be useful to convert quarterly GDP in the natural log to the growth rate of GDP and rewrite (3.18) as:

$$\begin{aligned}\Delta^q y_t^q &= (1/3)(y_t^m + y_{t-1}^m + y_{t-2}^m) - (1/3)(y_{t-3}^m + y_{t-4}^m + y_{t-5}^m) & (3.19)\\ &= (1/3)(\Delta y_t^m + 2\Delta y_{t-1}^m + 3\Delta y_{t-2}^m + 2\Delta y_{t-3}^m + \Delta y_{t-4}^m), & (3.20)\end{aligned}$$

Rewriting (3.18) to (3.19) is a standard approach and is often referred to as GDP interpolation in the literature (see, for example, Marcellino et al. (2003) and Schumacher & Breitung (2008)). This yields the relationship between the quarterly GDP growth rate and the monthly GDP growth rate of $\Delta^q Y^q = A_y \Delta Y^m$ and implicitly defines the rows of A_y as

$$A_y = \frac{1}{3}\begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdots & 0 & 1 & 2 & 3 & 2 & 1 & 0 & 0 & 0 \\ \cdots & 0 & 0 & 0 & 0 & 1 & 2 & 3 & 2 & 1 \end{pmatrix}, \tag{3.21}$$

where $\Delta^q Y^q = (\cdots, \Delta^q y_{T-6}^q, \Delta^q y_{T-3}^q, \Delta^q y_T^q)'$ and the corresponding monthly GDP growth rates $\Delta Y^m = (\cdots, \Delta y_{T-2}^m, \Delta y_{T-1}^m, \Delta y_T^m)'$. If more timely monthly observations are available at the end of the sample than quarterly GDP series, the rows of mapping matrix A_y corresponding to missing values of the quarterly GDP must be removed, and the final three columns of A_y will become zero vectors.

Example 3: Business Survey Index as Quarterly I(0) Stock Variable

The quarterly business survey index is a stock variable as it is a point-in-time figure at the end of each quarter, measuring how confident business managers are given current economic conditions. The range of the quarterly business survey index is relatively small—usually between 45 and 55 in our study. It is an I(0) process as verified by unit root tests.[12] If the quarterly series is a stock variable in the I(0) process, it can be treated as a monthly series with missing observations occurring in the first and second months of each quarter. The corresponding mapping matrix A_i is then:

$$A_i = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & 0 & 1 \end{pmatrix} \tag{3.22}$$

.

3.2.3 Forecasting Models

Our choice of forecasting models was guided by two considerations. First, our interest is to assess whether large-dimensional factor models with ragged-edge and mixed-frequency datasets that work well for Western countries can also improve forecasting of China's macroeconomy. Second, we aim to investigate whether univariate and multivariate models can perform as well as factor models. We settled on the following set of models.

Mean: In a mean forecast approach, h-month-ahead predictions are equal to the arithmetic mean of all available observations up to and including the period in which the forecast is made.

AR: The second univariate model is the AR model.[13] The number of AR lags can be determined using the sequential downward t test or information criteria such as the

[12]The unit root test is performed using the adf.test function in R.
[13]Detailed discussion of the AR model can be found in many textbooks (e.g., Stock et al. (2003)).

AIC and the BIC. The h-month-ahead AR forecasting model is:

$$\widehat{y}_{t+h} = \widehat{\alpha}_h + \sum_{j=1}^{p} \widehat{\gamma}_{jh} y_{t-j+h}. \tag{3.23}$$

We propose two variants of the AR model. For the first, (AR(p)), we use the AICc to determine the number of lags with maximum allowed orders of 12.[14] For the second, (AR(2)), the lag order of the AR term is (arbitrarily) set equal to 2.[15] The distribution of the lag length chosen by AICc is presented in Figure C.3.1 of Appendix C.

Autoregressive moving average (ARMA): Extending the AR model to an ARMA model that considers the autocorrelation in error terms is relatively straightforward. The h-step-ahead ARMA forecast is:

$$\widehat{y}_{t+h} = \widehat{\alpha}_h + \sum_{j=1}^{p} \widehat{\gamma}_j y_{t-j+h} + \sum_{i=1}^{q} \widehat{\delta}_j \varepsilon_{t-i+h} \tag{3.24}$$

We fit two types of ARMA models. For the first type, we select the order of p and q based on the AICc with maximum allowed orders of 12, denoting this as ARMA(p,q). For the second type, we arbitrarily set the value of p and q equal to 2 and denote it as ARMA(2,2).

VAR: The VAR model is one of the most successful, flexible, and easy-to-use models for the analysis of multivariate time series. It has proved especially useful for describing

[14] AIC is commonly known to over-parameterised the model and choose long lags, which would, in general, lead to inferior forecasting performance (Findley & Wei 2002, Konishi & Kitagawa 2008). This is particularly problematic when the sample size is small. AICc can be used to correct the bias lag selection of small sample sizes. Robustness checks using AIC, AICc, BIC, the Hannan-Quin criterion and the Schwarz information criterion show similar forecasting results.

[15] AR(p) model is the most commonly used univariate model for macroeconomic forecasting, see for example Stock & Watson (2002*a*, 1998), Gupta & Kabundi (2011), Artis et al. (2005), Schumacher (2007) A specific lag (i.e. lag=2) is often used for robustness check.

the dynamic behaviour of economic time series and for undertaking macroeconomic forecasting (Lütkepohl 2014, Stock & Watson 2017). We follow the recommendations of Stock & Watson (2002*b*) by specifying the vector to contain an indicator for inflation, an indicator for real economic activity and an indicator for monetary policy instruments. We use the growth rate of money supply M2, rather than the interest rate, in our VAR model. Such a choice regarding the monetary policy instrument is based on findings by Higgins et al. (2016), He et al. (2013) and Chen et al. (2016). Chen et al. (2016) and Higgins et al. (2016) found that the Chinese government is effective at controlling bank credit via M2 growth to influence investment and, therefore, GDP growth. He et al. (2013) found that the repurchase rate, the benchmark lending rate and a market-based monetary stance have little effect on the Chinese economy, but that non–market based measures such as growth rates of total loans and money supply are effective in adjusting the real economy and price level. Therefore, to forecast an inflation variable, the vector Y_t consists of the variable to be forecast, the industrial production index and the growth rate of M2. To forecast a real economic variable, the vector Y_t includes the variable to be predicted, CPI and the growth rate of M2. Multi-step forecasts are computed by iterating the VAR forwards. The general formula for the h-month-ahead point forecast is:

$$\widehat{z}_{t+h} = \widehat{\phi}_0 + \widehat{\Phi}_1 z_t ... + \widehat{\Phi}_p z_{t-p} \tag{3.25}$$

where $\widehat{z}_{t+h}$ is the vector of k variables (in our study k=3) for h-step-ahead forecasts, p is the lag order of the VAR, $\widehat{\phi}_0$ is a $(k \times 1)$ vector of constants, $\widehat{\Phi}$ are $(k \times k)$ coefficient matrices. We consider two implementations of the VAR model. For the first implementation, the order of the VAR model is set equal to 2. For the second implementation, the order is chosen by an AICc with a maximum allowed order of 4.[16]

Multivariate Leading Indicator: Leading indicators are useful for predicting future economic conditions and have been used widely to predict output and inflation in the economic forecasting literature (e.g., Stock & Watson (1989, 2002*b*), Granger et al. (2006), Stock & Watson (2008)). The multivariate leading indicator forecasts have the form:

$$\widehat{y}^h_{T+h|T} = \widehat{\delta}_{h0} + \sum_{j=1}^{m} \widehat{\delta}'_{hi} W_{T-j+1} + \sum_{j=1}^{p} \widehat{\gamma}_{hj} y_{T-j+h}, \tag{3.26}$$

where W_t is a vector of leading indicators that have been featured in our forecasting applications and $\widehat{\delta}_{hj}, j = 0, ..., m$ are ordinary least squares (OLS) coefficient estimates.

For the real activity forecasts, the vector of leading indicators W_t consists of the Organisation for Economic Co-operation and Development (OECD) composite leading indicator, the change of railway freight turnover, the growth rate of fixed asset investment and the growth rate of imports. For the inflation forecasts, the W_t includes the change in the outstanding loans, the OECD composite leading indicators, the change in retail sales of consumer goods and the one-year benchmark deposit rate set by the People's Bank of China. In all cases, the leading indicators are transformed such that W_t is $I(0)$ and screened for potential outliers. We propose two variants of multivariate leading indicator forecasts. For the first variant, MLD(2), the lag orders are arbitrarily set to 2. For the second variant, MLD(p), the lag orders in each variable to be forecast are determined by a recursive AICc with maximum

[16] We use AICc rather than AIC to avoid the issue of over-fitting the VAR model. A robustness analysis is done using AIC, AICc, BIC, the Hannan-Quin criterion and the Schwarz information criterion. These information criteria show similar forecasting results.

allowed orders of 6.

FA-AR: The factor models we use are based on the diffusion index developed by Stock & Watson (2002*b*, 1998). The distinctive feature of this index is that it adds latent factors to a standard AR model. The general formula for the h-month-ahead point forecast model is:

$$\widehat{y}_{t+h|t} = \widehat{\alpha}_h + \sum_{j=1}^{m} \widehat{\beta}'_{hj} \widehat{F}_{T-j+1} + \sum_{j=1}^{p} \widehat{\gamma}_{hj} y_{t-j+h}, \tag{3.27}$$

where $\widehat{y}_{t+h|t}$ is the h-step-ahead variable to be forecast, the constant term is introduced explicitly, $\widehat{F}_t$ are factors, y_{t-j+1} is a lag of y_t, $\widehat{\gamma}$ and $\widehat{\beta}$ are the coefficients associated with the lags and the subscript t reflects the time horizon of the variable. The number of factors is determined using information criteria developed by Bai & Ng (2002), with the maximum allowed number of factors set to 6. We consider three variants of (3.27). The first and the second, denoted as models FA-AR(p) and FA-AR(2), include contemporaneous $\widehat{F}_t$ and a lag of $\widehat{y}_t$ (i.e., m = 1). The lag order p in the first variant FA-AR(p) is chosen by the AICc with the maximum order of 6, whereas in the second variant FA-AR(2), it is arbitrarily set to be 2. The third, denoted as DI, includes only contemporaneous $\widehat{F}_t$ (so m = 1 and p = 0).[17]

FA-VAR: Our study proposes a stochastic process to capture linear interdependencies among variables to be predicted and latent factors. Bernanke et al. (2005) suggested a VAR framework for factors and factor loadings wherein the vector contains only a single observable variable plus unobservable and latent factors. Let r denote the number of static factors and q denote the number of dynamic factors, we represent the FA-VAR model as in Stock & Watson (2016) so that the factors and factor loadings can be estimated by least squares and all the restrictions can been seen

[17]Bai & Ng (2002) estimate the number of factor to be 1.

clearly as follows:

$$X_t = \Lambda \begin{pmatrix} Y_t \\ F_t \end{pmatrix} + u_t,$$

$$F_t^+ = \Phi(L)F_{t-1}^+ + G\eta_t, \quad \text{where } F_t^+ = \begin{pmatrix} Y_t \\ F_t \end{pmatrix}, \tag{3.28}$$

$$\eta_t = H\varepsilon_t.$$

where $\Phi(L)$ is the lag operator, $\overset{r\times q}{G}$ and $\overset{r\times r}{H}$ are restriction matrices, $\overset{q\times 1}{\varepsilon_t}$ is the vector of structural shock, $\overset{q\times 1}{\eta_t}$ is the vector of innovations to factors. The h-months-ahead FA-VAR forecasting models is then:

$$\widehat{z}_{t+h} = \widehat{\phi}_0 + \widehat{\Phi}_1 z_t ... + \widehat{\Phi}_p z_{t-p} \tag{3.29}$$

where $\widehat{z}_t$ is the vector of variables to be forecasted and latent factors, p is the lag order of the FA-VAR, $\widehat{\phi}_0$ is a vector of constants, $\widehat{\Phi}_i$ are coefficient matrices. We fit the FA-VAR model with only one factor for two reasons. First, the information criteria developed by Bai & Ng (2002) suggests that $k = 1$ (one factor) minimises the penalty function.[18] Second, we tend to avoid the problem of overfitting the VAR model. Since most VAR models are estimated using symmetric lags (i.e., the same lag length is used for all variables in all equations of the model), a large number of variables would result in a significantly large number of parameters to be estimated. Jolliffe (1993) indicated that overfitting the VAR model causes an increase in the mean squared forecasting error. Therefore, we use only one factor and also restrict

[18]Selection of the number of factors for FAVAR models could be problematic, especially in the presence of structural instability. Mao Takongmo & Stevanovic (2015) documented that, in the presence of structural instability, many selection methods typically overestimate the number of factors.

the maximum number of p and q to 4. Two variants of the FA-VAR are considered: FA(1)-VAR(2), wherein the lag orders p and q are arbitrarily set equal to 2, and FA(1)-VAR(p), wherein the lag orders p and q are selected using an AICc with a maximum allowable lag of 4.

We generate FA-VAR forecasts directly rather than forecasts of the common and idiosyncratic components separately. As Boivin & Ng (2005) suggested, when data-generating processes and their dynamic structures are known, no single method stands out to be systematically superior. However, when they are unknown, direct and unrestricted factor forecasts that impose fewer constraints and estimate a smaller number of auxiliary parameters appear to be less vulnerable to misspecification compared to indirect and restricted factor forecasts—leading to improved forecasts.

Targeted FA-AR and FA-VAR: A critique of factor models is that not all series in X_t are relevant for predicting y_{t+h}. Instead of estimating factor space based on hundreds of predictors, Bai & Ng (2008*a*) suggested that a preselected subset of X_t that are relevant for forecasting y_{t+h} yield a better forecasting performance. Bai & Ng (2008*a*) proposed two ways to construct the subset X_t^*:

- Hard threshold (OLS):

$$y_t = \alpha + \sum_{j=0}^{m} p_j y_{t-j} + \beta_i X_{i,t} + \epsilon_t, \tag{3.30}$$

$$X_t^* = \{X_i \in X_t | t_{Xi} > t_c\}. \tag{3.31}$$

- Soft threshold (LASSO):

$$\beta^{LASSO} = argmin_{\beta}[RSS + \lambda \sum_{i=1}^{N} |\beta_i|], \tag{3.32}$$

$$X_t^* = X_i \in X_t | \beta_i^{LASSO} \neq 0. \tag{3.33}$$

where in (3.33) the RSS refers to the sum of squared residual from a regression of y_t on all available regressors X_t. In hard thresholding, a regression is conducted for each predictor X_t. Then, the subset X_t^* is obtained by gathering those series whose coefficients have t-statistic larger than critical value. In soft thresholding, the least absolute shrinkage and selection operator (LASSO) technique is used to select X_t^* by regressing y_t on all elements of X_t and discarding uninformative predictors (this process is also known as variable selection).

We follow the precedent of Bai & Ng (2008*a*) to pre-screen and select (judgementally) our targeted predictors. First, we remove variables with missing observations from our full-panel dataset, reducing the number of variables from 285 to 140. Then, we combine hard and soft thresholds, and judgements to obtain 41 monthly time series from December 2001 to June 2018. These series represent 12 main categories of macroeconomic time series: industrial output and manufacturing production, retail and private consumption, international trade, investment, financial market, exchange rates, interest rates, money supply and loan, price indexes, government expenditure and revenue, business and consumer confidence index, and transportation sector. The description of data for targeted predictors are presented in Table B.0.2 of Appendix B. Finally, we denote forecasts of factor models from selected predictors as targeted DI, targeted FA-AR and targeted FA-VAR.

In Appendix C, we present the stability of predictors selection for the targeted FA-

AR and the FA-VAR in Figure C.4.1. It shows the type of series that was selected (or not) selected by hard thresholding with the critical value being 1.67 over the whole out-of-sample period. We found that hard thresholding can consistently select predictors over the whole out-of-sample period for price series, whereas the selection for investment and consumption varies over time.

3.3 China's Data

Having discussed the factor estimation, performance evaluation and forecasting methodology in the previous section, we now provide an empirical application using China's dataset.

3.3.1 Composition of China's Data

Chinese official economic data is characterised by two critical features. First, the available data from the NBSC in a particular period may differ substantially from the data released in the previous month due to regular benchmark revisions. The size of revisions tends to be relatively large compared to that in Western economies and the reasons for revisions are not always clear (Holz 2014, Koch-Weser 2013). Second, some particular data are not available in January and February due to Chinese New Year. One typical example is that of industrial production. The NBSC used to publish the industrial production series for January and February; however, due to the significant effect of Chinese New on the seasonal pattern of the data, the NBSC discontinued the release of industrial production data for both January and February in 2007. Therefore, every year, the firsthand information on Chinese industrial production becomes available to the public in mid-March in the form of a growth rate of production measuring accumulated production in January and February compared with the previous year. This break in the data dissemination pattern can greatly affect the accuracy of forecasts due to the scarcity of information available over that specific time window. Thus, we do not consider forecasting industrial production as a measure of real economic activity.

The full dataset used to estimate the factors includes 251 important monthly macroeconomic variables and 34 quarterly variables for China, spanning from December 2001 to June 2018.[19] Due to the rapid institutional and structural change in China's economy, data from the 1990s might (arguably) be of limited informative value regarding the current state of the Chinese economy (He et al. 2013, Fernald et al. 2014, Lin & Wang 2013). Our sample period covers the sharp downturn during the Global Financial Crisis as well as China's rapid subsequent recovery, which reflects the most important information about China's macroeconomy to date. The data include the following categories of macroeconomic variables: taxes and government expenditure, industrial sales and industrial production, energy production and consumption, exchange rates and stock indexes, interest rates and money supply, commodity price index and agricultural price index, exports and imports, business and consumer confidence indexes, and various survey data. We source our Chinese data from Thomson Reuters Datastream. All data are also freely available to download at the NBSC website. We report the description of data for the full-panel dataset in Table B.0.1 of Appendix B.

The theory outlined in Section 3.2 states that stationarity of X_{it} is a moment condition for factor models (Bai & Ng 2008*b*, Stock & Watson 2002*a*). As such, all monthly and quarterly variables are subjected to four preliminary steps: possible deseasonalisation, possible transformation by taking either the difference or the percentage change, screening for possible outliers, and standardisation. The deseasonalisation process removes the

[19]Studies using mixed-frequency methods to estimate factors for forecasting the Chinese macroeconomy are limited compared to studies into the macroeconomy of Western countries. Quarterly variables provide additional information over monthly variables and may lead to improved forecasts.

seasonal patterns of China's macroeconomic data, especially the effect of Chinese New Year on many industrial sector–related series. We use *X-13ARIMA-SEATS* in *R* to perform the deseasonalisation. Developed by the United States Census Bureau, the *X-13* supports user-defined holiday variables such as Chinese New Year or Indian Diwali. It returns deseasonalised series if seasonality is detected and returns original series if seasonality is not detected. The decision to take the difference or the percentage change is made judgementally, mainly based on inspection of time series plots and unit root tests. Generally, the difference is taken for those series that are already in the index or percentage terms, and the percentage change is implemented to series with actual quantities. We then clean the data by removing outliers for which the first difference deviates from the median of the first difference by more than five times the interquartile range and replacing them with missing data. Finally, all series are standardised to mean zero and unit variance according to Stock & Watson (2002*a*,*b*).

3.3.2 The Five Variables of Interest

We consider two categories of monthly economic indicators: the measures of inflation and the measures of real economic activity. The measures of inflation that we forecast are the CPI and the RPI, while the measures of real economic activity are nominal investment, nominal consumption and the railway traffic freight. The CPI and the RPI are two common indicators for measuring inflation and are standard in economic forecasting literature. Nominal investment and consumption account for more than 80% of total GDP according to the aggregated expenditure approach. Using the railway traffic freight as a real economic indicator has two benefits. First, it is not on the government target list; that is, the statistical authorities do not have incentives to purposefully falsify the railway cargo data to meet certain growth rate targets or to please the media and public

(Koch-Weser 2013, Holz 2014). Second, the information about the actual cargo volume is relatively easier to collect than the GDP. This means that the margin of collection error is relatively small.

CPI and RPI:

The data for the two price indexes are available from January 2002 to June 2018. Since the CPI and the RPI variables are collected in index terms with the index value of the same month in the previous year being 100, subtracting 100 provides a proxy of the inflation rate. The CPI and the RPI are annualised year-on-year rate. After deseasonalisation and screening for outliers, the transformation equations for the CPI and the RPI are:

$$\begin{aligned} y_{t|CPI} &= \mathrm{CPI}_t - 100, \\ y_{t|RPI} &= \mathrm{RPI}_t - 100. \end{aligned} \tag{3.34}$$

Investment, Consumption and Railway Traffic Freight:

Nominal investment, nominal consumption and railway traffic freight are expressed in nominal terms in billions of yuan and billions of tonnes and are available from December 2001 to June 2018. Similarly to the CPI and the RPI, we seasonally adjust them and screen for outliers. Since there exists a unit root for all three real variables, as shown in Table 3.1, we model these as first difference of logarithms. Since we treat all three real variables identically, consider investment, which we transform as:

$$y_t = 1200 * \ln(\mathrm{Investment}_t/\mathrm{Investment}_{t-1}). \tag{3.35}$$

Time series plots of the five variables are displayed in the first four images of Figure 3.1, and time series plots of transformed variables are presented the last four images of

Figure 3.1. Note that plots for raw series are of the data after outliers have been removed and seasonal adjustment has been applied.

The results of augmented Dickey–Fuller tests for unit roots are presented for each variable in Table 3.1. In each case, we set the maximum lag order to 6. For untransformed variables in Panel 1, we find evidence against a unit root existing for the CPI and the RPI, but not for nominal investment, consumption and railway cargo freight. Once transformed, the unit root hypothesis is rejected for all variables, as shown in Panel 2 of Table 3.1.

Table 3.1: Augmented Dickey–Fuller Test Statistics with Trend and Drift

Panel 1: Untransformed Variables

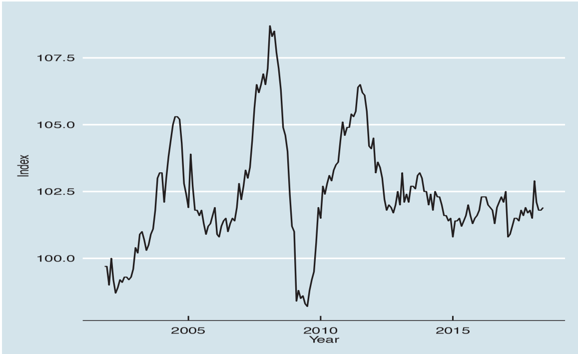
	CPI	RPI	Investment	Consumption	Railway Cargo
Lag Order	6.00	6.00	6.00	6.00	6.00
Test Statistics	–4.09	–3.81	–2.40	–2.40	–1.56
p-value	0.01	0.02	0.41	0.41	0.76

Panel 2: Transformed Variables

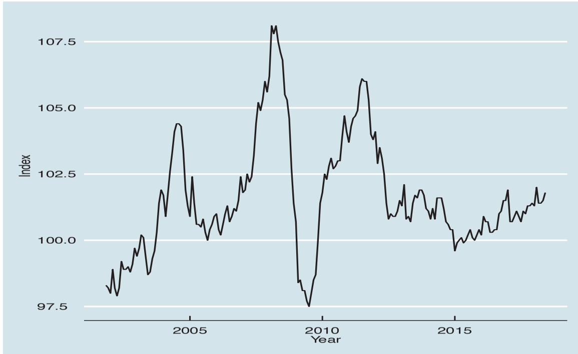
	CPI	RPI	Investment	Consumption	Railway Cargo
Lag Order	6.00	6.00	6.00	6.00	6.00
Test Statistics	–4.07	–3.73	–3.61	–5.37	–5.38
p-value	0.01	0.02	0.03	0.01	0.01

Figure 3.1: Time Series Plot of Variables

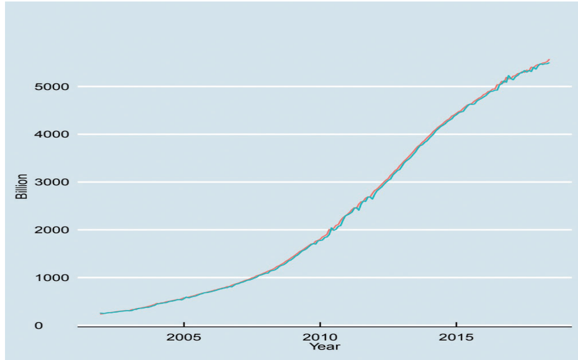
(a) Raw CPI



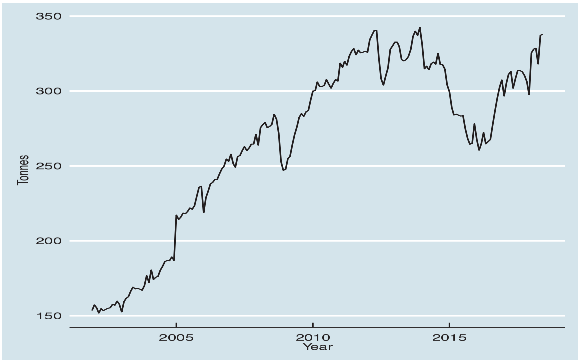
(b) Raw RPI



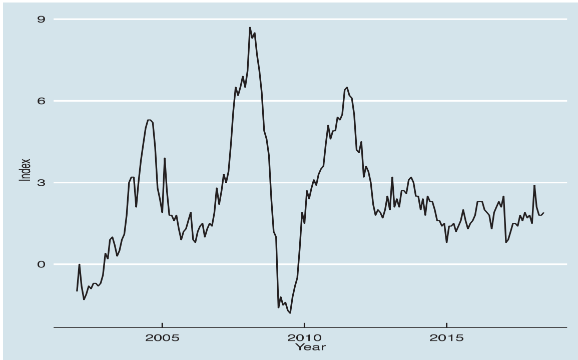
(c) Raw Investment and Consumption



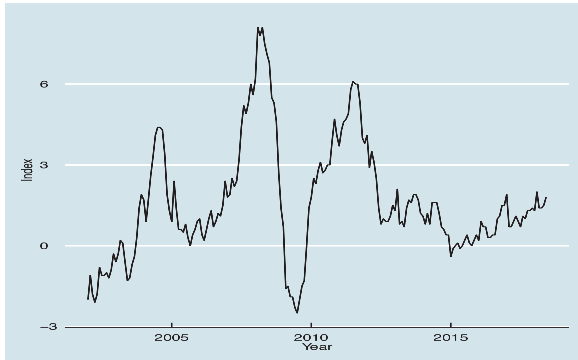
(d) Raw Railway Cargo



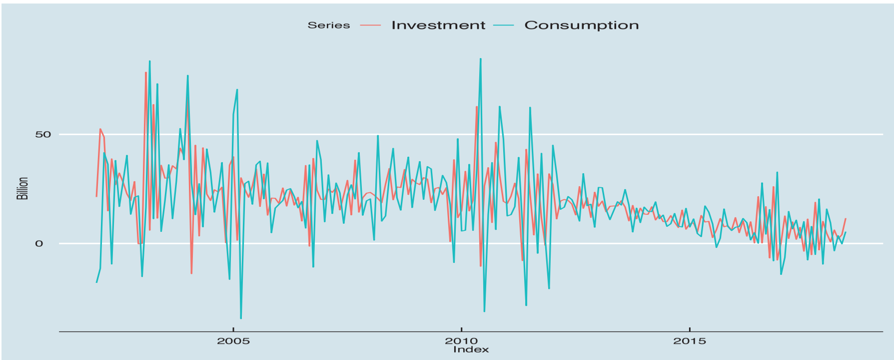
(e) Transformed CPI



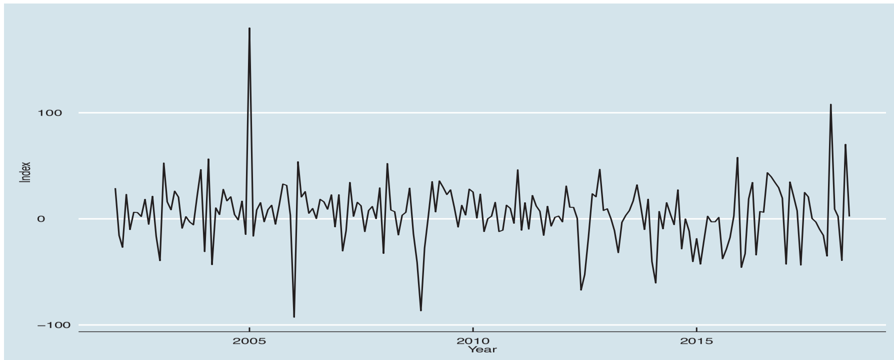
(f) Transformed RPI



(g) Transformed Investment and Consumption



(h) Transformed Railway Cargo



3.3.3 Forecasting Design

Our study forecasts five Chinese monthly economic indicators h months ahead, and we set h equal to 1, 3, 6, 9, and 12 months. All forecasts are performed in a simulated out-of-sample framework wherein all statistical calculations are conducted using a fully recursive methodology. To ensure a sufficient number of within-sample and out-of-sample forecasting periods, we consider the total sample size in a 50:50 per cent split between in-sample and out-of-sample periods.[20] We apply the fixed rolling window method, in which we use a fixed number of observations to estimate model parameters in every within-sample period. The first h-month-ahead forecast was made for May 2010. To construct this forecast, we take the first 100 observations from the raw dataset—that is, observations from January 2002 to April 2010—and remove outliers, seasonally adjust, standardise, estimate factors, select the model orders and run forecasting models using only the data in this subsample. After constructing forecasts from all models at each horizon, we drop the first observation from the previous subsample and add an extra observation of raw data to the end of the previous subsample. Then, we repeat the same steps for the second subsample—that is, take observations from February 2002 to May 2010 and construct the second h-month-ahead forecast for June 2010. We continue this process, dropping the first observation, adding an extra observation to the end of subsample, screening for outliers, standardising data, adjusting seasonality, estimating factors and running forecasting models using the data in the subsample, computing forecasts until the final forecast is made for June 2018.

In total, we produce a sequence of 99 1-month-ahead forecasts from April 2010 to June 2018, a sequence of 97 3-month-ahead forecasts from June 2010 to June 2018, a

[20] When the number of time observations is large, the literature tends to split the estimation sample and forecast sample evenly to ensure a sufficient number of within-sample and out-of-sample periods, see for example Stock & Watson (2002*b*), Kotchoni et al. (2019), Goulet et al. (2019), Faust & Wright (2013).

sequence of 94 6-month-ahead forecasts from September 2010 to June 2018, a sequence of 91 9-month-ahead forecasts from December 2010 to June 2018, and a sequence of 88 12-month-ahead forecasts from March 2011 to June 2018.

3.4 Assessing Forecast Results

In this section, we assess the forecasting performance of competing models. We evaluate the forecasts' performance using the relative mean squared error (MSE). Evaluation of the relative MSE in out-of-sample forecasting exercises is commonly used in economic forecasting literature. Let y_t be the variable to be forecast at time t. Let $\hat{y}_{t+h|m}$ denote the h-step-ahead forecasts for variable by model m and define y_{t+h} as the actual value of the variable at time $t+h$. The forecasting error at time $t+h$ for model m is computed as $\Delta_h y_{t+h|m} = \hat{y}_{t+h|m} - y_{t+h}$. The MSE by model m for each variable is then calculated as:

$$\frac{\sum_1^{T-P}(\Delta_h y_{t+h|m})^2}{T-P}. \tag{3.36}$$

where in (3.36) T refers to the total number of time observations, and P refers to the number of within-sample periods. The relative MSE is computed by setting the MSE of the benchmark AR(p) model to be 1.[21] For example, to calculate the relative MSE for the FA-VAR model, we divide the MSE of the FA-VAR model by the MSE of the AR(p) model. A relative MSE greater than 1 means the model performs worse than the benchmark model and vice versa. The choice of the benchmark model is guided by two

[21] Dufour & Stevanović (2013) advised caution when choosing either the AR model or the ARMA model as a benchmark. As the AR model uses a finite order and only approximates to the process of X_{it}, a long AR model is needed if the MA polynomial has roots close to the non-invertibility region. This could cause over-fitting of the AR model. However, the AR model has been widely used as the benchmark in the context of large datasets and factor models, see for example Stock & Watson (1998, 2002*b*), Bai & Ng (2008*a*), Gupta & Kabundi (2011), Artis et al. (2005), Schumacher (2007), Schumacher & Breitung (2008). Even Dufour & Stevanović (2013) set AR(p) as the benchmark model.

considerations: (1) a well-defined AR model is commonly used as the benchmark model in economic forecasting literature (e.g., Stock & Watson (2002*b*), Schumacher & Breitung (2008), Schumacher (2007), Li et al. (2015), Gupta & Kabundi (2011), Artis et al. (2005)) and (2) Lin & Wang (2013), Higgins et al. (2016), He & Fan (2015), Mehrotra & Sánchez-Fung (2008) found that AR models perform relatively well compared to ARMA models, diffusion index models, Bayesian VAR models and Phillips curves for forecasting China's inflation and GDP. If factor models based on mixed-frequency datasets with missing observations can outperform the AR(p), this means strong predictability.

In addition to the relative MSE, we employ the conditional predictive test described by Giacomini & White (2006) to test null hypotheses of MSE equality between mixed-frequency factor models and corresponding targeted factor models. Giacomini & White (2006)'s test had a null of equal expected loss evaluated at the current parameter estimates as in (3.37) and, therefore, is conditional on the information set available at time t. Giacomini & White (2006) also claimed that the conditional test can capture the effect of estimation uncertainty on relative forecast performance when forecasting models may be mis-specified.

$$H_0 : E[L(f_{1t+1|t}(\hat{\beta}_{1t}), y_{t+1})] = E[L(f_{2t+1|t}(\hat{\beta}_{2t}, y_{t+1})]. \tag{3.37}$$

Estimates of the relative mean squared forecasting errors for inflation are presented in Table 3.2, while forecasting results for real variables are shown in Table 3.3. Pairwise conditional predictive test results for inflation are presented in Table 3.4 and for real variables in Table 3.5.

3.4.1 Out-of-Sample Forecasting Results

We now address the results for inflation in Table 3.2. For the CPI inflation rate, the benchmark AR(p) model produces the best one-month-ahead forecasts out of sample. For all other forecast horizons, VAR models, targeted FA-AR(2) models and mixed-frequency FA-AR(2) models generate superior forecasts to the benchmark AR(p) model and other competing models. The relative improvements are more modest at horizons of three months ahead and six month ahead compared to nine months ahead and twelve months ahead. In some cases, the improvements over the benchmark forecasts are substantial; for example, the FA-AR(2) model with factors estimated using mixed-frequency data with missing observations has a forecast error variance that is 61% that of the AR(p) model. Other factor models perform poorly at horizons of three and six months, but perform well at horizons of nine and twelve months.

The results for the RPI inflation rate are similar to those for the CPI inflation rate. The VAR forecasts, multivariate leading indicator forecasts, FA-VAR forecasts and FA-AR forecasts generally outperform the benchmark AR(p) forecasts at any horizon, regardless of whether lag is predetermined at 2 or is selected by AICc and whether factors are estimated using mixed-frequency datasets or targeted predictors. The relative performance improves substantially as the forecast horizon grows. Among all competing models, VAR(2) and mixed-frequency FA-AR(2) are the two best performing models for horizons of three, six, nine and twelve months. Therefore, judged solely on point forecasts, there appears to be gain in using factors estimated from mixed-frequency datasets and the M2 growth rate to forecast Chinese CPI and RPI inflation at horizons longer than one month.

When the differences between targeted and full-panel factor forecasts are considered, the use of mixed-frequency datasets with missing observations yields substantial improvements for FA-VAR(p) forecasts and FA-VAR(2) forecasts while yielding marginal or no improvements for DI forecasts, FA-AR(p) forecasts and FA-AR(2) forecasts for the CPI inflation rate at all horizons. In some cases, improvements are noticeable. For example, the mixed-frequency FA-VAR(p) model produces 49% less forecast error variance than the targeted FA-VAR(p) model. For the RPI inflation rate, the use of mixed-frequency datasets can only improve DI forecasts at horizons of one, three, six and nine months and FA-VAR(p) forecasts at horizons of one, three and six months. For sampling variability, the results of pairwise conditional predictive tests between mixed-frequency factor models and targeted factor models are presented in Table 3.4. Only p-values of the DI model for the RPI inflation rate are smaller than 5% at h = 1, h = 3, h = 6 and h = 9, indicating that DI models based on mixed-frequency datasets generally have better predictive performance than corresponding DI models based on targeted predictors at 5% levels of significance. For other factor models, mixed-frequency datasets with missing observations could not yield better predictive performance at any horizons.

One observed oddity from Table 3.2 is sharp declines in relative MSEs of the mean forecast, the targeted DI forecast, and the mixed-frequency DI forecast for price series as forecast horizons increase. As it stands, the benchmark AR(p) model may perform extremely poorly at longer horizons, so do other models since the performance of competing models are assessed using relative terms. Figure C.2.1 plots the AR(p) forecast, the mean forecast, the targeted DI forecast, and the mixed-frequency DI forecast for the CPI and the RPI at h=1, 3, 6, 9, and 12. At h=1, 3, and 6, the AR(p) forecast are superior to DI forecast, regardless of whether the factors are estimated using preselected targeted predictors or full-panel dataset. At the longer horizon, the performance of DI forecast improved, especially after the year 2013.

The forecasting results for the real variables are presented in Table 3.3. For nominal investment, the performance of mixed-frequency FA-VAR(p) modes, targeted FA-VAR(p) models and ARMA(p,q) models is usually better than that of the benchmark model across all horizons, except for targeted FA-VAR(p) models at one month ahead. The relative improvements are more modest at h = 6, h = 9 and h = 12 than at h = 1 and h = 3. In contrast, VAR models, mean models, targeted DI models and mixed-frequency DI models perform poorly compared to the benchmark AR(p) model.

For nominal consumption at a forecasting horizon of one month, all models considered are inferior to the benchmark AR(p) model. For a forecasting horizon of three-months-ahead, only mixed-frequency FA-VAR(p) models provide smaller mean squared forecasting error than the benchmark model. At longer horizons, the relative performance of the FA-VAR(p) model improves, regardless of whether factors are estimated based on mixed-frequency datasets or preselected targeted predictors. The multivariate models and FA-VAR(2) model perform significantly worse than the benchmark model, whereas the univariate models, DI models and FA-AR model produce forecasts that are similar or slightly worse than benchmark model.

As with the railway freight, the results are less clear. Overall, competing models can barely outperform the benchmark model at any horizon. The level of relative improvement is marginal, with maximum error reduction being 11% for targeted DI models at h = 1 and minimum error reduction being 1% for the mixed-frequency FA-VAR(2) model at h = 6. Of all competing models, FA-VAR models and VAR models perform relatively well, regardless whether lag orders are predetermined at p = 2 or selected by AICc and whether factors are estimated on mixed-frequency datasets or targeted predictors. Multivariate leading indicators perform poorly at all horizons, with these forecasts resulting in, on average, 51.8% more forecasting error than the benchmark forecasts.

For the real variables, forecasting performance is usually not improved when the empirical factors from the mixed-frequency dataset are used compared to those from the targeted predictors. We report results of pairwise conditional predictive tests between mixed-frequency factor forecasts and targeted factor forecasts in Table 3.5. Overall, we find no strong evidence that the use of the mixed-frequency dataset yields statistical improvements over the targeted predictors when a 5% level of significance is considered. Occasionally, the conditional test finds evidence in favour of mixed-frequency factor models. For example, the p-values of the FA-AR(2) model for investment yield less than a 5% level of significance at h = 1, h = 3 and h = 6. Therefore, we do not consider that using mixed-frequency datasets with missing observations can provide evidence of superior forecasting power for real variables as the relative improvements are very small.

Table 3.2: Inflation: Simulated Out-of-Sample Standardised Mean Squared Forecast by Horizon

	CPI					RPI				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Mean	9.87	3.78	1.91	1.13	0.74	12.30	3.81	1.69	1.02	0.70
AR(p)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(2)	1.24	1.06	1.12	1.01	0.94	1.07	1.01	1.01	0.94	0.84
ARMA(p,q)	1.23	1.11	0.98	0.80	0.77	0.99	0.98	0.98	0.94	0.89
ARMA(2,2)	1.16	1.02	0.93	0.87	0.82	1.05	1.20	1.39	1.39	1.29
MLD(p)	1.25	1.20	1.17	0.99	0.89	0.97	0.95	0.90	0.74	0.68
MLD(2)	1.14	1.03	1.00	0.84	0.71	0.92	0.84	0.74	0.63	0.57
VAR(p)	1.28	1.01	0.90	0.75	0.68	0.97	0.80	0.65	0.55	0.54
VAR(2)	1.44	1.23	1.04	0.79	0.67	1.11	1.04	0.84	0.63	0.55
Targeted DI	9.41	3.57	1.82	1.07	0.72	11.97	3.71	1.65	1.00	0.69
Targeted FA-AR(p)	1.30	1.22	1.10	0.82	0.74	1.00	1.01	0.93	0.77	0.71
Targeted FA-AR(2)	1.32	1.06	0.92	0.71	0.63	0.98	0.87	0.71	0.55	0.50
Targeted FA-VAR(p)	1.42	1.10	1.12	0.80	0.67	1.07	0.86	0.76	0.57	0.57
Targeted FA-VAR(2)	1.20	1.05	1.00	0.71	0.59	0.91	0.80	0.70	0.55	0.54
Mixed-frequency DI	7.68	3.02	1.62	1.02	0.70	8.76	2.77	1.27	0.80	0.61
Mixed-frequency FA-AR(p)	1.21	1.14	1.10	0.86	0.77	0.95	0.95	0.92	0.77	0.71
Mixed-frequency FA-AR(2)	1.20	0.95	0.89	0.71	0.61	0.95	0.79	0.67	0.53	0.47
Mixed-frequency FA-VAR(p)	1.16	0.96	0.92	0.75	0.67	0.92	0.78	0.65	0.51	0.48
Mixed-frequency FA-VAR(2)	1.29	0.97	0.87	0.68	0.59	0.98	0.80	0.65	0.51	0.48

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the point forecasts of the relative mean squared forecasting error expressed as a ratio of the mean squared forecasting error of the benchmark model for the same forecast horizon. The benchmark model is the AR(p) model.

Table 3.3: Real Variables: Simulated Out-of-Sample Standardised Mean Squared Forecast by Horizon

	Investment					Consumption					Railway Cargo				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Mean	1.69	1.56	1.56	1.46	1.28	1.28	1.26	1.21	1.16	1.07	0.96	0.96	1.00	1.00	1.00
AR(p)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(2)	1.01	1.04	0.96	0.96	0.95	1.06	1.00	0.97	0.94	0.91	1.00	1.00	1.00	1.00	1.00
ARMA(p,q)	0.97	0.98	0.87	0.86	0.80	1.05	1.05	1.04	1.09	1.04	1.01	0.97	0.99	1.01	1.02
ARMA(2,2)	1.05	1.03	0.95	0.95	0.94	1.22	1.07	1.06	1.15	1.02	1.12	1.02	0.99	0.98	1.00
MLD(p)	1.12	1.14	1.02	1.02	0.96	1.68	1.28	1.12	1.23	1.32	1.46	1.59	1.54	1.39	1.52
MLD(2)	1.80	1.62	1.54	1.46	1.28	1.65	1.30	1.12	1.23	1.30	1.39	1.54	1.53	1.42	1.50
VAR(p)	3.39	2.90	2.81	2.29	1.82	1.92	1.86	1.83	1.58	1.34	0.93	0.93	0.97	0.97	0.94
VAR(2)	3.38	2.90	2.81	2.29	1.82	1.92	1.86	1.83	1.58	1.34	0.93	0.93	0.97	0.97	0.94
Targeted DI	1.71	1.58	1.56	1.44	1.27	1.29	1.24	1.21	1.15	1.07	0.89	0.93	1.08	0.99	1.05
Targeted FA-AR(p)	1.08	1.11	1.02	1.00	0.96	1.20	1.02	1.10	1.10	1.05	0.92	0.95	1.07	1.00	1.05
Targeted FA-AR(2)	1.70	1.63	1.60	1.48	1.30	1.43	1.13	1.21	1.16	1.08	0.95	0.93	1.08	1.00	1.05
Targeted FA-VAR(p)	1.15	0.85	0.74	0.62	0.63	1.27	1.10	0.86	0.87	0.93	0.99	0.93	0.98	0.96	0.94
Targeted FA-VAR(2)	1.68	1.09	0.95	0.93	0.99	1.68	1.28	1.39	1.40	1.30	0.95	0.91	0.97	0.97	0.94
Mixed-frequency DI	1.60	1.50	1.46	1.38	1.21	1.29	1.23	1.18	1.15	1.06	0.91	0.96	0.96	0.99	1.06
Mixed-frequency FA-AR(p)	1.06	1.11	1.02	1.02	0.97	1.16	1.02	1.07	1.10	1.05	0.94	0.97	0.97	0.98	1.04
Mixed-frequency FA-AR(2)	1.66	1.58	1.45	1.37	1.20	1.41	1.12	1.16	1.14	1.05	0.95	0.95	0.97	0.98	1.05
Mixed-frequency FA-VAR(p)	0.90	0.85	0.71	0.60	0.64	1.08	0.98	0.88	0.88	0.96	1.01	1.02	0.99	0.97	0.94
Mixed-frequency FA-VAR(2)	1.67	1.09	0.97	0.95	1.02	1.73	1.27	1.43	1.43	1.32	0.95	0.93	0.97	0.97	0.94

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the point forecasts of the relative mean squared forecasting error expressed as a ratio of the mean squared forecasting error of the benchmark model for the same forecast horizon. The benchmark model is the AR(p) model.

Table 3.4: Inflation: Pairwise Conditional Predictive Test Results by Horizon

	CPI					RPI				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Diffusion Index	**0.03**	0.17	0.38	0.65	0.80	**0.00**	**0.02**	**0.05**	**0.05**	0.24
FA-AR(p)	0.30	0.11	0.97	0.08	0.26	0.48	0.09	0.77	0.94	0.88
FA-AR(2)	0.12	**0.05**	0.39	0.88	0.33	0.49	**0.02**	0.12	0.42	0.22
FA-VAR(p)	0.37	0.28	0.17	0.12	0.15	0.77	1.00	0.50	0.25	0.17
FA-VAR(2)	0.54	0.33	0.09	0.26	0.27	0.43	0.92	0.44	0.75	0.62

Table 3.5: Real Variables: Pairwise Conditional Predictive Test Results by Horizon

	Investment					Consumption					Railway Cargo				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Diffusion Index	**0.01**	0.07	0.07	0.23	0.23	0.93	0.60	0.17	0.93	0.67	0.54	0.54	**0.01**	0.91	0.89
FA-AR(p)	0.75	0.83	0.82	0.45	0.64	0.18	0.94	0.19	0.84	0.68	0.53	0.62	**0.01**	0.72	0.96
FA-AR(2)	0.47	0.18	**0.00**	**0.03**	**0.05**	0.67	0.41	0.12	0.41	0.37	0.94	0.64	**0.01**	0.66	0.95
FA-VAR(p)	**0.04**	0.98	0.65	0.55	0.33	0.18	0.25	0.55	0.82	0.19	0.85	**0.01**	0.46	0.57	0.59
FA-VAR(2)	0.81	0.73	0.44	0.53	0.31	0.49	0.66	**0.04**	0.11	**0.04**	0.94	0.10	0.72	0.20	0.71

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the p-values of the pairwise conditional predictive test between mixed-frequency factor models and corresponding targeted factor models (i.e., mixed-frequency FA-AR(p) model v. targeted FA-AR(p) model). P-values displayed with a bold font mean that we reject null hypotheses of equal expected loss at a 5% level of significance and conclude that the mixed-frequency factor model has better expected loss than the targeted factor models.

3.4.2 The Effects of the Global Financial Crisis

In response to the Global Financial Crisis, the Chinese government injected an unprecedented four trillion yuan to stimulate the economy in 2009. These investments aimed to promote households consumption, increase job opportunity, and stabilise economic growth until the end of 2011. This undoubtedly helped stabilise the double-digit real GDP growth in the aftermath of the financial crisis. Fears of the Global Financial Crisis and follow-up investments by the Chinese government have had persistent effects on inflation and output expectations. Several existing studies have argued that AR models and random walk, which are gold-standard models for forecasting United States inflation and output, might be less effective during economic recessions with unstable financial conditions (Alessandri & Mumtaz 2017, Pain et al. 2014, Faust & Wright 2013, Gilchrist & Zakrajšek 2012, Liu & Moench 2016, Berge 2015). Contrastingly, Kotchoni et al. (2019) and Goulet et al. (2019) found that factor structure–based models perform well in National Bureau of Economic Research recession periods; this may be explained by the fact that these models are flexible enough to accommodate the faster-than-usual changes in economic variables during recessions. In this regard, this paper also considers assessing the performance of the factor models during the Global Financial Crisis.

In our study, the Global Financial Crisis is defined as the period from November 2008 to December 2011.[22] The process of forecast construction is the same as in the out-of-sample period. All forecasts are made based on a simulated framework whereby all calculations are conducted using a fully recursive methodology. The first h-month-

[22]The National Bureau of Economic Research defines the GFC recession period as lasting from December 2007 to June 2009, but this definition was for the United States. We define the recession period for China based on how the State Council responded to the GFC. In the end of 2008, the Chinese government announced a series of investment plans (about 4 trillion CNY) to promote consumption, increase job opportunities and stabilise economic growth until the end of 2011. Therefore, our definition of the GFC period is from November 2008 to December 2011.

ahead forecast in the Global Financial Crisis period is made for November 2008 using the raw dataset from January 2002 to October 2008. To compute this forecast, we remove outliers; seasonally adjust, standardise and estimate factors; select the model orders; and run forecasting models using only the data in the January 2002 to October 2008 subsample period. To compute next h-month-ahead forecast for December 2008, we delete the first observation from the previous sample, add an extra observation to the end of the previous sample and repeat the same steps for the second subsample period. We continue this process—at each subsample period, we screen for outliers, perform seasonal adjustments, estimate factors and run forecasting models—until the final forecast is made for December 2011.

The relative MSEs for inflation are reported in Table 3.6. For the CPI inflation rate, competing models generally perform less well than the benchmark AR(p) model at horizons of one, six, nine and twelve months. At a horizon of three months, mixed-frequency FA-VAR and ARMA models outperform the benchmark model. The relative improvements are substantial, reaching approximately 40% for mixed-frequency FA-VAR models and 30% for ARMA models. For the RPI inflation rate, the mixed-frequency FA-VAR(p) model generates superior one-, three-, six- and nine-months-ahead forecasts to other competing models. At 12 months ahead, the ARMA(2,2) model generates the lowest relative MSE.

For real variables in Table 3.7, results are less clear out. For nominal investment and consumption at horizons of one month, three months, six months and nine months, competing models can barely outperform the benchmark AR(p) model. In some cases, such as those of VAR models at all horizons, they even generate considerably worse relative MSEs than the benchmark. At horizons of 12 months, most competing models perform better than the benchmark, but the relative improvements are very small. For the railway freight at one month ahead, the mixed-frequency DI model and the mean model provide the lowest MSE of all competing models. At three months ahead, six months ahead, nine months ahead and twelve months ahead, competing models can hardly generate forecasts that are better than the benchmark forecast.

For sampling variability between mixed-frequency factor models and targeted factor models, we report p-values of conditional predictive tests for inflation in Table 3.8 and for real variables in Table 3.9. Overall, we find no strong evidence that the use of mixed-frequency datasets with missing observations yields statistical improvements over the targeted predictors when a 5% level of significance is considered for both inflation and real variables. Occasionally, the conditional test finds evidence in favour of the mixed-frequency factor models. For example, the p-values of the FA-VAR(2) model for investment and the p-values of the FA-AR models for RPI have less than a 5% level of significance at h = 6. Therefore, we do not consider that using mixed-frequency datasets with missing observations can provide evidence of superior forecasting power to preselected targeted predictors during the Global Financial Crisis period.

Table 3.6: Inflation: Standardised Mean Squared Forecast by Horizon during Global Financial Crisis

	CPI					RPI				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Mean	13.78	3.69	1.70	1.33	1.24	12.23	2.92	1.34	1.08	0.93
AR(p)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(2)	1.15	0.98	1.01	1.10	1.28	0.95	0.99	0.99	1.07	1.07
ARMA(p,q)	0.95	0.70	0.95	1.66	1.77	1.01	1.05	1.19	1.45	1.46
ARMA(2,2)	0.97	0.74	0.91	1.12	1.17	0.94	0.97	1.03	1.05	0.91
MLD(p)	1.13	0.99	1.18	1.75	2.76	0.98	0.97	0.99	1.17	1.34
MLD(2)	1.13	1.02	1.22	1.87	2.89	0.99	1.02	1.13	1.58	1.98
VAR(p)	1.26	1.21	1.38	1.91	2.65	1.16	1.26	1.09	1.79	2.06
VAR(2)	1.23	1.06	1.19	1.58	2.04	0.87	0.91	0.94	1.31	1.40
Targeted DI	13.67	3.57	1.62	1.27	1.20	12.55	3.21	1.54	1.28	1.06
Targeted FA-AR(p)	1.16	0.97	1.19	1.63	2.61	1.02	1.01	1.13	1.60	1.76
Targeted FA-AR(2)	1.19	0.94	1.10	1.54	2.27	0.96	0.90	0.98	1.26	1.50
Targeted FA-VAR(p)	1.44	1.27	1.38	1.82	2.43	1.00	1.03	1.10	1.32	1.40
Targeted FA-VAR(2)	1.32	1.09	1.21	1.53	1.97	0.96	1.00	1.07	1.28	1.36
Mixed-frequency DI	12.60	3.21	1.44	1.19	1.29	11.40	2.59	1.15	0.99	0.97
Mixed-frequency FA-AR(p)	1.10	0.85	1.05	1.45	2.14	0.85	0.86	1.03	1.36	1.61
Mixed-frequency FA-AR(2)	1.12	0.94	1.14	1.60	2.31	0.86	0.89	1.02	1.32	1.55
Mixed-frequency FA-VAR(p)	1.00	0.62	1.01	1.24	1.66	0.75	0.63	0.86	0.95	1.05
Mixed-frequency FA-VAR(2)	1.01	0.60	1.52	3.45	7.96	0.77	0.58	1.20	2.34	4.16

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the point forecasts of the relative mean squared forecasting error expressed as a ratio of the mean squared forecasting error of the benchmark model for the same forecast horizon. The benchmark model is the AR(p) model.

Table 3.7: Real Variables: Standardised Mean Squared Forecasts by Horizon during Global Financial Crisis

	Investment					Consumption					Railway Cargo				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Mean	1.32	1.06	1.00	0.98	0.91	1.28	1.10	1.02	1.01	0.95	0.81	1.00	1.00	1.00	1.00
AR(p)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AR(2)	1.06	1.02	0.99	0.98	0.91	1.00	1.00	0.99	1.00	0.95	1.00	1.00	1.00	1.00	1.00
ARMA(p,q)	1.16	1.03	0.98	0.98	0.91	0.93	0.94	0.92	1.01	0.96	1.11	1.01	1.00	1.00	1.00
ARMA(2,2)	1.16	1.04	0.98	1.01	0.93	0.96	0.97	1.00	0.99	0.94	1.11	0.99	0.98	1.01	1.01
MLD(p)	1.14	1.02	0.97	0.94	0.87	1.51	1.02	0.89	1.21	1.28	0.94	2.70	2.12	1.83	1.54
MLD(2)	1.25	1.11	0.96	1.01	0.94	1.50	1.01	0.90	1.21	1.26	1.11	2.79	2.27	1.89	1.60
VAR(p)	4.81	3.71	3.35	3.10	2.75	2.43	2.11	1.90	1.80	1.63	0.85	1.32	1.14	0.98	1.01
VAR(2)	4.81	3.71	3.35	3.10	2.75	2.43	2.11	1.90	1.80	1.63	0.85	1.32	1.14	0.98	1.01
Targeted DI	1.33	1.06	1.03	1.03	0.96	1.33	1.09	1.02	1.01	0.97	1.14	2.06	1.65	1.25	1.62
Targeted FA-AR(p)	1.22	1.06	1.00	1.03	0.93	1.05	1.01	0.98	1.01	0.98	1.27	2.02	1.66	1.24	1.61
Targeted FA-AR(2)	1.28	1.12	0.95	1.01	0.93	1.11	1.02	1.03	1.02	0.99	1.32	2.00	1.66	1.25	1.61
Targeted FA-VAR(p)	1.22	1.04	0.98	0.98	0.95	1.07	1.05	1.04	1.01	0.97	1.05	1.36	1.18	1.03	1.04
Targeted FA-VAR(2)	1.30	1.15	0.96	1.02	1.05	1.15	1.03	1.04	1.01	0.98	1.17	1.50	1.43	1.27	1.49
Mixed-frequency DI	1.39	1.10	1.01	1.04	0.93	1.29	1.10	1.02	1.01	0.96	0.80	1.02	1.02	0.98	1.03
Mixed-frequency FA-AR(p)	1.20	1.07	1.02	1.07	0.94	0.95	1.00	0.98	1.02	0.96	1.02	1.04	1.04	0.97	1.04
Mixed-frequency FA-AR(2)	1.28	1.20	0.96	1.05	0.93	1.05	1.04	1.03	1.01	0.96	1.06	1.05	1.04	0.97	1.04
Mixed-frequency FA-VAR(p)	1.20	1.17	1.01	0.98	0.91	1.12	1.06	1.03	1.01	0.96	1.02	0.98	1.00	0.99	1.00
Mixed-frequency FA-VAR(2)	1.23	1.16	0.97	0.99	0.91	1.05	1.05	1.03	1.01	0.96	1.05	0.99	0.99	0.97	1.01

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the point forecasts of the relative mean squared forecasting error expressed as a ratio of the mean squared forecasting error of the benchmark model for the same forecast horizon. The benchmark model is the AR(p) model.

Table 3.8: Inflation: Pairwise Conditional Predictive Test Results by Horizon during Global Financial Crisis

	CPI					RPI				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Diffusion Index	0.17	0.16	0.31	0.56	0.18	0.13	0.14	0.21	0.27	0.39
Diffusion Index	0.17	0.16	0.31	0.56	0.18	0.13	0.14	0.21	0.27	0.39
FA-AR(p)	0.32	0.16	0.27	0.55	0.14	**0.03**	0.13	0.40	0.45	0.47
FA-AR(2)	0.16	0.98	0.17	0.14	0.23	**0.04**	0.86	0.07	**0.02**	0.09
FA-VAR(p)	**0.04**	0.12	0.13	0.18	0.25	0.12	0.09	**0.04**	**0.05**	0.21
FA-VAR(2)	**0.05**	0.09	0.59	0.39	0.34	0.15	0.06	0.74	0.43	0.35

Table 3.9: Real Variables: Pairwise Conditional Predictive Test Results by Horizon during Global Financial Crisis

	Investment					Consumption					Railway Cargo				
	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12	h=1	h=3	h=6	h=9	h=12
Diffusion Index	0.21	0.37	0.53	0.74	0.66	0.06	0.12	0.78	0.70	0.12	0.07	0.08	0.07	0.20	**0.01**
FA-AR(p)	0.90	0.88	0.47	0.42	0.70	0.10	0.66	0.81	0.73	**0.04**	0.22	0.08	0.06	0.22	**0.01**
FA-AR(2)	0.96	0.18	0.73	0.26	0.94	0.15	0.25	0.83	0.66	**0.02**	0.21	0.08	0.06	0.20	**0.01**
FA-VAR(p)	0.86	0.12	**0.03**	0.96	0.33	0.54	0.72	0.73	0.31	0.11	0.85	0.19	0.26	0.33	0.16
FA-VAR(2)	0.37	0.86	0.64	0.39	0.30	0.17	0.39	0.92	0.51	0.17	0.45	0.23	0.28	0.31	0.27

The rows are the forecasting models described in Section 3.2.3. The columns are forecast horizons. The table elements are the p-values of the pairwise conditional predictive test between mixed-frequency factor models and corresponding targeted factor models (i.e., mixed-frequency FA-AR(p) model v. targeted FA-AR(p) model). P-values displayed with a bold font mean that we reject null hypotheses of equal expected loss at a 5% level of significance and conclude that the mixed-frequency factor model has better expected loss than the targeted factor models.

3.5 Concluding Remarks

Our paper considers forecasting five important Chinese monthly macroeconomic variables under a data-rich environment in a simulated out-of-sample forecasting exercise. To exploit the information contained in the big data era, we use the large-dimensional approximate factor model framework developed by Stock & Watson (2002*b*) and Bai & Ng (2008*b*). We also employ several univariate and multivariate forecasting models that are widely used for Western economies as competing models to assess whether simple forecasting models can perform better or worse than complicated large-dimensional factor models. We consider two types of dataset:(i) the full-panel dataset, which is a mixed-frequency dataset with missing observations, and (ii) the targeted dataset, which is single frequency (monthly) without missing observations and we preselect variables according to the work of Bai & Ng (2008*b*). Our dataset covers 16 main sectors of the Chinese macroeconomy, spanning from December 2001 to June 2018. We estimate factors and factor loadings simultaneously using the asymptotic principal components analysis. When data are unbalanced, we apply the EM algorithm of Stock & Watson (2002*b*). We assess the performance evaluation of forecasting results through the relative MSE by setting the MSE of AR(p) forecasts to 1. We also test sampling variability between mixed-frequency factor models and targeted factor models using the conditional predictive test of Giacomini & White (2006).

The main results are presented in Section 3.4. They provide clear guidance for analysts who are interested in forecasting Chinese macroeconomic variables. For the CPI inflation rate at one-month-ahead, the AR(p) models generate the lowest MSE. At three-, six-, nine- and twelve-months-ahead horizons, the mixed-frequency FA-VAR(2) model is the best performing of all competing models. For RPI inflation rate at any horizon, mixed-frequency FA-VAR(p) forecasts are the best forecasts. For nominal investment at all horizons, the FA-VAR(p) model generates superior forecasts to the benchmark model, regardless of whether factors are estimated from mixed-frequency datasets or targeted predictors. For nominal consumption at horizons longer than one month, mixed-frequency and targeted FA-VAR(p) models can improve forecasting performance, but the levels of relative improvements are more modest compared to those for nominal investment. For railway freight, competing models can barely outperform the benchmark AR(p) model. The forecasting results for the Global Financial Crisis period are markedly different. In most cases, the AR(p) model is hard to beat for real variables. For inflation, only the mixed-frequency FA-VAR(p) model provides superior forecasts for CPI at one month ahead and for RPI at one month ahead, three months ahead and six months ahead.

An important contribution of our paper is that, to the best of our knowledge, it is one of the first papers using a large dataset with mixed frequency and missing observations to forecast the Chinese macroeconomy. Our dataset contains 251 monthly variables and 34 quarterly variables, covering all important sectors of the Chinese macroeconomy. The number of predictors that has been used in the previously published literature for China is relatively smaller than that for the United States, which usually covers 40 variables or less. Also, we run a large number of training sets and validation sets. The range of size of the out-of-sample period in our paper is from 88 (for 12-month-ahead forecasts) to 99 (for one-month-ahead forecasts), compared with a range from five to 60 observations in the existing literature. Our paper finds evidence that large-dimensional factor models that are well established for forecasting Western macroeconomies also work well for China, especially the mixed-frequency FA-VAR model. Our paper also finds no strong evidence that the use of mixed-frequency datasets can provide statistical improvement over targeted predictors at a 5% level of significance for both inflation and real variables during our entire sample period, as well as during the Global Financial Crisis period.

Arguably, there may exist superior models or methods not considered in this paper. Such an argument is inevitable as there is a wide range of econometric models that could deal with a large number of variables. Our work finds evidence that large-dimensional factor models could substantially improve forecasting performance for the CPI, RPI, nominal investment and nominal consumption but not for railway freight. However, this does not imply that other big data-compatible econometric models could not further improve the forecasting performance. Compared to major Western countries, Chinese macroeconomic data are of a short duration with a period of significant structural change, have too many missing values and are of questionable quality. Considering this, researchers might consider econometric forecasting that is more capable of dealing with these stylised facts. Techniques that allow for time variation in the factor loadings might be useful to deal with model instability. Mixed-frequency and state-space VAR models might also be a valuable approach for mixed-frequency and ragged-edge datasets. Recently proposed machine learning algorithms that are compatible with non-stationary and non-parametric features might be useful. Bayesian approaches that allow sample information to be combined with structurally relevant prior information could be an alternative to the frequentist domain. Model combinations across a set of candidate forecasting models could be useful to improve forecasting performance over a single model. Thus, this paper provides considerable scope remains for future research.

Chapter 4

Nowcasting Chinese GDP Using Machine Learning Algorithms

4.1 Introduction

In this paper, we undertake the task of nowcasting gross domestic product (GDP) for mainland China using machine learning (ML) algorithms. Macroeconomic policy-making in real time often faces the problem of needing to monitor the current economic conditions with incomplete information, as some important macroeconomic variables are released with significant delays and are subject to frequent revisions. For example, a key indicator of real economic activity, GDP, is published only at a quarterly frequency with up to six-week delay. Nowcasting models have been widely used by many central banks and other institutions to mitigate some of these uncertainties (Hendry et al. 2016, Bańbura et al. 2013). Nowcasting refers to forecasting the current or most recent aggregate state of an economy based on a range of partially available indicators before the official economic figures are released. It can be undertaken at the stage of initial estimation of macroeconomic variables with incomplete data to better understand the current economic conditions and can be used to update estimates as more information becomes available (Bańbura et al. 2013, Daiane Marcolino de Mattos 2019). Nowcasting also can assist with early identification of turning points or significant shifts in an economy's momentum (Giannone et al. 2008, Zhang et al. 2018). There are a range of different approaches that can be taken to produce nowcasts of the current GDP, from judgement-based professional nowcasts to statistical models that link high-frequency predictors to low-frequency variables, such as the in-filling approach, the bridge equation, the mixed-data sampling (MIDAS) model and the multivariate dynamic models in the state-space representation. For complete survey papers of nowcasting techniques see, for example, Hendry et al. (2016) and Bańbura et al. (2013).

Developments in time series econometrics (mainly dynamic factor models and large Bayesian VAR models) over the past 20 years have made it possible to include arbitrarily many series in nowcasting and forecasting systems and to incorporate data release in real time. The current literature has shown large improvements in this regard (see, e.g., Stock & Watson (2002*b*) for forecasting based on a large number of predictors and Giannone et al. (2008) for improved nowcasting performance using high-dimensional time series). However, there is some evidence that the parametric restrictions (or priors) that make these models work discard potentially important information (Carrasco & Rossi 2016, Stock & Watson 2012). Instead, there is an exploitable nonlinear and non-parametric structure that could be revealed by modern ML algorithms (Stock & Watson 2017). ML algorithms have recently been proposed as new tools by which empirical economists could solve the prediction problem rather than the parameters estimation problem (Mullainathan & Spiess 2017, Athey 2018). They offer a solution for *the curse of dimensionality and regularisation* to standard regression analysis without having the need to transform variables to latent and unobservable factors (Athey 2018, Athey & Imbens 2019).

ML algorithms have been applied to macroeconomic forecasting and have been found effective in improving forecasting accuracy. For instance, Ahmed et al. (2010) conducted a large-scale comparison study of the major ML models for time series forecasting using the M3 competition dataset.[1] They found the best two methods to be the multilayer per-

[1]The competition was organised by the International Journal of Forecasting (Makridakis and Hibon,

ceptron and the Gaussian process regression. Chakraborty & Joseph (2017) introduced ML algorithms to the context of central banking and policy analysis and concluded that ML algorithms generally outperform traditional modelling approaches in the prediction task. Gogas et al. (2015) found that an ML framework is useful for improving the accuracy of forecasts of the inflation and the output gap. Additionally, Plakandaras et al. (2015) proposed several ML algorithms and found that they can reduce by half the mean squared error of the random walk model. Lin et al. (2012) used least square support vector regression to forecast foreign exchange rate and found that it can outperform the autoregressive integrated moving average (ARIMA) model. Further, Buchen & Wohlrabe (2011) found that boosting is a serious competitor for dynamic factor models in terms of forecasting United States industrial production. Zeng (2017) found that using the boosting technique to select disaggregated variables is a feasible and competitive approach in forecasting Euro macroeconomic key variables. The most comprehensive study to date was conducted by Goulet et al. (2019). They moved beyond the question of, 'Is machine learning useful for macroeconomic forecasting?' by adding the question, 'how' and they found that machine learning is useful for macroeconomic forecasting because it mostly captures important nonlinearities that arise in the context of uncertainty and financial frictions.

In the existing literature, limited studies address the problem of nowcasting China's GDP. Yiu & Chow (2010) used the dynamic factor model to nowcast the current quarter GDP growth rate and found that the dynamic factor model can outperform the benchmark random walk model. Barnett & Tang (2016) also found that monetary aggregate variables are more informative than other variables and thereby help the dynamic factor model produce the best available nowcasting results. Zhang et al. (2018) found that the Bayesian approach may be a viable option for nowcasting China's GDP. However, these studies are

2000) and attracted much attention.

insufficiently designed because the chosen benchmark models are not appropriate and the researchers only consider one model at a time. This creates an urgent need for evaluating a wide range of nowcasting models for China in a 'benchmark free' environment.

In response to this issue, we investigate the performance of different ML algorithms in obtaining accurate nowcasts of the Chinese GDP growth rate. We compare the predictive accuracy of ML algorithms with those of the dynamic factor model by Giannone et al. (2008) and the autoregressive (AR) model. Naturally constrained by data availability and relatively short samples for many series, we implement nowcasting algorithms with a relatively small-scale dataset (compared to Western economics) of monthly and quarterly indicators, which are manually selected to best represent the Chinese economy. Our dataset contains 17 quarterly series and 24 monthly series over the 1993Q1 to 2018Q2 period. To examine which type of data is more informative, we divide our dataset into two subsets: the soft dataset and the hard dataset. The soft dataset contains survey data, financial market variables and various price indexes, whereas the hard dataset contains variables used for GDP calculations, such as the industrial production index and international trade data.[2] We pay particular attention to soft data as they are released with a shorter delay than official GDP statistics and are not subject to benchmark revision. All data are released before the announcement of official GDP figures. We evaluate the nowcasting performance of each model using the model confidence set (MCS) developed by Hansen et al. (2011) and we report p-values of the MCS based on three datasets: the total dataset, the soft dataset and the hard dataset.[3]

[2]Sources: Federal Reserve Bank of St.Louis

[3]Model selection, and its impact on macroeconomic forecasting, has been a long-standing topic in applied macroeconometrics. The reason for this was that the most parsimonious model containing the data-generating process is generally unknown in empirical work. Hansen et al. (2011) developed the MCS that estimates a set of superior models from an initially chosen set, where superiority is defined by a user-specified loss function. It is constructed by first finding the best forecasting model, and then selecting the subset of models that are not significantly different from the best model at a desired confidence level. The MCS has been widely used in selecting the set of best forecasting models in the literature, see for example Kotchoni et al. (2019), Goulet et al. (2019).

This empirical paper joins the growing body of empirical application evaluating the relative success of ML models in macroeconomic forecasting literature. To the best of our knowledge, ours is the first empirical paper to consider nowcasting China's GDP growth rate using ML algorithms. This paper empirically contributes to the existing literature in two ways. First, we compare a large number of nowcasting algorithms that have been widely used by Western economies, including traditional time series, the dynamic factor model, ML algorithms and deep learning specifications. Second, our set of ML algorithms and deep learning specifications have the ability to capture nonlinear and non-parametric relationships. Contrastingly, existing studies such as those by Yiu & Chow (2010) and Zhang et al. (2018) typically consider one model at a time and only consider linear and parametric models.

The remainder of the paper is organised as follows. Section 4.2 presents nowcast evaluation methodology and describes the data used. Section 4.3 explains nowcasting algorithms. The empirical results are discussed in Section 4.4, and Section 4.5 provides concluding remarks.

4.2 Nowcast Evaluation and Data

4.2.1 Nowcast Evaluation Methodology

We evaluate the nowcasting performance using the MCS by Hansen et al. (2011). The MCS a set of models that contains the best model(s) from a collection of models with a given level of confidence, wherein the definition of 'best model(s)' can be user specified. The MCS is constructed from sample information about the relative performances of all competing models; therefore, it is a random data-dependent set of models that contains the best forecasting model(s), just like a standard confidence interval for the population parameter. Compared to the tests for superior predictive ability, there are two attractive features of the MCS approach. First, it acknowledges the limitations of the data. More informative data will result in a MCS that contains only the best model, whereas less informative data will yield a MCS that contains many (possibly all) models.[4] Second, the MCS is benchmark free and allows for more than one 'best' performing model to be selected. This is particularly useful in our empirical study since there are limited studies concerning superior nowcasting models for China's GDP in the current literature. In contrast, superior predictive ability tests such as those by Francis & Roberto (1995), Hansen (2005), Giacomini & White (2006) are designed to test whether a particular forecasting model can significantly outperform the benchmark model. The use of superior predictive ability tests could be less effective if there are no suggested appropriate benchmark models in the literature. When the superior predictive ability test is rejected, there is little guidance about which set of models is the best. It is possible that almost all models could be significantly better than the benchmark model. Similarly, when the superior predictive ability test is not rejected, it is possible that the benchmark model is the best model, but this may also apply to several other models.

We evaluate the performance of the nowcasting algorithms using an out-of-sample simulation. We adopt a fixed estimation window method in which we use the first R observations to estimate parameters of each model where R refers the number of within-sample periods and $T - R$ refers to the number of out-of-sample periods. Then, we generate the nowcast for the time period $R + 1$. To compute the nowcast for the period $R+2$, we drop the first observation from the previous estimation sample and add an extra observation to the end of the previous estimation sample. Parameters of each model

[4]For example, Hansen et al. (2011) replicated Stock & Watson (1999)'s study by computing a MCS across all the Stock and Watson inflation forecasting models. They found that the MCS of subsample period 1970–1983 contains only five forecasting models, while the MCS of subsample period 1984–1996 contains 14 models. This indicates that the earlier sample possessed useful information to differentiate the forecast, whereas the later sample was less informative.

are re-estimated, and the nowcast is computed based on estimated parameters. This process—updating the estimation sample, re-estimating model parameters and computing a nowcast for the next time period—will be continued until $T - R$ time period out-of-sample nowcasts have been computed. We consider the total sample size in a 70:30 per cent split between in-sample and out-of-sample periods to ensure a sufficient number of the in-sample period for model selection and evaluation.[5] Therefore, the within-sample period covers the period from 1993Q1 to 2018Q2. The first out-of-sample nowcast of China's nominal GDP is made for 2010Q4. To compute this nowcast, we deseasonalise all series, screen for potential outliers, standardise all series and estimate models using data only available from 1993Q1 to 2010Q3. To compute the next nowcast, we re-determine the value of hyperparameters and re-estimate models using data from 1993Q2 to 2010Q4. This process repeats until the final simulated out-of-sample nowcast is made for 2018Q2. The MCS is implemented using the MCSprocedure function from the MCS package by Catania & Bernardi (2019), and we use the MSE as the loss function in the MCS.

4.2.2 Data

We collect a panel of 17 quarterly series and 29 monthly series that we believe best represent China's macroeconomy, spanning from 1993:01 to 2018:06. Our sample period covers the initial transition of the Chinese economy since the reform and opening up, the Asian financial crisis, the sharp downturn during the Global Financial Crisis and China's rapid subsequent recovery. Our panel covers all important sectors of China's macroeconomy, including the government sector, international trade, the industrial sector, the service sector, the money and credit sector, and the financial sector. It is important to note that, naturally constrained by data availability, our dataset covers a smaller number

[5] We perform the robustness analysis using a 50:50 split and find similar nowcasting results.

of variables than other studies. For example, Yiu & Chow (2010) used 189 macroeconomic and financial series and Zhang et al. (2018) constructed a panel of 159 monthly variables spanning from 1999 to 2009. The trade-off in our circumstance is a large number of predictors but a small number of time observations or a large number of time observations but a small number of predictors. To ensure a sufficient number of within-sample periods and out-of-sample periods, we opt for the latter. The selection criteria for indicators is based on economic theory and judgement. Generally, we tend to select key indicators for certain sectors rather than import all available series. All series were released before the announcement of the official GDP figure. We source our Chinese data from Thomson Reuters Datastream and the Federal Reserve Bank of Atlanta - China's Macroeconomy: Time Series Data. The variables used to compute the machine learning and the dynamic factor model nowcasts are shown in Table B.0.3.

The stationarity of X_t is a moment condition for ML algorithms, the dynamic factor model and the AR model (Stock & Watson 2016, Athey 2018), so all series were subjected to four preliminary steps: possible deseasonalisation, possible transformation by taking either the difference or the percentage change, standardisation, and screening for possible outliers. Seasonal patterns are common in China's macroeconomic data and we use *X-13ARIMA-SEATS* in *R* platform to address them.[6] The decision to take the difference or the percentage change was made judgementally, mainly based on inspection of the time series plots and unit root tests. Generally, the difference is taken for those series that are already in the index or percentage terms, and the percentage change is implemented to series with actual quantities. We also clean the dataset by removing outliers for which the first difference deviates from the median of the first difference by more than five times the interquartile range and replacing them with missing data. Next, all monthly series

[6]Developed by the United States Census Bureau, the *X-13* supports user-defined holiday variables such as Chinese New Year or Indian Diwali. It returns deseasonalised series if seasonality is detected and returns original series if seasonality is not detected.

for the construction of the dynamic factor model are further standardised to mean zero and unit variance.

4.3 Nowcasting Algorithm

We consider algorithms that form predictions about the current growth rate of Chinese nominal GDP on the basis of the available historical and contemporaneous information. In total, there are 10 nowcasts computed from a range of different models. In the following, $\Omega_t = \sigma(r_t, r_{t-1}, ...)$ denotes the information set wherein r_t is a vector containing all independent variables in time period t. In this study, we train each algorithm over a fixed window procedure. This method implies that all models are estimated based on the within-sample (training) information set $\Omega_{train} = t|t = 1, 2, 3, ..., (T - 1)$. Next, trained algorithms are employed to construct a nowcast of the GDP growth rate at time period T. To obtain the nowcast at the next quarter $(T + 1)$, the training information set is moved to $\Omega_{train} = t|t = 2, 3, ..., T$ and the parameters of each model are re-estimated to make the nowcast. This procedure is repeated until the final nowcast is computed.

Model selection is performed using information criteria for the AR model forecasts and the dynamic factor model forecasts. Cross-validation is used for ML and deep learning algorithms.[7] More precisely, we use K-fold cross-validation by splitting the training dataset into 10 consecutive folds (i.e., K = 10). For $K = 1, 2, ..., 10$, each fold is used to evaluate the performance of the ML algorithm fitted on the remaining K-1 folds. This produces 10 estimations of the performance measure (in this case, the mean squared error [MSE]). Hyperparameters are then determined by the smallest average of the MSEs, and

[7] As Bergmeir et al. (2018) pointed out, the use of K-fold cross-validation for time series forecasting is appropriate only if model errors are not serially correlated. We performed a robustness analysis to check whether models have correlated errors. Based on the results of ADF tests, we found that nowcasting models have uncorrelated errors in most sample periods.

the final model is re-estimated on the training dataset.[8]

For each ML algorithm considered, there are myriad variations proposed in the literature, so it would be a challenge to employ all possible varieties. For example, common variations of artificial neural networks include feed-forward neural networks, radial basis function neural networks, multilayer perceptrons, convolutional neural networks, recurrent neural networks, modular neural networks and sequence-to-sequence models. Our strategy is to consider the basic version of each algorithm. The rationale is that most users are most likely to employ the basic form of ML algorithms, at least initially. In total, we consider eight ML algorithms, the dynamic factor model in the state-space representation, and the AR model. These algorithms are some of the most commonly used models in macroeconomic forecasting literature. We implement the computation of our ML algorithm in the *R* platform. Below are descriptions of the algorithms considered.

4.3.1 Machine Learning Algorithms

LASSO, Ridge and Elastic Net To introduce background information relevant to the discussion of the *curse of dimensionality*, we establish some concepts first. Let T be the number of time observations and N to be the number of predictors. We have:

$$Y = X\beta + U, \tag{4.1}$$

where Y is the $T \times 1$ vector of observed response, X is the $T \times N$ matrix of observed predictors, U is an unobserved zero mean error or nuisance term and β is the vector of the regression coefficients. The most straightforward solution to (4.1) is the ordinary least squares (OLS) of $\hat{\beta}_{OLS} = (X'X)^{-1}X'Y$, provided that rank (X)=N and $T \geq N$. In high-

[8] Hyperparameters refer to the parameters in the penalty function. Examples of hyperparameters are λ in a LASSO regression or the number of lags in an AR model.

dimensional situations in which the number of predictors is larger than the number of time observations, the matrix $X'X$ is not full rank and is not invertible; hence $\hat{\beta}_{OLS}$ does not exists. An additional difficulty arises when the matrix $X'X$ has eigenvalues that are close to zero. This situation could prevent a stable determination of the minimum-norm least-squares because small fluctuations in the response vector Y will be amplified by the effect of the small eigenvalues, which will cause uncontrolled fluctuations in the estimation of β (Reichlin et al. 2017). In time series forecasting, this implies that a good fit of the observed data can easily be obtained, but a good predictive power for responses corresponding to new observations cannot (good in-sample fit, poor out-of-sample forecasts) (Athey 2018, Mullainathan & Spiess 2017).

LASSO, ridge and elastic net are similar methods by which to avoid overfitting the regression model by incorporating different types of constraints or penalties on the vector of the regression coefficients β. Penalties that allow for variable selection by enforcing some of the regression coefficients as zero have been popularised in the ML and econometric literature under the name of *LASSO* regression. Such penalties yield a penalty term proportional to the L1-norm of β by incorporating penalties on the absolute value of the magnitude of the coefficients and shrinking some of them to zero (the variable selection). Estimation of $\hat{\beta}_{LASSO}$ can be achieved by minimising the following objective function:

$$\hat{\beta}_{LASSO} = \text{argmin}_{\beta}[\| Y - X\beta \|_2^2 + \lambda_1 \| \beta \|_1], \tag{4.2}$$

where $\| \beta \|_1 = \sum_{j=1}^{N} |\beta_j|$. The LASSO penalty provides a nonlinear shrinkage of the component $\hat{\beta}_{OLS}$. In contrast to $\hat{\beta}_{ridge}$, there is no closed-form expression of $\hat{\beta}_{LASSO}$ in terms of the general matrix of X, except for one special case of orthonormal regressors

$(X'X = \mathbf{I})$ where $\hat{\beta}_{LASSO}$ is given by:

$$\hat{\beta}_{LASSO} = S_{\lambda_1}([X'Y]_j), \qquad S_{\lambda_1} = \begin{cases} x + \lambda_1/2 & \text{if} \quad x \leq -\lambda_1/2 \\ 0 & \text{if} \quad |x| \leq \lambda_1/2 \\ x - \lambda_1/2 & \text{if} \quad x \geq \lambda_1/2 \end{cases}, \tag{4.3}$$

where S_{λ_1} is the 'soft thresholder'. In all other cases, the estimation of $\hat{\beta}_{LASSO}$ must be solved numerically as a convex optimisation problem. We use the glmnet function from the glmnet package by Friedman et al. (2010) to estimate the LASSO regression (by setting $\alpha = 1$ so that we have 100% L1-norms and no L2-norm types of penalty).

The ridge regression estimator consists of adding a penalty to the least square objective function in proportion to the size of β, measured by its squared L2-norm. Therefore, elements of the coefficient vector β estimated by the ridge estimator are shrunk uniformly towards zero. The ridge regression estimator is given by:

$$\hat{\beta}_{ridge} = \text{argmin}_{\beta}[\| Y - X\beta \|_2^2 + \lambda_2 \| \beta \|_2^2] \tag{4.4}$$

$$= (X'X + \lambda_2 I)^{-1} X'Y, \tag{4.5}$$

where $\| \beta \|_2^2 = \sum_{j=1}^{p} \beta_j^2$ and $\mathbf{I}$ is the identity matrix. The ridge regression is implemented with the glmnet function (by setting $\alpha = 0$) from the glmnet package by Friedman et al. (2010).

The elastic net regression incorporates a mixture of the L1-norm and L2-norm penalties so it can be interpreted as a hybrid version of the ridge and LASSO regressions. It can either shrink the coefficients close to zero or completely remove them. The elastic net regression estimator $\hat{\beta}_{en}$ is given by:

$$\hat{\beta}_{en} = \text{argmin}_{\beta}[\parallel Y - X\beta \parallel_2^2 + \lambda_1 \parallel \beta \parallel_1 + \lambda_2 \parallel \beta \parallel_2^2]. \tag{4.6}$$

The elastic net regression is fitted using the glmnet function from the glmnet package by Friedman et al. (2010) (setting $\alpha = 0.5$ allows one to achieve an equal weighting of L1-norm and L2-norm penalties). In the ridge, LASSO and elastic net regressions, the values of hyperparameters λ_1 and λ_2 control the strength of the penalty. A larger value leads to a greater penalty and a stronger effect of shrinkage and regularisation of regression coefficients. We select the value of hyperparameters by the K-Fold cross-validation of the training dataset.

Support Vector Regression The objective of the support vector machine (SVM) is to find a hyperplane in an N-dimensional space that distinctly classifies data points. For our nowcast modelling, we use the epsilon-insensitive SVM with a polynomial kernel trick (ϵ-SVM regression hereafter). More specifically, we train both the predictors and the response variable to fit a linear function that deviates from the response variable to an extent no greater than ϵ for each training point. The ϵ-SVM regression estimator can be achieved by minimising the following function:

$$\hat{\beta}_{SVR} = \text{argmin}_{\beta}[\parallel V(Y - X\beta) \parallel_2^2 + \frac{\lambda}{2} \parallel \beta \parallel_2^2], \tag{4.7}$$

where

$$V_{\epsilon}(u) = \begin{Bmatrix} 0 & \text{if} \quad |u| < \epsilon \\ |u| - \epsilon & \text{otherwise,} \end{Bmatrix} \tag{4.8}$$

and $\parallel \beta \parallel_2^2 = \sum_{j=1}^{p} \beta_j^2$. Based on the above equations, the support vector regression ignores nowcast errors with a size smaller than ϵ while still implementing penalties on the coefficients via the L2-norm. The ϵ-SVM is implemented through the ksvm function from the kernlab package by Karatzoglou et al. (2004). Two hyperparameters, ϵ and λ, are chosen by K-fold cross-validation of each training dataset for every within-sample period.[9]

Feed-Forward Neural Network Inspired by how the human brain works, the neural network is an information processing algorithm that resembles biological nervous systems. While there are several variations of the neural network, we consider the feed-forward neural network, which is perhaps the simplest type of artificial neural network.[10] One distinct feature of this network is that connections between nodes do not form a cycle. The model works forward only, by feeding datasets into the network through input neurons, which trigger the layers of hidden neurons and finally turn into output neurons. More precisely, each input is weighted and passed through an activation function to determine the value of nodes in the first layer of perceptrons.[11] Then, nodes from the first layer become new inputs (re-weighted) for the second layer and are passed through a new activation function to determine the value of nodes for the third layers. This process is repeated until the N_{th} layer is created. Finally, nodes in the last layer are re-weighted and passed through the final activation function to generate the output value. The key hyperparameters of the feed-forward neural network are the number of nodes and the number of hidden layers. We use a single hidden layer because the universal approximation theorem states that the feed-forward neural network with a single hidden layer and a finite number of neurons can approximate any continuous function for inputs within a specific

[9]The root mean squared error is used as a performance metric in the cross-validation process. Hyperparameters are tuned to optimise the out-of-sample mean squared error.

[10]The identity function is used as the activation function. The robustness check was performed using two and three hidden layers and a different number of neurons. The results indicate that the single hidden feed-forward neural network is the best nowcasting model among various feed-forward neural nets.

[11]A perceptron is an algorithm used for supervised learning of binary classifiers.

range (Hornik et al. 1989). The number of nodes is selected via K-fold cross-validation. We use the nnet function from the nnet package by Venables & Ripley (2002) to implement the feed-forward neural network.

K-Nearest Neighbour Regression K-nearest neighbour (KNN) regression is a non-parametric method that works by predicting the outcome of new data based on K most similar observations (neighbours) in the training dataset. Specifically, given a data point (in this case, the nominal GDP growth rate), we compute the Euclidean distance between that point and all points in the training dataset. Then, the nowcast of China's GDP growth rate is the average of the target output value by the closet K training data points. Let $\phi(x)$ be the set of KNNs. The nowcast is given by:

$$\hat{y} = \frac{1}{K} \sum_{m \in \phi(x)} y_m, \tag{4.9}$$

where y_m is the target output for training data point x_m. The KNN regression is fitted using the knnreg function from the caret package by Kuhn (2008). The hyperparameter is the value of K and is determined by K-fold cross-validation.

Decision Tree and Random Forest for Regression Generally, the decision tree for regression is based on a hierarchical tree-like partition of the input space wherein data are partitioned into smaller and smaller subsets that contain similar values. The tree consists of internal decision nodes and terminal leaves. For any given data point, the flow starts with the root node and then moves to the path along the tree until it reaches a leaf node by a sequence of tests and decision nodes. A prediction can be made according to the local model of the leaf node. In the regression algorithm, the tree uses within-node variance to split data according to paths that result in the highest variance reduction. The splitting of the data is performed recursively until the termination criterion is reached. The final prediction is typically the within-node sample mean of each leaf node. The hyperparameters are the maximum depth of the tree, the minimum number of samples required to split an internal node and the number of features to consider when searching for the best split. The decision tree is fitted using he rpart function from the rpart package by Therneau & Atkinson (2019).

The random forest for regression is an ensemble ML algorithm[12] that constructs a number of decision trees on subsamples of the dataset and makes a prediction based on the average of the leaf nodes to reduce model variance. Each tree is built on d features randomly selected from n elements of the vector of conditioning variables. The hyperparameters for the random forest include d, n, the number of trees built and the hyperparameters for the decision tree. More trees built leads to better performance but also requires more computational power. We select the value of the hyperparameters via cross-validation. The random forest is fitted using the randomForest function from the randomForest package by Liaw & Wiener (2002).

4.3.2 The Dynamic Factor Model

The dynamic factor model in the state-space representation is a formalised nowcasting method within a coherent statistical framework that started garnering attention in the academic literature after the works of Giannone et al. (2008) and Evans (2005). An important feature of Giannone et al. (2008)'s framework is that the model allows researchers to interpret and comment on various data released at different frequencies with different publication lags in terms of the signals they provide on current economic conditions. This is possible because the Kalman filter generates projections of missing observations for all variables and updates the nowcast as new data are released. Subsequent research has extended to using the dynamic factor model to nowcast current GDP for many countries

[12] An ensemble algorithm combines the predictions from multiple ML algorithms to make more accurate predictions than any individual model.

(see Lahiri & Monokroussos (2013) for the United States; Angelini et al. (2008), Bańbura & Rünstler (2011) and Camacho & Perez-Quiros (2010) for the aggregate Euro area; Barhoumi et al. (2010) for France; Matheson (2010) for New Zealand; and Siliverstovs & Kholodilin (2012) for Switzerland). The dynamic factor model has been widely used by a number of international institutions and central banks as a standard tool for up-to-date GDP nowcasting, some example users being the Federal Reserve Bank of Atlanta (Higgins 2014),[13] the Federal Reserve Bank of New York (Aarons et al. 2016),[14] the Bank of England (Hendry et al. 2016),[15] the Reserve Bank of Australia (Dungey et al. 2018),[16] the European Central Bank (Bańbura et al. 2013)[17] and the International Monetary Fund (Matheson 2011).

To motivate the dynamic factor model, we assume that a joint model for observed variables is specified and that it has a state-space representation as below:

$$X_t = \mu + \zeta(\theta)Z_t + G_t, G_t \sim i.i.d.N(0, \sum_G(\theta)), \tag{4.10}$$

$$Z_t = \varphi(\theta)Z_{t-1} + H_t, H_t \sim i.i.d.N(0, \sum_H(\theta)), \tag{4.11}$$

where (4.10) is the measurement equation that links the vector of observed variables X_t to a vector of state variables (possibly unobserved) Z_t, and (4.11) is the transition equation

[13]The Quantitative Economic Research Department in the Federal Reserve Bank of Atlanta frequently produces nowcasts of United States GDP before United States Bureau official releases. It nowcasts 13 separate expenditure components of the GDP to mimic the expenditure approach to calculating GDP using dynamic factor modelling.

[14]The Federal Reserve Bank of New York uses a dynamic factor approach to provide a model-based nowcast of GDP. However, it does not mimic a particular approach to calculating GDP.

[15]The Bank of England's Monetary Policy Committee uses a compilation of nowcasts from three different models (the MIDAS model, the dynamic factor model and judgements based on different industries) to form its initial view on the current state of the economy.

[16]The Reserve Bank of Australia uses the dynamic factor model to nowcast components of GDP using the aggregate expenditure approach.

[17]The European Central Bank also uses dynamic factor-based nowcasting models to inform its policy decisions. Its staff have released a number of working papers, which are cornerstones of the nowcasting literature.

that specifies the dynamics of the vector of state variables (see, e.g., Harvey (1990) for a comprehensive treatment of state-space models). It is important to note that a state-space model with representation (4.10)–(4.11) and the parameter θ can efficiently deal with any missing observations using the Kalman filter and Kalman smoother, which provide the conditional expectation of the state vector based on the information set Ω_v and associated precision:

$$Z_{t|\Omega} = E[Z_t|\Omega_\nu] \quad P_{t|\Omega_\nu} = E_\theta[(Z_t - E_\theta[Z_t|\Omega_\nu])(Z_t - E_\theta[Z_t|\Omega_\nu])']. \tag{4.12}$$

In practice, the issues of mixed-frequency data and publication delays can be considered the missing data problem, which can be easily solved by the Kalman filter and Kalman smoother apparatus (Lahiri & Monokroussos 2013, Giannone et al. 2008). When new information is released, it can be linked to nowcast revision. Different versions of the general model (4.10)–(4.11) have been considered in the current literature; the dynamic factor model proposed by Giannone et al. (2008) is the most commonly used.

The dynamic factor model is well suited to a context in which parameter proliferation of a parsimonious model should be avoided but, at the same time, the salient features of high-dimensional economic data should be captured. In the dynamic factor model framework, each economic variable can be divided by the sum of two orthogonal components: (1) a handful of latent and unobserved factors that capture the joint dynamics and (2) an idiosyncratic residual. Similarly to Giannone et al. (2008)'s framework, we specify that the high-frequency variable (in this case, monthly time series), X_t, has a factor structure following a VAR process:

$$X_t = \mu + \Lambda F_t + E_t, \quad E_t \sim i.i.d.N(0, \Sigma_E), \tag{4.13}$$

$$F_t = AF_{t-1} + Bu_t, \quad u_t \sim WN(0, I_q), \tag{4.14}$$

where r is the number of common factors, q is the number of macroeconomic shocks, X_t is $N \times 1$ vector of monthly series, Λ is $N \times r$ of the factor loading, F_t is $r \times 1$ vector of factors and E_t is an $N \times 1$ residual vector. B is a $r \times q$ matrix of full rank q, A is a $r \times r$ matrix, all roots of $\det(I_r - A_Z)$ lie outside the unit circle and u_t is a q dimensional vector of the macroeconomic stochastic shocks to the common factors. Giannone et al. (2008) further assumed that $\sum_E$ is diagonal.

Giannone et al. (2008) estimated the dynamic factor model (4.13)-(4.14) using a two-step procedure. In the first step, the factors and factor loadings are estimated using principal component analysis applied to a balanced panel of X_t by considering only the series for which all observations are observed. In the second step, factors are re-estimated by applying the Kalman smoother to the entire information set. Unlike in the work of Giannone et al. (2008), which used United States time series, we raise particular concerns about issues related to the availability of relevant Chinese macroeconomic variables, most of which started being published only recently and encounter issues related to missing observations during Chinese New Year.[18] These two issues could potentially eliminate the majority of variables in the first estimation procedure step. We adopt the first step used by Giannone et al. (2008), using an expectation–maximisation (EM) algorithm to estimate the factors and the factor loadings. Stock & Watson (2002*b*) demonstrated that the EM algorithm is an iterative algorithm for efficient estimation and can consistently estimate the factors and factor loadings when data irregularities are present. Then, we update the estimated factors and factor loadings using the Kalman smoother. The Kalman smoother exploits the dynamics of the common factors and the cross-sectional heteroscedasticity of the idiosyncratic components, thereby providing efficiency improvements over simple principal component analysis. The nowcasts are then obtained through a regression of

[18]In the replication file of the research by Giannone et al. (2008), the data only contain few missing observations at the end of the sample period.

GDP on temporally aggregated factors:

$$y_{t,1}^{k1} = \hat{\alpha} + \hat{\beta} F_{t|\Omega_v}^{k1} + e_t^{k1}, \quad t = k_1, 2k_1, ... \tag{4.15}$$

The key hyperparameters in the dynamic factor model are the number of common factors (r) and the number of shocks (q) in equation (4.13)–(4.14). We assume that the macroeconomic dynamics are affected by the real shock and monetary shock (i.e., q = 2), and we select the number of common factors (r) using Bai & Ng (2002)'s information criteria. The dynamic factor model is fitted using the nowcast function from the nowcasting package by Daiane Marcolino de Mattos (2019).[19]

4.3.3 Autoregressive Model

The AR model is a commonly used univariate model in the literature of nowcasting the Chinese GDP growth rate (Zhang et al. 2018, Zeng 2017, Yiu & Chow 2010). The key hyperparameter in the AR model is the lag order p. It can be determined by the sequential downward t test or information criteria such as the Akaike information criterion (AIC) and the Bayesian information criterion. In our application, we choose the lag order by the minimum AIC of the training dataset. We have:

$$\hat{y}_t = \hat{\alpha}_0 + \sum_{j=1}^{p} \hat{\alpha}_j y_{t-j} + u_t, \tag{4.16}$$

where $\hat{\alpha}_0$ is a constant term, $\sum_{j=1}^{p} \hat{\alpha}_j$ is the lag operator of coefficients and u_t is white noise.[20]

[19] The 'nowcasting' package in R provides the tools to make forecasts of monthly or quarterly economic variables using dynamic factor models. It is published in the R journal (a peer-reviewed journal).

[20] Compared to Western countries, data on GDP growth rate is released relatively faster by China. The National Bureau of Statistics of China usually releases GDP figures with a four- to six-week delay, so the direct AR approach that forecasts y_{t+1} using y_{t-1} is not necessary. In contrast, most Western countries release GDP figures with a three- to four-month delay, which causes a timing issue; hence, the direct AR

4.4 Empirical Results

In this section, we describe the main results of our analysis. We compute the MCS across all the competing models and set the level of significance to 10%. Thus, models with p-values greater than 10% are covered by the MCS, while models with p-values less than 10% are not covered by the MCS. We also calculate the relative MSE of each model. Table 4.1 documents the relative MSE and MCS p-value of each nowcasting algorithm for the simulated out-of-sample period 2010Q4 to 2018Q2.

Table 4.1: Relative MSE and MCS P-Value Results

	Total Dataset		Soft Dataset		Hard Dataset	
Models	rMSE	p-value	rMSE	p-value	rMSE	p-value
AR	1.00	0.01	1.00	0.05	1.00	0.05
Ridge	0.40	0.12*	2.78	0.03	0.45	1.00*
LASSO	0.32	1.00*	2.62	0.02	0.51	0.85*
Elastic Net	0.46	0.84*	2.87	0.01	0.49	0.80*
Support Vector Machine	0.45	0.79*	2.70	0.05	0.58	0.60*
K-Nearest Neighbour	1.96	0.00	2.35	0.00	0.71	0.58*
Feed-Forward Neural Network	2.87	0.00	3.15	0.03	2.92	0.00
Random Forest	1.16	0.02	3.05	0.02	0.82	0.71*
Decision Tree	0.47	0.55*	0.57	1.00*	0.46	0.93*
Dynamic Factor Model	4.58	0.00	4.38	0.03	4.20	0.00

Table 4.1 reveals that the MCS of the total dataset, the MCS of the soft dataset and the MCS of the hard dataset are strikingly different.[21] The MCS of the soft data contains only the regression tree model at the 90% level of confidence, $\widehat{M}^*_{90\%}$, whereas only the AR model, feed-forward neural network and dynamic factor model fail to enter into $\widehat{M}^*_{90\%}$ based on the hard dataset. Therefore, the soft dataset possesses useful information to differentiate the nowcast, whereas the hard dataset is less informative. Further, ridge,

approach must be considered.

[21]According to Federal Reserve Bank of St. Louis, the soft dataset contains survey data, financial market variables and various price indexes, whereas the hard dataset contains variables used for GDP calculations, such as the industrial production index and international trade data.

LASSO, elastic net and SVM are covered by the MCS at the 10% level of significance in the total dataset. This indicates that the linear shrinkage technique, regardless of whether it is L1-norm of penalty, L2-norm of penalty or a mixture of both, is the best model for the total dataset. The decision tree, which splits data into paths that result in the highest variance reduction and makes predictions from the sample mean of each leaf node, is also included in the $\widehat{M}^*_{90\%}$.

Finally, Table 4.1 supports a rejection of the dynamic factor model as it fails to enter into $\widehat{M}^*_{90\%}$ or even $\widehat{M}^*_{95\%}$ of the three datasets. The MCS p-values for the dynamic factor model are all very small: 0 for the total dataset, 0.03 for the soft dataset and 0 for the hard dataset. This is surprising, because our results are contrary to the actions of most central banks and International Monetary Fund policies, which use dynamic factor-based nowcasting models to inform their policy decisions. We find conclusive evidence that dynamic factor models are inferior to some ML algorithms for nowcasting China's GDP, such as LASSO and elastic net.

It is also important to analyse whether a nowcasting model can predict the direction of changes in a modelled time series and in which cases the direction of changes given by the nowcast agrees with the actual changes in the Chinese GDP. The success ratio, measured by the proportion of correctly predicted signs to the number of total predictions, provides the fraction of times the signs of the actual values are correctly predicted. Figure 4.1 reports the success ratio for ten nowcasting models. The results of the success ratio generally agree with those of the MCS. The ridge, LASSO, elastic net and support vector machine perform well, whereas the AR, the dynamic factor model, feed-forward neural network and K-Nearest neighbour perform relatively poorly in terms of the fractions of correctly predicting the direction of changes in the Chinese GDP growth rate.

Figures 4.2 plots the quarterly GDP growth rate and its nowcasts obtained from each algorithm over the out-of-sample period using the total dataset. Panel (g) illustrates that ridge, LASSO, elastic net and SVM successfully predict the sharp downturn in activity from 2010Q4 to 2011Q4, as well as most fluctuations in the rest of the sample period. Sometimes, these shrinkage techniques fail to capture nowcasting accuracy. For example, they overestimate the growth rate in 2011Q4, 2012Q1 and 2014 while underestimating the growth rate in 2017. Presented in Panel (a), (b), (c), and (d), the nowcasts made with KNN, feed-forward neural network, random forest and decision tree are greater than the actual growth rate of Chinese GDP, even through these four ML algorithms can successfully predict the overall trend for the sample period. The level of overprediction of the current GDP growth rate is more substantial for the KNN and neural network than for the random forest and decision tree. The AR nowcasts in Panel (f) demonstrate a similar historical movement to the actual rate, with a slight overprediction of the GDP rate from 2012Q2 to 2015Q4. As shown in Panel (e), the dynamic factor model in the state-space representation produces nowcasting trends that are opposite to actual numbers between 2012Q1 and 2015Q2. From 2016Q1 onward, the dynamic factor model predicts GDP growth rates that are larger than the actual rate.

Figure 4.1: The Success Ratio of Ten Nowcasting Models

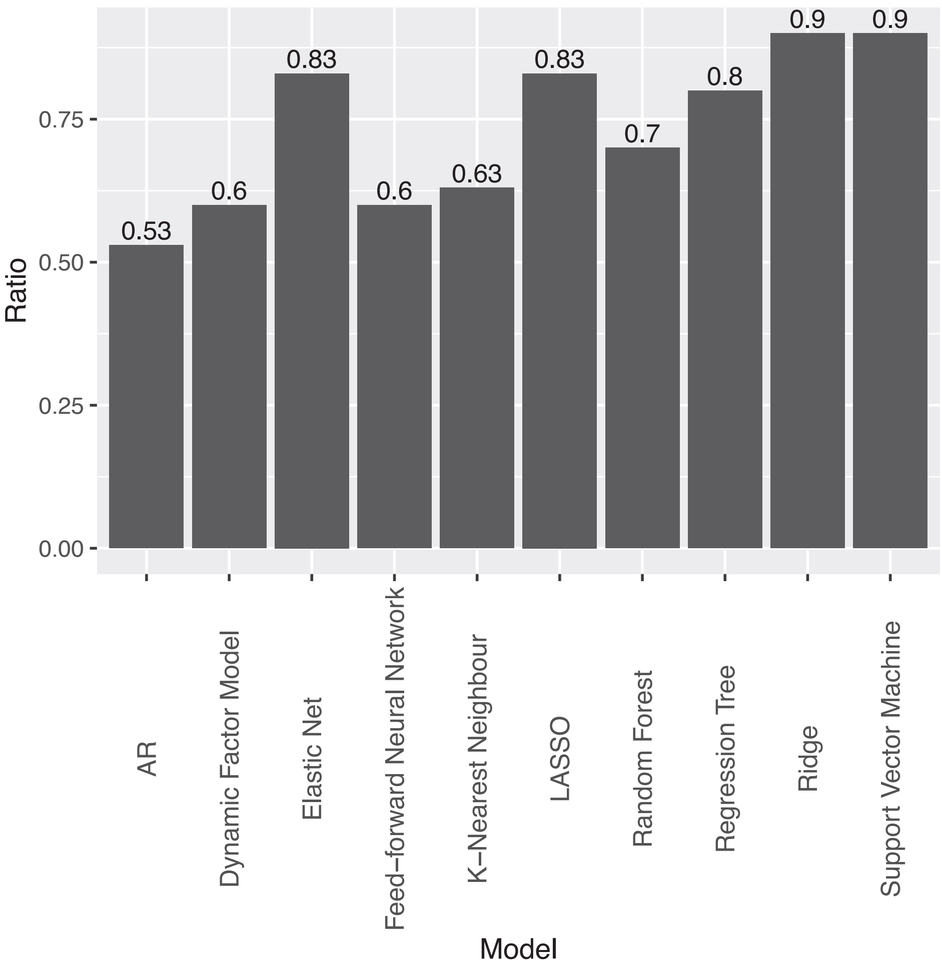
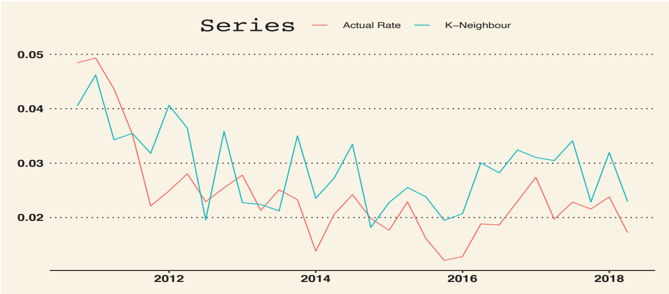
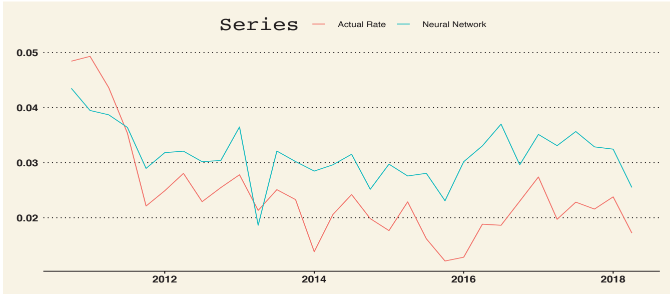


Figure 4.2: Nowcasts of GDP vs. Actual GDP

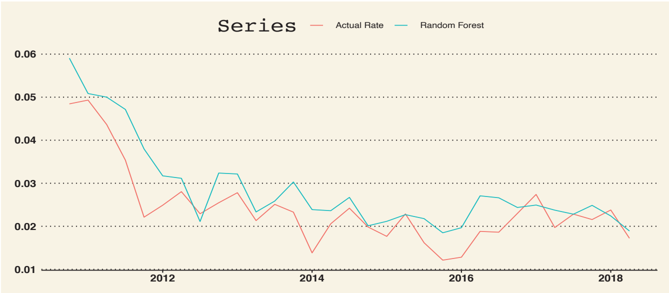
(a) K-Nearest Neighbour Regression



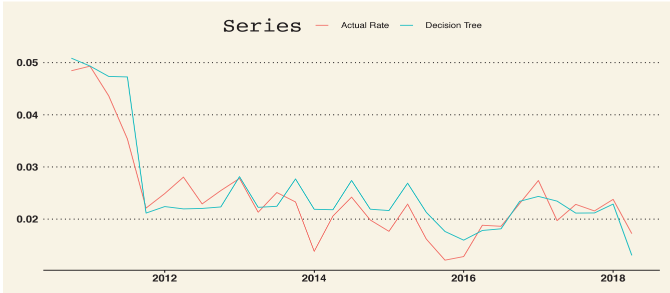
(b) Feed-Forward Neural Network



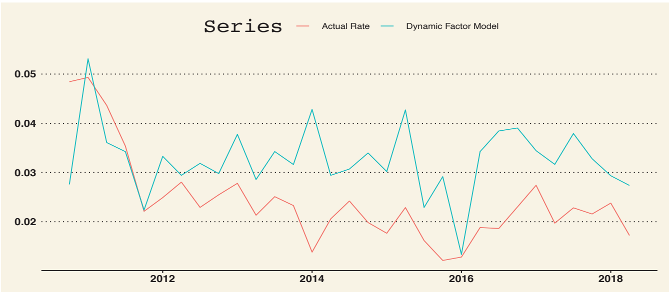
(c) Random Forest



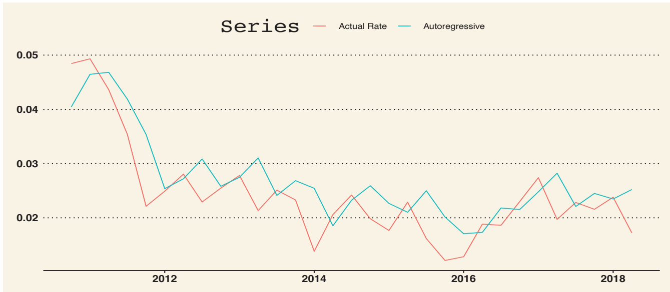
(d) Decision Tree



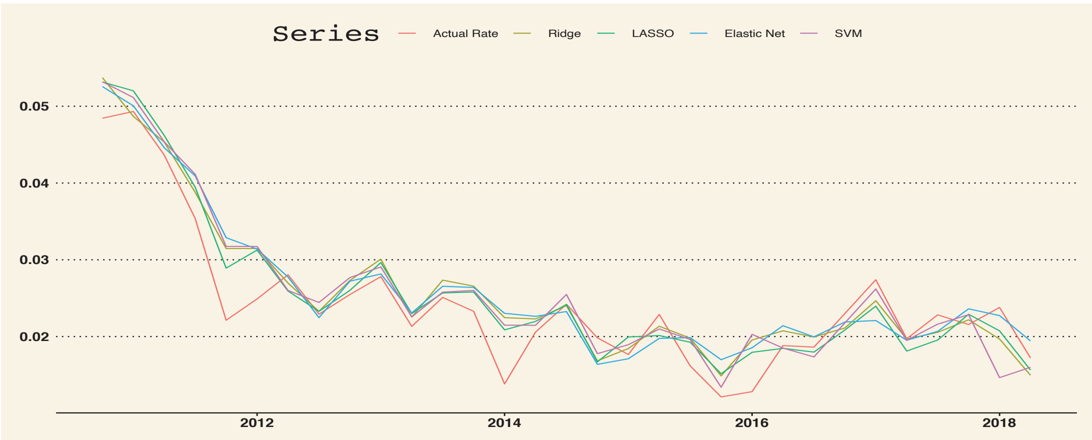
(e) Dynamic Factor Model



(f) Autoregressive



(g) Ridge, LASSO, Elastic Net and SVM



Note: Nominal GDP in percentage change, quarter-on-quarter rate.

4.5 Concluding Remarks

Nowcasting models have become an increasingly popular tool for mitigating uncertainties regarding the current state of the economy and have been widely used by policymakers at many central banks to nowcast the current GDP growth rate. However, the literature discussing nowcasting China's GDP is relatively limited compared to that of Western economies and the performance of various nowcasting models is not well examined. ML algorithms have been recently proposed as an alternative to time series regression models to forecast key macroeconomic variables. In this paper, we evaluate the performance of popular ML algorithms in obtaining accurate nowcasts of the nominal GDP growth rate for China. The ML algorithms used consist of shrinkage methods such as LASSO, ridge, elastic net and SVM regression with Kernel loss function; non-parametric methods such as KNN, feed-forward neural network and decision tree; and ensemble ML algorithms such as the random forest. We then compare the nowcasts obtained from these ML algorithms with the nowcasts generated by the AR model and the dynamic factor model. We collect data over the 1993Q1 to 2018Q2 period, covering all important sectors of China's macroeconomy. We train each ML algorithm based on a fixed estimation window method. The first nowcast of the current nominal GDP growth rate is made for 2010Q4 and the final nowcast is made for 2018Q2. We evaluate the performance of each nowcasting algorithm through the MCS presented by Hansen et al. (2011). We find two notable results. First, based on the total dataset, the ridge, LASSO, elastic net, SVM regression and decision tree are covered by the MCS at the 10% level of significance, indicating that they are the set of 'best models'. Second, compared to the hard dataset, the soft dataset contains more useful information and is more informative for nowcasting the Chinese GDP.

Our paper joins the growing body of literature that addresses the relative success of ML algorithms in the field of macroeconomic nowcasting and forecasting. To the best of our knowledge, our study is one of the first few papers to consider using ML algorithms to nowcast the current state of economic growth for China. Arguably, there may exist other variants of ML algorithms that are not included in this paper. This is unavoidable, as there is a wide range of ML algorithms that could be useful to improve nowcasting accuracy. Our choice of ML algorithms is based on the rationale that most economic forecasters who are not familiar with ML algorithms will be more likely to consider basic forms and select the value of hyperparameters via K-fold cross-validation. Our results show that regularisation techniques yield the set of 'best' models according to the MCS. This is an interesting result, as it provides a direction for future research investigating whether other types of penalties can generate better nowcasts than ridge and LASSO.

A serious competitor in nowcasting GDP is the mixed-data sampling (MIDAS). The MIDAS model has received a good deal of attention in macroeconomic nowcasting literature in recent years. It links the quarterly GDP growth rate with monthly variables and is a serious competitor to the dynamic factor model. The effectiveness of MIDAS in the context of nowcasting Chinese GDP is a promising avenue for future research. The non-parametric factor models require less restrictive assumptions than the parametric factor models and do not impose a prior structure on the underlying data-generating process. Therefore, non-parametric factor models that allow for data-driven flexibility are robust alternatives to parametric factor models and could be useful to improve forecasting performance when the structure of the data-generating process is unknown (Stock & Watson 2016). Overall, we believe that our study not only guides practitioners in selecting appropriate nowcasting models for China's economy, but also provides the research community with more feasible directions for ML algorithms in macroeconomic nowcasting and forecasting.

Chapter 5

Conclusion

In the past several decades, China has experienced tremendous economic progress—the country became the world's second-largest economy in 2010 and the living standards of approximately 400 million people have significantly improved. The connection between China and the rest of the world has grown increasingly closer and the country has entered a new era with the goal of turning into a great modern socialist country in all respects. China's growth and cyclical fluctuations have largely depended on China's macroeconomic policies. For example, Chinese governments have been promoting investment in heavy industries, such as real estate and infrastructure, which constitutes a driving force behind the double-digital growth rate and cyclical fluctuations. Considering this, there has been a lack of rigorous and systematic research in the evaluation of out-of-sample model-based forecasts of China's inflation and real economic variables. This paper aims to fill the gap in the research on macroeconomic forecasting in China.

Households, firms and policy-makers rely on macroeconomic data as inputs for forecasting models to make timely predictions and decisions. If contamination of the official macroeconomic data is a severe problem, forecasting models can be misspecified, possibly causing forecasting results to be biased. As such, Chapter 2 evaluates the reliability of China's official macroeconomic data through two aspects. First, the existing literature is analysed to explain the suspicion of China's economic data. During this search, we find evidence that NBSC may engage in intentional data falsification and manipulation. The Chinese government can intervene arbitrarily to influence the level of multiple key macroeconomic indicators. At the local government level, leaders in Beijing provide a set of incentives to reward local governors who achieve strong economic growth and to sideline those who did not. Thus, there are clear incentives for local officials to inflate output and tax revenue. Additionally, the NBSC faces several technical difficulties and weaknesses in the Chinese statistical system: the data collected by the NBSC through sample surveys are incomplete, the direct-reporting system only covers firms with revenues greater than 20 million yuan and the revisions are frequent, large and not always clearly justified. The second aspect analysed is the evidence for internal inconsistency and data contamination. The differences between GDP measured by the VA method and GDP measured the AE method increased since 2005 and the sum of the provincial output exceeded the national output figure for the VA and the AE approaches. Moreover, the sectoral discrepancies are severe for the industrial sector and for gross capital formation. As the NBSC has not disclosed how it calculates the implicit GDP deflators for each sector, alternative GDP deflators are constructed using various price indices. Based on this analysis, the estimated official real growth rate could be between 8.4% and 10.6%. This implies that the NBSC has a wide range of implicit deflators that can direct the level of real GDP to meet the pre-announced target.

This paper makes several contributions to the current literature. First, it provides a comprehensive review of the current literature for researchers and policy-makers who are interested in assessing the reliability of Chinese official statistics. Second, eight different price indices are used to construct the implicit GDP deflator for three economic sectors and to provide a range of estimated annual real GDP growth rate between 1979 and 2015. Third, the internal inconsistency within China is investigated using quantitative methods.

Forecasting Chinese macroeconomic variables using large-scale factor models with mixed-frequency data and missing observations is explored in Chapter 3. Developments in information technology provide access to thousands of economic variables in real-time at a reasonable computational cost. This raises the prospect of a new frontier in macroeconomic forecasting, in which a large number of predictors are used to forecast a few key macroeconomic variables, including measures of inflation and real output. Although the literature has shown the factor models by Stock & Watson (2002*b*) are useful to improve forecasting accuracy for Western countries, macroeconomic forecasting in a data-rich environment of China is under-researched. To fill gap in the research, factor forecasts of the CPI, RPI, nominal investment, nominal consumption and railway cargo are constructed using 251 monthly variables and 34 quarterly variables and compared to the forecasts produced by univariate and multivariate models. The forecasting performance for both the out-of-sample period and the global financial crisis period are assessed using relative MSE and Giacomini & White (2006)'s conditional predictability tests. Statistical evidence indicates that mixed-frequency factor models, especially mixed-frequency FA-VAR models, generate superior forecasts to the univariate and multivariate models for price series, nominal investment and nominal consumption, except for the CPI inflation rate and nominal consumption at one month ahead. Therefore, the results of this study provide clear guidance and important implications for academics, practitioners and the public who are interested in macroeconomic forecasting in China.

Chapter 3 is likely one of the first papers detailing the use of a large dataset with mixed-frequency and missing observations to forecast the Chinese macroeconomy. Two approaches differentiate the chapter from the existing literature. The first considers the factor models based on 251 monthly variables and 34 quarterly variables with missing observation components, as well as factor models based on 39 preselected targeted variables. The number of predictors used in most of the previous studies for forecasting China's monthly macroeconomic variables is relatively small—usually 40 variables or fewer. The second approach involves running a sufficiently large validation set, ranging from 88 (for forecasts 12 months ahead) to 99 (for forecasts 1 month ahead); other studies generally have 5 to 60 observations for the validation set. The major contribution is the identification of statistical evidence that confirms the large dimensional factor models that are suitable for forecasting in Western countries are also suitable for China.

Knowing the current state of the economy in real-time is an ongoing, arguably increasingly important responsibility for policy-makers. However, there are few studies considering macroeconomic monitoring in China. Chapter 4 considers the evaluation of out-of-sample model-based nowcasts of China's GDP growth. The most common method of the past decade was the dynamic factor model by Giannone et al. (2008), which allows for the inclusion of a large number of predictors in nowcasting models and the incorporation of economic data releases in real-time. A more ambitious question is whether there are exploitable nonlinear and non-parametric structures in macroeconomic datasets that could be reveal by modern ML methods. To assess whether ML algorithms are useful for improving macroeconomic monitoring in China, the nowcasts obtained from eight ML algorithms are compared with the nowcasts generated by the AR model and the dynamic factor model. We find evidence that shrinkage methods perform well and the dynamic factor models perform poorly, based on the results of the model confidence set by Hansen et al. (2011). Conversely, the current literature states that the central banks in several Western countries regard the dynamic factor model as the most effective method for macroeconomic monitoring. Importantly, this paper contributes to the growing literature on using ML algorithms for macroeconomic forecasting and monitoring. Further, it may be one of the first papers to consider using ML algorithms to nowcast China's GDP, which guides practitioners when selecting the appropriate nowcasting models for China's economy and provides considerable scope for future research.

Finally, a few of the limitations and research challenges for future macroeconomic forecasting in China must be considered. First, as the official economic data may not be reliable, a database for macroeconomic research in China should be constructed with the goal of establishing a convenient starting point for empirical analysis that requires big data. Higgins & Zha (2015) provide time series data for China's macroeconomy that covers annual, quarterly and monthly frequencies. However, as the dataset is only updated semi-annually, more timely updated data are urgently needed for macroeconomic forecasting. In recognising the over-reporting behaviour of local officials, future research could examine the agency problems between China's central and local governments when affecting the economy. As local governments must consider the trade between tax transfer to the central government and local economic growth, using the value-added tax revenue of the local government to estimate the size of over-reporting could be another future research topic. Studying the use of alternative indicators to test the reliability of official economic data and trade data from developed countries (which are more reliable) could also be fruitful.

Second, the underlying data-generating process for China's GDP and the predictors often undergo changes that are difficult to track accurately through time. Therefore, it is difficult to determine whether a single "best" forecasting (or nowcasting) model could always dominate competing models. As such, a viable alternative is to use forecast combinations or model combinations. Forecast combinations across a set of candidate forecasting models could be useful to improve forecasting performance over a single model[1]. Similarly, Bayesian techniques are useful for managing model instability over time. For instance, large Bayesian VAR models that allow sample information to be combined with structurally relevant prior information could serve as an alternative to the frequentist domain. Compressed VAR methods that use Bayesian model averaging with time variation

[1]Several studies have shown that forecasts based on model combinations perform superiorly to single model forecasts in Western countries. For example, the complete subset regression by Elliott et al. (2013) that constructs forecasts as a weight average from all possible linear regression models can outperform the factor models. Frequentist model averaging criteria that combine forecasts from many factor-augmented regressions is also useful to improve forecasting performance that is based on single factor-augmented regression. Ensemble ML algorithms that combine the predictions from multiple ML algorithms can produce more accurate predictions than any individual model.

parameters outperformed the factor models in the US, which might also be the case in China.

Although this thesis focuses on point forecasts, there are several opportunities for the future study of density forecasts. Density forecasts provide a confidence interval—an interval that should contain the future value with a pre-specified probability. Moreover, density forecasts are not tied to a particular loss function. As such, they are particularly useful when the loss function for a monetary policy-maker is heterogeneous and asymmetric (i.e., it could involve not only the magnitude of the forecast error but also the level of the outcome itself). Over-predictions of economic growth in states with low growth can be costly because the central banks would fail to apply an appropriately aggressive monetary policy. Thus, research projects that consider asymmetric loss function for Chinese monetary policy-makers can be fruitful. Unlike point forecasts, density forecasts also provide a summary of risk forecasts, which is invaluable when the economic outlook is uncertain. Several central banks and international organisations are providing density forecasts to the public. For example, the Bank of England and IMF report "fan chart" forecasts for inflation, which use different shades of colour to illustrate the range of quantiles, beginning with the median forecast and fanning out to cover an increasing level of confidence interval. The Chinese version of "fan chart" forecasts can be constructed to account for risks in macroeconomic forecasting and nowcasting.

Bibliography

Aarons, G., Caratelli, D., Cocci, M., Giannone, D., Sbordone, A. & Tambalotti, A. (2016), 'Just released: introducing the FRBNY nowcast', *Liberty Street Economics* .

Ahmed, N. K., Atiya, A. F., Gayar, N. E. & El-Shishiny, H. (2010), 'An empirical comparison of machine learning models for time series forecasting', *Econometric Reviews* **29**(5-6), 594–621.

Alessandri, P. & Mumtaz, H. (2017), 'Financial conditions and density forecasts for US output and inflation', *Review of Economic Dynamics* **24**, 66–78.

Angelini, E., Banbura, M. & Rünstler, G. (2008), 'Estimating and forecasting the Euro area monthly national accounts from a dynamic factor model'.

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

Aruoba, S. B. (2008), 'Data revisions are not well behaved', *Journal of Money, Credit and Banking* **40**(2-3), 319–340.

Athey, S. (2018), The impact of machine learning on economics, *in* 'The economics of artificial intelligence: An agenda', University of Chicago Press.

Athey, S. & Imbens, G. W. (2019), 'Machine learning methods that economists should know about', *Annual Review of Economics* **11**, 685–725.

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

Bai, J. & Ng, S. (2006), 'Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions', *Econometrica* **74**(4), 1133–1150.

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

Bai, J. & Ng, S. (2008*b*), *Large dimensional factor analysis*, Now Publishers Inc.

Bai, J. & Ng, S. (2013), 'Principal components estimation and identification of static factors', *Journal of Econometrics* **176**(1), 18–29.

Bailliu, J., Han, X., Kruger, M., Liu, Y.-H. & Thanabalasingam, S. (2019), 'Can media and text analytics provide insights into labour market conditions in China?', *International Journal of Forecasting* **35**(3), 1118–1130.

Bańbura, M., Giannone, D., Modugno, M. & Reichlin, L. (2013), Nowcasting and the real-time data flow, *in* 'Handbook of Economic Forecasting', Vol. 2, Elsevier, pp. 195–237.

Bańbura, M., Giannone, D. & Reichlin, L. (2010), 'Large Bayesian vector auto regressions', *Journal of Applied Econometrics* **25**(1), 71–92.

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**(2), 333–346.

Banerjee, A., Marcellino, M. & Masten, I. (2005), 'Leading indicators for Euro-area inflation and GDP growth', *Oxford Bulletin of Economics and Statistics* **67**(s1), 785–813.

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

Barnett, W. A. & Tang, B. (2016), 'Chinese divisia monetary index and GDP nowcasting', *Open Economies Review* **27**(5), 825–849.

Bell, V., Co, L. W., Stone, S. & Wallis, G. (2014), 'Nowcasting UK GDP growth', *Bank of England Quarterly Bulletin* p. Q1.

Berge, T. J. (2015), 'Predicting recessions with leading indicators: model averaging and selection over the business cycle', *Journal of Forecasting* **34**(6), 455–471.

Bergmeir, C., Hyndman, R. J. & Koo, B. (2018), 'A note on the validity of cross-validation for evaluating autoregressive time series prediction', *Computational Statistics & Data Analysis* **120**, 70–83.

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

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

Boivin, J. & Ng, S. (2005), Understanding and comparing factor-based forecasts, Technical report, National Bureau of Economic Research.

Buchen, T. & Wohlrabe, K. (2011), 'Forecasting with many predictors: is boosting a viable alternative?', *Economics Letters* **113**(1), 16–18.

Bulman, D. J. (2016), *Incentivized development in China: leaders, governance, and growth in China's counties*, Cambridge University Press.

Cai, Y. (2000), 'Between state and peasant: local cadres and statistical reporting in rural China', *The China Quarterly* **163**, 783–805.

Camacho, M. & Perez-Quiros, G. (2010), 'Introducing the Euro-sting: short-term indicator of Euro area growth', *Journal of Applied Econometrics* **25**(4), 663–694.

Carrasco, M. & Rossi, B. (2016), 'In-sample inference and forecasting in misspecified factor models', *Journal of Business & Economic Statistics* **34**(3), 313–338.

Carriero, A., Clark, T. E. & Marcellino, M. (2015), 'Bayesian VARs: specification choices and forecast accuracy', *Journal of Applied Econometrics* **30**(1), 46–73.

Catania, L. & Bernardi, M. (2019), *MCS: model Confidence Set Procedure in R*, R Foundation for Statistical Computing.
URL: *https://cran.r-project.org/web/packages/MCS/index.html*

Chakraborty, C. & Joseph, A. (2017), 'Machine learning at central banks'.

Chen, K., Higgins, P., Waggoner, D. F. & Zha, T. (2016), China pro-growth monetary policy and its asymmetric transmission, Technical report, Working Paper.

Chen, T. & Kung, J.-S. (2016), 'Do land revenue windfalls create a political resource curse? Evidence from China', *Journal of Development Economics* **123**, 86–106.

Chow, G. C. (1986), 'Chinese statistics', *The American Statistician* **40**(3), 191–196.

Cristadoro, R., Forni, M., Reichlin, L. & Veronese, G. (2005), 'A core inflation indicator for the Euro area', *Journal of Money, Credit and Banking* pp. 539–560.

Daiane Marcolino de Mattos, Pedro Costa Ferreira, S. d. V. G. B. G. (2019), *Nowcasting: predicting economic variables using dynamic factor models in R*, R Foundation for Statistical Computing.
URL: *https://cran.r-project.org/web/packages/nowcasting/index.html*

Döpke, J., Fritsche, U. & Pierdzioch, C. (2017), 'Predicting recessions with boosted regression trees', *International Journal of Forecasting* **33**(4), 745–759.

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

Dungey, M., Fazzone, P., Le, A., Mackay, L., Morley, J., Miscamble, R., Ray, N. & Van Bridges, A. (2018), 'Nowcasting Australia's GDP'.

Elliott, G., Gargano, A. & Timmermann, A. (2013), 'Complete subset regressions', *Journal of Econometrics* **177**(2), 357–373.

Evans, M. D. (2005), Where are we now? Real-time estimates of the macroeconomy, Technical report, National Bureau of Economic Research.

Faust, J. & Wright, J. H. (2013), Forecasting inflation, *in* 'Handbook of Economic Forecasting', Vol. 2, Elsevier, pp. 2–56.

Fernald, J. G., Hsu, E. & Spiegel, M. M. (2015), 'Is China fudging its figures? Evidence from trading partner data'.

Fernald, J., Malkin, I., Spiegel, M. et al. (2013), 'On the reliability of Chinese output figures', *FRBSF Economic Letter* **8**, 1–5.

Fernald, J., Spiegel, M. M. & Swanson, E. T. (2014), 'Monetary policy effectiveness in China: evidence from a FAVAR model', *Journal of International Money and Finance* .

Findley, D. F. & Wei, C.-Z. (2002), 'Aic, overfitting principles, and the boundedness of moments of inverse matrices for vector autotregressions and related models', *Journal of Multivariate Analysis* **83**(2), 415–450.

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

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2003), 'Do financial variables help forecasting inflation and real activity in the Euro area?', *Journal of Monetary Economics* **50**(6), 1243–1255.

Francis, X. & Roberto, S. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**(3).

Friedman, J., Hastie, T. & Tibshirani, R. (2010), 'Regularization paths for generalized linear models via coordinate descent', *Journal of Statistical Software* **33**(1), 1–22.
URL: *http://www.jstatsoft.org/v33/i01/*

Geweke, J. (1977), 'The dynamic factor analysis of economic time series', *Latent Variables in Socio-economic Models* .

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

Giannone, D., Lenza, M. & Primiceri, G. E. (2015), 'Prior selection for vector autoregressions', *Review of Economics and Statistics* **97**(2), 436–451.

Giannone, D. & Matheson, T. (2007), 'A new core inflation indicator for New Zealand'.

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

Gilchrist, S. & Zakrajšek, E. (2012), 'Credit spreads and business cycle fluctuations', *American Economic Review* **102**(4), 1692–1720.

Gogas, P., Papadimitriou, T., Matthaiou, M. & Chrysanthidou, E. (2015), 'Yield curve and recession forecasting in a machine learning framework', *Computational Economics* **45**(4), 635–645.

Goulet, P. C., Leroux, M., Stevanovic, D., Surprenant, S. et al. (2019), How is machine learning useful for macroeconomic forecasting?, Technical report, CIRANO.

Graff, M., Matheson, T. et al. (2004), Composite leading indicators of inflation in New Zealand, Technical report, Citeseer.

Granger, C. W. J., Elliott, G. & Timmermann, A. (2006), *Handbook of Economic Forecasting*, Vol. 1, Elsevier.

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

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

Hansen, P. R. (2005), 'A test for superior predictive ability', *Journal of Business & Economic Statistics* **23**(4), 365–380.

Hansen, P. R., Lunde, A. & Nason, J. M. (2011), 'The model confidence set', *Econometrica* **79**(2), 453–497.

Harvey, A. C. (1990), *Forecasting, structural time series models and the Kalman filter*, Cambridge university press.

He, Q. & Fan, C. (2015), 'Forecasting inflation in China', *Emerging Markets Finance and Trade* **51**(4), 689–700.

He, Q., Leung, P.-H. & Chong, T. T.-L. (2013), 'Factor-augmented VAR analysis of the monetary policy in China', *China Economic Review* **25**, 88–104.

Hendry, D., Castle, J. & Kitov, O. (2016), 'Forecasting and nowcasting macroeconomic variables: a methodological overview', *Handbook on Rapid Estimates* .

Higgins, P. C. (2014), 'GDPNow: a model for GDP'Nowcasting''.

Higgins, P. C. & Zha, T. (2015), 'China's macroeconomic time series: methods and implications', *Unpublished Manuscript, Federal Reserve Bank of Atlanta* .

Higgins, P., Zha, T. & Zhong, W. (2016), 'Forecasting China's economic growth and inflation', *China Economic Review* **41**, 46–61.

Holz, C. A. (2003), '"Fast, clear and accurate": how reliable are Chinese output and economic growth statistics?', *The China Quarterly* **173**, 122–163.

Holz, C. A. (2004), 'China's statistical system in transition: challenges, data problems, and institutional innovations', *Review of Income and Wealth* **50**(3), 381–409.

Holz, C. A. (2008), 'China's 2004 economic census and 2006 benchmark revision of GDP statistics: more questions than answers?', *The China Quarterly* **193**, 150–163.

Holz, C. A. (2013), 'Chinese statistics: Output data', *Available at SSRN 2230070* .

Holz, C. A. (2014), 'The quality of China's GDP statistics', *China Economic Review* **30**, 309–338.

Hornik, K., Stinchcombe, M., White, H. et al. (1989), 'Multilayer feedforward networks are universal approximators.', *Neural Networks* **2**(5), 359–366.

Hsien, C.-T. & Song, Z. (2019), 'A forensic examination of China's national accounts', *Brookings Papers on Economic Activity* .

Jia, R., Kudamatsu, M. & Seim, D. (2015), 'Political selection in China: the complementary roles of connections and performance', *Journal of the European Economic Association* **13**(4), 631–668.

Jiang, Y., Guo, Y. & Zhang, Y. (2017), 'Forecasting China's GDP growth using dynamic factors and mixed-frequency data', *Economic Modelling* **66**, 132–138.

Jolliffe, I. T. (1993), 'Introduction to multiple time series analysis'.

Kamal, L. (2013), 'Forecasting inflation in China', *Journal of Applied Business Research* **29**(6), 1825.

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

Karatzoglou, A., Smola, A., Hornik, K. & Zeileis, A. (2004), 'kernlab – an S4 package for kernel methods in R', *Journal of Statistical Software* **11**(9), 1–20.
URL: *http://www.jstatsoft.org/v11/i09/*

Kerola, E. (2019), 'In search of fluctuations: another look at China's incredibly stable GDP growth rates', *Comparative Economic Studies* **61**(3), 359–380.

Koch-Weser, I. N. (2013), 'The reliability of China's economic data: an analysis of national output', *US-China Economic and Security Review Commission Staff Research Project* **4**.

Konishi, S. & Kitagawa, G. (2008), *Information criteria and statistical modeling*, Springer Science & Business Media.

Koop, G. M. (2013), 'Forecasting with medium and large Bayesian VARs', *Journal of Applied Econometrics* **28**(2), 177–203.

Kotchoni, R., Leroux, M. & Stevanovic, D. (2019), 'Macroeconomic forecast accuracy in a data-rich environment', *Journal of Applied Econometrics* .

Kuhn, M. (2008), 'Building predictive models in R using the caret package', *Journal of Statistical Software, Articles* **28**(5), 1–26.
URL: *https://www.jstatsoft.org/v028/i05*

Lahiri, K. & Monokroussos, G. (2013), 'Nowcasting US GDP: the role of ISM business surveys', *International Journal of Forecasting* **29**(4), 644–658.

Landry, P. F. (2008), *Decentralized Authoritarianism in China: the Communist Party's control of local elites in the post-Mao era*, Vol. 1, Cambridge University Press New York.

Landry, P. F., Lü, X. & Duan, H. (2017), 'Does performance matter? evaluating political selection along the Chinese administrative ladder', *Comparative Political Studies* p. 0010414017730078.

Li, H. & Zhou, L.-A. (2005), 'Political turnover and economic performance: the incentive role of personnel control in China', *Journal of Public Economics* **89**(9-10), 1743–1762.

Li, X., Shang, W., Wang, S. & Ma, J. (2015), 'A MIDAS modelling framework for Chinese inflation index forecast incorporating google search data', *Electronic Commerce Research and Applications* **14**(2), 112–125.

Liaw, A. & Wiener, M. (2002), 'Classification and regression by randomForest', *R News* **2**(3), 18–22.
URL: *https://CRAN.R-project.org/doc/Rnews/*

Lin, C.-S., Chiu, S.-H. & Lin, T.-Y. (2012), 'Empirical model decomposition–based least squares support vector regression for foreign exchange rate forecasting', *Economic Modelling* **29**(6), 2583–2590.

Lin, C.-Y. & Wang, C. (2013), 'Forecasting China's inflation in a data-rich environment', *Applied Economics* **45**(21), 3049–3057.

Liu, W. & Moench, E. (2016), ‘What predicts US recessions?’, *International Journal of Forecasting* **32**(4), 1138–1150.

Lütkepohl, H. (2014), ‘Structural vector autoregressive analysis in a data rich environment: a survey’.

Ma, B., Song, G., Zhang, L. & Sonnenfeld, D. A. (2014), ‘Explaining sectoral discrepancies between national and provincial statistics in China’, *China Economic Review* **30**, 353–369.

Maddison, A. (2007), ‘Chinese economic performance in the long run’, *Edition, Deve* .

Mao Takongmo, C. & Stevanovic, D. (2015), ‘Selection of the number of factors in presence of structural instability: a monte carlo study’, *L’Actualité économique* **91**(1-2), 177–233.

Marcellino, M., Stock, J. H. & Watson, M. W. (2003), ‘Macroeconomic forecasting in the Euro area: country specific versus area-wide information’, *European Economic Review* **47**(1), 1–18.

Marcellino, M., Stock, J. H. & Watson, M. W. (2006), ‘A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series’, *Journal of Econometrics* **135**(1-2), 499–526.

Matheson, T. (2011), *New indicators for tracking growth in real time*, number 11-43, International Monetary Fund.

Matheson, T. D. (2010), ‘An analysis of the informational content of New Zealand data releases: the importance of business opinion surveys’, *Economic Modelling* **27**(1), 304–314.

McCracken, M. W. & Ng, S. (2016), 'Fred-md: A monthly database for macroeconomic research', *Journal of Business & Economic Statistics* **34**(4), 574–589.

Mehrotra, A. & Sánchez-Fung, J. R. (2008), 'Forecasting inflation in China', *China Economic Journal* **1**(3), 317–322.

Mei, C. & Wang, X. (2017), 'Political incentives and local policy innovations in China', *Journal of Chinese Political Science* **22**(4), 519–547.

Meyer, D., Shih, V. C. & Lee, J. (2016), 'Factions of different stripes: gauging the recruitment logics of factions in the reform period', *Journal of East Asian Studies* **16**(1), 43–60.

Moser, G., Rumler, F. & Scharler, J. (2007), 'Forecasting Austrian inflation', *Economic Modelling* **24**(3), 470–480.

Movshuk, O. (2002), 'The reliability of China's growth figures: a survey of recent statistical controversies', *The Journal of Econometric Study of Northeast Asia* **4**(1), 31–45.

Mullainathan, S. & Spiess, J. (2017), 'Machine learning: an applied econometric approach', *Journal of Economic Perspectives* **31**(2), 87–106.

Naughton, B. J. & Yang, D. L. (2004), *Holding China together: diversity and national integration in the post-Deng era*, Cambridge University Press.

Pain, N., Lewis, C., Dang, T.-T., Jin, Y. & Richardson, P. (2014), 'OECD forecasts during and after the financial crisis'.

Plakandaras, V., Gupta, R., Gogas, P. & Papadimitriou, T. (2015), 'Forecasting the US real house price index', *Economic Modelling* **45**, 259–267.

Rawski, T. G. (1976), 'On the reliability of Chinese economic data', *The Journal of Development Studies* **12**(4), 438–441.

Rawski, T. G. (2001), 'What is happening to China's GDP statistics?', *China Economic Review* **12**(4), 347–354.

Reichlin, L. (2002), 'Factor models in large cross-sections of time series'.

Reichlin, L., De Mol, C., Gautier, E., Giannone, D., Mullainathan, S., van Dijk, H. & Wooldridge, J. (2017), 'Big data in economics: evolution or revolution?'.

Sargent, T. J., Sims, C. A. et al. (1977), 'Business cycle modeling without pretending to have too much a priori economic theory', *New Methods in Business Cycle Research* **1**, 145–168.

Schauberger, G. & Tutz, G. (2015), Regularization methods in economic forecasting, *in* 'Empirical Economic and Financial Research', Springer, pp. 61–80.

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

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**(3), 386–398.

Shih, V., Adolph, C. & Liu, M. (2012), 'Getting ahead in the communist party: explaining the advancement of central committee members in China', *American Political Science Review* **106**(1), 166–187.

Siliverstovs, B. & Kholodilin, K. A. (2012), 'Assessing the real-time informational content of macroeconomic data releases for now-/forecasting GDP: evidence for Switzerland', *Jahrbücher für Nationalökonomie und Statistik* **232**(4), 429–444.

Sinclair, T. M. (2019), 'Characteristics and implications of Chinese macroeconomic data revisions', *International Journal of Forecasting* **35**(3), 1108–1117.

Sokolov-Mladenović, S., Milovančević, M., Mladenović, I. & Alizamir, M. (2016), 'Economic growth forecasting by artificial neural network with extreme learning machine based on trade, import and export parameters', *Computers in Human Behavior* **65**, 43–45.

Stock, J. H. & Watson, M. W. (1989), 'New indexes of coincident and leading economic indicators', *NBER Macroeconomics Annual* **4**, 351–394.

Stock, J. H. & Watson, M. W. (1998), Diffusion indexes, Technical report, National bureau of economic research.

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

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

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

Stock, J. H. & Watson, M. W. (2006), 'Forecasting with many predictors', *Handbook of Economic Forecasting* **1**, 515–554.

Stock, J. H. & Watson, M. W. (2008), Phillips curve inflation forecasts, Technical report, National Bureau of Economic Research.

Stock, J. H. & Watson, M. W. (2011), 'Dynamic factor models', *Oxford Handbook of Economic Forecasting. Oxford University Press, USA* pp. 35–59.

Stock, J. H. & Watson, M. W. (2012), Disentangling the channels of the 2007-2009 recession, Technical report, National Bureau of Economic Research.

Stock, J. H. & Watson, M. W. (2016), Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics, *in* 'Handbook of Macroeconomics', Vol. 2, Elsevier, pp. 415–525.

Stock, J. H. & Watson, M. W. (2017), 'Twenty years of time series econometrics in ten pictures', *Journal of Economic Perspectives* **31**(2), 59–86.

Stock, J. H., Watson, M. W. et al. (2003), *Introduction to econometrics*, Vol. 104, Addison Wesley Boston.

Therneau, T. & Atkinson, B. (2019), *rpart: recursive partitioning and regression trees.* R package version 4.1-15.
URL: *https://CRAN.R-project.org/package=rpart*

Van Nieuwenhuyze, C. (2005), 'A generalized dynamic factor model for the Belgian economy: identification of the business cycle and GDP growth forecasts', *Journal of Business Cycle Measurement and Analysis* **2005**(2), 213–247.

Venables, W. N. & Ripley, B. D. (2002), *Modern applied statistics with S*, fourth edn, Springer, New York. ISBN 0-387-95457-0.
URL: *http://www.stats.ox.ac.uk/pub/MASS4*

Wang, J. (1998), 'Geng zhun, geng kuai, geng quan: tiaozheng tongji qing yu biao (accurate, fast and wider statistics', *China Daily* .

Wang, X. & Meng, L. (2001), 'A reevaluation of China's economic growth', *China Economic Review* **12**(4), 338–346.

Wang, Y. & Chandler, W. (2011), 'Understanding energy intensity data in China', *Policy Outlook, Carnegie Endowment for International Peace* .

Xi, T., Yao, Y. & Zhang, M. (2018), 'Capability and opportunism: evidence from city officials in China', *Journal of Comparative Economics* **46**(4), 1046–1061.

Xianchun, X. (2002), 'Study on some problems in estimating China's GDP', *Review of Income and Wealth* **48**(2), 205–215.

Xianchun, X. (2006), 'On the changes in the GDP estimation in the economic census year [j]', *Economic Research Journal* **2**.

Xiaolu, W. & Lian, M. (2001), 'A reevaluation of China's economic growth', *China Economic Review* **12**(4), 338–346.

Xiong, W. (2018), The Mandarin model of growth, Technical report, National Bureau of Economic Research.

Xiong, W. (2019), 'Discussion of "a forensic examination of China's national accounts" by wei chen, xilu chen, chang-tai hsieh, and zheng song', *Brookings Papers on Economic Activity* .

Xu, C. (2011), 'The fundamental institutions of China's reforms and development', *Journal of Economic Literature* **49**(4), 1076–1151.

Xu, X. (2009), 'The establishment, reform, and development of China's system of national accounts', *Review of Income and Wealth* **55**, 442–465.

Xu, X., Ye, Y. & Blades, D. (2000), 'National accounts for China: sources and methods'.

Yiu, M. S. & Chow, K. K. (2010), 'Nowcasting Chinese GDP: information content of economic and financial data', *China Economic Journal* **3**(3), 223–240.

Young, A. (2003), 'Gold into base metals: productivity growth in the People's Republic of China during the reform period', *Journal of Political Economy* **111**(6), 1220–1261.

Zeileis, A. & Grothendieck, G. (2005), 'zoo: S3 infrastructure for regular and irregular time series', *Journal of Statistical Software* **14**(6), 1–27.

Zeng, J. (2017), 'Forecasting aggregates with disaggregate variables: does boosting help to select the most relevant predictors?', *Journal of Forecasting* **36**(1), 74–90.

Zhang, Y., Cindy, L. Y., Li, H. & Hong, Y. (2018), 'Nowcasting China's GDP using a Bayesian approach', *Journal of Management Science and Engineering* **3**(4), 232–258.

Zhou, R., Wang, X. & Tong, G. (2013), 'Forecasting macroeconomy based on the term structure of credit spreads: evidence from China', *Applied Economics Letters* **20**(15), 1363–1367.

Appendices

Appendix A

History of Statistical Reform

Introduced in 1952, the Soviet Material Product System (SMPS) – an adopted method prevalent throughout the Soviet Union – was the primary statistical system used in China before the reform and opening up of the economy in the late 1970s. Many economists generally agreed that the SMPS was effective for China as a planned economy (Holz 2004, 2014, 2003, Rawski 1976, Xiaolu & Lian 2001, Chow 1986). However, as China implemented the reform and opening up, the SMPS left several weaknesses for statistical work under a market-oriented economy:

- The system was designed to measure physical inputs and outputs of a planned economy. Demand-side indicators for investment and consumption were not properly measured.
- The statistics for non-state sectors were either discounted or inaccurately measured. For example, the SMPS system created one category called non-material service that contains several types of services, including retail sales, the housing market, transport service and welfare benefits. However, transactions in rural regions and defence industry section were not accounted for.
- The data collection process was highly decentralized. State government relied heavily on local statistical bureaus to collect and report data up to the administrative hierarchy. At the central government level, each ministry collected and published their own data.

- The system also relied on direct reporting of output by state-owned and large enterprises. Output data from private and small size enterprises were low in quality and often came with a considerable amount of guesswork. Surveys and census, which are sophisticated forms of data gathering process in private sectors, were hardly used. The fast-growing economic activity outside the formal sector was not sufficiently monitored.

- There was a sufficiently large room for data manipulation and falsification. Local cadres and government officials were more preoccupied with pleasing their superiors than informing the general public while they had a fear of punishment for under-reporting grain output from the Maoist period. With a lack of independent surveys and census, such manipulation and falsification were hard to diagnose, especially in the early years.

To overcome these weaknesses, the NBSC and the State Council have implemented a series of reforms jointly since the 1980s. The NBSC has moved from a "complicated and comprehensive" reporting system to a two-streams data compilation system. In the first stream, the direct reporting system adopted a minimum threshold in 1998 to ensure that only state-owned firms and private firms with a sufficiently large amount of revenue would be included. The minimum thresholds were further increased in 2007 and 2011 accordingly. In the second stream, the NBSC has put more weight on census and sample survey to gather data for the rest of economy[1] (Koch-Weser 2013, Holz 2014). In order to perform economic census and sample survey effectively and correctly, the NBSC sent teams to conduct several types of surveys nationwide. These include weekly price data collection at markets, annual household surveys including income and expenditure in urban and rural areas, and revenues of those from non-state-owned and small size

[1] Before the 1990s, China had minimal experience using census and surveys

firms that are not included in the direct reporting system. Generally speaking, sample survey and census allowed the NBSC gathering and collecting data at broader internals. Examples of such survey and census include the census of the industrial sector in 1985 and 1995, census of the service sector in 1993, census of agriculture sector in 1996, national economic censuses in 2004, 2009, and 2014.

There were several data-manipulation related problems found in the early 1980s. This lead to a series of reforms in the National Statistics Law, aiming to reduce the level of data manipulation and falsification or even eliminate them. The first National Statistics Law was introduced in 1983 with the purpose of reducing the effect of data manipulation. In 1995, the NBSC found several problems in the rural industrial sector during the first industrial census; as a result, gross industrial output was revised downwards. Nationwide inspection of data falsification and manipulation were conducted by the NBSC along with the Ministry of Supervision and the Bureau of Legislative Affairs of the State Council in 1987, 1989, 1994 and 2001. 60,000 cases of violations of the Statistics Law were found in 2001 inspection, with almost 60% of cases were misreporting output figures by enterprises (Wang & Meng 2001, Wang 1998). Other types of violations include refusal of reporting and late reporting output data, and misbehaviour of statisticians and government departments were hardly seen with only very few officials were punished by their superiors. Even though inspection teams were sent out to local bureaus of statistics to investigate violations and unprofessional manners, the information about the result of internal inspection of the National Statistics System is unavailable to the general public (Cai 2000). By 2001, Zhongguo Tongji (Chinese Statistics), a journal managed by the NBSC and issued by Peking University Press, ceased a series of sensational reports on data falsification. These reports provided concrete evidence on data manipulation and falsification as well as reasons why data falsification occurs.

The State Council and the NBSC have been gradually improving the measurement and presentation of macroeconomic data. In 1995, the NBSC fully adopted the International System of National Accounting, which is a better method for the market-oriented economy and has been used by many Western economics (Holz 2004). Also the NBSC has codified the method of calculating GDP and presenting it to the general public. The NBSC has also developed a better technique to impute output statistics based on limited sample data (Xu 2009). Overall, the reform era has improved Chinese statistical framework and the quality of data in two ways: (1) the NBSC was able to better collect and gather data outside of the formal sectors through a series of national surveys and censuses, and (2) transiting from Soviet Material Product System to the International System of National Accounting, which allows the NBSC to capture activities outside of direct reporting system effectively. Table A.0.1 summaries the timeline of the statistical reform in China.

Table A.0.1: Brief Summary of Chinese Statistical Reforms

Year	Event
1983	National Statistics Law was introduced in order to reduce the effects of data manipulation.
1984	Offices were set up at the provincial and municipal level to collect data for CPI.
1985	The NBSC began to compel the tertiary sector and aggregate production statistics.
1989	The first time experimented with expenditure approach GDP statistics.
1991	First tertiary sector census (lasts from 1991 to 1995).
1992	Began to calculate GDP independently rather than obtain it from Net Material Product System.
1993	The NBSC ceased Net Material Product System and officially switched to System of National Accounts.
1995	Problems of rural industrial data found in the first industrial sector census. Downward revision for gross output value as a result.
1996	First revision of National Statistics Law reduces the role of the industrial enterprise reporting system and favour census and sample surveys.
1997	The NBSC publish national accounts for every province for the period 1952 to 1995.
	Issue of data manipulation at the local government level was investigated by State Council.
1998	Only large industrial enterprises are included in the direct reporting system. Definition of state-owned enterprises has been reformed.
1999	The NBSC increase monitoring of provincial-level statistical work.
2004	First national economic census with national GDP being revised up by 16 per cent.
2007	Adjustment to direct reporting system to include fewer firms.
2009	Second national economic census with GDP being revised up by 4 per cent.
	Second revision of statistical law to increase punishments for manipulation of statistics.
2011	Direct reporting system was adjusted again to include fewer firms.
	Seasonally adjusted quarter-by-quarter growth rate was introduced.
2014	Third national economic census concluded that GDP in 2013 had been revised up by 3.4%.

Sources: National Bureau of Statistics of China Yearbooks and Koch-Weser (2013).

Appendix B

Data Description

This appendix includes the data description used throughout this thesis. In Table B.0.1, we present the full-panel dataset that we use to construct h-steps-ahead forecast in Chapter 3. The format is as follows: series number, series name, transformation technique used, frequency, and the number of missing observations. Next, we report the series name, frequency and transformation technique for our preselected targeted predictors in Table B.0.2, with selection criteria based on hard and soft thresholding and judgements. The time series used to compute the machine learning and the dynamic factor model nowcasts discussed in Chapter 4 are shown in Table B.0.3. The format is as follows: series number, series name, transformation technique used, and frequency. We downloaded these data sets from CEIC Data Manager.

Table B.0.1: Data Description: Full-panel Dataset

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
1	Govt Expenditure	Percentage Change	Monthly	0
2	Govt Revenue	Percentage Change	Monthly	0
3	Industrial Production: Crude Oil	Percentage Change	Monthly	0
4	Industrial Production: Natural Gas	Percentage Change	Monthly	0
5	Industrial Production: Salt	Percentage Change	Monthly	0
6	Industrial Production: Cement	Percentage Change	Monthly	0
7	Industrial Production: Crude Steel	Percentage Change	Monthly	0
8	Bond Depository: Bond in RMB	Percentage Change	Monthly	0
9	Bond Depository: Govt Bond	Percentage Change	Monthly	0
10	Bond Depository: Financial Bond	Percentage Change	Monthly	0
11	Fixed Asset Investment	Percentage Change	Monthly	0
12	FX Rate: PBOC: Month End: RMB to USD	Percentage Change	Monthly	0
13	Effective Exchange Rate Index: BIS: Real	Percentage Change	Monthly	0
14	Effective Exchange Rate Index: BIS: Nominal	Percentage Change	Monthly	0
15	Money Supply M0: sa	Percentage Change	Monthly	0
16	Money Supply M1: sa	Percentage Change	Monthly	0
17	Money Supply M2: sa	Percentage Change	Monthly	0
18	Loan: Short Term	Percentage Change	Monthly	0
19	Loan: Medium & Long Term	Percentage Change	Monthly	0
20	Loan: sa	Percentage Change	Monthly	0
21	Deposit: sa	Percentage Change	Monthly	0
22	Deposit: Saving: sa	Percentage Change	Monthly	0

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
23	Property Price: YTD Avg: Overall	Percentage Change	Monthly	0
24	Import Price: Iron Ore & Concentrate	Percentage Change	Monthly	0
25	Index: Shanghai Stock Exchange: Composite	Percentage Change	Monthly	0
26	Index: Shanghai Stock Exchange: A Share	Percentage Change	Monthly	0
27	Index: Shanghai Stock Exchange: B Share	Percentage Change	Monthly	0
28	Turnover: Value: Shanghai SE: Total	Percentage Change	Monthly	0
29	Index: Shenzhen Stock Exchange: Composite	Percentage Change	Monthly	9
30	Index: Shenzhen Stock Exchange: A Share	Percentage Change	Monthly	9
31	Index: Shenzhen Stock Exchange: B Share	Percentage Change	Monthly	9
32	Turnover: Value: Shenzhen SE: Total	Percentage Change	Monthly	9
33	Market Capitalization: Shenzhen SE: Total	Percentage Change	Monthly	9
34	Market Capitalization: Shanghai SE: Stock	Percentage Change	Monthly	0
35	Market Capitalization: Shanghai SE: A Share	Percentage Change	Monthly	0
36	Market Capitalization: Shanghai SE: B Share	Percentage Change	Monthly	0
37	Market Capitalization: Shenzhen SE: Stock	Percentage Change	Monthly	0
38	Market Capitalization: Shenzhen SE: A Share	Percentage Change	Monthly	0
39	Market Capitalization: Shenzhen SE: B Share	Percentage Change	Monthly	0
40	Commodity Retail Sales: Above Designated Size Enterprise	Percentage Change	Monthly	0
41	Commodity Retail Sales: Food, Beverage, Tobacco & Liquor (FB)	Percentage Change	Monthly	0
42	Commodity Retail Sales: FB: Beverage	Percentage Change	Monthly	0
43	Commodity Retail Sales: FB: Tobacco & Liquor	Percentage Change	Monthly	0
44	Commodity Retail Sales: Clothing, Shoes, Hats & Textile (CT)	Percentage Change	Monthly	0
45	Commodity Retail Sales: Cosmetics	Percentage Change	Monthly	0
46	Commodity Retail Sales: Gold, Silver and Jewelry	Percentage Change	Monthly	0

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
47	Commodity Retail Sales: Daily Use Goods	Percentage Change	Monthly	0
48	Commodity Retail Sales: Other	Percentage Change	Monthly	0
49	Commodity Retail Sales: Automobile	Percentage Change	Monthly	0
50	Commodity Retail Sales: Petroleum & Related Product	Percentage Change	Monthly	0
51	Commodity Retail Sales: Chinese & Western Medicine (CM)	Percentage Change	Monthly	0
52	Commodity Retail Sales: Household Electric & Video Appliance	Percentage Change	Monthly	0
53	Commodity Retail Sales: Cultural & Office Goods	Percentage Change	Monthly	0
54	Commodity Retail Sales: Furniture	Percentage Change	Monthly	0
55	Commodity Retail Sales: Communication Appliance	Percentage Change	Monthly	0
56	Energy Production: Crude Oil	Percentage Change	Monthly	0
57	Energy Production: Natural Gas	Percentage Change	Monthly	0
58	Energy Production: Refined/Processed Crude Oil	Percentage Change	Monthly	0
59	Energy Production: Gasoline	Percentage Change	Monthly	0
60	Energy Production: Diesel Fuel	Percentage Change	Monthly	0
61	Energy Production: Kerosene	Percentage Change	Monthly	0
62	Energy Production: Fuel Oil	Percentage Change	Monthly	0
63	Energy Production: Liquefied Petroleum Gas (LPG)	Percentage Change	Monthly	0
64	Energy Production: Petroleum Pitch	Percentage Change	Monthly	0
65	Energy Production: Coke	Percentage Change	Monthly	0
66	Energy Production: Coal Gas	Percentage Change	Monthly	0
67	Energy Production: Electricity	Percentage Change	Monthly	0
68	Energy Production: Electricity: Thermal	Percentage Change	Monthly	0
69	Energy Production: Electricity: Hydro	Percentage Change	Monthly	0
70	Energy Production: Electricity: Nuclear	Percentage Change	Monthly	0

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
71	Consumer Confidence Index	Percentage Change	Monthly	0
72	Consumer Satisfactory Index	Percentage Change	Monthly	0
73	Consumer Expectation Index	Percentage Change	Monthly	0
74	Import CIF	Percentage Change	Monthly	0
75	Export FOB	Percentage Change	Monthly	0
76	Export: Live Pig (excl for Breeding)	Percentage Change	Monthly	0
77	Export: Live Poultry	Percentage Change	Monthly	9
78	Export: Aquatic Product	Percentage Change	Monthly	9
79	Export: Grain Food: Cereal & Cereal Flour	Percentage Change	Monthly	12
80	Export: Fresh Egg	Percentage Change	Monthly	12
81	Export: Vegetable	Percentage Change	Monthly	12
82	Export: Edible Oil Seed	Percentage Change	Monthly	0
83	Export: Edible Vegetable Oil	Percentage Change	Monthly	0
84	Import: Grain Food: Cereal & Cereal Flour: Wheat	Percentage Change	Monthly	0
85	Import: Edible Vegetable Oil	Percentage Change	Monthly	0
86	Import: Copper Ore & Concentrate	Percentage Change	Monthly	0
87	Import: Aluminium Oxide	Percentage Change	Monthly	0
88	Import: Crude Petroleum Oil	Percentage Change	Monthly	0
89	Import: Refined Petroleum Product	Percentage Change	Monthly	3
90	Import Price: Crude Oil	Percentage Change	Monthly	3
91	Export Price: Crude Oil	Percentage Change	Monthly	3
92	Transaction Price: Diesel Oil, No 0: Beijing	Percentage Change	Monthly	0
93	Transaction Price: Diesel Oil, No 0: Tianjin	Percentage Change	Monthly	0
94	Transaction Price: Diesel Oil, No 0: Shijiazhuang	Percentage Change	Monthly	3

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
95	Transaction Price: Diesel Oil, No 0: Hohhot	Percentage Change	Monthly	3
96	Transaction Price: Diesel Oil, No 0: Shenyang	Percentage Change	Monthly	3
97	Transaction Price: Diesel Oil, No 0: Dalian	Percentage Change	Monthly	0
98	Transaction Price: Diesel Oil, No 0: Shanghai	Percentage Change	Monthly	0
99	Transaction Price: Diesel Oil, No 0: Nanjing	Percentage Change	Monthly	0
100	Transaction Price: Diesel Oil, No 0: Hangzhou	Percentage Change	Monthly	0
101	Transaction Price: Diesel Oil, No 0: Ningbo	Percentage Change	Monthly	0
102	Transaction Price: Diesel Oil, No 0: Hefei	Percentage Change	Monthly	0
103	Transaction Price: Diesel Oil, No 0: Fuzhou	Percentage Change	Monthly	0
104	Transaction Price: Diesel Oil, No 0: Xiamen	Percentage Change	Monthly	0
105	Transaction Price: Diesel Oil, No 0: Nanchang	Percentage Change	Monthly	0
106	Transaction Price: Diesel Oil, No 0: Jinan	Percentage Change	Monthly	0
107	Transaction Price: Diesel Oil, No 0: Qingdao	Percentage Change	Monthly	0
108	Transaction Price: Diesel Oil, No 0: Wuhan	Percentage Change	Monthly	0
109	Transaction Price: Diesel Oil, No 0: Changsha	Percentage Change	Monthly	0
110	Transaction Price: Diesel Oil, No 0: Guangzhou	Percentage Change	Monthly	0
111	Transaction Price: Diesel Oil, No 0: Nanning	Percentage Change	Monthly	0
112	Transaction Price: Diesel Oil, No 0: Haikou	Percentage Change	Monthly	0
113	Transaction Price: Diesel Oil, No 0: Chengdu	Percentage Change	Monthly	0
114	Transaction Price: Diesel Oil, No 0: Guiyang	Percentage Change	Monthly	21
115	Transaction Price: Diesel Oil, No 0: Kunming	Percentage Change	Monthly	21
116	Transaction Price: Diesel Oil, No 0: Xian	Percentage Change	Monthly	21
117	Transaction Price: Diesel Oil, No 0: Lanzhou	Percentage Change	Monthly	21
118	Transaction Price: Diesel Oil, No 0: Yinchuan	Percentage Change	Monthly	21

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
119	Transaction Price: Diesel Oil, No 0: Urumqi	Percentage Change	Monthly	21
120	Transaction Price: Diesel Oil, No 0: Beijing	Percentage Change	Monthly	21
121	Transaction Price: Diesel Oil, No 0: Tianjin	Percentage Change	Monthly	21
122	Transaction Price: Diesel Oil, No 0: Shijiazhuang	Percentage Change	Monthly	21
123	Transaction Price: Diesel Oil, No 0: Taiyuan	Percentage Change	Monthly	21
124	Transaction Price: Diesel Oil, No 0: Hohhot	Percentage Change	Monthly	21
125	Transaction Price: Diesel Oil, No 0: Shenyang	Percentage Change	Monthly	21
126	Transaction Price: Diesel Oil, No 0: Dalian	Percentage Change	Monthly	21
127	Transaction Price: Diesel Oil, No 0: Shanghai	Percentage Change	Monthly	21
128	Transaction Price: Diesel Oil, No 0: Nanjing	Percentage Change	Monthly	21
129	Transaction Price: Diesel Oil, No 0: Hangzhou	Percentage Change	Monthly	21
130	Transaction Price: Diesel Oil, No 0: Ningbo	Percentage Change	Monthly	9
131	Transaction Price: Diesel Oil, No 0: Hefei	Percentage Change	Monthly	9
132	Transaction Price: Diesel Oil, No 0: Fuzhou	Percentage Change	Monthly	9
133	Transaction Price: Diesel Oil, No 0: Xiamen	Percentage Change	Monthly	9
134	Transaction Price: Diesel Oil, No 0: Nanchang	Percentage Change	Monthly	9
135	Transaction Price: Diesel Oil, No 0: Jinan	Percentage Change	Monthly	9
136	Transaction Price: Diesel Oil, No 0: Qingdao	Percentage Change	Monthly	9
137	Transaction Price: Diesel Oil, No 0: Wuhan	Percentage Change	Monthly	9
138	Transaction Price: Diesel Oil, No 0: Changsha	Percentage Change	Monthly	9
139	Transaction Price: Diesel Oil, No 0: Guangzhou	Percentage Change	Monthly	9
140	Transaction Price: Diesel Oil, No 0: Nanning	Percentage Change	Monthly	9
141	Transaction Price: Diesel Oil, No 0: Haikou	Percentage Change	Monthly	9
142	Transaction Price: Diesel Oil, No 0: Chengdu	Percentage Change	Monthly	9

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
143	Transaction Price: Diesel Oil, No 0: Guiyang	Percentage Change	Monthly	9
144	Transaction Price: Diesel Oil, No 0: Kunming	Percentage Change	Monthly	9
145	Transaction Price: Diesel Oil, No 0: Xian	Percentage Change	Monthly	0
146	Transaction Price: Diesel Oil, No 0: Lanzhou	Percentage Change	Monthly	0
147	Transaction Price: Diesel Oil, No 0: Yinchuan	Percentage Change	Monthly	0
148	Construction Material Production: Cement	Percentage Change	Monthly	0
149	Construction Material Production: Plated Glass	Percentage Change	Monthly	0
150	Construction Material Production: Aluminium Product	Percentage Change	Monthly	0
151	Transport: Passenger Traffic	Percentage Change	Monthly	6
152	Transport: Passenger Turnover	Percentage Change	Monthly	0
153	Transport: Freight Traffic	Percentage Change	Monthly	0
154	Transport: Freight Turnover	Percentage Change	Monthly	0
155	Railway: Passenger Traffic	Percentage Change	Monthly	0
156	Railway: Passenger Turnover	Percentage Change	Monthly	0
157	Railway: Freight Traffic	Percentage Change	Monthly	0
158	Railway: Freight Turnover	Percentage Change	Monthly	0
159	Highway: Passenger Traffic	Percentage Change	Monthly	0
160	Highway: Passenger Turnover	Percentage Change	Monthly	0
161	Highway: Freight Traffic	Percentage Change	Monthly	0
162	Highway: Freight Turnover	Percentage Change	Monthly	0
163	Waterway: Passenger Traffic	Percentage Change	Monthly	0
164	Waterway: Passenger Turnover	Percentage Change	Monthly	0
165	Waterway: Freight Traffic	Percentage Change	Monthly	20
166	Waterway: Freight Turnover	Percentage Change	Monthly	20

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
167	Air: Passenger Traffic	Percentage Change	Monthly	16
168	Air: Passenger Turnover	Percentage Change	Monthly	9
169	Air: Freight Traffic	Percentage Change	Monthly	10
170	Air: Freight Turnover	Percentage Change	Monthly	8
171	Postal: No of Remittance	Percentage Change	Monthly	17
172	Postal: No of Express	Percentage Change	Monthly	14
173	Postal: No of Package	Percentage Change	Monthly	6
174	Postal: No of Letter	Percentage Change	Monthly	8
175	Postal: No of Subscribed Magazine	Percentage Change	Monthly	6
176	Postal: No of Subscribed Newspaper	Percentage Change	Monthly	13
177	Postal: Business Volume: Express	Percentage Change	Monthly	8
178	APPI: Rural Market Fair: Long Grained Nonglutinous Rice: Medium	Percentage Change	Monthly	4
179	APPI: Rural Market Fair: Round Grained Rice: Medium	Percentage Change	Monthly	12
180	APPI: Rural Market Fair: Wheat: Medium	Percentage Change	Monthly	8
181	APPI: Rural Market Fair: Maize: Medium	Percentage Change	Monthly	6
182	APPI: Rural Market Fair: Soybean: Medium	Percentage Change	Monthly	8
183	Consumer Price Index: MoM	First Difference	Monthly	8
184	CPI: MoM: Food, Tobacco & Liquor: Food	First Difference	Monthly	15
185	CPI: MoM: Clothing	First Difference	Monthly	6
186	CPI: MoM: Transportation and Communication (TC)	First Difference	Monthly	14
187	CPI: MoM: Residence	First Difference	Monthly	4
188	CPI: Food, Tobacco & Liquor: Food: Vegetable: Fresh Vegetable	First Difference	Monthly	4
189	CPI: Food, Tobacco & Liquor: Food: Aquatic Product	First Difference	Monthly	6
190	CPI: Food, Tobacco & Liquor: Food: Egg	First Difference	Monthly	6

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
191	CPI: Food, Tobacco & Liquor: Food: Meat of Livestock	First Difference	Monthly	11
192	CPI: Food, Tobacco & Liquor: Food: Grain	First Difference	Monthly	11
193	CPI: Food, Tobacco & Liquor: Food	First Difference	Monthly	32
194	Retail Price Index	First Difference	Monthly	20
195	Retail Price Index: Food	First Difference	Monthly	16
196	Retail Price Index: Beverage, Tobacco & Liquors	First Difference	Monthly	9
197	Retail Price Index: Cosmetics	First Difference	Monthly	8
198	Retail Price Index: Daily Sundry Articles	First Difference	Monthly	10
199	Retail Price Index: Fuels	First Difference	Monthly	8
200	Retail Price Index: Building, Hardware & Electric Materials	First Difference	Monthly	17
201	Corporate Goods Price Index: Overall	First Difference	Monthly	14
202	Producer Price Index(PPI)	First Difference	Monthly	6
203	Producer Price Index: Producer Goods	First Difference	Monthly	8
204	Producer Price Index: Producer Goods: Mining and Quarrying	First Difference	Monthly	6
205	Producer Price Index: Producer Goods: Raw Material	First Difference	Monthly	13
206	Producer Price Index: Producer Goods: Manufacturing	First Difference	Monthly	8
207	Producer Price Index: Consumer Goods	First Difference	Monthly	4
208	Producer Price Index: Consumer Goods: Food	First Difference	Monthly	12
209	Producer Price Index: Consumer Goods: Clothing	First Difference	Monthly	8
210	Producer Price Index: Consumer Goods: Daily Sundry Article	First Difference	Monthly	6
211	Producer Price Index: Consumer Goods: Durable	First Difference	Monthly	8
212	Purchasing Price Index	First Difference	Monthly	8
213	Purchasing Price Index: Fuel and Power	First Difference	Monthly	15
214	Purchasing Price Index: Ferrous Metal Material	First Difference	Monthly	6

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
215	Purchasing Price Index: Non Ferrous Metal and Electric Wire	First Difference	Monthly	14
216	Purchasing Price Index: Chemical Material	First Difference	Monthly	4
217	Purchasing Price Index: Timber and Paper Pulp	First Difference	Monthly	4
218	Purchasing Price Index: Building Material	First Difference	Monthly	6
219	Purchasing Price Index: Industrial Raw Material and Semi Finished Product	First Difference	Monthly	6
220	Purchasing Price Index: Farm and Sideline Product	First Difference	Monthly	11
221	Purchasing Price Index: Textile Material	First Difference	Monthly	11
222	Retail Price Index: Textiles	First Difference	Monthly	12
223	Retail Price Index: Clothings , Shoes and Hats	First Difference	Monthly	12
224	Required Reserve Ratio	None	Monthly	12
225	Rediscount Rate	None	Monthly	2
226	Central Bank Benchmark Interest Rate: Loan to FI: 1 Year	None	Monthly	0
227	Central Bank Benchmark Interest Rate: Loan to FI: 6 Month or Less	None	Monthly	0
228	Central Bank Benchmark Interest Rate: Loan to FI: 3 Month or Less	None	Monthly	0
229	Interbank Offered Rate: Weighted Avg: 3 Month	None	Monthly	0
230	Interbank Offered Rate: Weighted Avg: 1 Month	None	Monthly	0
231	Interbank Offered Rate: Weighted Avg: 21 Day	None	Monthly	0
232	Interbank Offered Rate: Weighted Avg: 2 Month	None	Monthly	0
233	Real Estate Inv: ytd	None	Monthly	0
234	Real Estate Inv: ytd: Residential Building	None	Monthly	0
235	Real Estate Inv: ytd: Office Building	None	Monthly	0
236	Real Estate Inv: ytd: Commercial Building	None	Monthly	0
237	Real Estate Inv: ytd: Other Building	None	Monthly	0
238	Central Bank Benchmark Interest Rate: Loan to FI: Less Than 20 days	None	Monthly	0

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
239	Household Savings Deposits Rate: Demand	None	Monthly	2
240	Household Savings Deposits Rate: Time: 3 Month	None	Monthly	2
241	Household Savings Deposits Rate: Time: 2 Year	None	Monthly	0
242	Household Savings Deposits Rate: Time: 3 Year	None	Monthly	0
243	Enterprise Deposits Rate: Demand	None	Monthly	0
244	Enterprise Deposits Rate: Time: 3 Month	None	Monthly	0
245	Enterprise Deposits Rate: Time: 6 Month	None	Monthly	0
246	Enterprise Deposits Rate: Time: 1 Year	None	Monthly	0
247	Nominal Lending Rate: Within 1 Year (Including 1 Year)	None	Monthly	0
248	Nominal Lending Rate: 1-5 Year (Including 5 Year)	None	Monthly	0
249	Nominal Lending Rate: Over 5 Year	None	Monthly	0
250	Household Savings Deposits Rate: Time: 6 Month	None	Monthly	0
251	Household Savings Deposits Rate: Time: 1 Year	None	Monthly	0
252	GDP	Percentage Change	Quarterly	132
253	GDP: Primary Industry	Percentage Change	Quarterly	132
254	GDP: Secondary Industry(SI)	Percentage Change	Quarterly	132
255	GDP: SI: Industry	Percentage Change	Quarterly	132
256	GDP: SI: Construction	Percentage Change	Quarterly	132
257	GDP: TI: Transport, Storage and Post	Percentage Change	Quarterly	132
258	GDP: TI: Wholesale and Retail Trade	Percentage Change	Quarterly	132
259	GDP: Tertiary Industry(TI)	Percentage Change	Quarterly	132
260	GDP: ow: Agriculture, Forestry, Animal Husbandry and Fishery (Incl. Services)	Percentage Change	Quarterly	132
261	GDP: TI: Other	Percentage Change	Quarterly	132
262	GDP: TI: Financial Intermediation	Percentage Change	Quarterly	132

Series Number	Variable Name	Transformation	Frequency	Number of missing observations
263	GDP: TI: Real Estate	Percentage Change	Quarterly	132
264	GDP: TI: Accommodation and Catering Trade	Percentage Change	Quarterly	132
265	Monetary Authority: Asset: Foreign Asset	Percentage Change	Quarterly	132
266	Monetary Authority: Asset: Foreign Asset: Gold	Percentage Change	Quarterly	132
267	Monetary Authority: Asset: Foreign Asset: Other	Percentage Change	Quarterly	132
268	Monetary Authority: Asset: Claims on Government	Percentage Change	Quarterly	132
269	Monetary Authority: Asset: Claims on Non Financial Institution	Percentage Change	Quarterly	132
270	Finance Company: Asset: Claims on Non Financial Institution	Percentage Change	Quarterly	132
271	Finance Company: Asset: Claims on Other Financial Corporation	Percentage Change	Quarterly	132
272	Finance Company: Asset: Claims on Government	Percentage Change	Quarterly	132
273	Finance Company: Asset: RA: Deposit with Central Bank	Percentage Change	Quarterly	132
274	Finance Company: Asset: Reserve Asset (RA)	Percentage Change	Quarterly	132
275	Diffusion Index: 5000 Industrial Enterprises Survey: Price Level of Sales	First Difference	Quarterly	132
276	Diffusion Index: 5000 Industrial Enterprises Survey: Profitability	First Difference	Quarterly	132
277	Diffusion Index: 5000 Industrial Enterprises Survey: Lending Attitude of Bank	First Difference	Quarterly	132
278	Diffusion Index: 5000 Industrial Enterprises Survey: Cash Inflow From Sales	First Difference	Quarterly	132
279	Diffusion Index: 5000 Industrial Enterprises Survey: Fund Turnover	First Difference	Quarterly	132
280	Diffusion Index: 5000 Industrial Enterprises Survey: Overseas Order Level	First Difference	Quarterly	132
281	Diffusion Index: 5000 Industrial Enterprises Survey: Domestic Order Level	First Difference	Quarterly	132
282	Diffusion Index: 5000 Industrial Enterprises Survey: Inventory Level	First Difference	Quarterly	132
283	Diffusion Index: 5000 Industrial Enterprises Survey: Production Capacity Utilization	First Difference	Quarterly	132
284	Diffusion Index: 5000 Industrial Enterprises Survey: General Business Condition	First Difference	Quarterly	132
285	Diffusion Index: 5000 Industrial Enterprises Survey: Fixed Asset Investment	First Difference	Quarterly	132

Table B.0.2: Data Description: Targeted Predictors

Series Number	Variable Name	Transformation	Frequency
1	Rediscount Rate	None	Monthly
2	Nominal Lending Rate: Within 1 Year (Including 1 Year)	None	Monthly
3	Consumer Price Index	First Difference	Monthly
4	Retail Price Index	First Difference	Monthly
5	Retail Price Index: Urban	First Difference	Monthly
6	Retail Price Index: Rural	First Difference	Monthly
7	Producer Price Index: Agricultural Input: Overall	First Difference	Monthly
8	Producer Price Index(PPI)	First Difference	Monthly
9	Household Savings Deposits Rate: Demand	None	Monthly
10	Interbank Offered Rate: Weighted Avg: Overnight	None	Monthly
11	Effective Exchange Rate Index: BIS: Real	First Difference	Monthly
12	Effective Exchange Rate Index: BIS: Nominal	First Difference	Monthly
13	Consumer Confidence Index	First Difference	Monthly
14	Consumer Confidence Index: Satisfactory	First Difference	Monthly
15	Consumer Confidence Index: Expectation	First Difference	Monthly
16	Coincident Index	First Difference	Monthly
17	Leading Index	First Difference	Monthly
18	Govt Revenue	Percentage Change	Monthly
19	Govt Expenditure	Percentage Change	Monthly
20	Industrial Production: Crude Oil	Percentage Change	Monthly
21	Industrial Production: Iron Ore	Percentage Change	Monthly
22	Energy Production: Electricity	Percentage Change	Monthly
23	Export FOB	Percentage Change	Monthly
24	Import CIF	Percentage Change	Monthly
25	Money Supply M2	Percentage Change	Monthly
26	Money Supply M1	Percentage Change	Monthly
27	Official Reserve Asset: Foreign Reserve(FR)	Percentage Change	Monthly

Series Number	Variable Name	Transformation	Frequency
28	Loan	Percentage Change	Monthly
29	Index: Shanghai Stock Exchange: Composite	Percentage Change	Monthly
30	Index: Shenzhen Stock Exchange: Composite	Percentage Change	Monthly
31	Official Reserve Asset: Gold: Gold Reserve	Percentage Change	Monthly
32	Money Supply M0	Percentage Change	Monthly
33	Deposit	Percentage Change	Monthly
34	Transport: Freight Traffic	Percentage Change	Monthly
35	Railway: Freight Traffic	Percentage Change	Monthly
36	Highway: Freight Traffic	Percentage Change	Monthly
37	Waterway: Freight Traffic	Percentage Change	Monthly
38	Air: Freight Traffic	Percentage Change	Monthly
39	NomInvestment	Percentage Change	Monthly
40	NomConsumption	Percentage Change	Monthly
41	InvestmentPrice	Percentage Change	Monthly

Table B.0.3: Data Description: Chapter 4

Series Number	Variable Name	Transformation	Frequency
1	GeneralBusinessConditionIndex	First Difference	Quarterly
2	ProductionCapacitiyIndex	First Difference	Quarterly
3	CPI	First Difference	Quarterly
4	RetailPriceIndex	First Difference	Quarterly
5	FAIPriceIndex	First Difference	Quarterly
6	GFCFPriceIndex	First Difference	Quarterly
7	R3mDeposit	None	Quarterly
8	LendingRatePBC1year	None	Quarterly
9	NominalRetailGoodsC	Percentage Change	Quarterly
10	NominalExportsGoods	Percentage Change	Quarterly
11	NominalImportsGoods	Percentage Change	Quarterly
12	NominalHHC	Percentage Change	Quarterly
13	NominalGovtC	Percentage Change	Quarterly
14	NominalGCF	Percentage Change	Quarterly
15	BankLoansTotal	Percentage Change	Quarterly
16	M2	Percentage Change	Quarterly
17	Industrial production index	First Difference	Quarterly
18	Consumer Confidence Index	First Difference	Monthly
19	Consumer Confidence Index: Satisfactory	First Difference	Monthly
20	Consumer Confidence Index: Expectation	First Difference	Monthly
21	CPI	First Difference	Monthly
22	RPI	First Difference	Monthly
23	Effective Exchange Rate Index: Real	None	Monthly
24	Effective Exchange Rate Index: Nominal	None	Monthly
25	Export	Percentage Change	Monthly
26	Import	Percentage Change	Monthly
27	Industrial Production: Crude Oil	Percentage Change	Monthly

Series Number	Variable Name	Transformation	Frequency
28	Industrial Production: Natural Gas	Percentage Change	Monthly
29	Energy Production: Crude Oil	Percentage Change	Monthly
30	Energy Production: Electricity	Percentage Change	Monthly
31	Construction Material Production: Cement	Percentage Change	Monthly
32	Construction Material Production: Plated Glass	Percentage Change	Monthly
33	NomInvestment	Percentage Change	Monthly
34	NomConsumption	Percentage Change	Monthly
35	M2	Percentage Change	Monthly
36	Govt Revenue	Percentage Change	Monthly
37	Govt Expenditure	Percentage Change	Monthly
38	Corporate Goods Price Index: Overall	Percentage Change	Monthly
39	Index: Shanghai Stock Exchange: Composite	First Difference	Monthly
40	Index: Shanghai Stock Exchange: A Share	First Difference	Monthly
41	Index: Shanghai Stock Exchange: B Share	First Difference	Monthly
42	Index: Shenzhen Stock Exchange: Composite	First Difference	Monthly
43	Index: Shenzhen Stock Exchange: A Share	First Difference	Monthly
44	Index: Shenzhen Stock Exchange: B Share	First Difference	Monthly
45	Industrial production index	First Difference	Monthly
46	DepositRate1YBench	None	Monthly

Appendix C

Additional Figures

C.1 The Graphical Presentation of Tables in Section 2.4.2

This section includes the graphical presentation of sectoral discrepancies by the VA approach, sectoral contribution rate by the VA approach, sectoral Discrepancies by the AE Approach and sectoral contribution rate by the AE approach. This appendix plots Table 2.2 into Figure C.1.1, Table 2.3 into Figure C.1.2, Table 2.4 into Figure C.1.3, and Table 2.5 into Figure C.1.4.

Figure C.1.1: Plot of Sectoral Discrepancies by the VA Approach

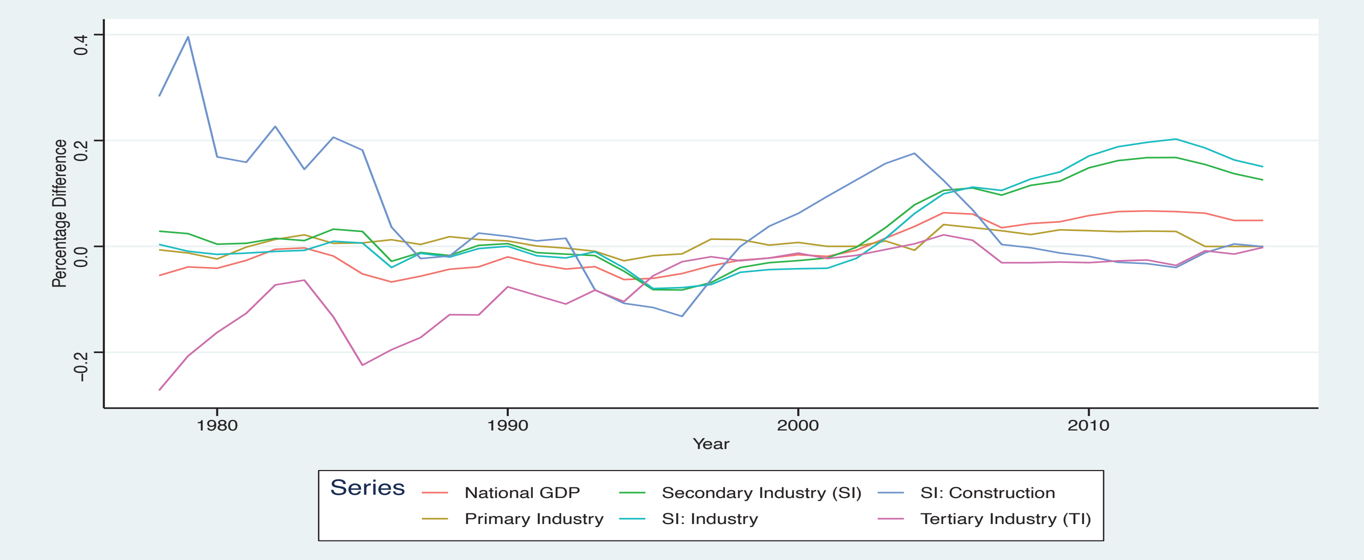


Figure C.1.2: Plot of Sectoral Contribution Rate by the VA Approach

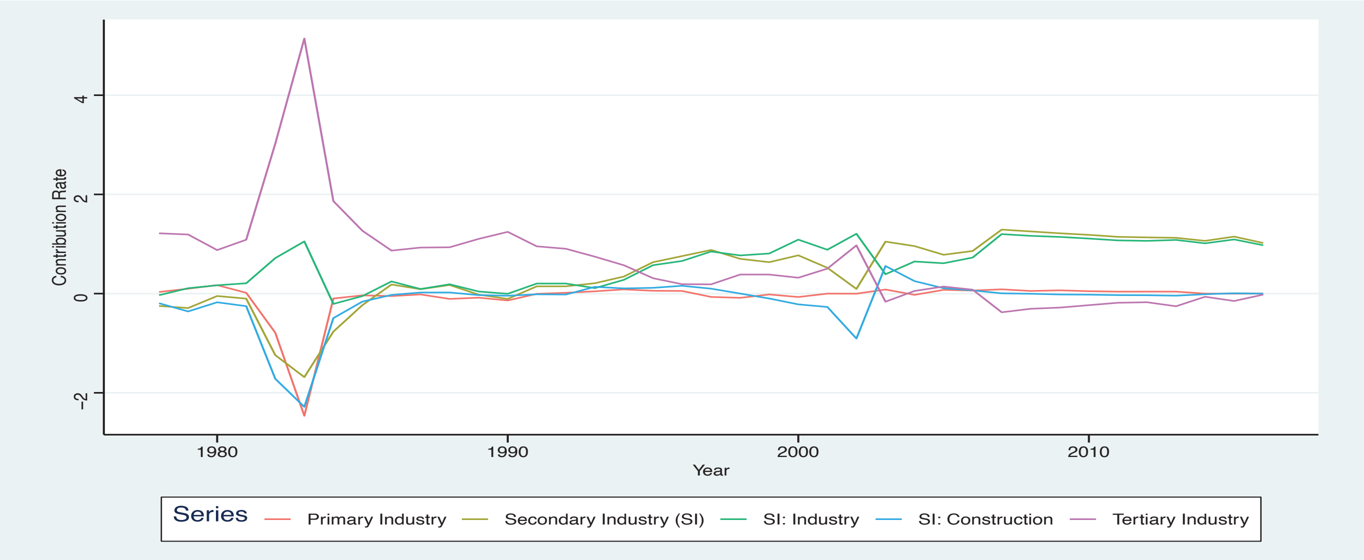


Figure C.1.3: Plot of Sectoral Discrepancies by the AE Approach

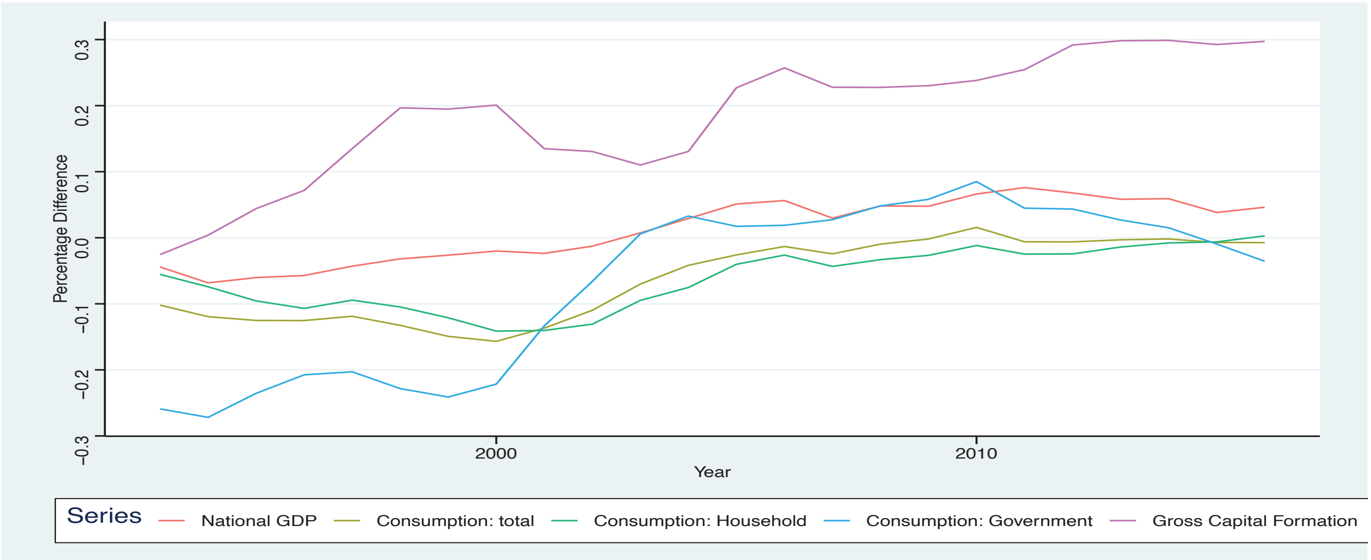
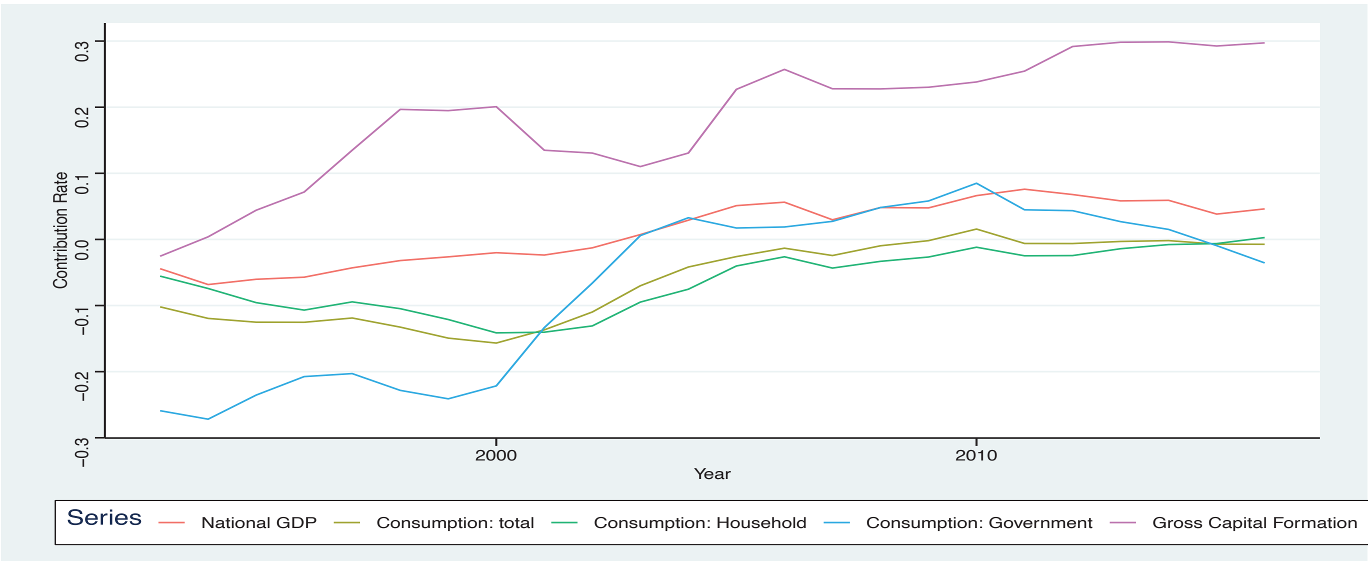


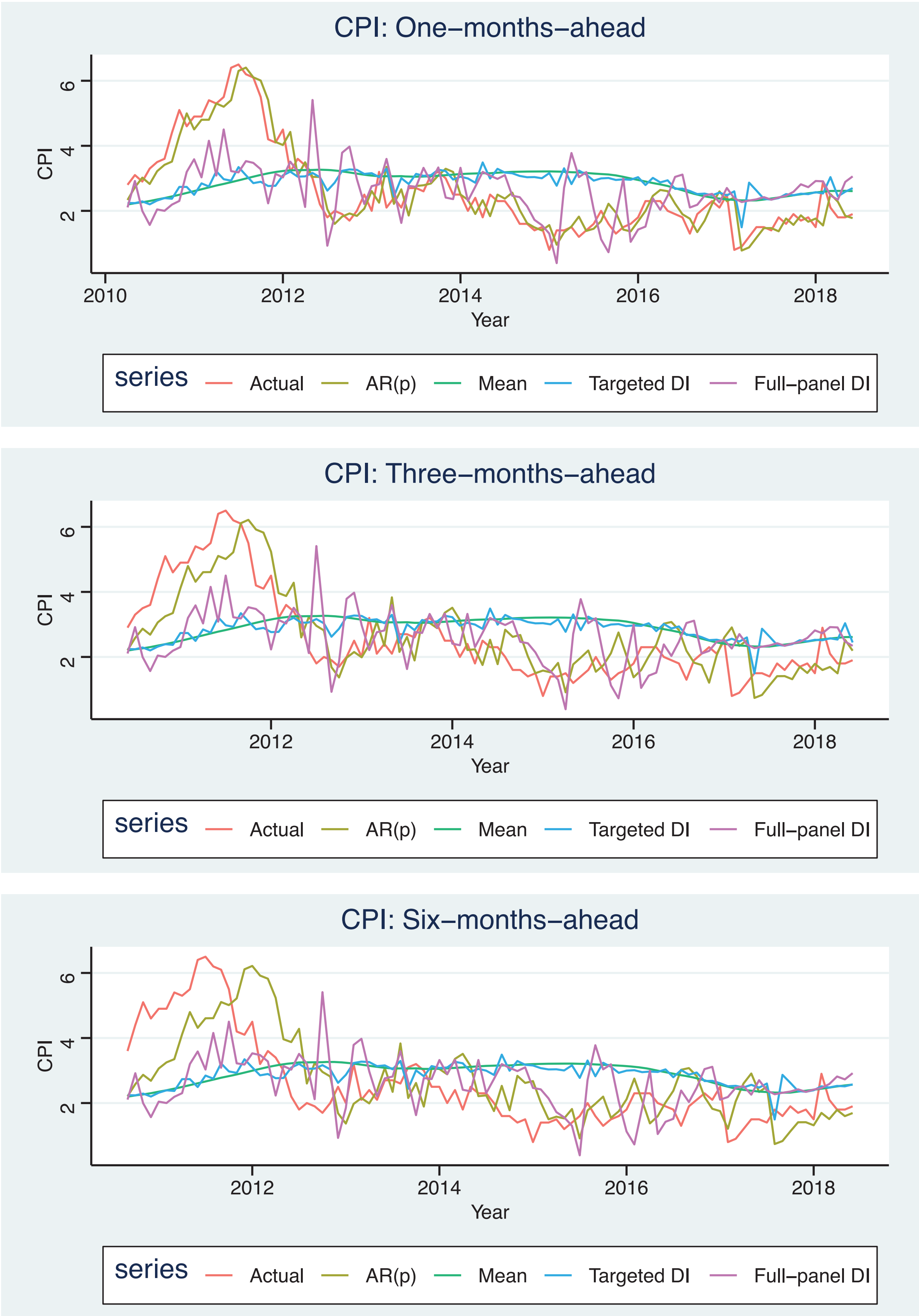
Figure C.1.4: Plot of Sectoral Contribution Rate by the AE Approach

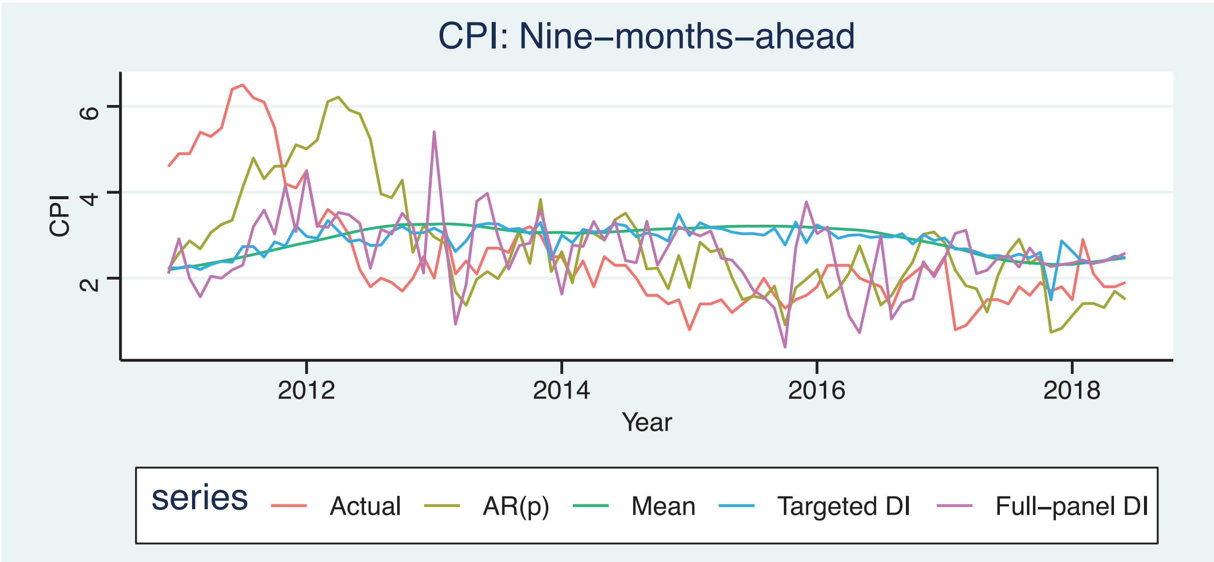


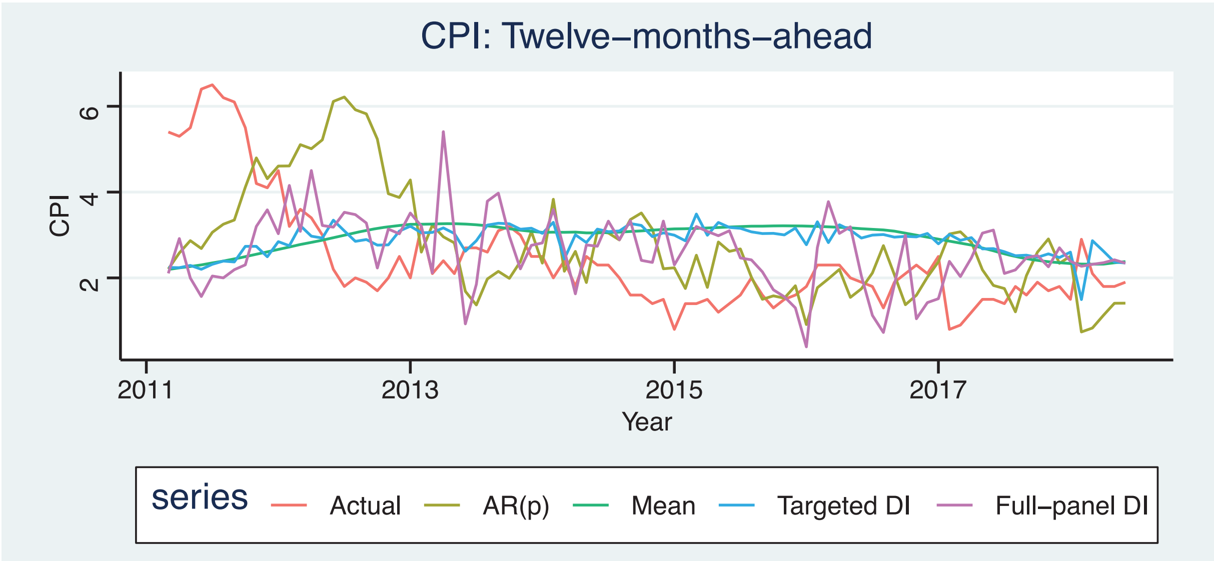
C.2 The Graphical Presentation of Forecasts for CPI and RPI

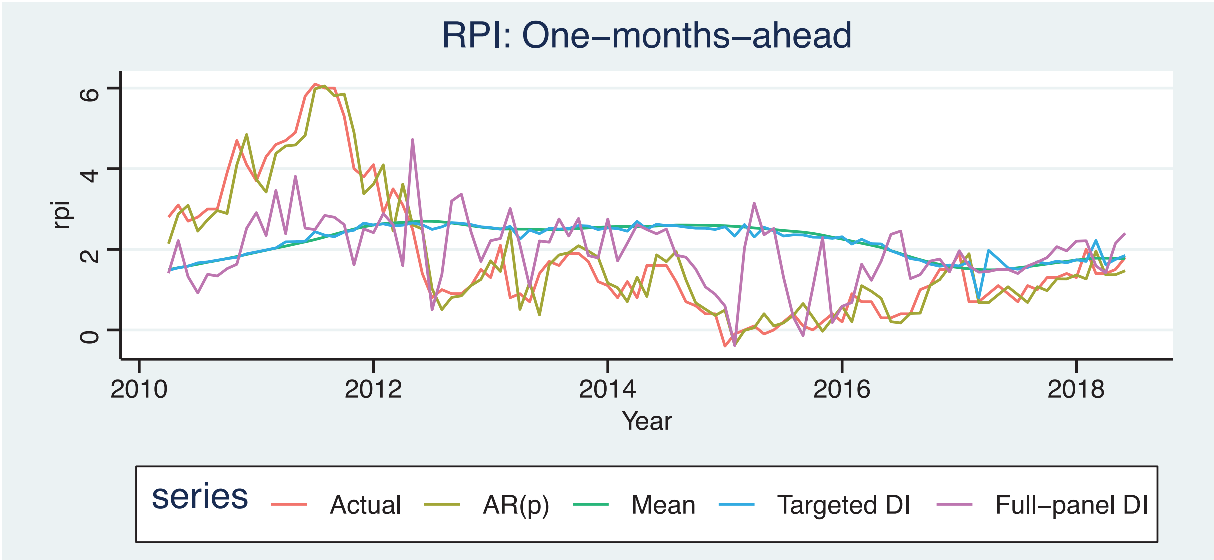
One observed oddity from Table 3.2 is the sharp declines in relative MSEs of the mean forecast, the targeted DI forecast, and the mixed-frequency DI forecast for the CPI and RPI as forecast horizons increase. As it stands, it is possible that the benchmark AR(p) model perform extremely badly at longer horizons, so do other models since the performance of competing models are assessed using relative terms. This section graphically shows the AR(p) forecast, mean forecast, targeted DI forecast, and mixed-frequency DI forecast for CPI and RPI at h=1,3,6,9, and 12.

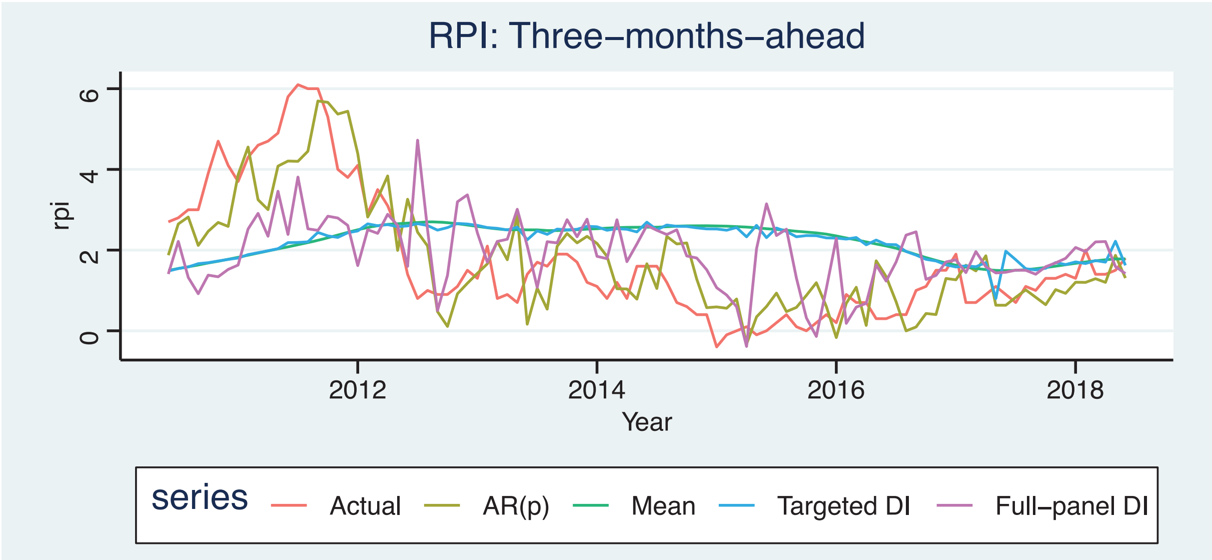
Figure C.2.1: Time Series Plot of CPI and RPI vs. AR(p) forecast, Mean forecast, Targeted DI forecast, and full-panel DI forecast

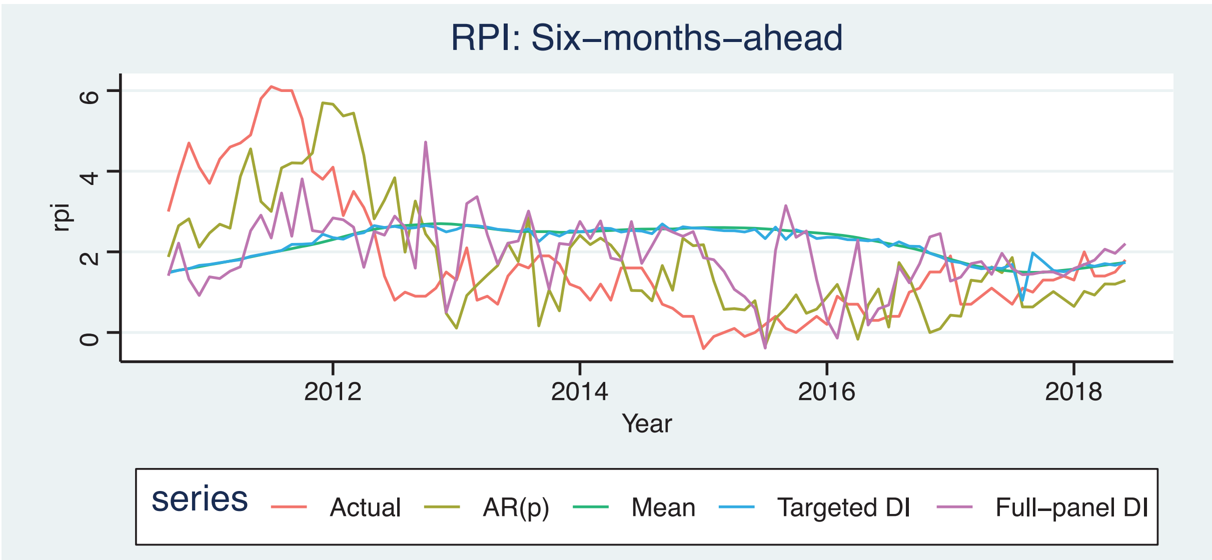


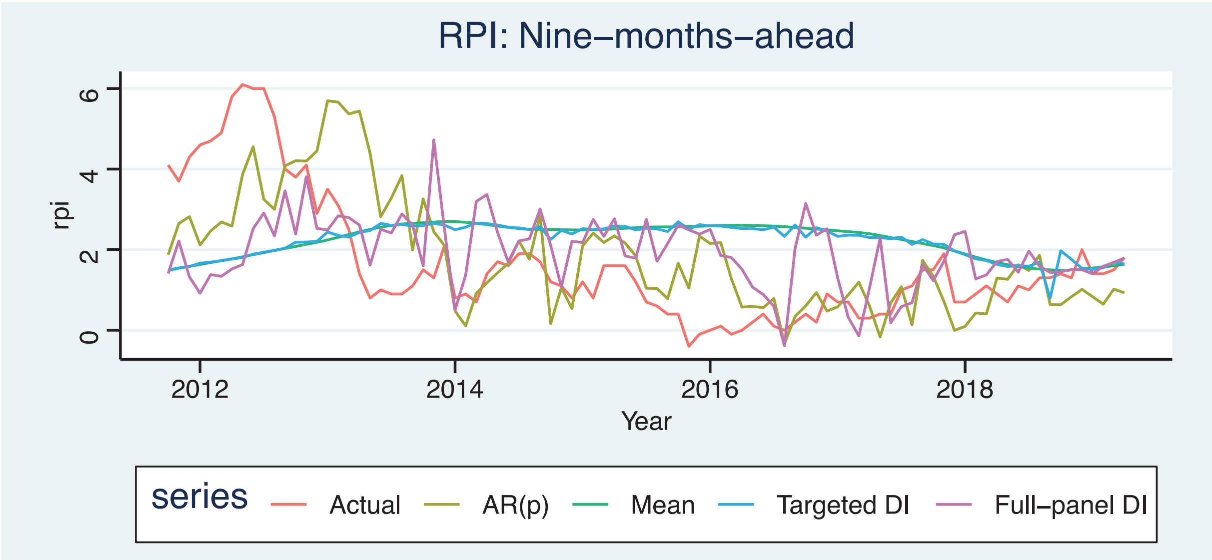

CPI: Nine–months–ahead
CPI
6
4
2
2012
2014
2016
2018
Year
series
Actual
AR(p)
Mean
Targeted DI
Full–panel DI


CPI: Twelve–months–ahead
CPI
6
4
2
2011
2013
2015
2017
Year
series
Actual
AR(p)
Mean
Targeted DI
Full–panel DI

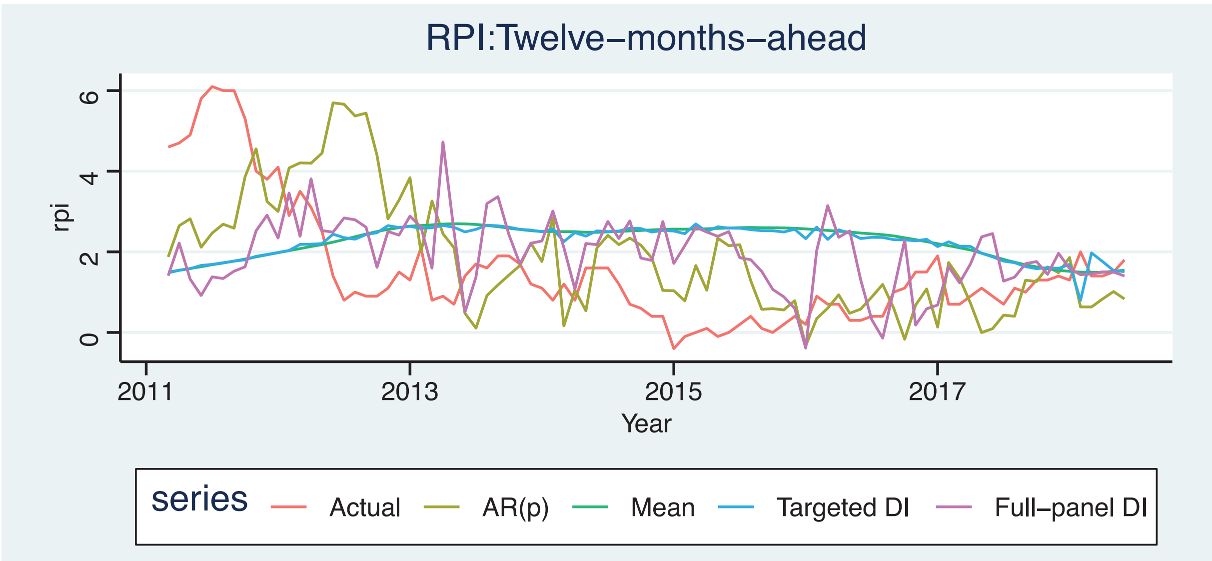

RPI: One–months–ahead
rpi
6
4
2
0
2010
2012
2014
2016
2018
Year
series
Actual
AR(p)
Mean
Targeted DI
Full–panel DI


RPI: Three–months–ahead
rpi
0
2
4
6
2012
2014
2016
2018
Year
series
Actual
AR(p)
Mean
Targeted DI
Full–panel DI


RPI: Six–months–ahead
rpi
0
2
4
6
2012
2014
2016
2018
Year
series
Actual
AR(p)
Mean
Targeted DI
Full–panel DI



RPI: Nine−months−ahead
rpi
0
2
4
6
2012
2014
2016
2018
Year
series
Actual
AR(p)
Mean
Targeted DI
Full−panel DI



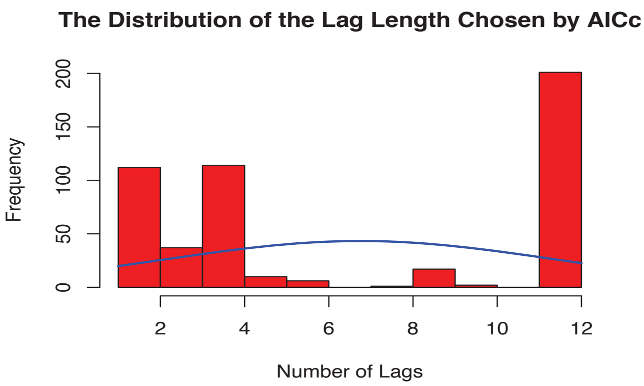
RPI:Twelve−months−ahead
rpi
0
2
4
6
2011
2013
2015
2017
Year
series
Actual
AR(p)
Mean
Targeted DI
Full−panel DI

C.3 The Distribution of the Lag Length Chosen by Information Criteria for the AR Model

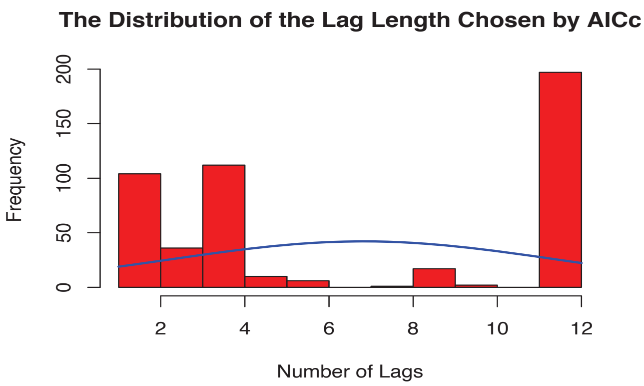
This section presents the distribution of the lag length chosen by information criteria for AR model at h=1,3,6,9, and 12.

Figure C.3.1: The distribution of the lag length chosen by information criteria for AR model

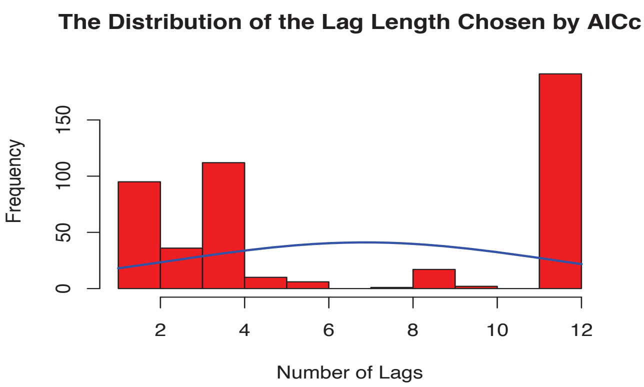
(a) One-month-ahead



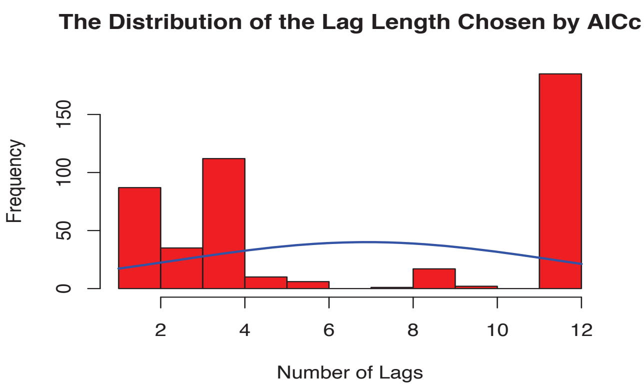
(b) Three-months-ahead



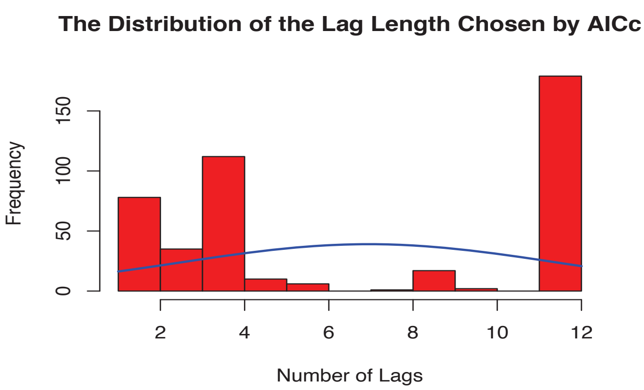
(c) Six-months-ahead



(d) Nine-months-ahead



(e) Twelve-months-ahead

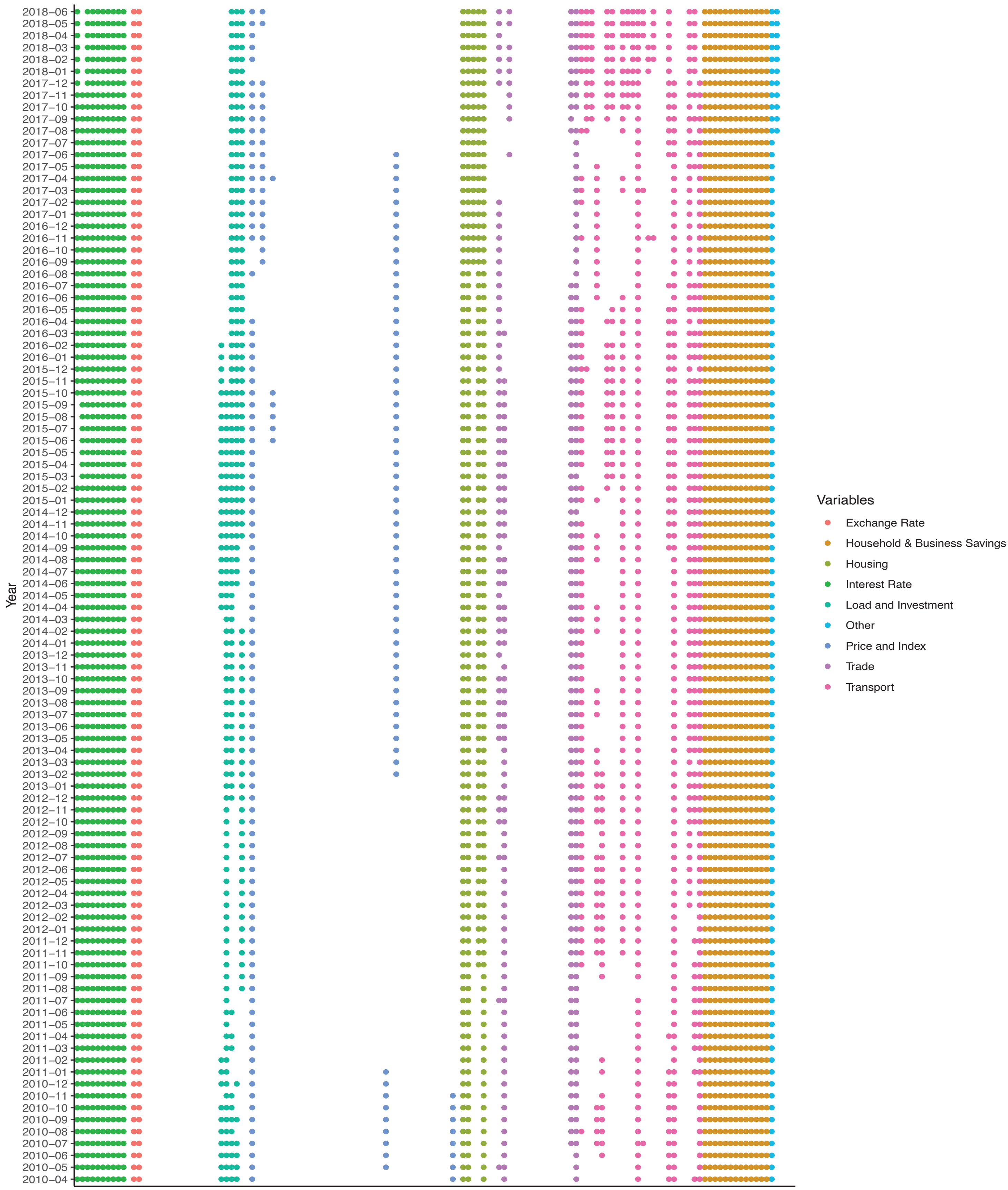


C.4 Stability of Predictors Selection by Hard Thresholding for Targeted FA-AR and FA-VAR

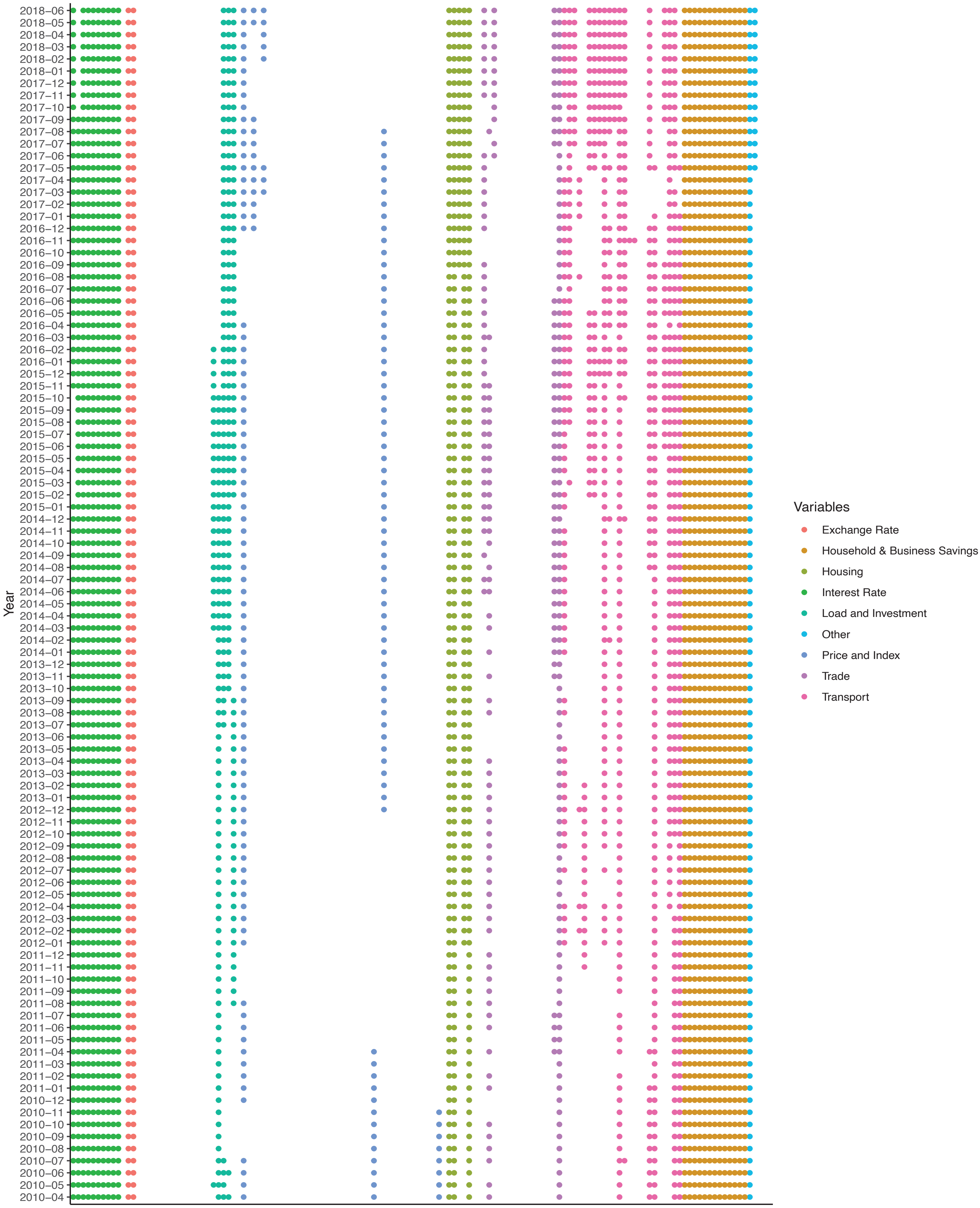
Figure C.4.1 shows the type of series selected or not by hard thresholding with critical Value being for targeted FA-AR and FA-VAR over the whole out-of-sample period. We group the data as into nine categories: exchange rate, household and business savings, housing, interest rate, loan and investment, price and index, trade, transport, and other. The probability that a particular predictor will be consistently selected is higher for some groups and depends on the series being predicted. For instance, indicators in interest rate, household and business savings, housing and load and investment are consistently present when predicting CPI, and RPI over the whole out-of-sample. A similar pattern is observed in the case of railway cargo. Contrastingly, there is vast instability in predictor selection for investment and consumption where the selection criteria change dramatically over time.

Figure C.4.1: Series Preselected by Hard Thresholding with Critical Value Being 1.67 for Targeted FA-AR and FA-VAR

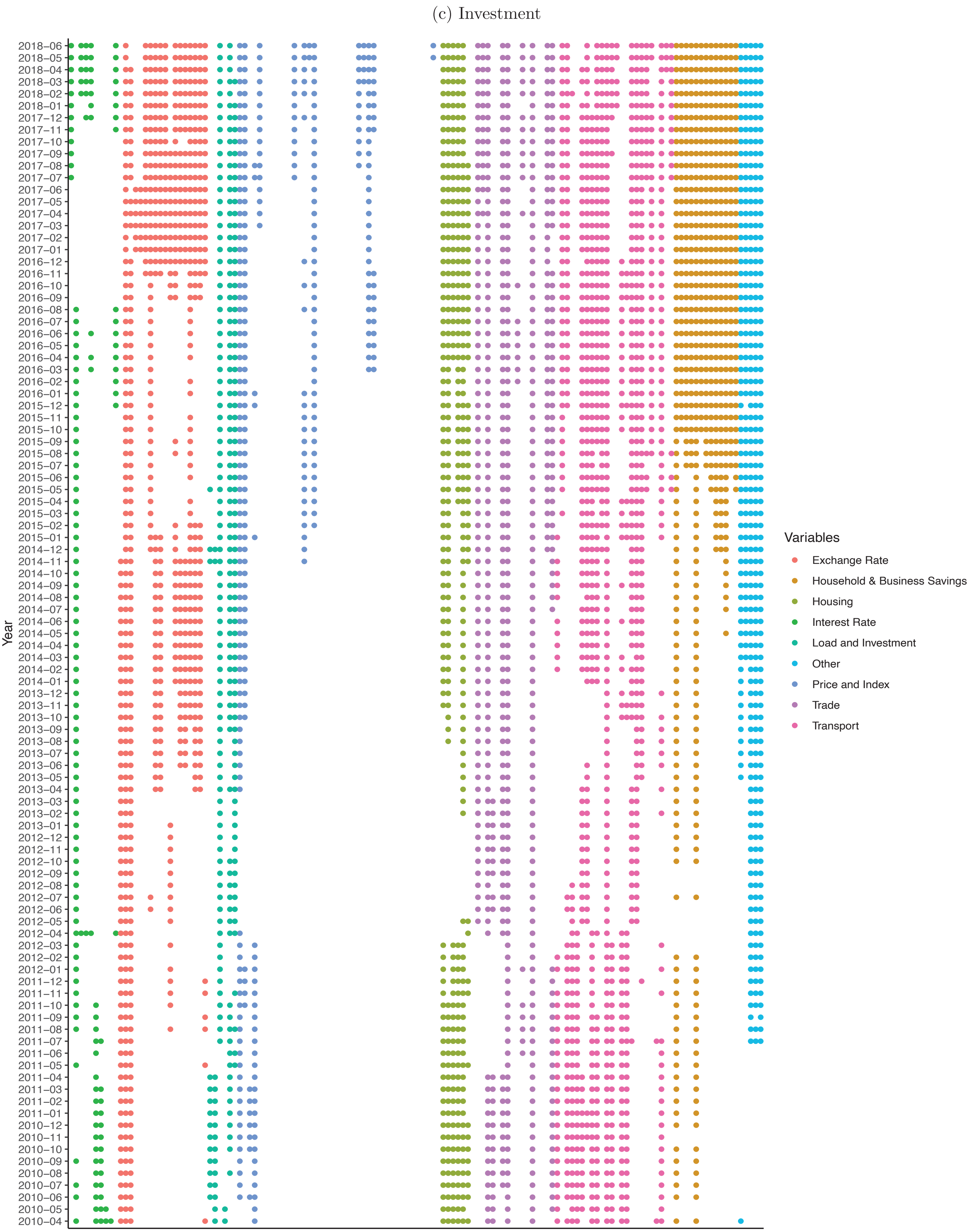
(a) CPI



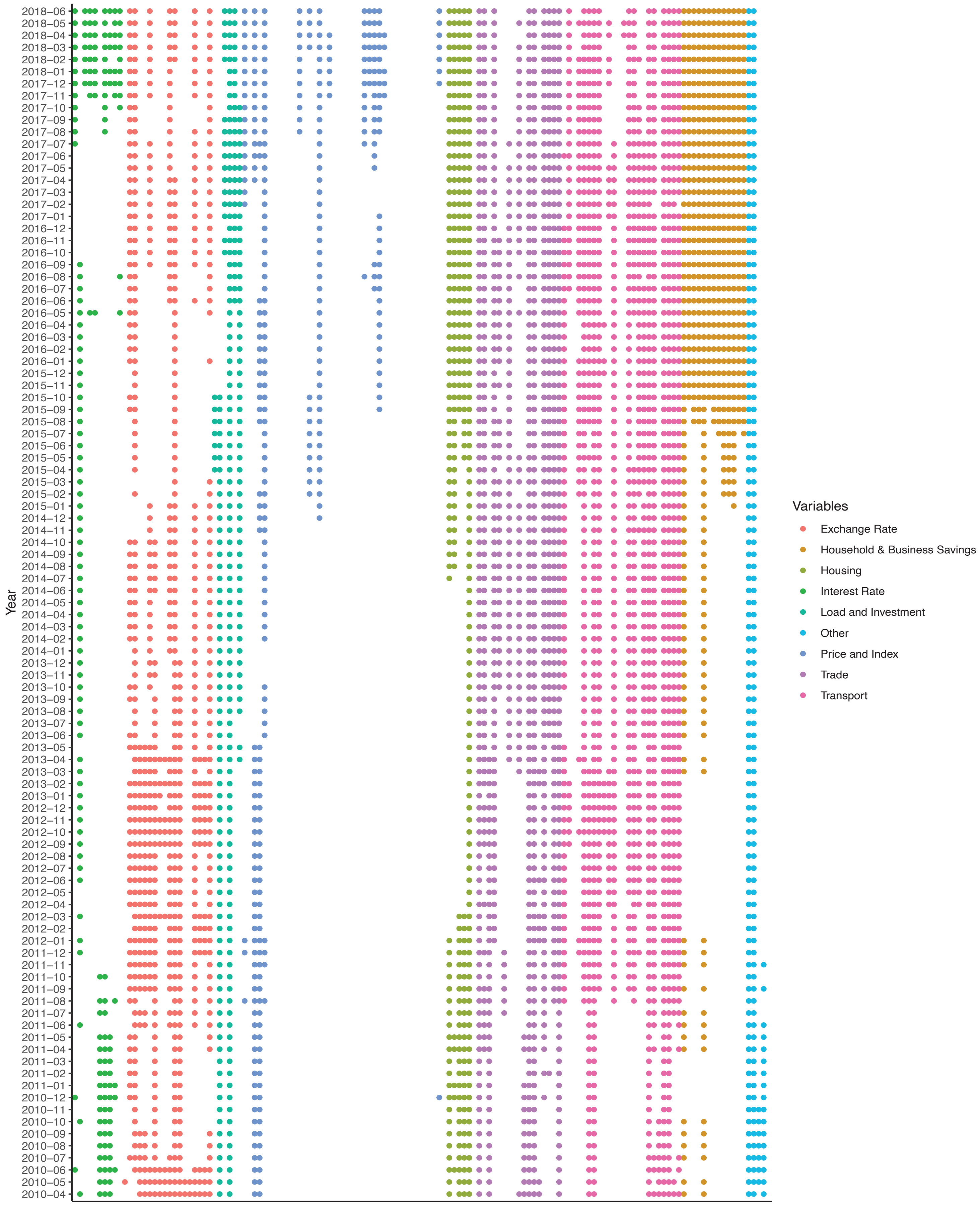
(b) RPI



Year
2018–06
2018–05
2018–04
2018–03
2018–02
2018–01
2017–12
2017–11
2017–10
2017–09
2017–08
2017–07
2017–06
2017–05
2017–04
2017–03
2017–02
2017–01
2016–12
2016–11
2016–10
2016–09
2016–08
2016–07
2016–06
2016–05
2016–04
2016–03
2016–02
2016–01
2015–12
2015–11
2015–10
2015–09
2015–08
2015–07
2015–06
2015–05
2015–04
2015–03
2015–02
2015–01
2014–12
2014–11
2014–10
2014–09
2014–08
2014–07
2014–06
2014–05
2014–04
2014–03
2014–02
2014–01
2013–12
2013–11
2013–10
2013–09
2013–08
2013–07
2013–06
2013–05
2013–04
2013–03
2013–02
2013–01
2012–12
2012–11
2012–10
2012–09
2012–08
2012–07
2012–06
2012–05
2012–04
2012–03
2012–02
2012–01
2011–12
2011–11
2011–10
2011–09
2011–08
2011–07
2011–06
2011–05
2011–04
2011–03
2011–02
2011–01
2010–12
2010–11
2010–10
2010–09
2010–08
2010–07
2010–06
2010–05
2010–04
Variables
Exchange Rate
Household & Business Savings
Housing
Interest Rate
Load and Investment
Other
Price and Index
Trade
Transport


(c) Investment
Year
2018–06
2018–05
2018–04
2018–03
2018–02
2018–01
2017–12
2017–11
2017–10
2017–09
2017–08
2017–07
2017–06
2017–05
2017–04
2017–03
2017–02
2017–01
2016–12
2016–11
2016–10
2016–09
2016–08
2016–07
2016–06
2016–05
2016–04
2016–03
2016–02
2016–01
2015–12
2015–11
2015–10
2015–09
2015–08
2015–07
2015–06
2015–05
2015–04
2015–03
2015–02
2015–01
2014–12
2014–11
2014–10
2014–09
2014–08
2014–07
2014–06
2014–05
2014–04
2014–03
2014–02
2014–01
2013–12
2013–11
2013–10
2013–09
2013–08
2013–07
2013–06
2013–05
2013–04
2013–03
2013–02
2013–01
2012–12
2012–11
2012–10
2012–09
2012–08
2012–07
2012–06
2012–05
2012–04
2012–03
2012–02
2012–01
2011–12
2011–11
2011–10
2011–09
2011–08
2011–07
2011–06
2011–05
2011–04
2011–03
2011–02
2011–01
2010–12
2010–11
2010–10
2010–09
2010–08
2010–07
2010–06
2010–05
2010–04
Variables
Exchange Rate
Household & Business Savings
Housing
Interest Rate
Load and Investment
Other
Price and Index
Trade
Transport

(d) Consumption



(e) Railway Cargo