ARTICLE IN PRESS

Economic Modelling xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/econmod

Forecasting China's GDP growth using dynamic factors and mixed-frequency data☆

Yu Jiang, Yongji Guo, Yihao Zhang*

Department of Finance and Insurance, Nanjing University, Nanjing, Jiangsu 210093, China

ARTICLE INFO

JEL classification:
C53
E17

Keywords:
GDP growth forecast
Dynamic factor model
MIDAS regression

ABSTRACT

Forecasting GDP growth is important and necessary for Chinese government to set GDP growth target. To fully and efficiently utilize macroeconomic and financial information, this paper attempts to forecast China's GDP growth using dynamic predictors and mixed-frequency data. The dynamic factor model is first applied to select dynamic predictors among large amount of monthly macroeconomic and daily financial data and then the mixed data sampling regression is applied to forecast quarterly GDP growth based on the selected monthly and daily predictors. Empirical results show that forecasts using dynamic predictors and mixed-frequency data have better accuracy comparing to traditional forecasting methods. Moreover, forecasts with leads and forecast combination can further improve forecast performance.

1. Introduction

Macroeconomic forecast, especially GDP growth forecast, has always been an important yet difficult task for the government of a country. GDP growth forecast becomes even more important for Chinese government because China is the only major economy that sets rigid target for annual GDP growth in the world.[1] However, the domestic and international economic situation has been becoming increasingly complicated since 2000s. The global economic growth has continued to decline and the fluctuations of global financial markets have become more intense and frequent than ever. On the other hand, the rapid but extensive economic growth since the Reform and Opening in 1980s resulted in serious structural imbalances of Chinese economy which causes the continuous economic slowdown over the past decade. These uncertainties caused by external shocks and internal adjustments lead to the increasing difficulty in China's GDP growth forecast. To formulate more reasonable GDP growth target, rational and efficient technology and methodology need to be developed and applied to China's GDP growth forecast.

GDP growth forecast has always been one of the most popular and important research topics and many forecasting models have been developed by economists, econometricians and statisticians. To make accurate GDP growth forecast, forecasting models need to address the following two issues, reasonable selection of predictors and efficient utilization of data with different frequencies (Koenig et al., 2003; Armesto et al., 2010; Andreou et al., 2013).

The trend of GDP growth is considered to be affected by many economic variables, such as macroeconomic variables and financial variables. Determining predictors of GDP growth forecast using a large number of economic variables is not easy and straightforward. Stock and Watson (1989, 2002) point out that macroeconomic fluctuations is transferred and diffused through a series of economic activities instead of a single economic variable and such process is complicated and covers all macroeconomic aspects. Therefore, they argue that macroeconomic forecast should make full use of economic variables that contain extensive information, including coincident variables such as industrial output, income per capita, trade volume, employment time, etc. and leading variables such as interest rates, exchange rates, stock returns, long-term bond yields, etc.

The diversity of affecting factors requires GDP growth forecast to focus on the comprehensive and efficient use of economic information, which means that predictors need to be reasonably determined using numerous economic variables without losing much information contained in these economic variables. Based on the work of Sargent and Sims (1977) and Engle and Watson (1981), Stock and Watson (2003, 2006) apply the method of factor analysis to decompose a large number

☆ This research is supported by the National Natural Science Foundation of China (Grant nos. 71301072 and 71672079). We would like to thank an anonymous referee and the editor, Sushanta Mallick, for many helpful suggestions and constructive comments.

* Corresponding author.
E-mail addresses: yujiang@nju.edu.cn (Y. Jiang), guoyongji1221@126.com (Y. Guo), zhangyh@nju.edu.cn (Y. Zhang).

[1] The target of China's GDP growth was set to be 7.5% and 7% in 2014 and 2015, respectively. The GDP growth target in 2016 is set to be between 6.5% and 7%, the first time of using an interval form since 1995.

http://dx.doi.org/10.1016/j.econmod.2017.06.005
Received 6 April 2017; Received in revised form 28 May 2017; Accepted 7 June 2017
0264-9993/

of economic variables into common and heterogeneous components and then extract a small number of common factors from the common component as predictors. Using the method of factor analysis to determine predictors can effectively reduce the dimension of data and solve the problem of over-fitting and over-parametric caused by the inclusion of a larger number of variables, while still capture a large amount of information and avoid the structural instability in the forecasting (Stock and Watson, 2010).

Although a small number of factors can be extracted from a large number of economic variables as predictors, GDP growth forecast still need to deal with the issue of different data frequencies. Generally speaking, the data of GDP is released quarterly; the data of price index, consumption, investment, etc. is released monthly; the data of stock market, bond market, futures market, etc. can be obtained daily. Traditional forecasting models require the same data frequency and certain methodologies are applied to transfer daily and monthly data of higher frequency into quarterly data of lower frequency, such as averaging, bridging (Diron, 2008) and temporal aggregation (Silvestrini and Veredas, 2008). However, these methods of frequency conversion may result in the loss of a considerable part of information contained in the high-frequency data such as the fluctuations of high-frequency data and thus reduce the utilization efficiency of sample information to a certain extent.

To avoid the information loss in frequency conversion, Ghysels et al. (2004) develop the mixed data sampling (MIDAS) method that uses the original mixed-frequency data directly in the forecasting model without any artificial processing. MIDAS method has the following advantages compared with the traditional methods of frequency conversion. First, it can take full use of high-frequency data and avoid substantial loss of sample information, thereby enhancing the forecast accuracy. Second, it can perform real-time forecasts of quarterly GDP growth, using the latest revealed high-frequency financial and economic data between quarters. Due to its specific features, MIDAS method is well suited for GDP growth forecast based on daily financial or monthly macroeconomic data (e.g. Ghysels and Wright, 2009; Andreou et al., 2011; Frale and Monteforte, 2011; Liu et al., 2012; Monteforte and Moretti, 2012; Dias et al., 2015).

In summary, the method of factor analysis allows the determination of finite number of factors that effectively summarize the common movement of a large number of economic variables affecting GDP growth, while the method of MIDAS allows the efficient use of economic data with different frequencies in the forecasting model. The procedure of GDP growth forecast using these two methods follows two steps naturally. First, the method of factor analysis is applied to extract a small number of factors from a large number of macroeconomic and financial variables of different data frequencies. Second, take the extracted factors as predictors and apply the method of MIDAS to make in-sample estimation and out-of-sample forecast. GDP growth forecast using these two methods not only takes use of massive amounts of economic data, but also avoids complicated and time-consuming computations. Moreover, the feature of low information loss of both methods can improve the forecast accuracy.

The engine of China's rapid GDP growth had been the investments and the exports for a long time since 1980s. During the period of 1980–2010, the growth of investments and net exports was faster than the growth of consumption, and the proportion of investments and net exports in GDP was higher than the proportion of consumption. However, as the external demand continued to decline and the domestic overcapacity became increasingly serious in recent years, China's GDP growth has been keeping slowing down. China's economy has entered the ‘New Normal’ and need to restructure and upgrade. Domestic demand, especially the consumption, gradually starts to take on the task of maintaining GDP growth. The constantly evolving macroeconomic situation of China in recent years implies that the static selection of dominant predictors is unreasonable and the predictor determinations for forecasting China's GDP growth should have dynamic characteristics. On the other hand, as the world's second largest economy, the determinants of China's GDP growth are complicated and a large number of macroeconomic and financial variables should be considered, which implies that the forecasting model should be able to deal with mixed-frequency data such as monthly macroeconomic data and daily financial data.

According to above considerations, we choose to forecast China's GDP growth using dynamic factors and mixed-frequency data in order to improve the forecast accuracy. To be more specific, we first apply the dynamic factor model (DFM) to extract a small number of dynamic factors out of a large amount of monthly and daily panel data. Second, using the extracted dynamic factors as predictors, we make in-sample estimation out-of-sample forecasts using the mixed data sampling (MIDAS) regression model. Third, we compare forecast performance with different predictors, leads and forecast combinations, and test for predictive abilities.

Forecasting GDP growth using dynamic factors and mixed-frequency data meets the complex and dynamic feature of China's GDP growth and also can take comprehensive and efficient use of economic data without subjective interference, which therefore leads to the improvement of forecast accuracy compared with traditional forecasting models. China is currently in the period of economic structure upgrading and economic growth mode transformation. Our forecasting model would provide useful references for Chinese government to set reasonable economic growth targets and implement timely and effective macroeconomic regulation policies, in order to prevent the impacts of economic growth uncertainty during such a critical period. Meanwhile, this paper also provides a useful and practical method for general macroeconomic forecasts.

The remainder of this paper is organized as follows. Section 2 provides the details of the methodology used in this paper. The description of data and the determination of dynamic factors are given in Section 3. Section 4 reports the results and comparisons between GDP growth forecasts of different cases. Section 5 concludes.

2. Methodology

2.1. Dynamic factor model

The dynamic factor model (DFM) is applied to extract dynamic factors as predictors from large amounts of macroeconomic and financial data. The DFM has two advantages. First, the idiosyncratic parts of the DFM are allowed to be autocorrelated and have heteroskedasticity in both the time and the cross-section dimension, which is especially suitable for the financial time series. Second, the DFM allows factor loadings to change with time, which can effectively solve the issue of instability that may exist in the data due to the nonstationary or the existence of structural breaks of the time series. The DFM is given as follows:

$$\begin{aligned}\mathbf{X}_t &= \mathbf{\Lambda}(L)\mathbf{f}_t + \mathbf{u}_t\\ \mathbf{f}_t &= \mathbf{\Phi}(L)\mathbf{f}_{t-1} + \boldsymbol{\varepsilon}_t{}'\end{aligned} \tag{1}$$

where $\mathbf{X}_t = (X_{1t}, X_{2t}, \cdots, X_{Nt})'$ is an $N \times 1$ vector containing the observations of N economic variables at time t $(t = 1, \cdots, T)$, $\mathbf{\Lambda}(L)$ is an $N \times k$ lag polynomial matrix of degree p, $\mathbf{f}_t$ is a $k \times 1$ vector containing k dynamic factors at time t, $\mathbf{u}_t = (u_{1t}, \cdots, u_{Nt})'$ is an $N \times 1$ vector of error terms, $\mathbf{\Phi}(L)$ is an $k \times k$ lag polynomial matrix of degree q, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \cdots, \varepsilon_{kt})'$ is a $K \times 1$ vector of error terms. The ith lag polynomial $\boldsymbol{\lambda}_i(L)$ is called the dynamic factor loading for the ith time series X_{it}, $\boldsymbol{\lambda}_i(L)\mathbf{f}_t$ is called the common component of the ith time series, and u_{it} is called the idiosyncratic component of the ith time series. $\mathbf{f}_t$ is the dynamic factor at time t that drives the co-movements of a large number of economic variables and can be used as the predictor of GDP growth forecast.

Let $\mathbf{F}_t = (\mathbf{f}_t{}', \cdots, \mathbf{f}_{t-p}{}')'$ be a $r \times 1$ vector of static factors and