HONG KONG INSTITUTE FOR MONETARY RESEARCH

NOWCASTING CHINESE GDP: INFORMATION CONTENT OF ECONOMIC AND FINANCIAL DATA

Matthew S. Yiu and Kenneth K. Chow

HKIMR Working Paper No.04/2011

January 2011

Nowcasting Chinese GDP: Information Content of Economic and Financial Data

Matthew S. Yiu^*

Hong Kong Monetary Authority

and

Kenneth K. Chow

Hong Kong Institute for Monetary Research

January 2011

Abstract

This paper applies the factor model proposed by Giannone, Reichlin, and Small (2005) on a large data set to nowcast (i.e. current-quarter forecast) the annual growth rate of China's quarterly GDP. The data set contains 189 indicator series of several categories, such as prices, industrial production, fixed asset investment, external sector, money market and financial market. This paper also applies Bai and Ng's criteria (2002) to determine the number of common factors in the factor model. The identified model generates out-of-sample nowcasts for China's GDP with smaller mean squared forecast errors than those of the Random Walk benchmark. Moreover, using the factor model, we find that interest rate data is the single most important block in estimating current-quarter GDP in China. Other important blocks are consumer and retail prices data and fixed asset investment indicators.

Keywords: Large Data Set, Pseudo Real Time Estimates, Factor Model, Kalman Filtering, Nowcasting, Information Content

JEL Classification: C33, C53, E32, E37

^ Contact information: Matthew S. Yiu, HKIMR, 55/F, Two International Finance Centre, 8 Finance Street, Central, Hong Kong, e-mail: msfyiu@hkma.gov.hk.

* We gratefully acknowledge the contribution from Domenico Giannone for making programs used by Giannone, Reichlin and Small (2007) available to us

1. Introduction

Mainland China opened its economy around the end of the 1970s and since then its economic performance has been extraordinary by any standards. China's GDP has grown at an average rate above 10 percent over the past thirty years and, as the economy has grown at such a fast pace, the role of the Chinese economy in the world has become more and more important, particularly in recent years. In 2007, its PPP-valued GDP was ranked in second place after the US and its total external trade accounted for more than 10 percent of the world trade.

In the Asian region, the Chinese economy is now one of the two most important economies and the growth of the region will rely more on China's growth in the years to come. Furthermore, in the global financial turmoil, caused initially by the subprime problem in the US, China is expected to be one of the very few economies in the world that could emerge as a more important economy after the storm. More significantly, it is the most important economy to Hong Kong. Mainland China is Hong Kong's largest export market, accounting for more than 40 percent of Hong Kong's total exports. At the same time, about 40 percent of Hong Kong's imports come from the Mainland. Hong Kong also serves as the entrepot and financial centre between the Mainland and the rest of the world. Yet, despite China's increasing importance for the world economy and the region, the number of studies on real-time estimating or forecasting its economic activity is still surprisingly small.

For policy making purposes, in spite of the difficulties of obtaining reliable and accurate estimates of current economic activity, a clear understanding of the state of macroeconomic activity is highly desirable. Some policy decisions, including monetary policy, need to be made in "real time" and are based on assessments of current and future economic conditions. However, GDP data, the broadest measure of economic activity, are available only with a long lag and are subject to several rounds of revisions. Thus, government agencies and central banks devote a considerable amount of resources to reconstructing current-quarter GDP from timely economic and financial series before the figure is disseminated (the so-called nowcasts of GDP) to gauge the overall macroeconomic conditions in real time.

To central banks, estimated current-quarter GDP figures are also important because they are often used as relevant inputs for banks' model-based longer-term forecasting exercises. In particular, Cervena and Schneider (2010) and Giannone *et al*. (2010) incorporate the monthly nowcasts into the standard, quarterly DSGE models to improve the accuracy of the forecasts from the models.

Since the seminal work of Stock and Watson (1989, 1991 and 1998), the single factor model on a small number of economic and financial series has been used to construct forecasts or (almost) real time estimates of the GDP of many economies, for example Gerlach and Yiu (2005) for the Hong Kong economy and Curran and Funke (2006) for the Chinese economy (one of the small number of studies on forecasting China's GDP). Although the single factor model approach developed by Stock and Watson provides a modern statistical framework for nowcasting or forecasting economic activity, one drawback is that predictions are estimated from a small number of economic and financial series. This drawback arises because the parameters of the single factor model are estimated by the maximum likelihood method with the model expressed in state space form. The number of parameters increases with the number of series and computational difficulties make it necessary to abandon information on many other economic and financial series even though they are readily available.

The result of the study by Giannone, Reichlin, and Sala (2004) shows that a large data set helps in nowcasting and forecasting. Other researchers, such as Marcellino, Stock, and Watson (2003) and Boivin and Ng (2005), also obtain similar evidence. Since the mid-1990s, China has improved its data provision and dissemination, especially in the coverage of macroeconomic and financial data. Given the availability of these data, it is desirable to have a model that can utilise a large data set to nowcast the Chinese economy.

In this paper, we use the factor model developed by Giannone, Reichlin, and Small (2005) to produce nowcasts (current quarter estimates) of the annual growth rate of China's quarterly GDP from a large panel of macroeconomic and financial series. To form the large panel, we select 189 series from the *CEIC, IFS and Bloomberg* databases. The size and scope of the panel allows us to describe the Chinese economy on a broad basis as it contains data of quite a number of different sectors, such as the external sector and production sector.

In the Giannone-Reichlin-Small approach, they assume that the large panel of series can be represented by a factor structure whereby the dynamics of each series is split in two orthogonal components—one capturing the bulk of cross-sectional co-movements and driven by few common factors, and the other being composed of the poorly cross-correlated idiosyncratic components. The parameters and common factors are estimated by a two-step method utilising asymptotic principal components and regression, and the Kalman filter. The use of the Kalman filter allows for unbalanced panels as different blocks of data are released on different dates.

One technical issue of the factor model is the determination of the appropriate number of common factors. In other words, this is a model selection problem. A standard approach is based on the degree of variance in the data set explained by the first few principal components. Another approach is the method developed by Bai and Ng (2002) to estimate consistently the appropriate number of common factors by utilising some penalty criteria under the assumption of large cross-section and large time dimension. In this paper, we use both approaches to determine the number of factors driving the Chinese economy.

In addition to generating nowcasts, Giannone, Reichlin, and Small (2005) evaluate the marginal impact of different blocks of data in the US economy and find that the survey information carries the most information about concurrent GDP growth. On the other hand, using the same methodology as Giannone, Reichlin, and Small, Aastveit and Trovik (2007) find that equity prices block is the single most important block of data to improve estimates of current quarter GDP in Norway. The other slightly less important blocks of data are labour market and industrial production. Aastveit and Trovik attribute the strong impact from financial data to an ability of the market clearing process to impart information about the real activity in Norway, which is a small open economy depending heavily on natural resources like oil.

This paper also uses the same methodology to investigate the marginal impact of different blocks of data in the Chinese economy and to find out the important blocks of data in the estimation of current quarter GDP in China. We would expect the results to be quite different from those in the previous two studies as the Chinese economy is not as mature as the US and not as open as Norway. Moreover, on the capital accounts, China still has many restrictions on the inward and outward movements of capital.

The paper proceeds as follows. Section 2 describes the (parametric) factor model. The model parameters and the common factors are estimated by a two-step approach involving principal components, regression and the Kalman filter to exploit both factor dynamics and idiosyncratic heteroskedasticity. Section 3 delineates the criteria developed by Bai and Ng (2002) to determine the "correct" number of common factors from a large data panel. Section 4 describes the panel used in the empirical work of the paper and the ordering of its data blocks. Section 5 presents the empirical results of the determination of the "correct" number of common factors, the estimation of the factor model on the large panel and the marginal impacts of different blocks of the data set on the out-of-sample nowcasts of China's GDP. The last section concludes.

2. The Econometric Methodology

The methodology we use in this paper is based on the (parametric) factor model developed by Giannone, Reichlin, and Small (2005). We assume that each series of the large data panel is assumed to have two orthogonal components: the co-movement component, which is a linear combination of a few common factors, and the idiosyncratic component. The dynamics of the common factors are further assumed to be represented by an autoregressive process of order one[1] (AR(1) process) driven by a small number of macroeconomic shocks, drawing an analogy to the real and monetary shocks. Once the parameters of the model are estimated consistently from asymptotic principal components and regression, the Kalman filter is used to generate the more efficient estimates of the common factors and nowcasts are provided by simple regression projections. We describe the details of the model below.

[1] We have tried different AR processes, such as AR2, for the dynamics of the underlying common factors. However, the empirical results are not much different. Thus, for simplicity, we use the AR1 process for all the common factors in this study.

2.1 The Model

We assume that the n macroeconomic and finance indicator series, after suitable transformation and standardisation, have a few common factors and an idiosyncratic component as follows:

$$y_{it} = \gamma_i F_t + \varepsilon_{it} \qquad i = 1,...,n \text{ and } t = 1,...,T \tag{2.1}$$

where $\gamma_i F_t \equiv \zeta_{it}$ and ε_{it} are two orthogonal unobserved stochastic processes. In matrix notation, we have:

$$Y_t = \Gamma F_t + \Xi_t \tag{2.2}$$

where $Y_t = (y_{1t},...,y_{nt})'$, $\Xi_t = (\varepsilon_{1t},...,\varepsilon_{nt})'$ and $\Gamma = (\gamma_1,...,\gamma_n)'$. The $n \times 1$ process ζ_{it} (the common component) is assumed to be a linear combination of a few unobserved common factors F_t that reflect most of the co-movements in the economy. Thus, the common factors are capable of summarising the "fundamental" state of the economy as they are assumed to be present in all the indicator series.

Specifically, the common factors are assumed to follow the following vector autoregressive (VAR) process:

$$F_t = AF_{t-1} + Bu_t \qquad \text{with } u_t \sim WN(0, I_q) \tag{2.3}$$

where B is an $r \times q$ matrix of full rank q, A is an $r \times r$ matrix and all roots of $\det(I_r - Az)$ lie outside the unit circle, and u_t are the macroeconomic stochastic shocks to the common factors. In order to capture the lead and lag relations among series along business cycles, the number of common factors, r, is set to be large relative to the number of macroeconomic shocks, q.[2]

2.2 Identification and Parameters Estimates

For identification, it is assumed that when the number of series in the panel data set is increasing, the common factors remain as the main source of variation and the effects of the idiosyncratic factors will not propagate to the whole data set but only confine to a particular group of series. Therefore, under these

[2] In the empirics of this study, q is set to be two as we assume there are only two macroeconomic shocks, drawing an analogy to the real shocks and monetary shocks.

two assumptions, the common factors can be consistently estimated by asymptotic principal components.[3]

To estimate the parameters of the factor model and the common factors, we follow the two-step procedure developed by Doz *et al*. (2007). The first step of the two-step procedure is to estimate the model parameters from an ordinary least squared (OLS) on the r largest principal components of the panel data. In order to have the principal components, we need to calculate the following sample correlation matrix of the series:

$$S = \frac{1}{T}\sum_{t=1}^{T} Y_t Y_t' \tag{2.4}$$

Then the r largest principal components are extracted from the sample correlation matrix.

Denote by D the $r \times r$ diagonal matrix with diagonal elements given the largest r eigenvalues of S and denote by V the $n \times r$ matrix of the corresponding eigenvectors subject to the normalisation $V'V = I_r$. We approximate the common factors as follows:

$$\tilde{F}_t = V'Y_t \tag{2.5}$$

Given the common factors, $\tilde{F}_t$, we can estimate the factor loadings, Γ, and the covariance matrix of the idiosyncratic components, Π, by regressing the data series on the estimated common factors as follows:

$$\hat{\Gamma} = \sum Y_t \tilde{F}_t' (\tilde{F}_t \tilde{F}_t')^{-1} = V \tag{2.6}$$

and

$$\hat{\Pi} = diag(S - VDV) \tag{2.7}$$

The parameters of Eq. 2.3, A and B, are also estimated by running a VAR on the common factors $\tilde{F}_t$.

3 For details, see Forni, Hallin, Lippi, and Reichlin (2000) and Stock and Watson (2002)

The estimates, $\hat{\Gamma}, \hat{\Pi}, \hat{A}$ and $\hat{B}$, can be shown to be consistent as $n,T \rightarrow \infty$. Under the above two assumptions, this is proven by Forni *et al*. (2000) and, under slightly different assumptions, it is proven by Stock and Watson (2002), Bai and Ng (2003) and Giannone, Reichlin, and Sala (2004).

In the second step, as the estimates, $\hat{\Gamma}, \hat{\Pi}, \hat{A}, \hat{B}$, are available, the Kalman filter is employed to re-estimate the underlying common factors.[4] The re-estimates of the common factors from the Kalman filter are more efficient than using the principal component method because the filter performs the best linear projection utilizing all the information on the present and past observations.

Having obtained the consistent estimates of the common factors conditional on all available information up to t, the nowcast is produced as a simple linear projection, i.e. the quarterly GDP growth is regressed on the common factors using OLS.

Before the above model can be used in empirical applications, three crucial parameters of the model namely q, r, p, have to be decided. As discussed before, p is assumed to be unity as the common factors are represented by an AR(1) process, whereas q is usually assumed to be two because of the assumption of the real shocks and monetary shocks in the macro-economy. Thus, the only parameter that needs to be determined is r, i.e. the number of common factors.

3. Number of Common Factors

A common question of using the factor model proposed by Giannone, Reichlin, and Small (2005) is the specification of the number of common factors, i.e. r. A standard approach is based on the degree of variance in the data set explained by the first few principal components. Usually, the number of factors is selected when the marginal explanation of the next consecutive factor is less than 10 percentage points. Although this approach seems quite practical, it has however been criticised for not having a solid theoretical basis.

Bai and Ng (2002) develop a methodology to estimate consistently the “correct” number of common factors by proposing some penalty criteria under the assumption of large cross-section, n, and large time dimension, T. In their methodology, the common factors in a large panel are estimated by asymptotic principal components. Then the “correct” number of common factors, r, of a factor model is estimated by minimising the following loss function:

[4] The details of the estimation procedure are discussed in Doz, Giannone, and Reichlin (2007)

$$V(k,\hat{F}^k)+kg(N,T) \tag{3.1}$$

$V(k,\hat{F}^k)$ is the sum of squared residuals from time series regressions of the data on the k common factors and $kg(N,T)$ penalises over-fitting. Because of different choices of the penalty function, Bai and Ng propose the following three criteria to determine the "correct" number of common factors:

$$IC_{p1}(k)=\ln(V(f,\hat{F}^k))+k(\frac{N+T}{NT})\ln(\frac{NT}{N+T}) \tag{3.2}$$

$$IC_{p2}(k)=\ln(V(f,\hat{F}^k))+k(\frac{N+T}{NT})\ln C_{NT}^2 \tag{3.3}$$

$$IC_{p3}(k)=\ln(V(f,\hat{F}^k))+k(\frac{\ln C_{NT}^2}{C_{NT}^2}) \tag{3.4}$$

where $C_{NT}=\min\{\sqrt{N},\sqrt{T})$.

The decision rule is to select k to minimise the above three criteria. However, since the criteria are constructed for the factor model in static form only, the "correct" number of common factors determined by the criteria here only indicates an upper bound of the true number of dynamic factors.[5]

4. Data and Ordering of Blocks

We have collected a panel of 189 macroeconomic and financial series from the commercial data providers *CEIC, IFS and Bloomberg* for the factor model to nowcast China's GDP growth. In addition to Chinese data, the panel contains interest rate, equity price and dividend yield series of China's major trading partners, such as the US, Japan and East Asian economies. The span of the panel is from January 1998 to June 2009. However, data from the year of 2008 onwards is reserved for the evaluation of out-of-sample nowcasts. Thus, the dimension of the data panel used in the estimation is only 108×189.

The dataset is described in Table 1 and all series are monthly, except for the GDP deflator for which monthly measures are derived from linear interpolation. In Table 1, Column 1 displays the name of the series and Column 2 lists the type of transformation applied on the series. Most series are transformed to either a 12-month growth rate or a 12-month difference to induce stationarity. Moreover, the two

5 For details, see the Concluding Remarks section of Bai and Ng (2002).

transformations are appropriate for the purpose of producing nowcasts for the annual growth rate of China's quarterly GDP. Details on data transformation are described in the footnote of the table.

The data panel is composed of 16 blocks, namely interest rates, equity market, foreign financials, consumer and retail prices, external one, buildings, government sector, fixed and direct investment, exchange rates, producer prices, world commodity prices, external two, retail sales, economic climate indicators, industrial production and money. The sixteen blocks are released on eight dates throughout the months. The number of series in each block varies from 36 in "foreign financials" to only two in "external one". The released dates are shown in Figure 1 as well.

The blocks are ordered in terms of their release dates. The time lag of all the selected economic and financial series is less than one month. We do not include any series that has a time lag longer than one month since the dissemination of the Chinese GDP figures is quite timely, with a time lag about two to three weeks after the end of the corresponding quarter. Therefore, the GDP or production data of China's major trading partners are not included in the data panel since most of them are released with a time lag longer than one month.

For the financial variables, such as interest rates and equity prices, though in principle they are available on any day during the month, we make an ad hoc assumption for simplicity that the monthly data of these variables are only available on the first day of the next month. This will simplify the block structure in the model albeit possibly understating the importance of the financial variables.

In the next section, we not only have the panel of sixteen blocks but also two more panels, one with fifteen blocks and another one with fourteen blocks. We take this approach in order to investigate the marginal impacts on nowcast performance when some blocks are excluded from the panel.

5. Empirical Results

As discussed in Section 3, we have to select an "appropriate" number of common factors for the factor model. Table 2 shows the total variance explained by up to the first ten principal components. In both the 16-block panel and the 15-block panel (which excludes the block of producer prices), the first five principal components explain about 60%, a non-trivial fraction, of the total variance in the data set. This indicates that most co-movements in the Chinese economy can be captured by a small number of underlying factors. If we use the cut-off for the marginal explanation of the next consecutive factor of less than ten percentage points, the choice of the number of factors for the two panels will be two or three. The marginal explanation of the third factor is around 12 to 11 percentage points, just slightly more than 10 percentage points.

We also use Bai and Ng's criteria to select the "correct" number of common factors in the data panels. The three criteria show different results in Table 3. While the first two criteria IC_{p1} and IC_{p2} show no convergence, the third criterion IC_{p3} has a minimum at $k = 2$ in both data panels. This may be due to the different characteristics of the three penalty functions and the noise in the data panel. The results, however, from Bai and Ng's third criterion are similar to the results looking at the marginal contribution of the principal components to the total variation. It is worth noting that, strictly speaking, both methods do not relate to the predictive power of the common factors to relative GDP growth.

Giannone *et al*. (2005) do not use any criteria to parameterise the factor model but just assume both the number of common factors and the number of stochastic shocks to be two. Similarly, using the principal component rule on Norwegian data, Aastveit and Trovik (2007) find the number of common factors is two. Furthermore, Koop and Potter (2004) investigate alternative methods for choosing factors that incorporate the predictive power and find that the optimal number of common factors chosen is two on average. Since this paper aims to develop a nowcasting model for the Chinese economy and to investigate the marginal impacts of different data blocks on the nowcasting performance, we also parameterise the factor model with two common factors and two types of stochastic shocks: one type for real shocks and another type for monetary shocks.

To examine the nowcasting performance of the factor model, we perform two sets of exercises. In the first one, we study the effect of each release during the quarter on the nowcasting accuracy. This means that we analyse the evolution of the nowcasting in relation to the flow of information throughout the quarter by computing the one-step-ahead, out-of-sample nowcasts that reflect the information content of each category in the data panel. In the second exercise, we provide the evaluation of the overall one-step-ahead out-of-sample nowcasting performance of the model using the information available at the end of the three different months of each quarter. The in-sample estimation of the common factors uses the data in the period from January 1998 to December 2007, while the out-of-sample period is from January 2008 to June 2009, i.e. the last six quarters data in the sample.

To measure the accuracy of the nowcasts from the factor model, we use two typical measures in the forecasting literature: the mean squared forecast error (MSFE) and the mean absolute forecast error (MAFE). The forecast error in both measures is the difference between our nowcast and the subsequent realization of the GDP growth in that quarter. Furthermore, in order to gauge how well the nowcasts perform, we compare the MSFE and MAFE with those of the Random Walk (RW) benchmark in which we use the realised GDP growth of the last quarter as the forecast of the current quarter.

In Figures 2 and 3, we see how the relative nowcasting errors (both MSFE and MAFE) are reduced as new information becomes available throughout the quarter in the 16-block panel. The bars represent the factor model and the dotted line is the RW model benchmark, i.e. naïve nowcast (RW). In the beginning of the month, the “interest rates” category gains about a 20% to 30% reduction against the naïve nowcast. Furthermore, when the release comes to the “foreign financials” category, the factor model obtains a quite large marginal gain. However, when the “producer prices” block is released, the marginal impact becomes worse and the factor model continues to produce worse nowcasts till the last block. Apparently some blocks seem to harm the nowcasting performance of the factor model. In particular, the “producer prices” and “external 2” blocks seem to contain more noise than information in the first two months, but become less noisy in the third month.

In order to search for a better nowcasting factor model, we drop the “producer prices” block from the original data panel to form a 15-block panel. We depict both relative nowcasting errors (MSFE and MAFE) against the RW benchmark in Figures 2 and 3. Without the “producer prices” block, the factor model produces much better nowcasts which consistently outperform the RM model benchmark in all three months throughout the quarter. The “interest rates”, “consumer and retail prices” and “fixed and direct investments” blocks have larger marginal gains than other blocks. On the other hand, the last five blocks in the data panel appear to be quite noisy, so that the relative nowcasting performance becomes worse, though their performance is still below the RW benchmark.

The results from the 16-block and 15-block panels are not quite as expected given that the “interest rates” block turns out to be a very informative block about the real growth of China. Similarly, Aastveit and Trovik (2007) find that equity data of the Norwegian stock market has a strong marginal impact on nowcasting performance. They attribute the result to the ability of the market clearing process to impart information about the real activity in the country in a timely manner since the Norwegian economy is a small open economy depending heavily on natural resources like oil. However, Giannone *et al*. (2005) do not find similar effects of financial data using the same methodology on US data. They find that the marginal impact of “financials” in the US is negligible while surveys are the most important type of data in addition to labour market data. To examine the importance of the “interest rates” block, we exclude it from the 15-block panel to form a 14-block panel. As shown in Figures 2 and 3, the nowcasting performance without the “interest rates” category is much worse than before. The marginal performance outperforms the RW model benchmark only when the data release comes to the “fixed and direct investments” block. We suspect that the importance of the “interest rates” block in the nowcasting performance may be due to the fact that, in the sample period, the Chinese government uses interest rates primarily as a policy instrument for the accommodative monetary policy in addition to credit expansion by the state-owned banks. Furthermore, public and private sectors increase fixed asset investments to provide an impetus to growth. So, both blocks are able to impart information on the real activity in China.

Figure 4 plots China’s quarterly GDP, the nowcasts from the factor model and the nowcasts from the RW model. The nowcasts of the factor model are produced from the 15-block data panel at the end of the second month of each quarter. The use of end of second month information, instead of end of third month which is supposed to contain all the information about the quarter, is due to the time lag of our data and the dissemination date of the Chinese GDP data. As previously mentioned, the Chinese government releases the GDP figures in a relative timely manner, usually around the second to third week of the first month in the next quarter. Thus, we can only use all the information in the second month or partial information in the third month to nowcast the current quarter GDP. So, in this study, we use the data panel up to the end of the second month to produce nowcasts. Quantitatively, the nowcasts from the factor model outperform the RW benchmark by 25% and 21% in terms of MSFE and MAFE respectively.

6. Conclusion

Since preliminary estimates of GDP in China, like many other economies, are released with a time lag and are subject to revisions, it is of interest to explore ways of obtaining pseudo real-time estimates of the state of the real economy.[6] This is particularly relevant to Hong Kong in light of the close link between the two economies. In this paper, we have applied the factor model proposed by Giannone, Reichlin, and Small (2005) on a large Chinese data panel. Overall, the empirical work suggests that a lower dimensional factor model, e.g. with only two common factors, is likely to produce useful nowcasts for China's GDP.

The research reported in this paper is a preliminary investigation on using a large data set to estimate the current state of the Chinese economy.[7] Given the release schedule of China's GDP, using the 15-block data of the second month in each quarter, will produce most practical nowcasts, outperforming the RW benchmark in the out-of-sample period from 2008Q1 to 2009Q2. We believe that these nowcasts could be of use for policy makers if they need to have a timely and reasonably good current estimate of the real activity in China.

About the information content of the 15 categories in the data panel, from the marginal impact analysis of each category, we find that the most important categories are the "interest rates, "fixed and direct investments" and "consumer and retail prices" blocks. The importance of the "interest rates" block may be due to the fact that the Chinese government over the sample period consistently uses interest rates as an instrument for monetary policy to stimulate economic growth, in addition to credit expansion and fiscal stimuli. Using the marginal impact analysis further, we should be able to find the so-called "optimal" panel

[6] Although the issue of data revisions does not apply to financial variables, most macro variables in our data panel are subject to revision. So, the nowcasts in this study are not the true real-time nowcasts.

[7] Nowcasting is a rather new method to continuously apply updated high-frequency data to forecast low-frequency variables, such as in our study using monthly data to forecast quarterly GDP figures. Two questions have been raised concerning the ways to deal with an unbalanced panel data set and to tackle the difference in frequency between the forecasting and predictive variables. The framework discussed in this paper, is suitable to deal with the two difficulties because of the use of the Kalman filter. The detailed description of the general methodology can be found in Aastveit and Trovik (2007) and Rossiter (2010).

on which we could produce nowcasts with the smallest MSFE and MAFE. However, we leave this to future research.

The nowcasts of the quarterly GDP growth are projected from the regression on the two underlying common factors only. An alternative specification is to include lagged values of quarterly GDP growth as a dependent variable. This might improve the model's ability to capture dynamics in quarterly GDP growth. However, it would make the marginal impact of new information less transparent. As the aim of this study is on prediction (nowcasting), not hypotheses testing, we do not include lagged GDP for the sake of the investigation on the information content of the constructed data panel.[8]

Finally, this study could be extended in one more direction. The attractive aspect of the proposed factor model framework is that it can be used to generate forecasts with associated confidence bands, and it therefore seems desirable to explore the forecasting ability, particularly in the multi-period-ahead, of the model.

[8] In particular, the marginal impact of the blocks following a new release of lagged GDP becomes hazardous to interpret.

References

Aastveit, K. and T. Trovik (2007), “Nowcasting Norwegian GDP: The Role of Asset Prices in a Small Open Economy,” Working Paper 9/2007, Norges Bank.

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

Bai, J. and S. Ng (2007), “Determining the Number of Primitive Shocks in Factor Model,” *Journal of Business and Economic Statistics*, 25: 52-60.

Boivin, J. and S. Ng (2006), “Are More Data Always Better for Factor Analysis?” *Journal of Econometrics*, 132: 169-94.

Cervena, M. and M. Schneider (2010), “Short-Term Forecasting GDP with a DSGE Model Augmented by Monthly Indicators,” Working Paper 163, Oesterreichische NationalBank

Curran, D. and M. Funke (2006), “Taking the Temperature: Forecasting GDP Growth for Mainland China,” BOFIT Discussion Papers 2006-06, Institute for Economies in Transition, Bank of Finland.

Doz, C., D. Giannone and L. Reichlin (2007), “A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering,” Discussion Paper No.6043, Centre for Economic and Policy Research.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), “The Generalised Dynamic Factor Model: Identification and Estimation,” *Review of Economics and Statistics,* 82: 540-54.

Gerlach, S. and M. Yiu (2005), “A Dynamic Factor Model of Economic Activity in Hong Kong,” *Pacific Economic Review*, 10: 279-92.

Giannone, D., L. Reichlin and L. Sala (2004), “Monetary Policy in Real Time,” in M. Gertler and K. Rogoff, eds., *NBER Macroeconomic Annual 2004*, Cambridge MA: MIT Press: 161-200.

Giannone, D., L. Reichlin and D. Small (2005), “Nowcasting GDP and Inflation: the Real Time Information Content of Macroeconomic Data Release,” Finance and Economic Discussion Series 2005-42, Board of Governors of the Federal Reserve System (U.S.).

Giannone, D., F. Monti and L. Reichlin (2009), “Incorporating Conjunctural Analysis in Structural Models,” in W. Volker, ed., *The Science and Practice of Monetary Policy Today,* Springer: 41-57.

Koop, G. and S. Potter (2004), “Forecasting in Dynamic Factor Models Using Bayesian Model Average,” *Econometrics Journal*, 7: 550-65.

Marcellino, M., J. Stock and M. Watson (2003), “Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-wide Information,” *European Economic Review*, 47: 1-8.

Rossiter, J. (2010), “Nowcasting the Global Economy,” Discussion Paper No 2010-12, Bank of Canada.

Stock, J. and M. Watson (1989), “New Indexes of Coincident and Leading Economic Indicators,” in O. Blanchard and S. Fischer, eds., *NBER Macroeconomic Annual 1989,* Cambridge MA: MIT Press: 315-94.

Stock, J. and M. Watson (1991), “A Probability Model of the Coincident Economic Indicators,” in K. Lahiri and G. Moore, eds., *Leading Economic Indicators New Approaches and Forecasting Records*, New York: Cambridge University Press: 63-85.

Stock, J. and M. Watson (1998), “Business Cycle Fluctuations in US Macroeconomic Times Series,” NBER Working Paper No.6528, Cambridge MA: National Bureau of Economic Research.

Stock, J. and M. Watson (2002), “Forecasting Using Principal Components from a Large Number of Predictors,” *Journal of American Statistical Association*, 97: 1167-79.

Table 1. Data Series

Name of Series	Transformation
National Interbank Offered Rate: Weighted Avg: NIBFC: 7 Days	2
National Interbank Offered Rate: Weighted Avg: NIBFC: 30 Days	2
National Interbank Offered Rate: Weighted Avg: NIBFC: 60 Days	2
National Interbank Offered Rate: Weighted Avg: NIBFC: 90 Days	2
Central Bank Base Interest Rate: Required Reserve	2
CN: Central Bank Benchmark Interest Rate: Reserve Requirement: Excess	2
Central Bank Base Interest Rate: Annual	2
Base Lending Rate: Working Capital: 1 Year	2
Savings Deposits Rate	2
Time Deposits Rate: 1 Year	2
Index: Shanghai Stock Exchange: Composite	1
Index: Shanghai Stock Exchange: Industrial	1
Index: Shanghai Stock Exchange: Utilities	1
Index: Shenzhen Stock Exchange: Composite	1
PE Ratio: Shanghai Stock Ex: All Shares: Wgt Avg by Issued Volume	2
Market Capitalization: Shanghai Stock Exchange: Stocks	1
Dividend Yield: Shanghai Stock Exchange: Composite	2
Dividend Yield: Shanghai Stock Exchange: Industrial	2
Dividend Yield: Shanghai Stock Exchange: Utilities	2
LIBOR: BBA: USD: 3 Months : Frequency Transform	2
LIBOR: BBA: Euro: 3 Months : Frequency Transform	2
HIBOR: Hong Kong Bankers Association: 3 Months : Frequency Transform	2
DOW JONES INDUSTRIAL AVERAGE	1
FTSE 100 INDEX	1
DAX INDEX	1
CAC 40 INDEX	1
FTSE MIB Index	1
OMX STOCKHOLM 30 INDEX	1
The Bombay Stock Exchange Sensitive Index (Sensex)	1
Nikkei-225 Stock Average	1
The KOSPI Index	1
Index: SGX Strait Times	1
Kuala Lumpur Stock Exchange Composite Index	1
The Bangkok SET Index	1
Jakarta Stock Price Index	1
The Philippine Stock Exchange PSEi Index	1
TWSE Index	1
Hang Seng Index	1
The Australian All Ordinaries Index	1
The New Zealand All Ordinaries Index	1
Dividend Yield: DOW JONES INDUSTRIAL AVERAGE	2
Dividend Yield: FTSE 100 INDEX	2
Dividend Yield: DAX INDEX	2
Dividend Yield: CAC 40 INDEX	2
Dividend Yield: OMX STOCKHOLM 30 INDEX	2
Dividend Yield: Nikkei-225 Stock Average	2
Dividend Yield: The KOSPI Index	2
Dividend Yield: Kuala Lumpur Stock Exchange Composite Index	2
Dividend Yield: The Bangkok SET Index	2
Dividend Yield: Jakarta Stock Price Index	2
Dividend Yield: The Philippine Stock Exchange PSEi Index	2
Dividend Yield: TWSE Index	2
Dividend Yield: Hang Seng Index	2
Dividend Yield: The Australian All Ordinaries Index	2

Name of Series	Transformation
Dividend Yield: The New Zealand All Ordinaries Index	2
CPI: Food (sa)	1
CPI: Clothing (sa)	1
CPI: Household Facilities (sa)	1
CPI: Medical Care (sa)	1
CPI: Traffic, Communications (sa)	1
CPI: Recreation, Education, Cultural (sa)	1
CPI: Residence (sa)	1
CPI: Aggregate (sa)	1
Retail Price Index, Aggregate _sa	1
GDP Deflator, sa	1
Exports fob	1
Imports cif	1
Commodity Building Sold: Total	1
CN: Commodity Building Sold: Residential: Total	1
CN: Commodity Bldg Selling Price	1
CN: Commodity Bldg Selling Price: Residential	1
CN: Floor Space Started: Total	1
CN: Floor Space under Construction: Commodity Bldg (CB)	1
CN: Floor Space Completed: Commodity Bldg (CB)	1
CN: Floor Space Completed: CB: Residential: Total	1
Government Surplus or Deficit	1
Government Expenditure	1
Fixed Asset Investment: Levels, Total (revised) S.A.	1
Real Estate Investment, Total (SA)	1
Foreign Direct Investment Utilized	1
FDI: Utilized: Joint Ventures	1
FDI: Utilized: Cooperative Ventures	1
FDI: Utilized: Foreign Enterprises	1
Spot Exchange Rate: Period Avg: SAFE: RMB to US Dollar	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Hong Kong Dollar	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Japanese Yen	1
Spot Exchange Rate: Period Avg: SAFE: RMB to British Pound	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Canadian Dollar	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Swiss Franc	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Swedish Krone	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Norway Krone	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Danish Krone	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Singapore Dollar	1
Spot Exchange Rate: Period Avg: SAFE: RMB to Australian Dollar	1
Nominal effective exchange rate	1
Real effective exchange rate, sa	1
Producer Price Index: Industrial Products (yoy%)	3
Producer Price Index: Industrial Products: Light Industry (yoy%)	3
Producer Price Index: Industrial Products: Heavy Industry (yoy%)	3
Producer Price Index: Industrial Products: Producer Goods (yoy%)	3
Producer Price Index: Industrial Products: Producer Goods: Excavation (yoy%)	3
Producer Price Index: Industrial Products: Producer Goods: Raw Material (yoy%)	3
Producer Price Index: Industrial Products: Producer Goods: Manufacturing (yoy%)	3
Producer Price Index: Industrial Products: Consumer Goods (yoy%)	3
Producer Price Index: Industrial Products: Consumer Goods: Durable (yoy%)	3
Producer Price Index, Consumer Goods: Food (yoy%)	3
Producer Price Index, Consumer Goods: Clothing, (yoy%)	3
Producer Price Index, Consumer Goods: Daily Use Articles (yoy%)	3
Purchasing Price Index, Raw Materials (yoy%)	3
Index of Market Prices: All Primary Commodities Index	1

Name of Series	Transformation
Index of Market Prices: Non Fuel Primary Commodities Index	1
Index of Market Prices: Food	1
Index of Market Prices: Beverages	1
Index of Market Prices: Agricultural Raw Materials	1
Index of Market Prices: Metals	1
Index of Market Prices: Energy Index	1
Index of Market Prices: Petroleum, average crude price	1
Index of Market Prices: Sugar Caribbean (N.Y.) Free Market	1
Index of Market Prices: Uranium Index	1
Exports: Primary Products	1
Exports: Manufactures	1
Exports: Animal and Vegetable Oils, Fats and Waxes (AVFW)	1
Exports: Beverages and Tobacco (BT)	1
Exports: Chemicals and Related Products (CRP)	1
Exports: Commodities Not Classified Elsewhere	1
Exports: Crude Materials, Inedible, Except Fuels (CM)	1
Exports: Food and Live Animals Chiefly For Food (FLA)	1
Exports: Machinery and Transport Equipment (MTE)	1
Exports: Manufactured Goods Chiefly by Materials (MG)	1
Exports: Mineral Fuels, Lubricants and Related Materials (MFLM)	1
Exports: Miscellaneous Manufactured Articles (MMA)	1
Imports: Primary Products	1
Imports: Manufactures	1
Imports: Animal and Vegetable Oils, Fats and Waxes (AVFW)	1
Imports: Beverages and Tobacco (BT)	1
Imports: Chemicals and Related Products (CRP)	1
Imports: Commodities Not Classified Elsewhere	1
Imports: Crude Materials, Inedible , Except Fuels (CM)	1
Imports: Food and Live Animals Chiefly For Food (FLA)	1
Imports: Machinery and Transport Equipment (MTE)	1
Imports: Manufactured Goods Chiefly by Materials (MG)	1
Imports: Mineral Fuels, Lubricants and Related Materials (MFLM)	1
Imports: Miscellaneous Manufactured Articles (MMA)	1
Retail Sales of Consumer Goods, Total (revised) S.A.	1
Retail Sales of Consumer Goods: Wholesale and Retail Trade	1
Retail Sales of Consumer Goods: Catering Trade	1
Retail Sales of Consumer Goods: Other	1
CN: Leading Index	1
CN: Coincident Index	1
CN: Business Cycle Signal	1
Value Added of Industry, TOTAL (sa, 2004 price)	1
Industrial Production: Bicycles	1
Industrial Production: Sewing Machines	1
Industrial Production: Camera	1
Industrial Production: Television Sets: Colour	1
Industrial Production: Hi Fi	1
Industrial Production: Household Washing Machines	1
Industrial Production: Household Refrigerator	1
Industrial Production: Air Conditioner	1
Industrial Production: Cloth	1
Industrial Production: Total Energy Production	1
Industrial Production: Processed Crude Oil	1
Industrial Production: Gasoline	1
Industrial Production: Kerosene	1
Industrial Production: Diesel Oil	1
Industrial Production: Lubricant Oil	1

Name of Series	Transformation
Industrial Production: Fuel Oil	1
Industrial Production: Power Generated	1
Industrial Production: Steel	1
Industrial Production: Steel Products	1
Industrial Production: Iron Alloy	1
Industrial Production: Cement	1
Industrial Production: Automobiles	1
Industrial Production: Automobiles: Buses and Coaches	1
Industrial Production: Automobiles: Cars	1
Industrial Production: Automobiles: Loading Vehicles	1
Industrial Production: Computer	1
Industrial Production: Micro Computer	1
Industrial Production: Semiconductor Integrated Circuit	1
Industrial Production: Motor Cycles	1
Financial Institution Loans	1
Financial Institution Deposits	1
Quasi Money, Seasonally Adjusted	1
Savings Deposits, Seasonally Adjusted	1
Money Supply M0	1
Money Supply M1	1
Money Supply M2	1
CN: Required Reserve Ratio	2
Foreign Reserves	1

* Type of transformation: 1, 2, 3 stand for 12-month growth rate, 12-month difference and no transformation respectively.

Table 2. Percentage of Total Variance Explained by the First *r* Static Principal Components based on the Sample from 1999 to 2007

No. of Common Factors	1	2	3	4	5	…	10
16-block dataset	0.21	0.34	0.45	0.54	0.61		0.76
15-block dataset	0.18	0.32	0.44	0.52	0.59		0.75

Table 3. Bai and Ng's Criteria

A: Sixteen Blocks Dataset

No. of Common Factors	$IC_{p1}(k)$	$IC_{p2}(k)$	$IC_{p3}(k)$
1	8.5585	8.5650	8.5403
2	8.5643	8.5774	**8.5279**
3	8.5898	8.6095	8.5352
4	8.6186	8.6449	8.5458
5	8.6573	8.6901	8.5663
6	8.7029	8.7424	8.5937
7	8.7526	8.7986	8.6252
8	8.8039	8.8565	8.6583
9	8.8575	8.9167	8.6937
10	8.9130	8.9787	8.7310
11	8.9704	9.0428	8.7703
12	9.0288	9.1077	8.8104
13	8.5585	8.5650	8.5403
14	8.5643	8.5774	8.5279
15	8.5898	8.6095	8.5352

B: Fifteen Blocks Dataset

No. of Common Factors	$IC_{p1}(k)$	$IC_{p2}(k)$	$IC_{p3}(k)$
1	8.6304	8.6376	8.6110
2	8.6375	8.6518	**8.5986**
3	8.6643	8.6858	8.6060
4	8.6943	8.7229	8.6165
5	8.7343	8.7701	8.6370
6	8.7812	8.8241	8.6645
7	8.8322	8.8822	8.6960
8	8.8847	8.9419	8.7291
9	8.9397	9.0040	8.7646
10	8.9964	9.0679	8.8019
11	9.0551	9.1338	8.8412
12	9.1147	9.2005	8.8813
13	8.6304	8.6376	8.6110
14	8.6375	8.6518	8.5986
15	8.6643	8.6858	8.6060

Figure 1. Structure of Data Release During a Quarter

		Quarter	2007Q4																2008Q1															
		Month	2007m11								2007m12								2008m1								2008m2							
Block	Number of	Release Date	1	12	17	19	21	24	25	30	1	12	17	19	21	24	25	30	1	12	17	19	21	24	25	30	1	12	17	19	21	24	25	30
Number	Series	Block Name																																
1	10	Interest Rates	7m10								7m11								7m12															
2	9	Stock Markets	7m10								7m11								7m12															
3	36	Foreign Financial	7m10								7m11								7m12															
4	10	Price Index		7m10								7m11								7m12														
5	2	External 1		7m10								7m11								7m12														
6	8	Building		7m10								7m11								7m12														
7	2	Government		7m10								7m11								7m12														
8	6	Investment			7m10								7m11								7m12													
9	13	Exchange Rates				7m10								7m11								7m12												
10	13	Producer Price Index				7m10								7m11								7m12												
11	10	World Commodity Prices					7m10								7m11								7m12											
12	24	External 2						7m10								7m11								7m12										
13	4	Retail Sales						7m10								7m11								7m12										
14	3	Economic Climate Indicator							7m10								7m11								7m12									
15	30	Industrial Production							7m10								7m11								7m12									
16	9	Monetary								7m10								7m11								7m12								
	189																																	
		Release of GDP:	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	7q4	8q1	8q1	8q1	8q1	8q1	8q1	8q1	8q1	8q1	8q1	8q1
Legend																																		
t-1																																		
t-2																																		
t-3																																		

Figure 2. Marginal Change in MSFE of Data Blocks in the 16-Block, 15-Block and 14-Block Panels

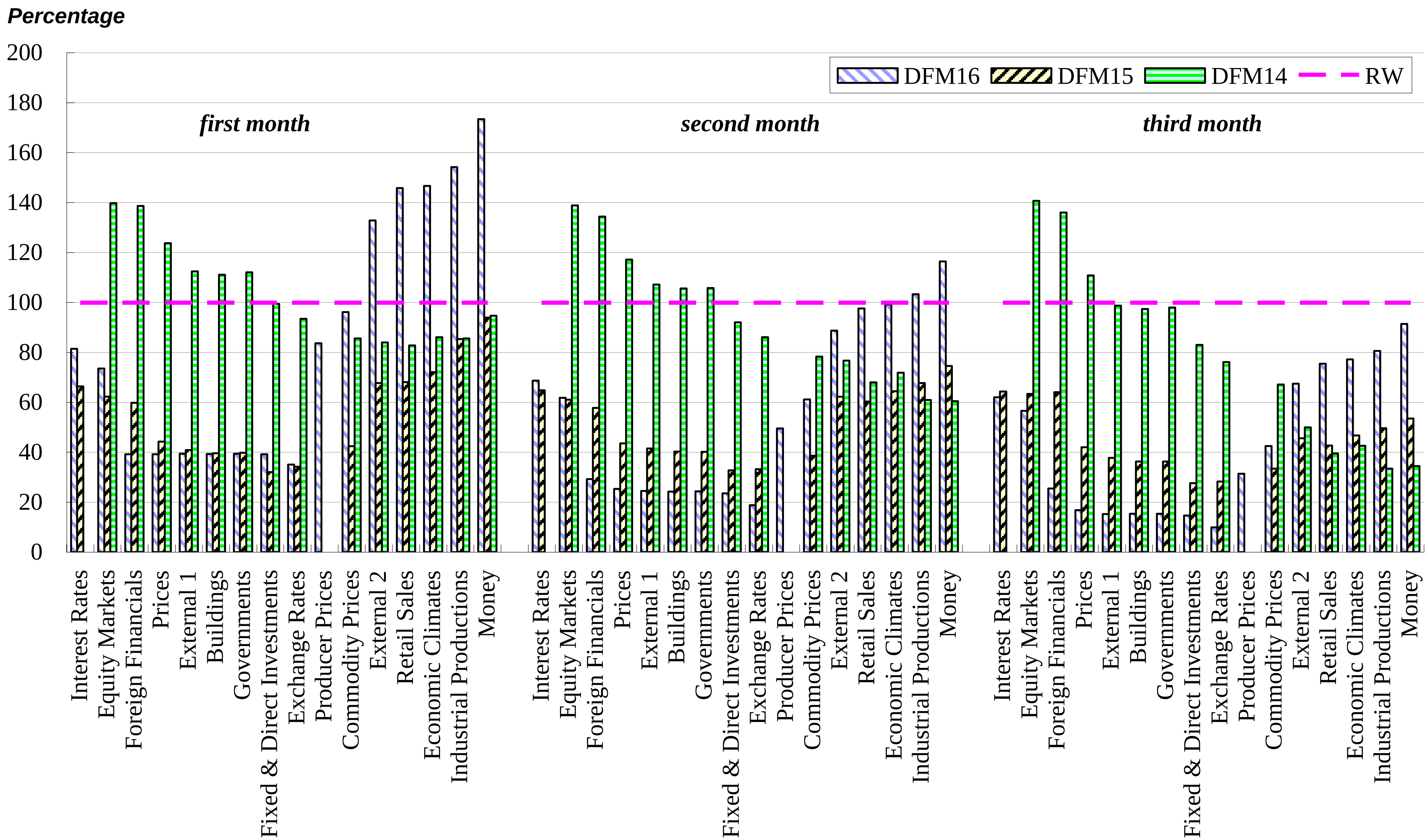

Figure 3. Marginal Change in MAFE of Data Blocks in the 16-Block, 15-Block and 14-Block Panels

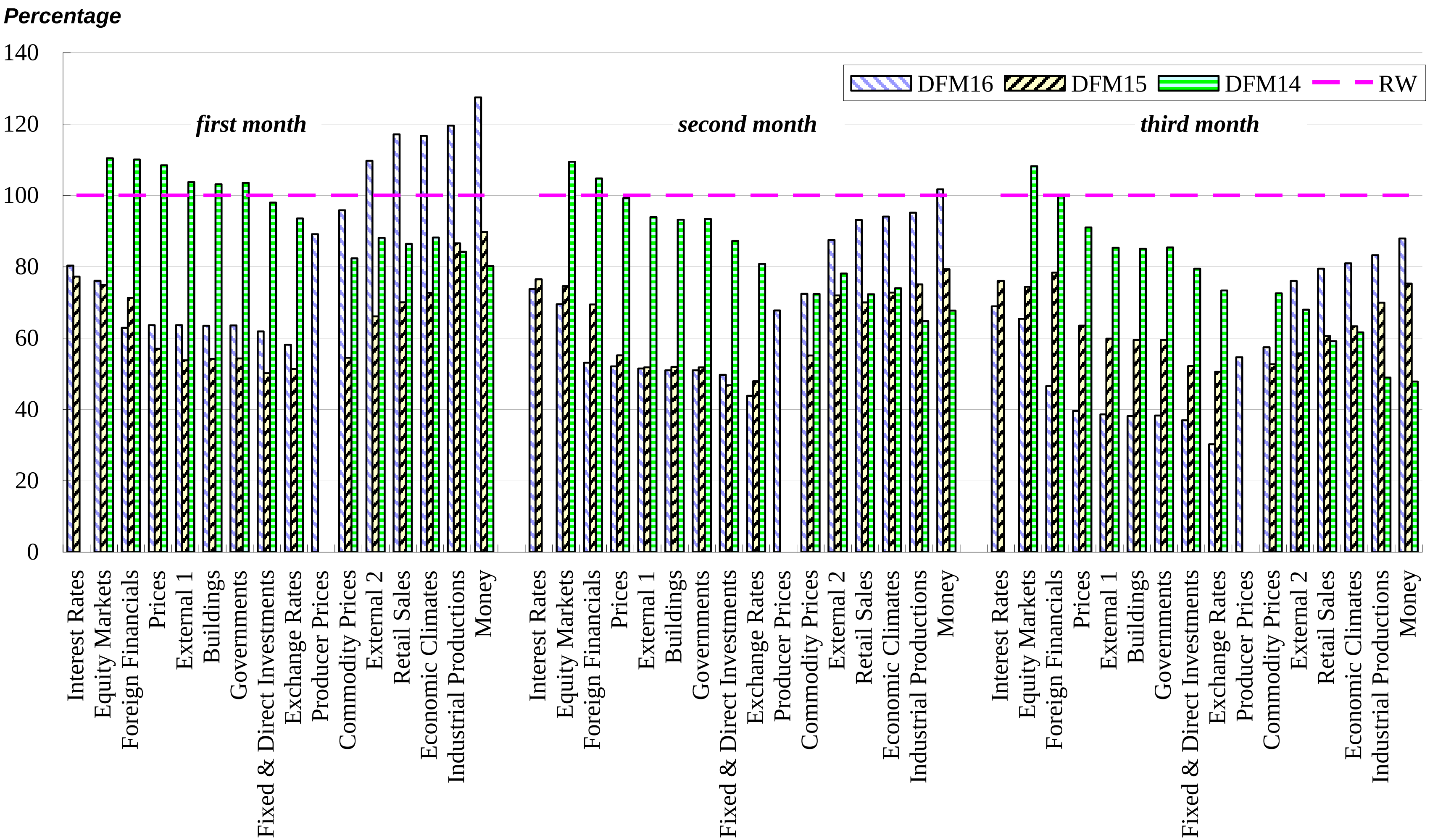

Figure 4. Out-of-Sample Nowcasts on the 15-Block Panel from Factor Model, RW Model and China's GDP Growth