

Mikosch, Heiner; Zhang, Ying

Working Paper

Forecasting Chinese GDP growth with mixed frequency data: Which indicators to look at?

KOF Working Papers, No. 359

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Mikosch, Heiner; Zhang, Ying (2014) : Forecasting Chinese GDP growth with mixed frequency data: Which indicators to look at?, KOF Working Papers, No. 359, ETH Zurich, KOF Swiss Economic Institute, Zurich,
<https://doi.org/10.3929/ethz-a-010184765>

This Version is available at:

<https://hdl.handle.net/10419/102958>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

KOF Working Papers

Forecasting Chinese GDP Growth with Mixed Frequency Data:
Which Indicators to Look at?

Heiner Mikosch and Ying Zhang



KOF

KOF Swiss Economic Institute
ETH Zurich
LEE G 116
Leonhardstrasse 21
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Forecasting Chinese GDP Growth with Mixed Frequency Data: Which Indicators to Look at?*

Heiner Mikosch[†], Ying Zhang[‡]

First version: December 01, 2013

This version: July 10, 2014

Abstract

Building on a mixed data sampling (MIDAS) model we evaluate the predictive power of a variety of monthly macroeconomic indicators for forecasting quarterly Chinese GDP growth. We iterate the evaluation over forecast horizons from 370 days to 1 day prior to GDP release and track the release days of the indicators so as to only use information which is actually available at the respective day of forecast. This procedure allows us to detect how useful a specific indicator is at a specific forecast horizon relative to other indicators. Despite being published with an (additional) lag of one month the OECD leading indicator outperforms the leading indicators published by the Conference Board and by Goldman Sachs. Albeit being smaller in terms of market volume, the Shenzhen Composite Stock Exchange Index outperforms the Shanghai Composite Stock Exchange Index and several Hong Kong Stock Exchange indices. Consumer price inflation is especially valuable at forecast horizons of 11 to 7 months. The reserve requirement ratio for small banks proves to be a robust predictor at forecast horizons of 9 to 5 months, whereas the big banks reserve requirement ratio and the prime lending rate have lost their leading properties since 2009. Industrial production can be quite valuable for now- or even forecasting, but only if it is released shortly after the end of a month. Neither monthly retail sales, investment, trade, electricity usage, freight traffic nor the manufacturing purchasing managers' index of the Chinese National Bureau of Statistics help much for now- or forecasting. Our results might be relevant for experts who need to know which indicator releases are really valuable for predicting quarterly Chinese GDP growth, and which indicator releases have less predictive content.

JEL classifications: C53, E27

Keywords: Forecasting, mixed frequency data, MIDAS, China, GDP growth

*We thank Eric Ghysels for providing us with his MIDAS codes. We further thank Dirk Drechsel, Stefan Neuwirth, Samad Sarferaz, Boriss Siliverstovs and Jan-Egbert Sturm for helpful suggestions.

[†]Corresponding author. KOF Swiss Economic Institute, ETH Zurich, Weinbergstrasse 35 WEH, CH-8092 Zurich, Switzerland, +41 44 632 42 33, e-mail: mikosch@kof.ethz.ch.

[‡]Partners Group AG, Zugerstrasse 57, CH-6341 Baar, Switzerland, e-mail: ying.zhang@partnersgroup.com.

1 Introduction

China is today one of the major economies in the world. It ranks second in terms of gross domestic product and it contributes about 36 percent to world growth.¹ Due to the size of its economy, Chinese business cycle fluctuations potentially affect economies worldwide. Forecasting Chinese economic activity seems thus a worthwhile task both from a Chinese and an international perspective. In the press we often read projection statements of the following sort:

Growth is set to pick up in most large economies during the first half of 2014 [...] according to leading indicators published Monday. The leading indicator for [...] China also rose, indicating [China] will experience a pick up in growth over coming months. (The Wall Street Journal, January 13, 2014)

China's electricity consumption, an indicator of economic activity, rose 9.5 percent year-on-year in October, official data showed on Thursday. [The] slower electricity consumption in October is in line with market expectation of lower economic growth in the final quarter of the year. (China Daily, November 15, 2013)

A professional forecaster would approach the task of forecasting economic growth quite differently than a journalist. Rather than observing just one or two indicators, he or she would collect a large number of different time series, employ several alternative forecasting models and finally condense everything into a forecast, may it be a point or density forecast or a full-fledged forecast scenario. Importantly however, when the day comes to present and justify the forecast to the public, the peers, the superiors etc., the forecaster cannot simply state, "I used a large amount of data and various different models and this is what came out". Rather, he or she is required to build a coherent story around the forecast and cite, just like a journalist, selected indicators whose recent development supports the forecast. Often however, there exist other indicators whose recent development does not quite support or even contradict the forecast. (Reality is often ambiguous.) The forecaster then has to argue why the latter indicators are either not relevant in the present context or not relevant at all. Here lies the main contribution of our paper. We evaluate and compare the predictive power of a variety of monthly macroeconomic indicators which are generally considered to have predictive content and/or which get public attention when it comes to forecasting quarterly Chinese GDP growth.² We iterate the evaluation for a whole range of forecast horizons, namely 370 days to 1 day prior to GDP release. We track the release days of the indicators so as to only use data which is actually available at the respective day of forecast (pseudo-real time setup). This procedure allows us to detect how useful a specific macroeconomic indicator is for forecasting quarterly Chinese GDP growth at a specific horizon as compared to

¹Average over 2011–2013.

²Mehrotra and Pääkkönen (2011, p. 411) state that "an avenue of further research could be an evaluation of the forecasting ability of the various leading indicators for China. Such analysis would have important implications for economic policymaking." This is exactly what we do.

alternative indicators. Our results might be relevant for experts who need to know which new data release is really valuable, and which new data release has less predictive content or contains just old information. To our best knowledge, there exists no previous study that compares the *individual* predictive power of a broad set of macroeconomic indicators for forecasting Chinese GDP growth. Rather, the literature has concentrated on analyzing the *joint* predictive power of indicators both via factor analysis or forecast averaging (see the literature cited below). To be sure, we also complement our indicator-wise analysis by evaluating the joint predictive power of a variety of indicators via factor analysis. Our day-specific pseudo-real time setup allows us to deal with ragged edge issues and to study how forecast errors evolve as ever new data get released over time.

This are the main findings of our paper: It strongly depends on the exact forecast horizon (= the number of days before GDP release) whether and to which extend a monthly indicator provides new information that helps improve quarterly (year-over-year) GDP growth forecasts. Notwithstanding this general lesson, we find substantial differences between individual indicators. First, the OECD leading indicator (trend restored version in year-over-year growth rates) adds substantial predictive power. Its value is highest at forecast horizons of 5 to 4 months, where it reduces forecast errors by 31% to 37% as compared to an ADL benchmark model. Importantly, the OECD leading indicator outperforms its counterparts published by the Conference Board (CB) and by the China Economic Monitoring and Analysis Center (CEMAC) in cooperation with Goldman Sachs (GS). A surprise additional finding is that the OECD leading indicator and the CEMAC-GS leading indicator have even better – or at least not worse – *nowcasting* properties than several coincident indicators published by OECD, CEMAC-GS or CB. Second, the Shenzhen Stock Exchange Composite Index and the Shanghai Stock Exchange Composite Index are quite valuable predictors; and despite covering a lower share of the Chinese economy, the former index tends to outperform the latter. In contrast, Hong Kong stock exchange indices are only of very little help. Third, except for the nowcast range consumer price inflation adds substantial predictive power for forecasting quarterly GDP growth. Its value is highest 11 to 6 months prior to GDP release with forecast error reductions of 26% to 36% compared to an ADL benchmark model. In contrast, producer price inflation does not seem to help much. Fourth, while China’s key short-term interest rate, the so called prime lending rate, as well as the reserve requirement ratio for big banks used to be quite valuable for forecasting in the past, their leading properties have vanished since 2009. In contrast, the reserve requirement ratio for small banks, which is uncoupled from the big banks reserve requirement ratio since 2009 only, proves to still be a robust predictor at forecast horizons of 9 to 5 months. Fifth, money supply M2 adds substantial predictive power, except for the nowcast range and for days directly after lagged GDP publications which render the predictive information in earlier M2 releases irrelevant. In contrast, money supply M0 and M1 turn out to be not very helpful. Sixth, a common factor series extracted from all aforementioned monetary variables turns out to be a robust and very valuable predictor at forecast horizons of 10 to 3.5 months. Seventh, industrial production (IP) growth can be quite valuable for now- or even forecasting

quarterly GDP growth. However, this is only the case if the *third* monthly IP growth observation of a quarter (hence, the value of March, June, September or December) is available before GDP growth of that quarter. In contrast, when GDP growth of a quarter is released before the third monthly IP growth observation of that quarter, it renders the predictive content in the latter observation irrelevant and depletes both the now- and the forecasting power of the IP growth series. Eighth, despite a remarkable 2-quarters lead for the GDP growth turning points in Q2 2009 and Q2 2010 the manufacturing purchasing managers' index of the Chinese National Bureau of Statistics is of little value in now- or forecasting quarterly GDP growth. Ninth, monthly retail sales, investment, exports and imports do not help nowcasting, neither individually nor jointly (common factor analysis). Tenth, while having received quite some attention lately, electricity usage and freight traffic turn out to be poor predictors. This notwithstanding the series might still be useful as *alternative* indicators for economic activity in China.

We forecast a low frequency variable (quarterly GDP growth) with higher frequency data (monthly macroeconomic indicators). To deal with this mixed frequency set up we follow three alternative approaches: mixed data sampling (MIDAS) proposed by Ghysels and co-authors ([Ghysels et al., 2007](#), [Andreou et al., 2010](#), e.g.), unrestricted mixed data sampling (U-MIDAS) proposed [Foroni et al. \(2012\)](#) and traditional bridge equations ([Golinelli and Parigi, 2007](#), [Diron, 2008](#), e.g.). MIDAS allows to parsimoniously exploit the information content of multiple high frequency data for predicting low frequency data and proves to be an efficient approach for forecasting and nowcasting ([Clements and Galvão, 2008, 2009](#), [Armesto et al., 2010](#), e.g.). U-MIDAS is especially efficient when the memory of time series tends to be short.³

There is a growing body of literature on forecasting Chinese GDP. [Klein and Mak \(2005\)](#) and [Mak \(2009\)](#) forecast GDP growth one quarter ahead using principal components analysis with a large set of economic variables. [Curran and Funke \(2006\)](#) construct a composite leading indicator of economic activity from a dynamic factor model with three time series: exports, a real estate climate index and the Shanghai Stock Exchange Composite Index. They show that the indicator performs well in forecasting GDP growth one quarter ahead. [Mehrotra and Pääkkönen \(2011\)](#) use a static factor model to produce a coincident indicator from a large number of economic indicators. The authors find that the coincident indicator closely matches the GDP dynamics with only very short periods of discrepancies. [Yiu and Chow \(2011\)](#) adopt the large scale factor model and Kalman filtering approach proposed by [Giannone et al. \(2008\)](#) to nowcast GDP growth. The interest rate data block turns out to have the highest predictive content, next to the consumer and retail prices data block and the fixed and direct investment data block. [Maier \(2011\)](#)

³An alternative for forecasting with mixed frequency data is a state space approach ([Mariano and Murasawa, 2003, 2010](#), [Giannone et al., 2008](#), e.g.). [Wohlrabe \(2009\)](#), [Banbura et al. \(2011\)](#) and [Foroni and Marcellino \(2013\)](#), e.g., provide overviews on forecasting with mixed frequency data. Comparative forecasting studies found neither (U-)MIDAS nor state space mixed frequency models to be generally superior ([Kuzin et al., 2011](#), [Bai et al., 2013](#), e.g.).

conducts monthly nowcasts and one quarter ahead forecasts of GDP growth by aggregating monthly observed indicators to quarterly values via bridge equations. The author evaluates three different forecasting methods: a model containing the China economic activity indicator of the Hong Kong Monetary Authority which condenses several indicators (Liu *et al.*, 2007), forecast averaging of various univariate indicator models following Stock and Watson (2004) and a static factor model using again various indicators. The latter method yields the strongest improvement over an autoregressive benchmark model. Pooling across all three methods improves the forecast accuracy even further. Franses and Mees (2013) show that the data generating process of quarterly cumulated nominal GDP can be approximated by the following simple rule: The year-over-year growth rate of the aforementioned series follows a random walk; shocks occur only in the first quarter of a year, whereas the error variance is not statistically different from zero in all other quarters of the year. Orlik (2012) provides an in-depth overview of China's most important economic statistics.

2 Econometric model and forecasting procedure

Let y_t denote a quarterly variable (= a variable which is observed once every quarter) and let $x_{t-j*1/3}$ denote a monthly variable (= a variable which is observed once every month). The autoregressive distributed lag mixed data sampling (ADL-MIDAS) forecasting model is given by

$$y_{t+q} = \alpha_0 + \sum_{i=1}^I \alpha_i y_{t-i} + \beta \sum_{j=m}^{J+m-1} \omega_{j-m+1} x_{t+q-j*1/3} + \epsilon_{t+q} \quad (1)$$

$\forall t = 1, \dots, T$, where $q \in \{0, 1, \dots, Q\}$, $m \in \{0, 1, \dots, M\}$, $\alpha_0, \alpha_1, \dots, \alpha_I$ and β are unknown parameters, ω_{j-m+1} denotes an unknown weight that is a function of the unknown parameter vector Θ and ϵ_{t+q} is an error term.⁴

We model the weight ω_j either as [1] a beta probability density function, [2] an exponential Almon lag polynomial or [3] a non-exponential Almon lag polynomial (Ghysels *et al.*, 2007, p. 57ff, Ghysels, 2012a, p. 48f, e.g.) In the first case, we let ω_j be either [1.1] an unrestricted (normalized) beta probability density function with non-zero last lag,

$$\omega_j(\Theta) = \omega_j(\theta_1, \theta_2, \theta_3) = \frac{z_j^{\theta_1-1} (1 - z_j)^{\theta_2-1}}{\sum_{j=1}^J z_j^{\theta_1-1} (1 - z_j)^{\theta_2-1}} + \theta_3, \quad (2)$$

where $z_j = (j-1)/(J-1)$, or [1.2] a restricted beta probability density function with non-zero last lag, $\omega_j = \omega_j(1, \theta_2, \theta_3)$, or [1.3] an unrestricted beta probability density

⁴ $q = 0$ is the nowcast case. Naturally, when $q = p$, then m equals either $3p + 1$ (third month of quarter), $3p + 2$ (second month of quarter) or $3p + 3$ (first month of quarter) $\forall p = 0, \dots, K$. Note however that this is not a necessary pattern. For instance, it may be that the statistical office decides to regularly publish y_t after $x_{t+1-3*1/3} = x_t$, and we want to employ the forecasting model of Equation (1) at a date between the publication dates of y_t and x_t . In this case, for a forecast of y_{t+1} , we will employ Equation (1) with $m = 3$ to include all available data, namely x_t but not y_t .

function with zero last lag, $\omega_j = \omega_j(\theta_1, \theta_2, 0)$, or [1.4] a restricted beta probability density function with zero last lag, $\omega_j = \omega_j(1, \theta_2, 0)$. In the second case,

$$\omega_j(\Theta) = \omega_j(\theta_1, \theta_2) = \frac{e^{\theta_1*j + \theta_2*j^2}}{\sum_{j=1}^J e^{\theta_1*j + \theta_2*j^2}}.$$

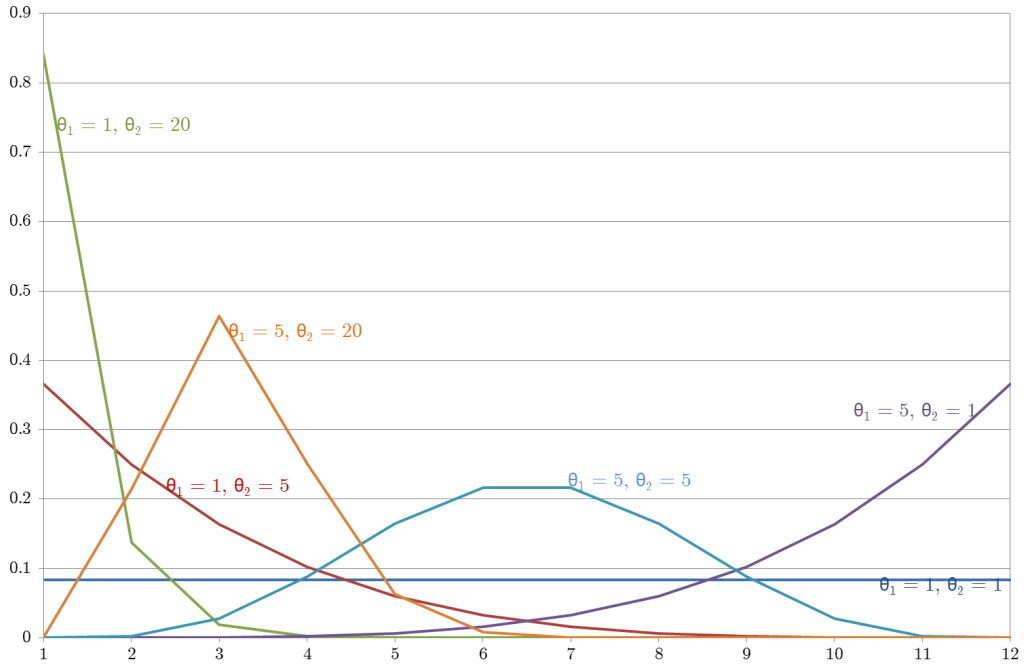
In the third case,

$$\beta\omega_j(\Theta) = \beta\omega_j(\theta_0, \dots, \theta_P) = \sum_{p=0}^P \theta_p j^p, \quad (3)$$

where P denotes the order of the Almon lag polynomial. We choose P to be either 1, 2, 3 or 4.

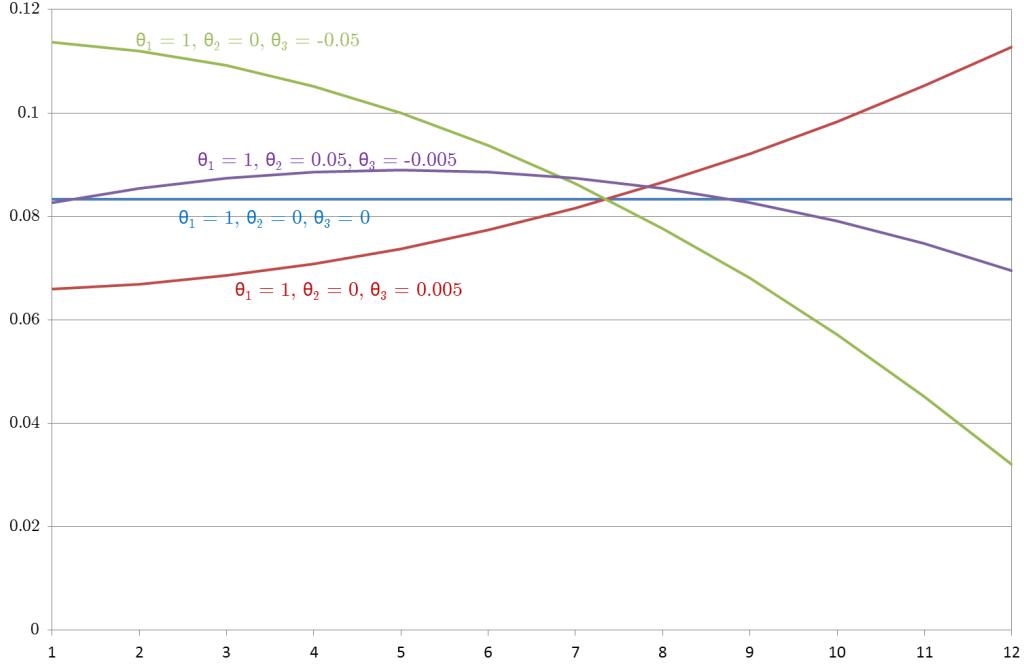
Why are the weights ω_j modeled as such complicated functions? The above weight functions combine two features. First, the form of the function ω_j with respect to $j = 1, \dots, J$ (not with respect to Θ) is very flexible depending on the choice of Θ . Figures 1 and 2 illustrate this graphically for the beta probability density function and the non-exponential Almon lag polynomial.

Figure 1: beta probability density function



Notes: See Equation (2). y -axis: beta probability density function weight, $\omega_j(\theta_1, \theta_2, \theta_3)$, attached to monthly variable, $x_{t+q-j*1/3}$, conditional on different values of θ_1 and θ_2 (θ_3 always set to zero). x -axis: $j = 1, \dots, 12$ (= first to twelve lag).

Figure 2: Non-exponential Almon lag polynomial



Notes: See Equation (3). y -axis: Non-exponential Almon lag polynomial weight of order 2, $\omega_j(\theta_1, \theta_2, \theta_3)$, attached to monthly variable, $x_{t+q-j*1/3}$, conditional on different values of θ_1 , θ_2 and θ_3 . x -axis: $j = 1, \dots, 12$ (= first to twelve lag). Sum of weights normalized to 1.

Second, the weight functions depend only on a small number of parameters (1, 2 or 3 in the first case, 1 or 2 in the second, 1, 2, 3 or 4 in the third case). As a consequence, it is possible to bring together two goals which are usually in a trade-off position to each other: First, ensure a flexible model, i.e. let the relative importance of any observation $x_{t-h*1/3}$ as compared to any other observation $x_{t-j*1/3} \forall h \neq j$ be determined by the data, but not pre-determined by the model itself.⁵ Second, prevent parameter proliferation or overfitting. In sum, MIDAS solves the classical dilemma between flexibility and parsimony.⁶

All this notwithstanding, it is a priori not clear whether the above specifications are superior to simpler forecasting approaches, namely [4] a fully unconstrained model or [5] bridge equations. A fully unconstrained model results from leaving each weight ω_j in Equation (1) unrestricted. The resulting linear regression equation with J unknown parameters, $\beta\omega_1, \dots, \beta\omega_J$, can be estimated via OLS. This is the unrestricted MIDAS (U-MIDAS) approach proposed by [Forni *et al.* \(2012\)](#). Albeit being very flexible, the U-MIDAS model is not parsimonious when the number of x -lags, J , is large. Thus, U-MIDAS will be superior (inferior) to alternative weighting schemes when J is sufficiently small (large) (see the evaluation in *ibid.*). The bridge

⁵This includes allowing ω_j to be non-zero even for large j .

⁶The above specifications merit several further remarks. Interested readers are referred to Appendix 6.1.

equation approach is traditionally used to aggregate variables of different frequencies to a common low frequency (Golinelli and Parigi, 2007, Diron, 2008, e.g.). Following this approach, we take quarterly averages of the monthly variable $x_{t-j*1/3}$ to build a quarterly variable x_t^q . Following Maier (2011) we deal with incomplete quarters as follows: If the last available observation of $x_{t-j*1/3}$ falls on the first month of a quarter, the quarterly value is set identical to the value recorded for the first month. Further, if the last available observation of $x_{t-j*1/3}$ falls on the second month of a quarter, the value for the last month of the quarter is set identical to the value recorded for the second month. Thus, the bridge equation forecasting model writes

$$y_{t+q} = \lambda_0 + \sum_{i=1}^I \lambda_i y_{t-i} + \sum_{k=0}^K \mu_k x_{t-k}^q + \zeta_{t+q}, \quad (4)$$

$\forall t = 1, \dots, T$, where $q \in \{0, 1, \dots, Q\}$, $\lambda_0, \lambda_1, \dots, \lambda_I$ and $\mu_0, \mu_1, \dots, \mu_K$ are unknown parameters, ζ_{t+q} is an error term and where

$$x_t^q = 1/3 \left(\sum_{j=3*k}^{3*k+2} x_{t-j*1/3} \right) \quad \forall k = 0, 1, \dots, K$$

and where

$$x_t = x_{t-1/3} = x_{t-2/3}$$

if $x_{t-2/3}$ is the last available observation and where

$$x_t = x_{t-1/3}$$

if $x_{t-1/3}$ is the last available observation.

Our benchmark to the above forecasting models is the ADL model

$$y_{t+q} = \gamma_0 + \sum_{i=1}^I \gamma_i y_{t-i} + \eta_{t+q} \quad (5)$$

$\forall t = 1, \dots, T$, where $q \in \{0, 1, \dots, Q\}$, $\gamma_0, \gamma_1, \dots, \gamma_I$ are unknown parameters and η_{t+q} is an error term. This is a natural choice as the latter model is nested in the former ones.

We apply a rolling window model selection and forecasting procedure: First, we select the last $T - \tau$ observations of y_{t+q} as the out-of-sample range. Second, we determine the optimal lag-length I^* of the ADL benchmark model (5) based on Bayesian in-sample information criteria analysis where the in-sample period range covers periods $t = 1, \dots, T - \tau - 1$. Third, we determine, again based on Bayesian in-sample information criteria analysis, the optimal model out of 132 alternative model specifications, namely [1.1–1.4] MIDAS with restricted or unrestricted beta probability density function with zero or non-zero last lag, [2] MIDAS with exponential Almon lag polynomial, [3.1–3.4] MIDAS with non-exponential Almon lag

polynomial of order 1 to 4, [4] U-MIDAS and [5] bridge equations, each with $J = 1$ to 12 x -lags and I^* y -lags. Fourth, we calculate the forecasts of the ADL benchmark model and the optimal model for period $T - \tau$ as well as the respective forecast errors, $e_{T-\tau}^{BM}$ and $e_{T-\tau}^{OM}$ (BM stands for benchmark model, OM stands for optimal model). Notably, for $q > 1$ direct forecasts are used (as compared to iterated forecasts). Thereafter, we repeat the four steps τ times for the rolling-window in-sample period ranges $t = 1 + l, \dots, T - \tau - 1 + l$ and respective forecast periods $T - \tau + l$ with $l = 1, \dots, \tau$ resulting in forecast errors, $e_{T-\tau+1}^{BM}, \dots, e_T^{BM}$ and $e_{T-\tau+1}^{OM}, \dots, e_T^{OM}$. This allows us to calculate the root mean squared forecast error (RMSFE) for the ADL benchmark model and the optimal model as well as the difference between the RMSFE of the optimal model and the benchmark model in percent of the RMSFE of the benchmark model,

$$\begin{aligned} \Delta RMSFE &= 100 * \left(\frac{RMSFE^{OM} - RMSFE^{BM}}{RMSFE^{BM}} \right) \\ &= 100 * \left(\frac{\sqrt{\sum_{t=T-\tau}^T (e_t^{OM})^2} - \sqrt{\sum_{t=T-\tau}^T (e_t^{BM})^2}}{\sqrt{\sum_{t=T-\tau}^T (e_t^{BM})^2}} \right). \end{aligned} \quad (6)$$

We refer to $\Delta RMSFE$ as the relative change in RMSFE. Clearly, the *more negative* $\Delta RMSFE$ is, the *better* performs the optimal model (ADL-MIDAS, ADL-U-MIDAS or bridge) relative to the ADL benchmark model in terms of predictive power. Finally, we repeat the whole routine for selected days d within 370 days to 1 day before release of y_{t+q} always including only data which is available at day d .⁷ Appendix 6.2 describes the day-specific rolling window model selection and forecasting procedure in more detail. We combine the day-specific procedure with a harmonization of release dates over different quarters and years. This is discussed in Appendix 6.3. Next to RMSFE comparisons we test whether the forecast performance of the optimal model is significantly better than the forecast performance of the ADL benchmark model using the classical test proposed by Diebold and Mariano (1995) and Giacomini and White (2006). Alternatively, we employ the Modified-Diebold-Mariano test proposed by Harvey *et al.* (1997), the forecast encompassing test proposed by Harvey *et al.* (1998) and the nested model test proposed by Clark and West (2007). The Diebold-Mariano test delivers the most conservative results, i.e. it generally rejects the null hypothesis of equal forecast performance less often than the alternative tests. A detailed explanation of the tests is deferred to Appendix 6.4.

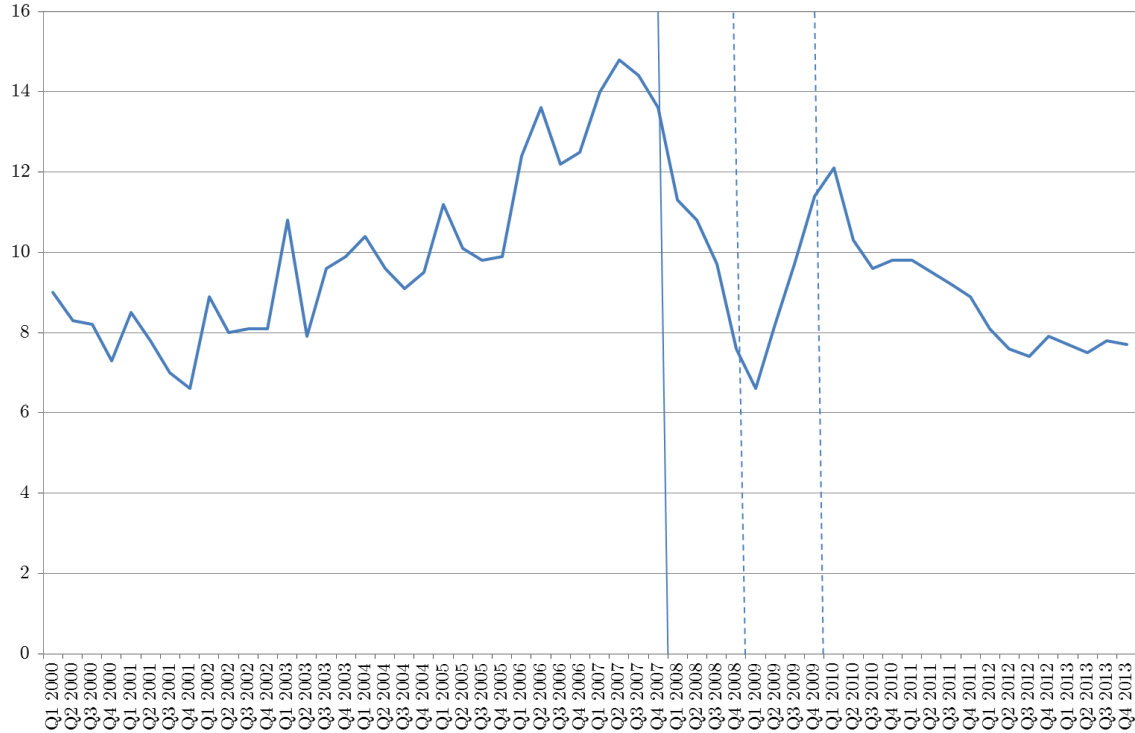
3 Data

We forecast the (non-cumulative) year-over-year growth rate of Chinese real GDP as published by the Chinese National Bureau of Statistics (NBS). The variable is released on a quarterly basis since the year 2000 and is today the main reference

⁷It would not make sense to forecast at each day $d = 1, \dots, D$, because not every day new data are available and, hence, the forecasts will not change every day.

series for quarterly Chinese GDP growth.⁸ The time series used in this paper covers the periods $Q1\ 2000$ to $Q4\ 2013$ and is displayed in Figure 3. We choose the out-of-sample range to start in $Q1\ 2008$. Thus, the out-of-sample range covers the last 24 out of 56 quarterly periods and entails substantial variation. Specifically, it includes more than two-thirds of the Great Recession induced downturn, the following recovery and the downward trend thereafter.⁹ For means of robustness, we repeat the analysis with out-of-sample ranges starting in $Q1\ 2009$ (= after the Great Recession) and $Q1\ 2010$ (= after the recovery following the Great Recession), respectively. During $Q1\ 2008$ to $Q4\ 2013$, the NBS released GDP growth of quarter t between the 13th and 24th day of quarter $t + 1$. In line with our conservative harmonization strategy outlined in Appendix 6.3 we choose the 13th day of quarter $t + 1$ as the harmonized release date for GDP growth of quarter t . For the x -variable in Equation (1) we use a variety of monthly macroeconomic indicators. Table 1 summarizes all data used in this paper. A more detailed description of the individual times series comes together with the results in Section 4.

Figure 3: Year-over-year growth of Chinese real GDP (in percent)



⁸Next to this series the NBS publishes a quarterly cumulative year-over-year (or year-to-date) real GDP growth rate which used to be the main reference series. From $Q4\ 2010$ onwards the NBS also publishes a quarterly (non-cumulative) quarter-over-quarter GDP growth rate; the series is yet too short for time series analysis. The NBS publishes quarterly nominal GDP in levels, but neither quarterly real GDP in levels nor a quarterly GDP deflator.

⁹Stock and Watson (2012, p. 605) recommend the out-of-sample range to cover 10–15% of the total sample. Our out-of-sample range covers 43% of total sample size which deems us appropriate given the short time series available.

Chinese GDP data get only rarely revised. The latest revisions occurred in 2001 and 2007. As far as the monthly indicators get nor revised either (see the notes in Table 1), this paper provides a real-time forecasting analysis.

The quality of China’s official statistics is sometimes contested in the media. A more subtle discussion occurs also in academia (Chow, 2006, Koch-Weser, 2013, e.g.). We see three reasons to take the data as they are. First, there simply do not exist more accurate estimates of Chinese GDP growth than the estimates provided by the NBS. Second, the results of Mehrotra and Pääkkönen (2011) indicate that there is no systematical bias in China’s real GDP growth data and that the data are in fact rather reliable. Third, even if there would be systematical or occasional bias in the data, an analysis seems worthwhile as long as the NBS releases provide at least *some* information about “true GDP”. After all, it should be noted that in no country of the world official statistics always perfectly reflect the true state of the economy.¹⁰ For this paper we care less about the accuracy of the monthly indicators that we use as regressors in our forecasting exercise. For our purpose a time series is good if it helps forecasting GDP growth no matter how accurately the time series reflects whatever it shall reflect.

4 Results

In Sections 4.1 to 4.6 we evaluate the individual predictive power of a variety of monthly macroeconomic indicators which are generally considered to have predictive content and/or which get public attention when it comes to forecasting quarterly Chinese GDP growth. In line with the forecasting procedure outlined in Section 2 we iterate the evaluation for a whole range of forecast horizons between 370 days and 1 day prior to GDP release. We track the monthly release days of the indicators so as to only use data which is actually available at the respective forecast horizon. This procedure allows us to detect how useful a specific indicator is for forecasting quarterly Chinese GDP growth at a specific horizon.

4.1 Leading and coincident indicators: OECD vs. Conference Board vs. Goldman Sachs

There exist several alternative leading and coincident indicators for China, the most cited among them being provided by the Organisation for Economic Co-operation and Development (OECD), the Conference Board (CB) and the China Economic Monitoring and Analysis Center (CEMAC), an affiliate of the NBS, in cooperation with Goldman Sachs (GS).¹¹ Which of the indicators is most helpful for forecasting Chinese GDP growth?

¹⁰Admittedly, an issue with the Chinese GDP and national accounts data is that they are, from an accounting perspective, sometimes not coherent to each other.

¹¹See www.oecd.org/std/leading-indicators, www.conference-board.org and www.cemac.org.cn.

Table 1: Data

TIME SERIES	TRANSFORMATIONS	PERIODICITY	REVISIONS	SOURCE	NOTES
<i>Real GDP</i>	(non-cumulative) y-o-y	quarterly	2001, 2007	Chinese National Bureau of Statistics (NBS)	
<i>OECD Composite Leading Indicator, trend restored</i>	y-o-y, m-o-m	monthly	yes	Organisation for Economic Co-operation and Development (OECD)	
<i>OECD Composite Leading Indicator, amplitude adjusted</i>	levels, y-o-y, m-o-m	monthly	yes	OECD	
<i>OECD Composite Leading Indicator, normalized</i>	levels, y-o-y, m-o-m	monthly	yes	OECD	
<i>CEMAC-GS Leading Indicator</i>	levels, y-o-y, m-o-m	monthly	yes	China Economic Monitoring and Analysis Center (CEMAC)	
<i>CB Leading Economic Index</i>	y-o-y, m-o-m	monthly	yes	Conference Board (CB)	
<i>CB 1-month Diffusion Index of Leading Components</i>	levels	monthly	yes	OECD	
<i>OECD Industrial Confidence Indicator</i>	levels, y-o-y, m-o-m	monthly	yes	CEMAC-GS	
<i>CEMAC-GS Business Cycle Signal Indicator</i>	levels, y-o-y, m-o-m	monthly	yes	CEMAC-GS	
<i>CEMAC-GS Coincident Indicator</i>	levels, y-o-y, m-o-m	monthly	yes	CEMAC-GS	
<i>CB Coincident Economic Index</i>	y-o-y, m-o-m	monthly	yes	CB	
<i>CB 1-month Diffusion Index of Coincident Components</i>	levels	monthly	yes	CB	
<i>Shanghai SE Composite Index</i>	y-o-y, m-o-m	monthly end value	no	Shanghai Stock Exchange (SSE)	All 1003 stocks at SSE.
<i>Shanghai SE B Share Index</i>	y-o-y, m-o-m	monthly end value	no	SSE	53 listed B stocks. B stocks are stocks that can be purchased by foreign investors without restriction.
<i>Shenzhen SE Composite Index</i>	y-o-y, m-o-m	monthly end value	no	Shenzhen Stock Exchange (SZSE)	All 1620 stocks at SZSE.
<i>Shenzhen SE B Share Index</i>	y-o-y, m-o-m	monthly end value	no	SZSE	57 listed B stocks.
<i>Hang Seng Index ("HSI")</i>	y-o-y, m-o-m	monthly end value	no	Hong Kong Stock Exchange (SEHK)	Major index for SEHK. 50 companies representing about 70% of total market capitalization at SEHK.
<i>Hang Seng China Enterprises Index ("H shares")</i>	y-o-y, m-o-m	monthly end value	no	SEHK	All 40 companies with incorporation in mainland China and with primary listing on SEHK Main Board ("H-Share companies").
<i>Hang Seng China-Affiliated Corporations Index ("HSCCI")</i>	y-o-y, m-o-m	monthly end value	no	SEHK	All 25 mainland controlled companies with incorporation outside mainland China and with primary listing on SEHK Main Board ("red chip companies").
<i>Hang Seng Mainland 25 Index ("HSML25")</i>	y-o-y, m-o-m	monthly end value	no	SEHK	25 largest Hong Kong-listed mainland companies including H and red chip companies.
<i>Hang Seng China 50 Index ("HSML25")</i>	y-o-y, m-o-m	monthly end value	no	SEHK	50 largest China companies listed on Hong Kong and/or mainland China stock exchanges.
<i>Consumer price inflation</i>	–	monthly	no	NBS	y-o-y growth of consumer price index
<i>Producer price inflation</i>	–	monthly	no	NBS	y-o-y growth of producer price index for industrial sector
<i>Prime lending rate</i>	–	monthly end value	no	People's Bank of China (PBC)	
<i>Reserve requirement ratio for big banks</i>	–	monthly end value	no	PBC	
<i>Reserve requirement ratio for small banks</i>	–	monthly end value	no	PBC	
<i>Money supply M0</i>	y-o-y, m-o-m	monthly	no	PBC	
<i>Money supply M1</i>	y-o-y, m-o-m	monthly	no	PBC	
<i>Money supply M2</i>	y-o-y, m-o-m	monthly	no	PBC	
<i>Total social financing (TSF)</i>	y-o-y, m-o-m	monthly	yes	PBC	Principal measure for credit growth in China. TSF comprises all kinds of new credit issuance, funding or liquidity provision by the financial sector to the real economy during a given period of time.
<i>Real industrial production</i>	y-o-y	monthly	no	NBS	Composite index based on several subindices.
<i>Purchasing managers' index (PMI) for manufacturing sector</i>	levels, y-o-y, m-o-m	monthly	no	NBS	No monthly deflators available
<i>Nominal retail sales in consumer goods</i>	y-o-y, m-o-m	monthly	yes	NBS	Excludes rural households. No monthly deflators available
<i>Nominal cumulative investment in fixed assets</i>	y-o-y	monthly	yes	NBS	No monthly deflators available
<i>Nominal exports</i>	y-o-y, m-o-m	monthly	no	Chinese General Administration of Customs (GAC)	No monthly deflators available
<i>Nominal imports</i>	y-o-y, m-o-m	monthly	no	GAC	No monthly deflators available
<i>Electricity usage in industrial production</i>	y-o-y, m-o-m	monthly	no	NBS	Measured in billions of kilowatts per hour
<i>Railway freight transport turnover</i>	y-o-y, m-o-m	monthly	no	NBS	Measured either in metric tons or in metric tons per kilometre

• Notes on CB diffusion indices: CB also publishes a 6-month diffusion index of leading and coincident components. We disregard these indices in our pseudo-real time analysis as it does not get released until three months after the release of the other CB indicators.

• Notes on stock exchange indices: The Shanghai SE A Share Index and the Shanghai SE Composite Index (which includes A shares and B shares) are very highly correlated. We opted for using the former index. The Shanghai SE 50 and the Shanghai SE 180 (the former being a subset of the latter) are highly correlated with each other and with the Composite and the A Share index. We abstain from taking the Shanghai SE 50 and the Shanghai SE 180 into the analysis, because they start only in 2004, respectively. The Shenzhen SE Composite and the Shenzhen SE A Share are very highly correlated. The correlation with the Shenzhen SE B Share is less high. We take the Composite and the B Share index into the analysis. We did not include the Hang Seng Mainland 100 index in our analysis because it is only available from 2010 onwards for us.

A careful analysis has to take into account that the different indicators come with different release schedules. OECD indicators for month m are released between the 8th and the 18th of month $m + 2$. CEMAC-GS indicators for month m are released between the 20th of month $m + 1$ and the 18th of month $m + 2$. In contrast, CB indicators for month m are released between the 16th and the 25th of month $m + 1$ already.¹² We choose the 18th day of each month $m + 2$ as the harmonized release date of OECD and CEMAC-GS indicator values and the 18th day of each month $m + 1$ (!) as the harmonized release date of CB indicator values.¹³

To start with, we compare different leading indicator versions of the same source with each other: The OECD provides its leading indicator in three versions: trend restored, amplitude adjusted and normalized, where the latter two are very highly correlated. We find that the trend restored version generally adds more predictive power to forecasting GDP growth (in addition to the predictive power of past values of GDP growth) as compared to the other versions. This result corresponds to the OECD’s recommendation to compare the (year-over-year growth rate of the) trend restored indicator series with the (y-o-y growth rate of the) original GDP series and the (y-o-y growth rate of the) amplitude adjusted indicator series with the (y-o-y growth rate of the) de-trended GDP series. The CEMAC-GS leading indicator comes in an amplitude adjusted version, i.e. it moves up and down in accordance with the growth cycle. In contrast, the CB leading indicator comes in a trend restored version meaning that it trends upwards reflecting the upward trend in GDP. We find the trend restored OECD leading indicator in y-o-y growth rates to outperform the same indicator in month-over-month growth rates in terms of predictive power (levels would not make sense given that the indicator trends upwards). Likewise, the CEMAC-GS leading indicator in m-o-m growth rates proves to have higher predictive power than the same indicator in levels or y-o-y growth rates, and the CB leading indicator in y-o-y growth rates turns out to dominate the same indicator in m-o-m growth rates. Table 1 provides a complete list of all employed indicators.¹⁴

Figure 4 displays the relative RMSFE changes of the ADL-(U-)MIDAS model including *either* the trend restored OECD leading indicator, the CEMAC-GS leading indicator *or* the CB leading indicator.¹⁵ During 370 to 237 days prior to GDP growth release none of the indicators adds predictive power to forecasting GDP growth (in addition to the predictive power of past values of GDP growth). Subsequently, the

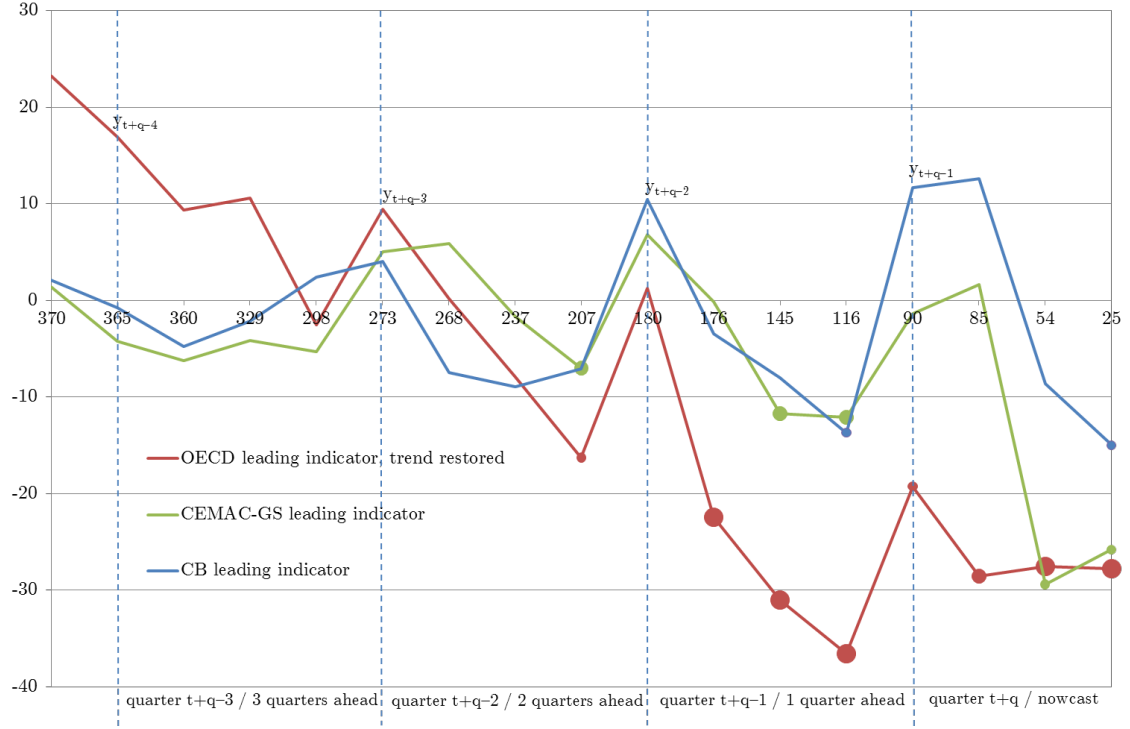
¹²In January 2012 the Conference Board switched from a $m+2$ publication lag to a $m+1$ publication lag.

¹³Choosing the 18th, but not the 25th for CB indicator values deviates marginally from our conservative strategy outlined in Appendix 6.3, but facilitates joint presentation of OECD, CEMAC-GS and CB indicators.

¹⁴CB also publishes a 1-month diffusion index of leading components. The index measures the share of leading indicator components which contribute positively to the indicator over a span of 1 month (cf. [the CB webpage](#) for a more detailed description). We find the CB 1-month diffusion index to be dominated by the CB leading indicator itself. CB further publishes a 6-month diffusion index of leading components. We disregard this index in our pseudo-real time analysis as it does not get released until three months after the release of the CB leading indicator.

¹⁵As bridge equations are never selected as the optimal model (cf. Section 2) we use the term ADL-(U-)MIDAS here and in the following.

Figure 4: Predictive power of leading indicators for GDP growth: Relative change in RMSFE



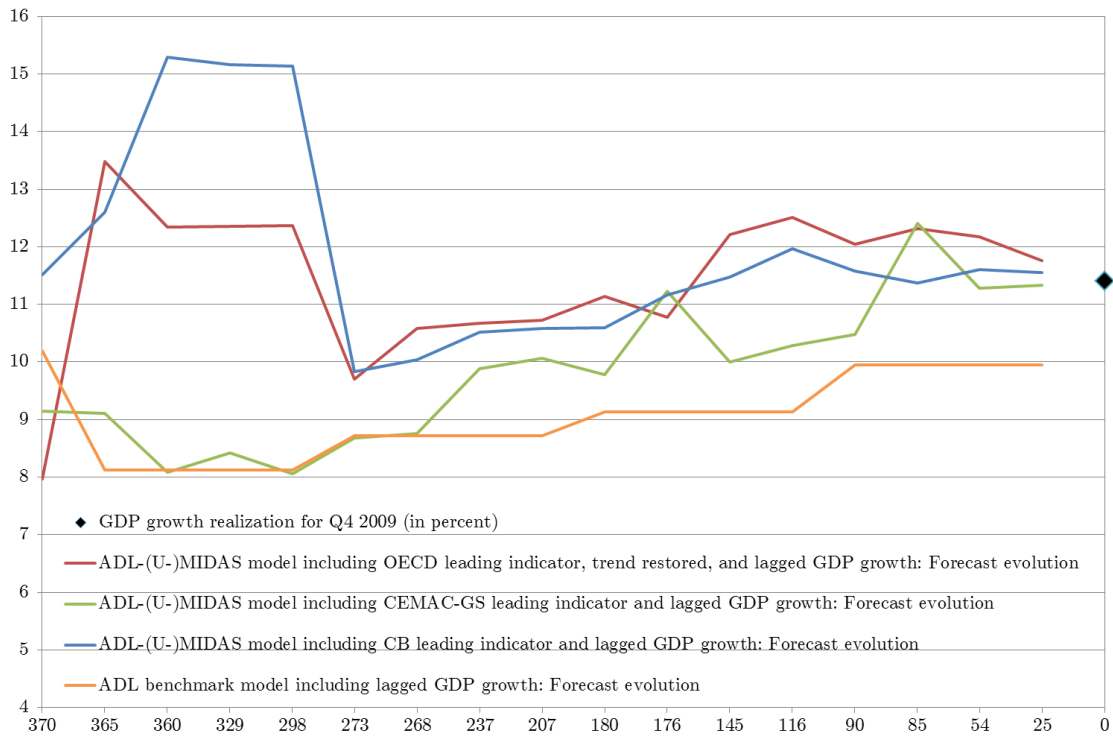
x -axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including *either* trend restored OECD leading indicator in y-o-y growth rates, CEMAC-GS leading indicator in m-o-m growth rates *or* CB leading indicator in y-o-y growth rates and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth are indicated by y_{t+q-i} . Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 3 for shares of different weight schemes in the 408 optimal models.

relative RMSFE changes show a characteristic cyclical pattern: They drop as new leading indicator observations come in over the course of the quarter, but increase again as 2-quarters ahead GDP growth gets published at day 180. The fact that the relative RMSFE changes are at values of zero (or even above) at day 180 implies that the predictive content of 2-quarters ahead GDP growth makes the predictive content in any prior leading indicator observation irrelevant. The cyclical pattern repeats for the subsequent forecast quarters, albeit on different levels. Notably, the OECD leading indicator clearly outperforms the other two indicators. For day 116 the RMSFE of the ADL-(U-)MIDAS model including the OECD leading indicator is as much as 37 percent lower than the RMSFE of the ADL benchmark model. The CEMAC-GS leading indicator releases of days 54 and 25 also add substantial predictive power. The above results are robust to iterating the analysis for out-of-sample ranges starting Q1 2009 (= after the Great Recession) or Q1 2010 (= after the recovery following the Great Recession) instead of Q1 2008. Figure 5 shows how the GDP growth forecast for a particular period, Q4 2009, evolves over time. Q4 2009 marks a temporary return to strong growth after the slump in 2008 (cf. Figure 3). It might be interesting to see whether/when the leading indicators anticipate

this rebound.¹⁶ Overall, the CB leading indicator tends to perform quite well here. The OECD leading indicator anticipates, from day 273 onwards, the rebound earlier than the other indicators, but delivers an exaggerated growth forecast from day 145 onwards. In contrast, the CEMAC-GS leading indicator underestimates GDP growth until day 90.

It may be argued that the aforementioned leading indicators are governed by a common latent process which is even more helpful for predicting GDP growth than the leading indicators themselves. To test this we extract the first common factor from the leading indicators and employ it in our MIDAS forecasting exercise. It turns out that the factor series never reduces the RMSFE by more than the OECD leading indicator series.

Figure 5: Forecast evolution for Q4 2009



x-axis: Days until GDP growth release for Q4 2009. y-axis: Forecasts for GDP growth of Q4 2009 at different forecast horizons (= different days prior to GDP growth release). Q4 2009 marks the start of a temporary return to strong growth after the slump in 2008 (cf. Figure 3). It might be interesting to see whether/when the leading indicators anticipate this rebound.

The aforementioned sources also provide a number of indicators on the current state of the economy which may be useful for nowcasting: the OECD Industrial Confidence Index, the so called CEMAC-GS Business Cycle Signal Indicator, the CEMAC-GS coincident indicator and the CB Coincident Economic Index. CB further publishes a 1-month diffusion index of coincident components.¹⁷ We find the OECD Industrial Confidence Index in levels and the CEMAC-GS coincident index

¹⁶Still, Q4 2009 is just an exemplary choice. Results for other forecast periods are available on request.

¹⁷The OECD does not provide a general coincident indicator. While capturing only the industrial

in m-o-m growth rates to have good and robust nowcasting properties for days 54 and 25 prior to GDP growth. Notably however, the aforementioned indicators prove to be less helpful for nowcasting GDP growth than the OECD *leading* indicator and the CEMAC-GS *leading* indicator.

Again, we extract the first common factor from the five nowcast series and employ it in our MIDAS forecasting exercise. Like the nowcast series themselves the common factor series turns out to be less valuable for nowcasting GDP growth than the OECD *leading* indicator or the CEMAC-GS *leading* indicator.

To sum up, the monthly OECD leading indicator (trend restored version in y-o-y growth rates) adds substantial predictive power for forecasting quarterly GDP growth in addition to the predictive power of past values of GDP growth. The additional value is highest 5 to 4 months prior to GDP release, where the inclusion of the OECD leading indicator time series improves the RMSFE by 31% to 37%. The OECD leading indicator outperforms its counterparts published by the Conference Board (CB) and by the China Economic Monitoring and Analysis Center (CEMAC), an affiliate of the NBS, in cooperation with Goldman Sachs (GS). Surprisingly, the OECD and CEMAC-GS leading indicators turn out to have even better – or at least not worse – nowcasting properties than several nowcast indicators published by OECD, CEMAC-GS or CB.

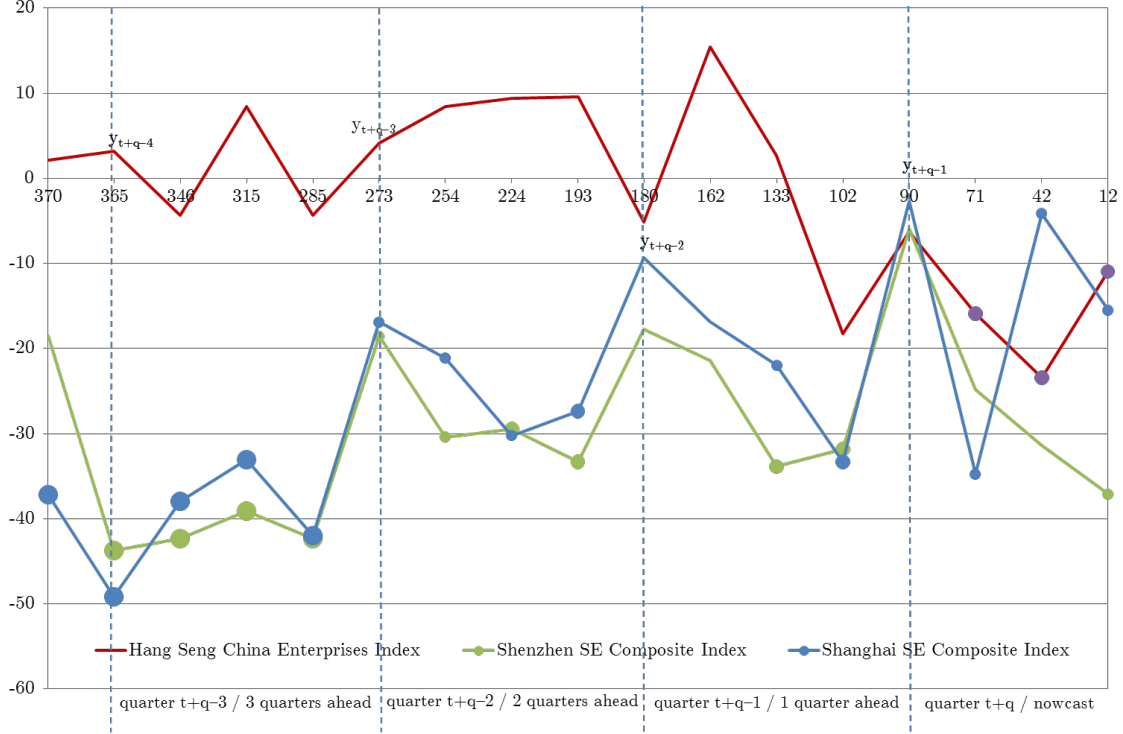
4.2 Stock market indices: Shanghai vs. Shenzhen vs. Hong Kong

The stock market commonly reflects macroeconomic developments. It would be interesting to know whether Chinese stock markets can help predicting GDP growth. China has two major stock exchanges, the Shanghai Stock Exchange (1003 listed companies with total market capitalization of about RMB 15 trillion by 2014) and the Shenzhen Stock Exchange (1620 listed companies with total market capitalization of about RMB 9 trillion by 2014). The Hong Kong Stock Exchange also lists many companies with business activities in mainland China. We test for the predictive power of several indices on the aforementioned stock exchanges. A complete list of the indices is provided in Table 1. All indices enter the analysis in year-over-year growth rates. Results on month-over-month growth rates are also briefly reported. In order to build monthly time series we always collect the closing value of the last trading day of a month.

Figure 6 shows the relative RMSFE changes of the ADL-(U-)MIDAS model including either the Shanghai SE Composite Index, the Shenzhen SE Composite Index or the Hang Seng China Enterprises Index which covers companies incorporated in mainland China and listed at the Hong Kong Stock Exchange. For most forecast

sector the OECD Industrial Confidence Index may be used for a general assessment of the current state of the economy. Like the CB 6-month diffusion index of leading components the CB 6-month diffusion index of coincident components is released with a lag of three months and, hence, is not included in our pseudo real-time analysis.

Figure 6: Predictive power of stock market indices for GDP growth: Relative change in RMSFE



x -axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including *either* Shanghai SE Composite Index, Shenzhen SE Composite Index *or* Hang Seng China Enterprises Index all in y -o- y growth rates and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth are indicated by y_{t+q-i} . Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure ?? for shares of different weight schemes in the 408 optimal models.

horizons, the Hang Seng index performs comparatively weak. This is in line with our expectations as the former index covers a relatively low share of the Chinese economy. Still, inclusion of the Hang Seng index makes the relative RMSFE ratio fall by more than 20 percent in the nowcast range. Both, the Shenzhen and Shanghai index add considerable predictive power to forecasting GDP growth; for most forecast horizons the former index tends to outperform the latter one. This is a notable finding because the Shenzhen index covers a lower share of the Chinese economy as compared to the Shanghai index, at least when measured by the market capitalization of the companies included in the indices. A possible explanation is that the former index includes more private medium sized companies which contribute less to overall growth, but more to overall growth *volatility*.¹⁸

¹⁸Further, we found the Shenzhen/Shanghai SE Composite Index to outperform the Shenzhen/Shanghai SE B Share Index. This is what we expect since the former indices encompass A shares (available to foreign investors with restrictions only) as well as B shares (available to foreign investors without restrictions) and, therefore, cover a higher share of the Chinese economy. We also tested a number of other Hang Seng indices, namely Hang Seng (Composite), Hang Seng China-Affiliated Corporations, Hang Seng Mainland 25 and Hang Seng China 50. None of these indices deliver better results.

Both the Shenzhen and the Shanghai index tend to yield larger RMSFE reductions three quarters ahead than two or one quarters ahead. Hence, 3-quarters ahead monthly stock market information contributes more predictive power to forecasting GDP growth *in addition to the predictive power of 4-quarters ahead GDP growth* than 2-quarters ahead (1-quarter ahead) monthly stock market information contributes to forecasting GDP growth *in addition to the predictive power of 3-quarters (2-quarters) ahead GDP growth*. Notably, this finding does not imply that 3-quarters ahead stock market information has *absolutely* more predictive power than 2- or 1-quarter ahead information. Both the Shenzhen and the Shanghai index show a characteristic zig-zag pattern: the RMSFE ratio drops as new stock market information gets released during a quarter but increases again as lagged GDP growth information becomes available because earlier indicator releases get less valuable with new GDP information.¹⁹ Thus, stock market information is most valuable for forecasting GDP growth *just before* the release of lagged GDP growth data.

Our findings are basically unchanged when iterating the analysis for out-of-sample ranges starting Q1 2009 (= after the Great Recession) or Q1 2010 (= after the recovery following the Great Recession). As mentioned above all indices have entered the analysis in monthly y-o-y growth rates. Results tend to degrade when repeating the analysis for monthly m-o-m growth rates. This may reflect that GDP growth itself is recorded in y-o-y growth rates. Arguably, the stock exchange indices are governed by a latent common process which might be even more helpful for predicting GDP growth than the individual indices. In a first step, we extract the first common (Hang Seng) factor from all five aforementioned Hang Seng indices. In a second step, we extract the first common factor from the Hang Seng factor, the Shenzhen index and the Shanghai index and apply it our MIDAS forecasting exercise. The RMSFE pattern strongly resembles the patterns of the Shanghai and Shenzhen indices shown in Figure 6. Notably, the common factor series never improves RMSFE by more than the Shenzhen index series.

To sum up, both the Shenzhen and Shanghai SE Composite Index are quite valuable for forecasting GDP growth; and despite covering a lower share of the Chinese economy, the former index tends to outperform the latter. In contrast, Hong Kong stock exchange indices do not help predicting GDP growth except for the nowcast range.

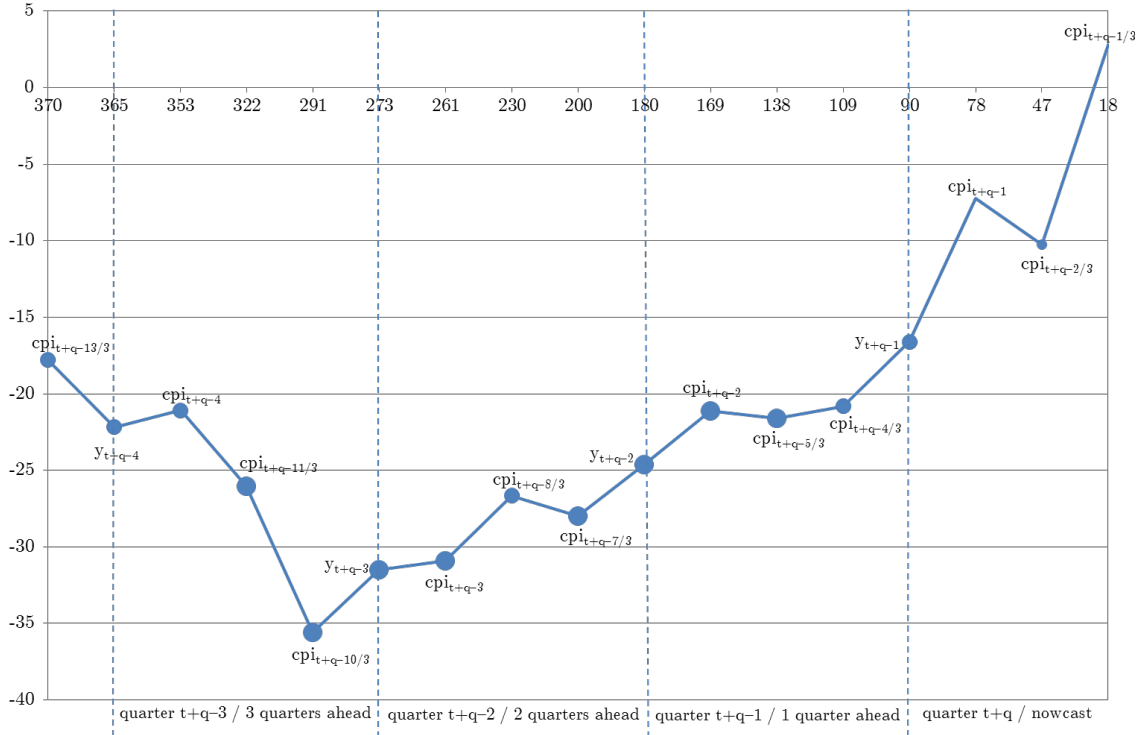
4.3 Consumer price inflation: A useful leading indicator for GDP growth

From a theoretical perspective the dynamics of GDP growth and inflation are endogenous to each other. It might be interesting to see whether inflation can indeed be of any help in forecasting growth.

¹⁹At day 90 the RMSFE ratios even increase to values near zero implying that the GDP growth release for quarter $q+t-1$ renders the predictive information in earlier indicator releases basically irrelevant.

The NBS releases the year-over-over growth rates in the consumer price index and the (industrial sector) producer price index of month m between the 9th and 25th day of month $m+1$. The time series are the main reference for consumer or producer price inflation in China; they receive Bloomberg relevance ratings of 96.8 or 90.3, which makes them the most observed macroeconomic time series for China next to GDP growth.²⁰ In line with our conservative strategy outlined in Appendix 6.3 we choose the 25th day of each month $m+1$ as the harmonized release date for inflation of month m .

Figure 7: Predictive power of consumer price inflation for GDP growth: Relative change in RMSFE



x -axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including consumer price inflation and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth and consumer price inflation are indicated by y_{t+q-i} or $cpi_{t+q-j+1/3}$. Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 4 for shares of different weight schemes in the 408 optimal models.

Figure 7 displays the relative RMSFE changes, $\Delta RMSFE$, of the ADL-(U-)MIDAS model including consumer price inflation and lagged GDP growth as compared to the ADL benchmark model including lagged GDP growth only. The figure reveals a cyclical pattern: The predictive content in a new (lagged) GDP growth release renders the predictive information in earlier inflation releases less valuable. As a consequence, the relative RMSFE change increases as lagged GDP growth data come

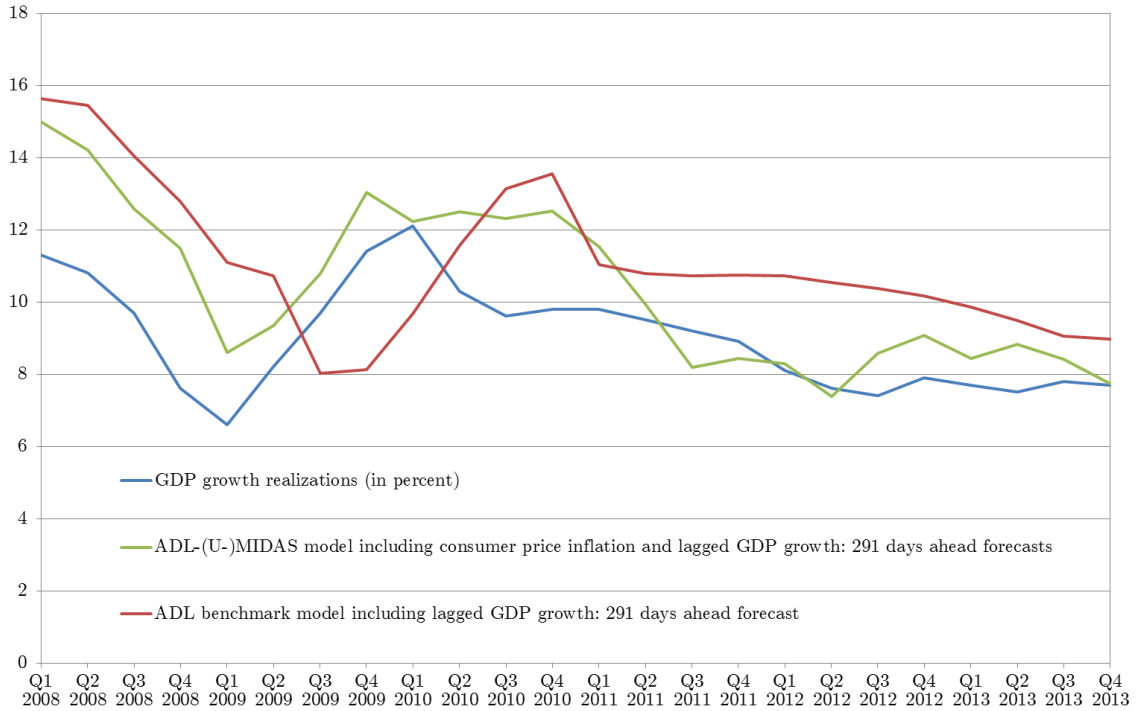
²⁰Bloomberg rates the importance of each time series on a range from 0 to 100 reflecting the number of alerts set by Bloomberg users for announcements on the respective time series as compared to the total number of alerts set for announcements on Chinese time series.

available (cf. y_{t+q-3} at day 273, y_{t+q-2} at day 180 and y_{t+q-1} at day 90; y_{t+q-4} at day 365 seems to be an exception). The relative RMSFE change increases further, or at least does not decrease, as inflation of the *third* month of a quarter gets available (cf. cpi_{t+q-4} at day 353, cpi_{t+q-3} at day 261, cpi_{t+q-2} at day 169 and cpi_{t+q-1} at day 78). Hence, as GDP growth of a quarter is already published, a monthly inflation release for the *same* quarter does not provide any additional predictive information. In contrast, monthly inflation releases for quarters for which GDP growth is not yet available do provide additional predictive information (cf. the drop in the relative RMSFE change for $cpi_{t+q-11/3}$, $cpi_{t+q-10/3}$, $cpi_{t+q-7/3}$, $cpi_{t+q-5/3}$ and $cpi_{t+q-2/3}$; $cpi_{t+q-8/3}$, $cpi_{t+q-4/3}$, $cpi_{t+q-1/3}$ seem to be exceptions here). The figure delivers two interesting further results: First, the RMSFE reductions tend to be higher between 322 and 200 days ahead of GDP release (hence, roughly 11 to 6 months prior to GDP release) than between 180 and 90 days ahead of GDP release. Hence, 3-quarters (2-quarters) ahead monthly inflation data contribute more predictive power to forecasting GDP growth *in addition to the predictive power of 4-quarters (3-quarters) ahead GDP growth* than 1-quarter ahead monthly inflation data contribute to forecasting GDP growth *in addition to the predictive power of 2-quarters ahead GDP growth*. Notably, this finding does not imply that 3- or 2-quarters ahead inflation data have *absolutely* more predictive power than 1-quarter ahead inflation data. Second, consumer price inflation helps relatively little for now-casting GDP growth: the RMSFE falls by only 10% or even less. Our findings make sense from a macroeconomic perspective: Changes in inflation need several months to affect economic activity directly or indirectly via changes in monetary or fiscal policy. Actually, the effect of inflation on GDP growth is *negative* as can be inferred from the signs of the β - and ω -coefficients in Equation (1) (not shown here). Thus, lower (higher) inflation today is associated with higher (lower) GDP growth in the future. As can be seen from Figure 7, consumer price inflation turns out to be most valuable 291 days before GDP release (relative RMSFE change of -35.6%).

Elaborating on this forecast horizon, Figure 8 displays GDP growth over the out-of-sample range together with the *291 days ahead* forecasts stemming from the ADL-(U)-MIDAS and the ADL benchmark model. Both models overpredict GDP growth in 2008, yet the ADL-(U)-MIDAS model predicts the turning point in Q1 2009 correctly, whereas the ADL benchmark model does not. Also, the former model forecasts GDP growth values of Q2 2011ff quite accurately, whereas the latter still overpredicts. Our findings prove to be robust to choosing alternative out-of-sample starts, namely Q1 2009 (= after the Great Recession) or Q1 2010 (= after the recovery following the Great Recession). The results get even better: the relative RMSFE change now stays at values below -10% during the full nowcast range.

Turning to producer price inflation, its predictive content for GDP growth is rather weak. For none of the forecast horizons, the relative RMSFE change is substantially below -1% . The negative finding is confirmed for alternative out-of-sample starts (Q1 2008, Q1 2009 or Q1 2010).

Figure 8: 291 days ahead forecasts



x-axis: 24 out-of-sample forecast periods (Q1 2008 to Q4 2013). *y*-axis: Actual GDP growth realization (in percent); 291 days ahead GDP growth forecast stemming from ADL-(U-)MIDAS model including consumer price inflation and lagged values of GDP growth, $\hat{y}_{t+q|t+q-4,t+q-11/3}$ (cf. Section 6.2); 291 days ahead GDP growth forecast stemming from ADL benchmark model including lagged values of GDP growth, $\hat{y}_{t+q|t+q-4}$ (cf. Section 6.2). 24 out-of-sample forecast periods imply 24 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 4 for shares of different weight schemes in the 24 optimal models.

To sum up, except for the nowcast range consumer price inflation adds substantial predictive power for forecasting quarterly GDP growth in addition to the predictive power of past values of GDP growth. The additional value is highest 11 to 6 months prior to GDP release, where the RMSFE gets reduced by 26% to 36%. In contrast, producer price inflation does not help much for forecasting GDP growth.

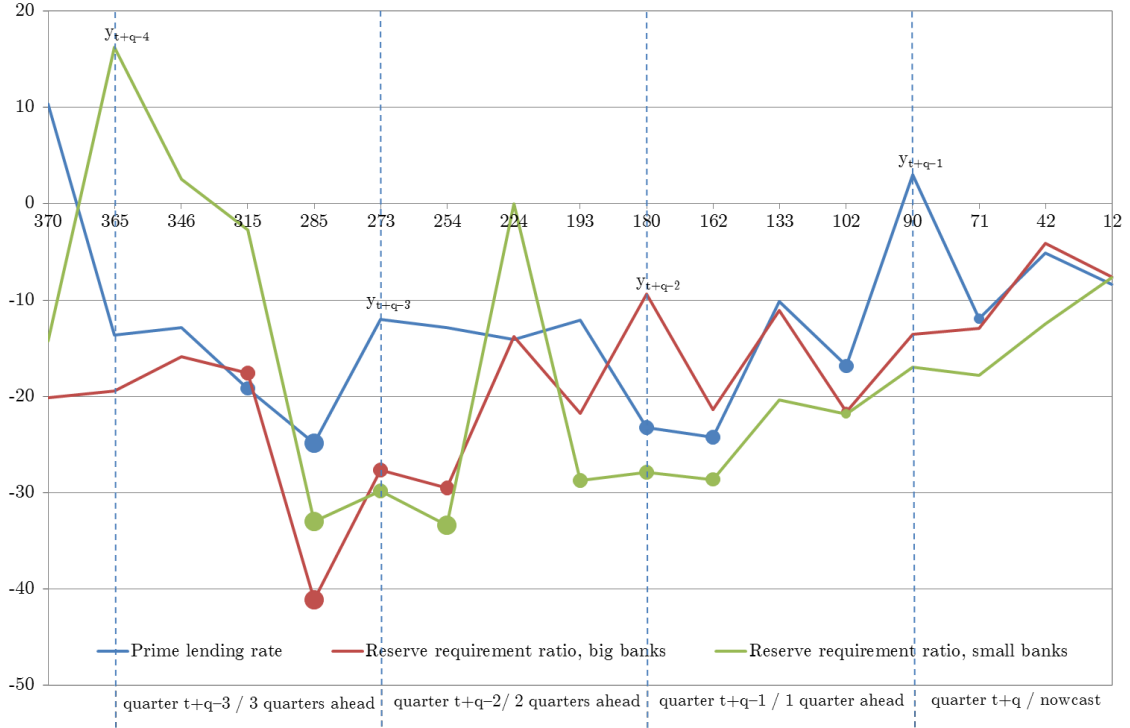
4.4 Interest rates, reserve requirements, money supply and credit growth

In a next step, we study the predictive power of several monetary policy instruments under surveillance or control of the People's Bank of China (PBC): the so called prime lending rate (= key short-term interest rate in China), the reserve requirement ratio (RRR) for big banks and the RRR for small banks.²¹ The two ratios are identical until end 2008 and then deviate from each other. We take both ratios into the analysis because the gap between them has attracted some attention lately. The PBC might change the prime lending rate and reserve requirement ratios at every day of a month. We always collect the monthly end values only. Further, the

²¹The four big state owned commercial banks are the Bank of China, the China Construction Bank, the Industrial and Commercial Bank of China and the Agricultural Bank of China. They hold 43% of total assets in the banking sector.

PBC publishes several money supply aggregates, M0, M1 and M2, on a monthly basis. Money supply M2 receives a Bloomberg relevance rating of 87.1 (on a range from 0 to 100), which makes it the most observed macroeconomic time series for China after inflation and GDP growth. M0 and M1 receive much less attention with Bloomberg relevance ratings below 50. Another monetary variable that has received a lot of attention lately is total social financing (TSF). TSF comprises all kinds of new credit issuance, funding or liquidity provision by the financial sector, including banks, security firms, insurance companies, banks' off-balance sheet items, credit, bond and equity markets, to the real economy during a given period of time. The PBC constructed TSF to account for the fact that bank loans, the traditional measure for credit growth, are getting relatively less important for financing of the real economy. We include M0, M1, M2 as well as TSF in year-over-year growth rates and month-over-month growth rates upon seasonal adjustment. Releases for month m occur between the 5th and 9th of month $m + 1$ leading us to choose the 9th of each month $m + 1$ as the harmonized release date. All variables are summarized in Table 1.

Figure 9: Predictive power of prime lending rate and bank reserve requirement ratios for GDP growth: Relative change in RMSFE

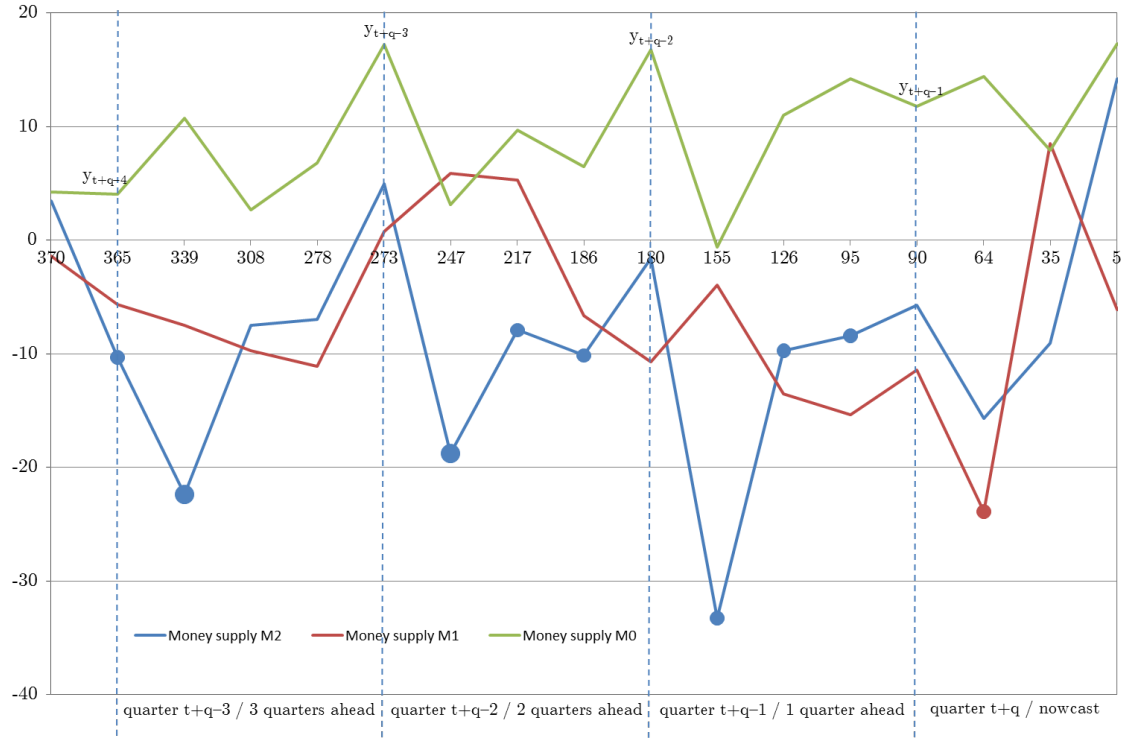


x -axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. y -axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including either the prime lending rate, the big bank reserve requirement ratio or the small bank reserve requirement ratio and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth are indicated by y_{t+q-i} . Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =) 408$ separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 5 for shares of different weight schemes in the 408 optimal models.

Figure 9 presents results for the out-of-sample range Q1 2008 to Q4 2013. The prime lending rate series adds substantial predictive power for forecasting quarterly

GDP growth during 10 to 9 months (315 and 285 days) and again 6 to 5 months (180 and 162 days) prior to GDP release. However, this finding proves not robust for out-of-sample ranges starting *Q1 2009* or *Q1 2010*: the RMSFE reductions get much lower and loose significance. Thus, the effect of interest rate conditions on economic activity seems to have weakened in recent times. The RRR for big banks adds substantial predictive power around 10 to 8 months (315 to 254 days) before GDP release. But again this finding vanishes partly or fully when the out-of-sample range starts *Q1 2009* or *Q1 2010*. The RRR for small banks also adds substantial predictive power during about 9 to 5 months (285 to 162 days) before GDP release. Importantly, this finding proves robust for out-of-sample ranges starting *Q1 2009* or *Q1 2010*. As one might expect the effects of the prime lending rate and the RRRs on GDP growth are *negative*.²² Thus, lower (higher) interest rates or RRRs today are associated with higher (lower) GDP growth in the future. Our results suggest that the PBC policy from 2009 onwards to allow small banks to hold lower reserves as compared to big banks had indeed a stimulating effect on the economy.

Figure 10: Predictive power of money supply for GDP growth: Relative change in RMSFE



x-axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. *y*-axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including either money supply M0, M1 or M2 and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth are indicated by y_{t+q-i} . Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (*Q1 2008* to *Q4 2013*) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 6 for shares of different weight schemes in the 408 optimal models.

Figure 10 presents results for money supply in y-o-y growth rates (which dominates results for m-o-m growth rates) for the out-of-sample range *Q1 2008* to *Q4 2013*.

²²This can be inferred from the signs of the β - and ω -terms in Equation (1) (not shown here).

Money supply M2 adds substantial predictive power except for days of lagged GDP releases (days 273, 180 and 90) and except for the nowcast range. The finding remains robust with out-of-sample ranges starting Q1 2009 or Q1 2010. The effect of money supply on GDP growth is *positive*, hence, higher (lower) money supply today is associated with higher (lower) GDP growth in the future. In contrast, neither M0 nor M1 help forecasting GDP growth.²³ Total social financing turns out to also be a poor predictor of GDP growth (results not shown).

As before it may be argued that the set of monetary variables studied here are governed by a common latent process which could prove even more helpful for predicting GDP growth. To test this, we extract the first common factor from all aforementioned time series and employ it in our MIDAS forecasting exercise. The factor series turns out to be indeed a very valuable predictor 10 to 3.5 months ahead of GDP release with relative RMSFE changes between -40% and -15% . RMSFE reductions are lower but still robust when the out-of-sample range starts Q1 2009 or Q1 2010.

To sum up, while China's key short-term interest rate, the so called prime lending rate, as well as the reserve requirement ratio for big banks used to be quite valuable for forecasting quarterly GDP growth in the past, their leading properties have vanished since 2009. In contrast, the reserve requirement ratio for small banks, which is uncoupled from the big banks reserve requirement ratio since 2009 only, proves to still be a robust predictor of GDP growth at forecast horizons of 9 to 5 months. Further, money supply M2 adds substantial predictive power for forecasting GDP growth, except for the nowcast range and for days directly after lagged GDP releases which render the predictive content of earlier money supply releases irrelevant.

4.5 Do monthly industrial production and PMI help nowcasting GDP growth? It depends.

The industry sector makes up about 48% of Chinese GDP (average over Q1 2000 to Q3 2013). The NBS publishes the year-over-year growth rate of real industrial production (IP growth) in its Monthly Report on Industrial Production Operation.²⁴ IP growth receives a Bloomberg relevance rating of 80.7 (on range from 0 to 100) making it the sixth most observed macroeconomic time series for China. While industrial production on a *quarterly* basis should not provide new information for quarterly GDP growth in addition to past values of quarterly GDP growth, the *monthly* IP growth time series might indeed help prediction. In particular, IP growth of month 1, 2 and 3 of quarter $t+q$ might contain valuable information for nowcasting GDP of quarter $t+q$. Further, past GDP growth affects current GDP growth via, amongst

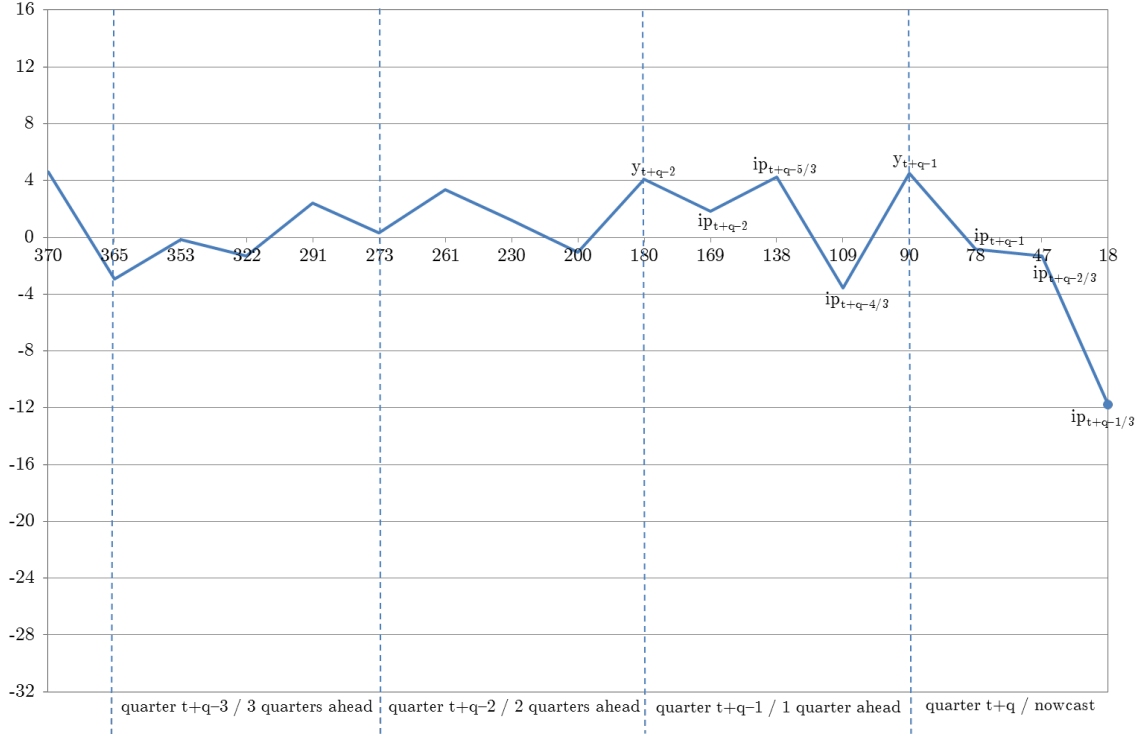
²³The comparatively strong RMSFE reductions for M1 during days 126 to 64 are not robust when using alternative out-of-sample ranges.

²⁴Next to the aforementioned time series, the NBS publishes the following other monthly industrial production series: monthly year-to-date real growth and monthly month-over-month real growth. We do not include these series in our analysis. The NBS publishes neither an industrial production series in real or nominal levels nor a real index of industrial production.

others, the amount of statistical overhang. As a consequence, the three monthly IP growth observations of quarter $t + q - i$ with $i = 1, \dots, \infty$ might also help predicting GDP of quarter $t + q$ as long as GDP of quarter $t + q - i$ is not yet available.

To see how different release schedules affect forecast results we choose two alternative release dates for IP growth of month m : the 9th and the 25th day of month $m + 1$, the former/latter being the earliest/latest release day during the past eight years. Importantly, IP growth of the *third* month of quarter $t + q - i$ is available before GDP growth of quarter $t + q - i$ *only under the early release schedule, but not under the late release schedule*.²⁵

Figure 11: Predictive power of late release IP growth for GDP growth: Relative change in RMSFE



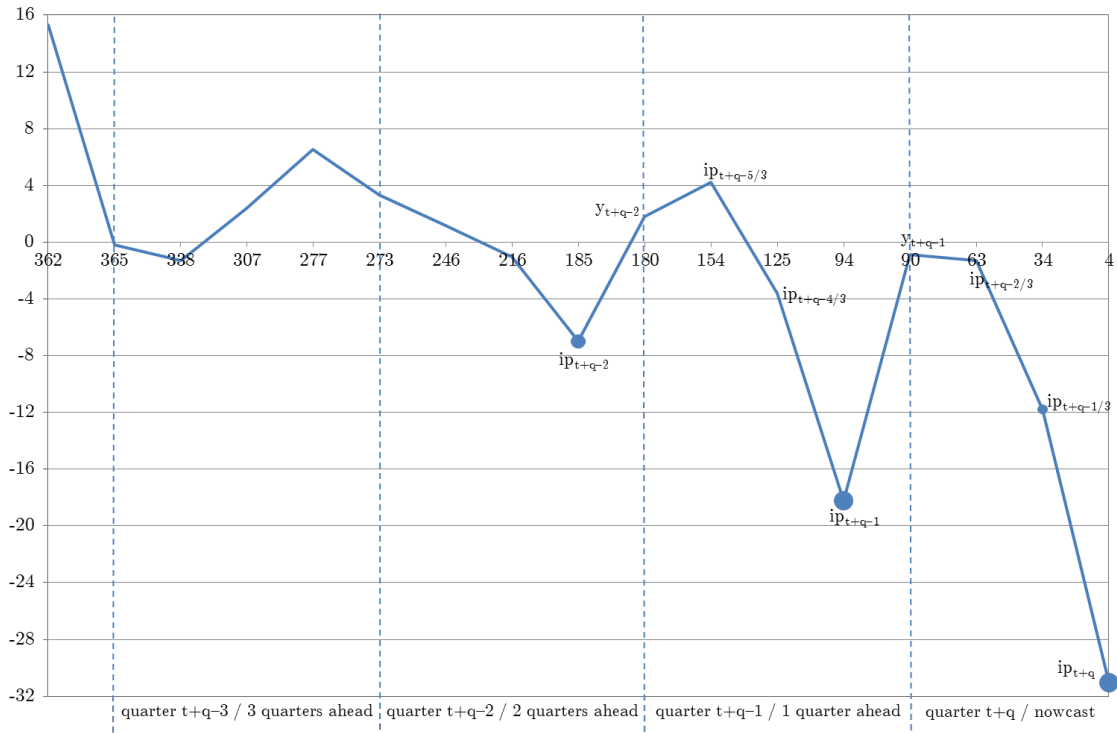
x-axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. *y*-axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including late release IP growth and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth and IP growth are indicated by y_{t+q-i} or $ip_{t+q-j*1/3}$. Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 7 for shares of different weight schemes in the 408 optimal models.

Figure 11 displays the relative RMSFE changes of the ADL-(U-)MIDAS model including monthly IP growth *based on the late release schedule*. Only IP growth of the second month of quarter $t + q$, $ip_{t+q-1/3}$, which gets released 18 days prior to GDP growth of quarter $t + q$ according to the late release schedule, seems to carry some

²⁵GDP growth of quarter $t + q - i$ always gets released on the 13th day of quarter $t + q - i + 1$ according to our harmonization (cf. Section 3).

valuable information in addition to past values of GDP growth (relative RMSFE change -11.8%). Notably, IP growth of the third month of quarter $t + q$, ip_{t+q} , is not included in the figure, because it is available only after the GDP growth release. We repeat the analysis for out-of-sample ranges starting Q1 2009 (= after the Great Recession) or Q1 2010 (= after the recovery following the Great Recession). For none of the two alternative samples the aforementioned finding proves to be robust (relative RMSFE changes 0% or -6% where Diebold and Mariano (1995) test does not reject null hypothesis of equal forecast accuracy at conventional levels of significance anymore).

Figure 12: Predictive power of early release IP growth for GDP growth: Relative change in RMSFE

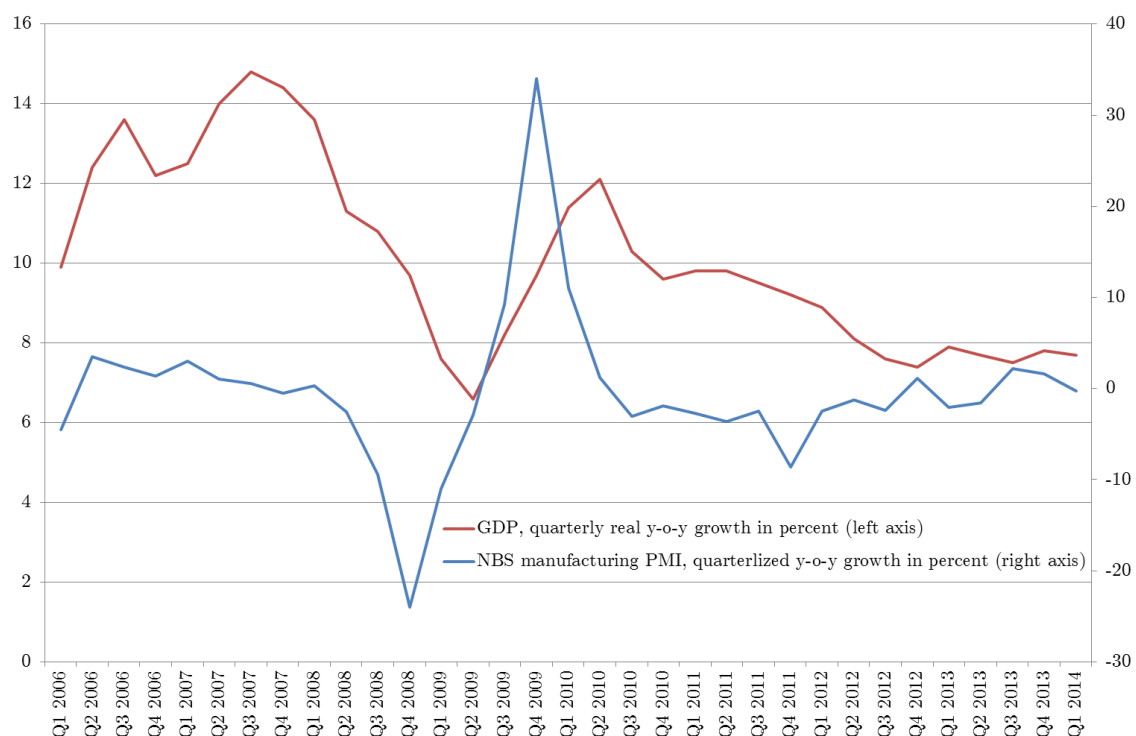


x-axis: Forecast horizon, i.e. day of forecast expressed in number of days prior to GDP growth data release. y-axis: Relative change in RMSFE, $\Delta RMSFE$, i.e. difference between (a) RMSFE of ADL-U-MIDAS model including early release IP growth and (b) RMSFE of ADL benchmark model in percent of RMSFE of benchmark model (cf. Equation (6)). Releases of (lagged) GDP growth and IP growth are indicated by y_{t+q-i} or $ip_{t+q-j \times 1/3}$. Small (medium size, big) dots indicate that, according to the Diebold and Mariano (1995) one-sided test, predictive power of ADL-U-MIDAS model is higher than predictive power of ADL benchmark model at 10% (5%, 1%) level of significance. Section 6.4 provides detailed explanation of the test. 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations for each of the three aforementioned indicators according to the rolling window model selection procedure outlined in Section 2. See Appendix 6.5, Figure 7 for shares of different weight schemes in the 408 optimal models.

Is monthly IP growth more valuable for forecasting quarterly GDP growth when it is released early? Figure 12 displays the relative RMSFE changes of the ADL-(U-)MIDAS model including monthly IP growth *based on the early release schedule*. IP growth of the third month of quarter $t + q$, ip_{t+q} , which gets released 4 days prior to the release of GDP growth of quarter $t + q$ according to the early release schedule, contains substantial predictive power (relative RMSFE change -31.0%). While ip_{t+q} can be used for *backcasting* only because it is in any case released after the end of the quarter, the early release schedule also changes the *forecasting prop-*

erties of the IP growth time series: Both, IP growth of the third month of quarter $t + q - 1$, ip_{t+q-1} , and IP growth of the third month of quarter $t + q - 2$, ip_{t+q-2} , carry valuable predictive information (relative RMSFE changed -18.2% or -7.0%). In contrast, when IP growth is released according to the late release schedule (see Figure 11), GDP growth of quarter $t + q - 1$ ($t + q - 2$) is released before ip_{t+q-1} (ip_{t+q-2}) rendering the predictive information in ip_{t+q-1} (ip_{t+q-2}) irrelevant and depleting the forecasting power of the IP growth series. Our early release schedule findings are even more accentuated when the out-of-sample range starts Q1 2009 or Q1 2010 instead of Q1 2008.

Figure 13: GDP and NBS manufacturing PMI



Since 2005 the NBS publishes a purchasing managers' index (PMI) for the manufacturing sector.²⁶ PMI for month m gets always released on the first day of month $m + 1$. Figure 13 depicts quarterlized PMI in y-o-y growth rates together with GDP growth. As can be seen from the figure, the PMI series announced the GDP growth turning points in Q2 2009 and Q2 2010 with a lead of 2 quarters. Despite this promising finding, our mixed frequency forecast exercise yields that neither the additional predictive power of the monthly PMI series in levels, y-o-y nor m-o-m growth rates is very strong (in addition to past GDP growth values). For none of

²⁶The PMI is a composite index based on several subindices (see [NBS PMI](#)). The NBS does not publish a service sector PMI. An alternative set of PMIs is published by HSBC in cooperation with Markit Economics (see [HSBC-Markit PMI](#)). The series is not openly available and disregarded in this analysis.

the forecast horizons, the relative RMSFE change is substantially below -5% . Also, the [Diebold and Mariano \(1995\)](#) test never rejects the null hypothesis that the ADL-(U-)MIDAS model and the ADL benchmark model forecast equally accurately at conventional levels of significance. The negative finding is confirmed for alternative out-of-sample starts (Q1 2008, Q1 2009 or Q1 2010).

To sum up, monthly industrial production growth data can be quite valuable for now- or even forecasting quarterly GDP growth. However, this is only the case if the *third* monthly IP growth observation of a quarter (hence, the value of March, June, September or December) is available before GDP growth of that quarter. In contrast, when GDP growth of a quarter is released before the third monthly IP growth observation of that quarter, it renders the predictive content in the latter observation irrelevant and depletes both the now- and the forecasting power of the IP growth series. Further, we find that – despite a remarkable 2-quarters lead for the GDP growth turning points in Q2 2009 and Q2 2010 – the monthly NBS manufacturing PMI series is of little help in now- or forecasting quarterly GDP growth.

4.6 Monthly information on consumption, investment, trade, electricity usage and freight traffic

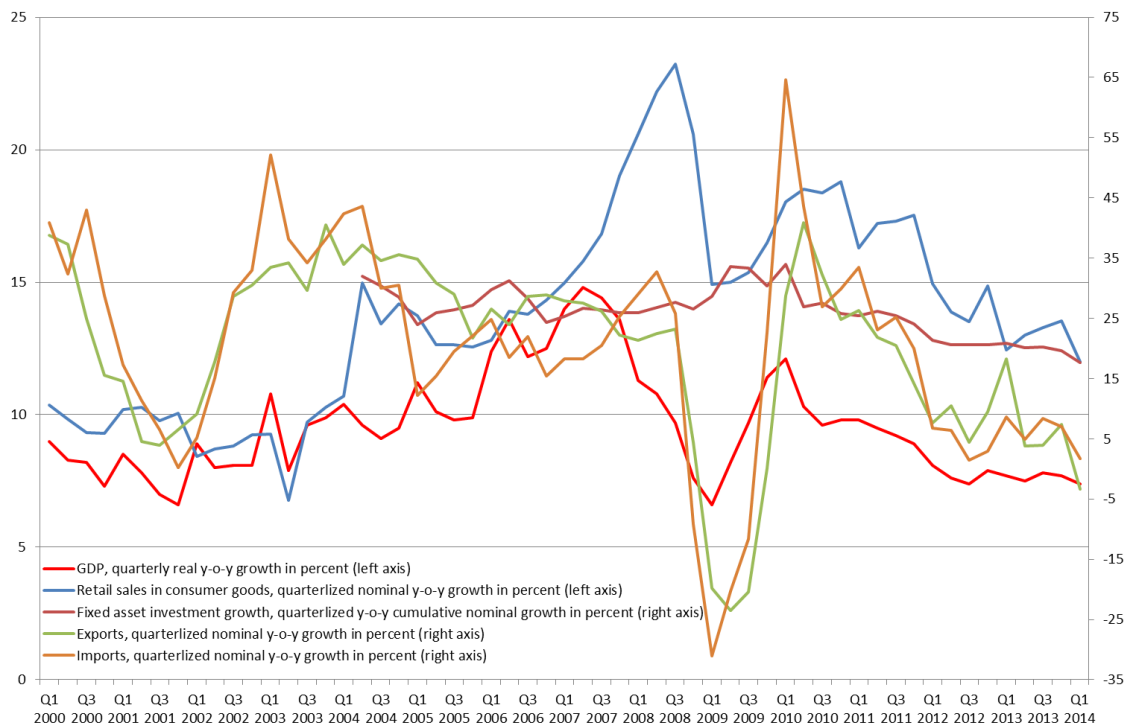
GDP consists of consumption, investment, exports and imports for each of which monthly indicators are available. It might be interesting to know whether this monthly information can help forecasting or nowcasting quarterly GDP growth. The NBS publishes, on a monthly basis, retail sales in consumer goods (in levels and year-over-year growth rates, nominal, not seasonally adjusted) as well as investment in fixed assets (in levels and y-o-y growth rates, nominal, cumulative, excluding rural households). To our knowledge, these are the best available indicators for monthly consumption or investment. The series receive Bloomberg relevance ratings of 77.4 or 58.1 (on a range from 0 to 100) indicating that market observers consider them as being of medium or rather low importance.²⁷ The release dates of the series correspond to the release dates of monthly IP growth. Monthly exports and imports (both nominal, not seasonally adjusted) come from the Chinese General Administration of Customs (Bloomberg relevance ratings of 67.7 or 64.5). Releases usually occur within the first week after the reference month. For all aforementioned series, we abstain from constructing real figures as appropriate (monthly) deflators are not available.

Figure 14 plots *quarterlized* y-o-y growth in retail sales, investment, exports and imports together with GDP growth.²⁸ All series have a certain correlation with

²⁷Surprisingly, transformation of the retail sales levels series into y-o-y growth rates does not always coincide with the y-o-y growth series provided by the NBS itself which leaves us with two alternative y-o-y growth series (cf. the discussion in [Maier, 2011](#)).

²⁸Here we display only the y-o-y retail sales growth series as provided by the NBS itself. Still, our forecasting exercise includes both alternative y-o-y retail sales growth series.

Figure 14: Growth in GDP, retail sales, investment, exports and imports



GDP growth.²⁹ But do the *monthly* time series really contain predictive information for quarterly GDP growth? We apply our MIDAS forecasting exercise outlined in Section 2 and find that none of the monthly series (robustly) improves forecasts or nowcasts of quarterly GDP growth. Repeating the analysis for month-over-month growth rates upon seasonal adjustment does not alter this negative finding.

According to U.S. diplomatic cables published by WikiLeaks in 2010, during a dinner with the U.S. ambassador in 2007 Chinese prime minister Li Keqiang — then the Communist Party Secretary of Liaoning province — held that China’s (provincial) GDP is “man made” and “for reference only”. He preferred to look at two alternative indicators to keep track of economic activity: monthly electricity usage in industrial production (in billions of kilowatts per hour) and monthly railway freight transport turnover (either in metric tons or in metric tons per kilometre).³⁰ The indicators have a certain correlation with GDP growth while being more choppy in times of down- and upturns (correlation coefficients of 0.62, 0.33 or 0.41). Might they be helpful in now- or forecasting quarterly GDP growth despite Li seeing them as alternative indicators? Applying our MIDAS forecasting exercise, we find that neither y-o-y growth in monthly electricity usage nor y-o-y growth in monthly freight traffic in metric tons or metric tons per kilometre robustly improve predictions. It-

²⁹Correlation coefficients are between 0.35 and 0.45. Notably, the cumulative nature of the investment series prevents it from moving down with the other time series during the Great Recession.

³⁰Li further mentioned bank loans. See the discussion on total social financing, which includes bank loans, in Section 4.4.

erating the analysis for m-o-m growth rates upon seasonal adjustment does not alter this finding.

It may be argued that the aforementioned GDP subcomponents – industrial production, retail sales, exports, imports, electricity usage and freight traffic – are governed by a common latent process which is more helpful for predicting GDP growth than the individual indicators themselves. To test this we extract the first common factor from the six indicators and employ it in our MIDAS forecasting exercise.³¹ It turns out that the monthly factor series does not (robustly) improve forecasts or nowcasts of quarterly GDP growth.

To sum up, despite certain correlation on a quarterly basis neither monthly retail sales in consumer goods, investment, export nor import figures prove to be valuable for now- or forecasting quarterly GDP growth. Equally, neither monthly electricity usage nor monthly freight traffic help predicting quarterly GDP growth. This notwithstanding the series might still be useful as *alternative* indicators for economic activity in China.

5 Conclusion

Professional forecasting comprises two distinct tasks. First, the forecaster collects a large set of data, develops and employs several alternative forecasting models and finally condenses everything into a forecast. Second, the forecaster presents and justifies the forecast to the public, his or her peers, his or her superiors etc. Notably, the forecaster cannot simply justify his forecast by stating, “I used a large amount of data and various different models and this is what came out”. Rather, he or she needs to build a coherent story around the forecast and cite selected indicators whose recent development supports the forecast. Often however, there exist other indicators whose recent development does not quite support or even contradict the forecast. The forecaster then has to argue why the latter indicators are either not relevant in the present context or not relevant at all. Here lies the main contribution of our paper. We evaluate and compare the predictive power of a variety of monthly macroeconomic indicators which are generally considered to have predictive content and/or which get public attention when it comes to forecasting quarterly Chinese GDP growth. We iterate the evaluation over forecast horizons from 370 days to 1 day prior to GDP release and track the release days of the indicators so as to only use information which is actually available at the respective day of forecast (pseudo-real time setup). The procedure allows us to detect how useful a specific indicator is at a specific forecast horizon as compared to other indicators. Our results might be relevant for experts who need to know which new data release is really valuable and which new data release has less valuable predictive content or contains just old information. We complement our indicator-wise analysis by evaluating the

³¹We abstain from including the investment series because it merely moves with the business cycle (due to its cumulative nature). Further, we include the metric tons per kilometre freight traffic series instead of the metric tons freight traffic series as the former performs a little bit better in the MIDAS forecasting exercise.

joint predictive power of a variety of indicators via factor analysis (see also [Curran and Funke, 2006](#), [Mehrotra and Pääkkönen, 2011](#), [Yiu and Chow, 2011](#) and [Maier, 2011](#)). Our day-specific pseudo-real time setup allows us to deal with ragged edge issues and to study how forecast errors evolve as ever new data get released over time.

It strongly depends on the exact forecast horizon (= the number of days before GDP release) whether and to which extent a monthly indicator provides new information that helps improve quarterly (year-over-year) GDP growth forecasts. Notwithstanding this general lesson, we find substantial differences between individual indicators. First, the OECD leading indicator (trend restored version in year-over-year growth rates) adds substantial predictive power. Its value is highest at forecast horizons of 5 to 4 months, where it reduces forecast errors by 31% to 37% as compared to an ADL benchmark model. Importantly, the OECD leading indicator outperforms its counterparts published by the Conference Board (CB) and by the China Economic Monitoring and Analysis Center (CEMAC) in cooperation with Goldman Sachs (GS). A surprise additional finding is that the OECD leading indicator and the CEMAC-GS leading indicator have even better – or at least not worse – *nowcasting* properties than several coincident indicators published by OECD, CEMAC-GS or CB. Second, the Shenzhen Stock Exchange Composite Index and the Shanghai Stock Exchange Composite Index are quite valuable predictors; and despite covering a lower share of the Chinese economy, the former index tends to outperform the latter. In contrast, Hong Kong stock exchange indices are only of very little help. Third, except for the nowcast range consumer price inflation adds substantial predictive power for forecasting quarterly GDP growth. Its value is highest 11 to 6 months prior to GDP release with forecast error reductions of 26% to 36% compared to an ADL benchmark model. In contrast, producer price inflation does not seem to help much. Fourth, while China’s key short-term interest rate, the so called prime lending rate, as well as the reserve requirement ratio for big banks used to be quite valuable for forecasting in the past, their leading properties have vanished since 2009. In contrast, the reserve requirement ratio for small banks, which is uncoupled from the big banks reserve requirement ratio since 2009 only, proves to still be a robust predictor at forecast horizons of 9 to 5 months. Fifth, money supply M2 adds substantial predictive power, except for the nowcast range and for days directly after lagged GDP publications which render the predictive information in earlier M2 releases irrelevant. In contrast, money supply M0 and M1 turn out to be not very helpful. Sixth, a common factor series extracted from all aforementioned monetary variables turns out to be a robust and very valuable predictor at forecast horizons of 10 to 3.5 months. Seventh, industrial production (IP) growth can be quite valuable for now- or even forecasting quarterly GDP growth. However, this is only the case if the *third* monthly IP growth observation of a quarter (hence, the value of March, June, September or December) is available before GDP growth of that quarter. In contrast, when GDP growth of a quarter is released before the third monthly IP growth observation of that quarter, it renders the predictive content in the latter observation irrelevant and depletes both the now- and the forecasting power of the IP growth series. Eighth, despite a remarkable 2-quarters lead for the GDP growth turning points in Q2 2009 and Q2 2010 the manufacturing purchasing

managers' index of the Chinese National Bureau of Statistics is of little value in now- or forecasting quarterly GDP growth. Ninth, monthly retail sales, investment, exports and imports do not help nowcasting, neither individually nor jointly (common factor analysis). Tenth, while having received quite some attention lately, electricity usage and freight traffic turn out to be poor predictors. This notwithstanding the series might still be useful as *alternative* indicators for economic activity in China.

Recently, the literature has made some progress in modeling mixed frequency vector autoregressions (MF-VAR) in a MIDAS or a state space framework ([Ghysels, 2012a](#), [Schorfheide and Song, 2013](#), e.g.). An avenue for further research could be to jointly forecast several Chinese macroeconomic time series, like GDP growth, inflation and interest rates, using MF-VAR.

References

- ANDREOU, E., GHYSELS, E. and KOURTELLOS, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, **158**, 246–261. [back to page 4]
- ARMESTO, M. T., ENGEMANN, K. M. and OWYANG, M. T. (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review*, **92(6)**, 521–36. [back to page 4]
- BAI, J., GHYSELS, E. and WRIGHT, J. H. (2013). State space models and MIDAS regressions. *Econometric Reviews*, **32(7)**, 779–813. [back to page 4]
- BANBURA, M., DOMENICO, G. and LUCREZIA, R. (2011). Nowcasting. In M. P. Clements and D. F. Hendry (eds.), *Oxford Handbook on Economic Forecasting*, Oxford University Press: Oxford. [back to page 4]
- CHOW, G. (2006). Are chinese official statistics reliable? *CESifo Economic Studies*, **52**, 396–414. [back to page 11]
- CLARK, T. E. and WEST, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, **138**, 291–311. [back to page 9, 41]
- CLEMENTS, M. P. and GALVÃO, A. B. (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics*, **26(4)**, 546–554. [back to page 4]
- and GALVÃO, A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*, **24**, 1187–1206. [back to page 4]
- CURRAN, D. and FUNKE, M. (2006). Taking the temperature: Forecasting GDP growth for mainland China. *BOFIT Discussion Papers*, **6/2006**, Bank of Finland. [back to page 4, 31]
- DIEBOLD, F. X. and MARIANO, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, **13(3)**, 253–263. [back to page 9, 14, 17, 19, 22, 23, 25, 26, 28, 39, 41]
- DIRON, M. (2008). Short-term forecasts of euro area real GDP growth: An assessment of real-time performance based on vintage data. *Journal of Forecasting*, **27**, 371–390. [back to page 4, 8]
- FORONI, C. and MARCELLINO, M. (2013). A survey of econometric methods for mixed-frequency data. *Norges Bank Working Paper*, **2013/06**, Norges Bank. [back to page 4]
- , — and SCHUMACHER, C. (2012). U-MIDAS: MIDAS regressions with unrestricted lag polynomials. *CEPR Discussion Papers*, **8828**, C.E.P.R. Discussion Papers. [back to page 4, 7]

- FRANSES, P. H. and MEES, H. (2013). Approximating the DGP of China’s quarterly GDP. *Applied Economics*, **45**(24), 3469–3472. [back to page 5]
- GHYSELS, E. (2012a). Macroeconomics and the reality of mixed frequency data. *mimeo*, University of North Carolina. [back to page 5, 32]
- (2012b). Matlab toolbox for mixed sampling frequency data analysis using MIDAS regression models. *mimeo*, **version 3.0**, University of North Carolina. [back to page 36]
- , SINKO, A. and VALKANOV, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, **26**(1), 53–90. [back to page 4, 5]
- GIACOMINI, R. and WHITE, H. (2006). Tests on conditional predictive ability. *Econometrica*, **74**(6), 1545–1578. [back to page 9, 39]
- GIANNONE, D., REICHLIN, L. and SMALL, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, **55**(4), 665–676. [back to page 4]
- GOLINELLI, R. and PARIGI, G. (2007). The use of monthly indicators to forecast quarterly GDP in the short run: An application to the G7 countries. *Journal of Forecasting*, **26**, 77–94. [back to page 4, 8]
- HARVEY, D. I., LEYBOURNE, S. J. and NEWBOLD, P. (1997). Testing the equality of prediction mean squared errors. *Journal of International Forecasting*, **13**, 281–291. [back to page 9, 40]
- , — and — (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, **16**(2), 254–259. [back to page 9, 40]
- KLEIN, L. R. and MAK, W. (2005). Initial steps in high-frequency modeling of China. *Business Economics*, pp. 11–14. [back to page 4]
- KOCH-WESER, I. N. (2013). *The Reliability of China’s Economic Data: An Analysis of National Output*. U.S.-China Economic and Security Review Commission Staff Research Project. [back to page 11]
- KUZIN, V., MARCELLINO, M. and SCHUMACHER, C. (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *International Journal of Forecasting*, **27**, 529–542. [back to page 4]
- LIU, L.-G., ZHANG, W. and SHEK, J. (2007). A real activity index for Mainland China. *Hong Kong Monetary Authority Working Paper*, **07/2007**. [back to page 5]
- MAIER, P. (2011). Mixed frequency forecasts for Chinese GDP. *Bank of Canada Working Paper*, **2011-11**. [back to page 4, 8, 28, 31]

- MAK, W. (2009). Forecasting the sustainability of China's economic performance: Early twenty-first century and beyond. In L. R. Klein (ed.), *The Making of National Economic Forecasts*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing. [back to page 4]
- MARIANO, R. S. and MURASAWA, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, **18**(4), 427–443. [back to page 4]
- and — (2010). A coincident index, common factors, and monthly real GDP. *Oxford Bulletin of Economics and Statistics*, **72**(1), 27–46. [back to page 4]
- MEHROTRA, A. and PÄÄKKÖNEN, J. (2011). Comparing China's GDP statistics with coincident indicators. *Journal of Comparative Economics*, **39**, 406–411. [back to page 2, 4, 11, 31]
- ORLIK, T. (2012). *Understanding China's Economic Indicators: Translating the Data into Investment Opportunities*. Upper Saddle River, New Jersey: FT Press Science. [back to page 5]
- SCHORFHEIDE, F. and SONG, D. (2013). Real-time forecasting with a mixed-frequency VAR. *NBER Working Papers*, **19712**, National Bureau of Economic Research. [back to page 32]
- STOCK, J. H. and WATSON, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, **23**, 405–430. [back to page 5]
- and — (2012). *Introduction to Econometrics*. Essex: Pearson Education, 3rd Global Edition. [back to page 10]
- WOHLRABE, K. (2009). *Forecasting with Mixed Frequency Time Series Models*. Ph.D. Thesis, Ludwig Maximilians University Munich. [back to page 4]
- YIU, M. S. and CHOW, K. K. (2011). Nowcasting Chinese GDP: Information content of economic and financial data. *HKIMR Working Paper*, **04/2011**, Hong Kong Institute for Monetary Research. [back to page 4, 31]

6 Appendix

6.1 Some remarks on the MIDAS weight functions

In case of the beta probability density function and in case of the exponential Almon lag polynomial, the ADL-MIDAS model in Equation (1) becomes highly non-linear and, hence, has to be estimated via a non-linear least squares approach.³² In contrast, in case of the non-exponential Almon lag polynomial, the ADL-MIDAS model in Equation (1) can be transformed in linear form and estimated via ordinary least squares (OLS). For this, re-write the sums in Equation (1) in vector form and re-write Equation (3) in matrix form:³³

$$\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \vdots \\ \vdots \\ \omega_J \end{bmatrix} = M\Theta = \begin{bmatrix} 1 & 1 & 1 & \dots & \dots & 1 \\ 1 & 2 & 4 & \dots & \dots & 2^P \\ 1 & 3 & 9 & \dots & \dots & 3^P \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & J & J^2 & \dots & \dots & J^P \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \vdots \\ \theta_P \end{bmatrix}. \quad (7)$$

Then, replace the ω -vector in the vector form of Equation (1) by the right-hand side of Equation (7) and multiply the high-frequency data vector (= x -vector) with matrix M . Now, we are left with a linear regression equation with P unknown parameters, $\theta_0, \dots, \theta_P$. This can be estimated via OLS.

6.2 Model selection and forecasting procedure

We select the last $T - \tau$ quarterly observations of y_t as the out-of-sample range. Then, we estimate parameters of the benchmark model (5) using OLS based on the in-sample observations 6³⁴ to $T - \tau - 1$ for $I = 1, 2, 3, 4$ and 5. Out of these five models, we select the one with the lowest in-sample Bayesian information criterion,³⁵

$$\ln \left(\frac{SSR(I)}{T - \tau - 1} \right) + (I + 1) \frac{\ln(T - \tau - 1)}{T - \tau - 1}, \quad (8)$$

where $SSR(I)$ denotes the sum of squared residual of the estimated I -lag ADL model. Then, we re-estimate the parameters of the selected model now based on the in-sample observations I^* to $T - \tau - 1$, where I^* denotes the lag length of the selected model (= the optimal lag length of the ADL term). Thereafter, we calculate

³²Drawing from Ghysels (2012b), we use the Matlab function *fminunc*, an algorithm that is suitable for unconstrained nonlinear optimization.

³³See Ghysels (2012b, p. 7). Note the difference between our M -matrix and Ghysels' M -matrix.

³⁴Why 6? Because information criterion comparison of different models should be based on always the same data set.

³⁵We employ the Bayesian information criterion (= Schwarz information criterion) instead of the Akaike information criterion. See Diebold (2007, p. 85f) for a comparison of the two criteria. In fact, Diebold recommends using the model selected by the Bayesian information criterion if the two criteria opt for different models.

the forecast,

$$\hat{y}_{t+q|t-1}^{BM} = \gamma_0 + \sum_{i=1}^{I^*} \gamma_i y_{t-i}$$

for $t+q = T-\tau$ and the respective residual, $e_{T-\tau}^{BM} = y_{T-\tau} - \hat{y}_{T-\tau|T-\tau-q-1}^{BM}$. Notably, for $q > 1$ direct forecasts are used (as compared to iterated forecasts). After this, we estimate, based on the in-sample observations I^* to $T-\tau-1$, the parameters of model (4) as well as the parameters and weights of model (1) for the 11 alternative models given in Section 2,³⁶ each for $J = 1$ to 12 where always $I = I^*$. Optimization occurs via a non-linear least squares approach or via OLS in cases of MIDAS with non-exponential Almon lag polynomial, U-MIDAS and bridge equations (see Footnote 32). Out of the 132 model specifications, we select the one with the lowest in-sample Bayesian information criterion³⁷ for which we then calculate the forecast,

$$\hat{y}_{t+q|t-1,t+q-m*1/3}^{OM} = \alpha_0 + \sum_{i=1}^{I^*} \alpha_i y_{t-i} + \beta \sum_{j=m}^{J+m-1} \omega_{j-m+1} x_{t+q-j*1/3} \quad (9)$$

for $t+q = T-\tau$ and the respective residual, $e_{T-\tau}^{OM} = y_{T-\tau} - \hat{y}_{T-\tau|T-\tau-q-1,T-\tau-m*1/3}^{OM}$. The forecast $\hat{y}_{t+q|t-1,t+q-m*1/3}^{OM}$ is $q+1$ quarters ahead of the latest y -variable publication and m quarters ahead of the latest x -variable publication. We iterate the above procedure for all other periods of the out-of-sample range, $T-\tau+1, \dots, T$ with a rolling window procedure. I.e. we apply the above procedure based on the in-sample observations $6+l$ to $T-\tau-1+l$ or I_t^*+l to $T-\tau-1+l$, resulting in the forecast errors $e_{T-\tau+l}^{OM}$ and $e_{T-\tau+l}^{BM}$ for $l = T-\tau+1, \dots, T$. Notably, the optimal lag length for the ADL term as well as the optimal weight function and the optimal lag length for the MIDAS/U-MIDAS/bridge term can be different for each iteration. Then, we calculate the root mean squared forecast errors (RMSFE) for the optimal model and the benchmark model as well as the relative ratio of both in percent,

$$100 * \left(\frac{\sqrt{\sum_{t=T-\tau}^T (e_t^{OM})^2} - \sqrt{\sum_{t=T-\tau}^T (e_t^{BM})^2}}{\sqrt{\sum_{t=T-\tau}^T (e_t^{BM})^2}} \right). \quad (10)$$

Finally, we repeat the whole procedure for selected days d within 370 days to 1 day before release of y_{t+q} .³⁸ Importantly, the choice of q and m is determined by d . For instance, suppose we want to forecast y_{t+q} and y_{t+q-1} is always released between 90 and 92 days before the release of y_{t+q} and $x_{t+q-3/3}$ is always released between 85 and 87 days before the release of y_{t+q} . Then, in case we forecast y_{t+q} at day 95 before

³⁶[1.1–1.4] MIDAS with restricted or unrestricted beta probability density function with zero or non-zero last lag, [2] MIDAS with exponential Almon lag polynomial, [3.1–3.4] MIDAS with non-exponential Almon lag polynomial of order 1 to 4, [4] U-MIDAS and [5] bridge equations.

³⁷See Equation (8) where $I+1$ is to be replaced by the total number of estimated parameters in the respective ADL-MIDAS, ADL-U-MIDAS or bridge model. Again, information criterion comparison of the different models is based on always the same data set.

³⁸It would not make sense to forecast at each day $d = 1, \dots, D$, because not every day new data are available and, hence, the forecasts will not change every day.

its release, we have in Equation (1) $q = 1$ and $m = 4$. If however we forecast y_{t+q} at day 89 before its release, we have in Equation (1) $q = 0$ and $m = 4$. Further, in case we forecast y_{t+q} at day 84 before its release, we have in Equation (1) $q = 0$ and $m = 3$. We combine this day-specific forecasting procedure with a harmonization of publication dates over different quarters and years. This is discussed in Section 6.3.

6.3 Dealing with Chinese release schedules

The NBS releases GDP growth of quarter t in the month that follows directly after quarter t (e.g., GDP growth of the first quarter is published in April). The release day differs from quarter to quarter and year to year. Equally, the release days of the various monthly indicators used in this study differ from month to month and year to year. This poses a problem for our forecasting exercise. To give an example, suppose we want to evaluate the predictive power of monthly consumer price inflation for quarterly GDP *at the 85th day before publication of GDP growth*. For some quarters the inflation release of the third month of quarter $t - 1$ will be available at this day, whereas for other quarters it will not yet be available leaving us with lagged values only. Hence, the result of our evaluation depends on how often the inflation release of the third month of quarter $t - 1$ was available at the 85th day before GDP growth release. As a consequence, slight modifications of the GDP growth release schedule and/or the inflation release schedule in the future might dramatically change the number of quarters for which the inflation publication of the third month of quarter $t - 1$ is available at the 85th day before GDP growth release. Thus, slight modifications of future release schedules might dramatically change the predictive power of the entire inflation time series *at day 85 before GDP growth release – without changing the predictive power of the inflation time series in general*. This is certainly not what we want. To circumvent the aforementioned issue, we harmonize the release dates in the following way: We select the *earliest* day at which quarterly GDP growth got released during the out-of-sample range (Q1 2008 – Q4 2013) as the uniform release day for all GDP growth releases. I.e. according to our harmonization, GDP of quarter t always got/gets released on the 13th day of quarter $t + 1$. Further, we select the *latest* day at which monthly indicator $n = 1, \dots, N$ got/gets released during the period range M1 2006 to M12 2013 as the uniform release day for this indicator.³⁹ Since we select the *earliest* GDP growth release and the *latest* indicator release, our harmonization strategy is rather conservative. Hence, our valuation of indicator n 's predictive power at day d before the release of GDP growth can be considered as a lower bound of the actual predictive power at day d . Possibly, the predictive power is higher, because more information is available at day d than the conservative harmonization strategy allows us to take into account.

³⁹Why M1 2006? Forecast GDP growth of Q1 2008 (= the first out-of-sample period) at 370 days before GDP growth release (= our longest forecast horizon) using possibly up to 12 lags of indicator n . For some indicators we could not find historical release schedules back to 2006. In this case we base our choice of the harmonized release day on the available historical release schedules.

6.4 Predictive Accuracy Tests

In a first step, we follow the test approach proposed by [Diebold and Mariano \(1995\)](#). Define the loss differential by

$$r_t = (e_t^{BM})^2 - (e_t^{OM})^2. \quad (11)$$

Further, let \bar{r} be the sample mean of the loss differential, $\frac{1}{T} \sum_{t=1}^T r_t$, and let a heteroskedasticity- and autocorrelation-consistent (HAC) estimator of T times the asymptotic variance, $var(\bar{r})$, be

$$\begin{aligned} \hat{\sigma}_{r,T}^2 &= \frac{1}{T} \sum_{\tau=-(T-1)}^{(T-1)} \sum_{t=|\tau|+1}^T 1 \left(\frac{\tau}{(q+1)-1} \right) (r_t - \bar{r})(r_{t-|\tau|} - \bar{r}) \\ &= \frac{1}{T} \sum_{t=1}^T (r_t - \bar{r})^2 + \frac{2}{T} \left[\sum_{\tau=1}^{(T-1)} \sum_{t=\tau+1}^T 1 \left(\frac{\tau}{(q+1)-1} \right) (r_t - \bar{r})(r_{t-\tau} - \bar{r}) \right], \end{aligned}$$

where $q+1$ is the number of quarters between y_{t+q} (= the variable to be forecasted) and y_{t-1} (= the first available y-lag) with $q \in \{0, \dots, Q\}$ and where the lag window is defined by

$$\begin{aligned} 1 \left(\frac{\tau}{(q+1)-1} \right) &= 1 \quad \text{for } \left| \frac{\tau}{(q+1)-1} \right| \leq 1 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

This choice is motivated by the “familiar result that optimal $(q+1)$ -step ahead forecast errors at most q -dependent [...] [and] q -dependence implies that only q sample autocovariances need to be used in the estimation” ([Diebold and Mariano, 1995](#), p. 254). $\hat{\sigma}_{r,T}^2$ turned out to be not always positive semidefinite, a reason being that our sample is rather small. This is why we resort to the Newey-West version of the estimator,

$$\frac{1}{T} \sum_{t=1}^T (r_t - \bar{r})^2 + \frac{2}{T} \left[\sum_{\tau=1}^{(T-1)} \left[1 - \frac{\tau}{T} \right] \sum_{t=\tau+1}^T 1 \left(\frac{\tau}{(q+1)-1} \right) (r_t - \bar{r})(r_{t-\tau} - \bar{r}) \right], \quad (12)$$

which is positive semidefinite by construction and has the same consistency properties as $\hat{\sigma}_{r,T}^2$. Equation (11) leads us to state: The null hypothesis that the ADL benchmark forecasting model and the optimal forecasting model (MIDAS, U-MIDAS or bridge) are equally accurate is equivalent to the null hypothesis that $E[r_t] = 0$. Our alternative hypothesis is that the benchmark model is less accurate than the optimal model. This is equivalent to the alternative hypothesis that $E[r_t] > 0$. [Diebold and Mariano \(1995\)](#) and [Giacomini and White \(2006, Section 3.4\)](#) show that the corresponding test is based on the statistic

$$z_{dm} = \sqrt{T} \frac{\bar{r}}{\hat{\sigma}_{r,T}}.$$

⁴⁰Actually, [Diebold and Mariano \(1995\)](#) define the loss differential more generally by $g(e_{1,t}) - g(e_{2,t})$. $g(e_{i,t}) = e_{i,t}^2$ for $i = 1, 2$ is a popular choice.

A (one-sided) α percent significance level test rejects the null hypothesis if $z_{dm} > z^\alpha$, where z^α is the $(1 - \alpha)$ percent quantile of a standard normal distribution, $\mathcal{N}(0, 1)$.

Harvey *et al.* (1997) assess the behavior of the Diebold-Mariano test statistic and propose a modified version (see also Harvey *et al.*, 1998, p. 257). Drawing on Harvey *et al.* (1997) we employ the following statistic:

$$t_{mdm} = \left[\frac{T + 1 - 2(q + 1) + T^{-1}(q + 1)((q + 1) - 1)}{T} \right]^{1/2} \sqrt{T} \frac{\bar{r}}{\hat{\sigma}_{r,T}}.$$

The authors recommend comparison of t_{mdm} with critical values from the t -distribution rather than the standard normal distribution, where the difference gets only negligible as $T \rightarrow \infty$. Thus, a (one-sided) α percent significance level test rejects the null hypothesis if $t_{mdm} > t^\alpha$, where t^α is the $(1 - \alpha)$ percent quantile of a t -distribution with $T - 1$ degrees of freedom, t_{T-1} .

For an alternative to the above test, we follow the forecast encompassing approach proposed by Harvey *et al.* (1998). Consider the composite forecast

$$\hat{y}_t^C = (1 - \lambda)\hat{y}_t^{BM} + \lambda\hat{y}_t^{OM}, \quad 0 \leq \lambda \leq 1, \quad \forall t \quad (13)$$

as a weighted average of the ADL benchmark model forecast and the optimal model forecast. If ϵ_t denotes the composite forecast error we can rewrite Equation (13) as

$$e_t^{BM} = \lambda(e_t^{BM} - e_t^{OM}) + \epsilon_t.$$

Further, define

$$c_t = e_t^{BM}(e_t^{BM} - e_t^{OM}),$$

let \bar{c} be the sample mean of c_t , $\frac{1}{T} \sum_{t=1}^T c_t$, and let a HAC estimator of T times the asymptotic variance, $var(\bar{c})$, be $\hat{\sigma}_{c,T}^2$, where $\hat{\sigma}_{c,T}^2$ is defined analogously to Equation (12). The above equations lead us to state: The null hypothesis that the optimal forecasting model (ADL-MIDAS, ADL-U-MIDAS or bridge) is not more accurate than the ADL benchmark forecasting model is equivalent to the null hypothesis that $\lambda = 0$ or that the covariance between e_t^{BM} and $(e_t^{BM} - e_t^{OM})$ is 0 or that $E[c_t] = 0$. Our alternative hypothesis is that the optimal model is more accurate than the benchmark model. This is equivalent to the alternative hypothesis that $\lambda > 0$ or that the covariance between e_t^{BM} and $(e_t^{BM} - e_t^{OM})$ is positive or that $E[c_t] > 0$. Harvey *et al.* (1998) show that a corresponding test is based on the statistic

$$t_{fe} = \sqrt{T} \frac{\bar{c}}{\hat{\sigma}_{c,T}}.^{41}$$

A (one-sided) α percent significance level test rejects the null hypothesis if $t_{fe} > t^\alpha$.

⁴¹This is the test statistic R_2 in Harvey *et al.* (1998, p. 257). We do not implement test statistic R_1 in *ibid*, p. 256.

For another alternative test, we follow the nested model approach proposed by [Clark and West \(2007\)](#). Let

$$\begin{aligned} f_t &= (e_t^{BM})^2 - [(e_t^{OM})^2 - (\hat{y}_t^{BM} - \hat{y}_t^{OM})^2] \\ &= (e_t^{BM})^2 - [(e_t^{OM})^2 - (e_t^{OM} - e_t^{BM})^2], \end{aligned} \quad (14)$$

let \bar{f} be the sample mean of f_t , $\frac{1}{T} \sum_{t=1}^T f_t$, and let a HAC estimator of T times the asymptotic variance, $var(\bar{f})$, be $\hat{\sigma}_{f,T}^2$, where $\hat{\sigma}_{f,T}^2$ is defined analogously to Equation (12). Following [Clark and West \(2007\)](#) we state: the null hypothesis that the ADL benchmark forecasting model and the optimal forecasting model (ADL-MIDAS, ADL-U-MIDAS or bridge) are equally accurate is equivalent to the null hypothesis that $E[f_t] = 0$. Our alternative hypothesis is that the benchmark model is less accurate than the optimal model. This is equivalent to the alternative hypothesis that $E[f_t] > 0$. The corresponding test is based on the statistic

$$t_{nm} = \sqrt{T} \frac{\bar{f}}{\hat{\sigma}_{f,T}}.$$

A (one-sided) α percent significance level test rejects the null hypothesis if $t_{nm} > t^\alpha$. The nested model approach differs from the [Diebold and Mariano \(1995\)](#) approach in two respect. First, [Clark and West \(2007\)](#) compare t_{nm} with critical values from the t -distribution rather than the standard normal distribution where the difference gets only negligible as $T \rightarrow \infty$. Second, and more important, the nested model testing approach centers around f_t , whereas the [Diebold and Mariano \(1995\)](#) testing approach centers around r_t . Why using f_t instead of r_t ? Combining Equation (14) and (11) we can write

$$f_t = r_t + (e_t^{OM} - e_t^{BM})^2$$

So, the question of why f_t instead of r_t is equivalent to the question of why adjusting r_t for $(e_t^{OM} - e_t^{BM})^2$? Following [Clark and West \(2007, p. 292\)](#), we can give the following intuition: The optimal model reduces to the benchmark model and both models have equal true (= population) RMSFE if the true (= population) MIDAS, U-MIDAS or bridge parameters in the optimal model are zero. In contrast, the *sample* RMSFE of the optimal model should be bigger than the *sample* RMSFE of the benchmark model if the MIDAS, U-MIDAS or bridge part in the optimal model does not help to improve prediction as compared to the more parsimonious benchmark model. A reason for this is that “the [more] parsimonious [benchmark] model gains efficiency by setting to zero parameters that are zero in population, while the [larger optimal] model introduces noise into the forecasting process that will, in finite samples, inflate its [RMSFE]”. And idem, p. 296 show that $(e_t^{OM} - e_t^{BM})^2$ is the “obvious adjustment” for the aforementioned noise such that the test statistic will have approximate zero mean under null hypothesis.⁴²

⁴²Note that $f_t = 2c_t$ which relates the nested model approach and the forecast encompassing approach to each other (see also the remarks in [Clark and West, 2007, p. 297](#)).

6.5 Shares of weight/model schemes

Table 2: Weight/model schemes for Figure 4

	OECD leading indicator	CEMAC-GS leading indicator	CB leading indicator
Unrestricted beta pdf with non-zero last lag	0	0	0
Restricted beta pdf with non-zero last lag	0	0	0
Unrestricted beta pdf with zero last lag	0	0	0
Restricted beta pdf with zero last lag	1.6%	1.6%	0
Exponential Almon lag	0.3%	0	2.4%
Non-exponential Almon lag, order 1	13.9%	17.6%	15.5%
Non-exponential Almon lag, order 2	3.2%	17.1%	12.6%
Non-exponential Almon lag, order 3	0.5%	0.3%	1.3%
Non-exponential Almon lag, order 4	0	0	0
U-MIDAS	80.5%	63.4%	68.2%
Bridge	0	0	0
Average number GDP growth lags	1.5	1.5	1.5
Average number indicator lags	2.4	4.3	3.4

17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of GDP growth and indicator lags.

Table 3: Weight/model schemes for Figure 6

	Hang Seng China Enterprises Index	Shenzhen SE Composite Index	Shanghai SE Composite Index
Unrestricted beta pdf with non-zero last lag	0	0	0
Restricted beta pdf with non-zero last lag	0	0	0
Unrestricted beta pdf with zero last lag	0	0	0.3%
Restricted beta pdf with zero last lag	12.3%	8.8%	1.3%
Exponential Almon lag	0.8%	0	2.4%
Non-exponential Almon lag, order 1	34.5%	58.6%	66.6%
Non-exponential Almon lag, order 2	1.6%	1.3%	4.3%
Non-exponential Almon lag, order 3	11.5%	4.3%	2.4%
Non-exponential Almon lag, order 4	4.0%	1.6%	0.3%
U-MIDAS	35.3%	25.4%	24.9%
Bridge	0	0	0
Average number GDP growth lags	1.5	1.5	1.5
Average number indicator lags	5.6	8.1	7.1

17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of GDP growth and indicator lags.

Table 4: Weight/model schemes for Figures 7 and 8

	Fig. 7: consumer price inflation	Fig. 8: consumer price inflation, 291 days ahead forecast
Unrestricted beta pdf with non-zero last lag	0	0
Restricted beta pdf with non-zero last lag	0	0
Unrestricted beta pdf with zero last lag	0	0
Restricted beta pdf with zero last lag	5.9%	4.2%
Exponential Almon lag	0	0
Non-exponential Almon lag, order 1	39.7%	41.7%
Non-exponential Almon lag, order 2	11.3%	0
Non-exponential Almon lag, order 3	0.7%	0
Non-exponential Almon lag, order 4	0	0
U-MIDAS	42.4%	54.2%
Bridge	0	0
Average number GDP growth lags	1.5	1.6
Average number indicator lags	6.0	5.7

Column 1: 17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of y- and x-lags.

Column 2: 1 forecast horizon (= 291 days ahead) at 24 out-of-sample forecast periods implies 24 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 24 optimal models along with the average number of GDP growth and indicator lags.

Table 5: Weight/model schemes for Figures 9

	Prime lending rate	Small bank reserve requirement ratio	Big bank reserve requirement ratio
Unrestricted beta pdf with non-zero last lag	0	0	0
Restricted beta pdf with non-zero last lag	0	0	0
Unrestricted beta pdf with zero last lag	0	0	0
Restricted beta pdf with zero last lag	0.2%	1.5%	3.4%
Exponential Almon lag	0	0	0
Non-exponential Almon lag, order 1	53.7%	46.6%	56.4%
Non-exponential Almon lag, order 2	2.0%	16.2%	8.6%
Non-exponential Almon lag, order 3	4.2%	0.5%	1.7%
Non-exponential Almon lag, order 4	0	0	0
U-MIDAS	40.0%	35.3%	29.9%
Bridge	0	0	0
Average number GDP growth lags	1.5	1.5	1.5
Average number indicator lags	5.6	7.7	7.3

17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of GDP growth and indicator lags.

Table 6: Weight/model schemes for Figures 10

	Money supply M0	Money supply M1	Money supply M2
Unrestricted beta pdf with non-zero last lag	0	0	0
Restricted beta pdf with non-zero last lag	0	0	0
Unrestricted beta pdf with zero last lag	0.5%	0.2%	0
Restricted beta pdf with zero last lag	10.3%	19.9%	0.7%
Exponential Almon lag	0.2%	0.2%	0
Non-exponential Almon lag, order 1	42.4%	49.0%	58.4%
Non-exponential Almon lag, order 2	12.7%	2.5%	0.7%
Non-exponential Almon lag, order 3	3.2%	1.0%	1.0%
Non-exponential Almon lag, order 4	1.4%	0.7%	0.2%
U-MIDAS	28.7%	26.5%	39.0%
Bridge	0	0	0
Average number GDP growth lags	1.5	1.5	1.5
Average number indicator lags	8.7	9.0	8.8

17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of GDP growth and indicator lags.

Table 7: Weight/model schemes for Figures 11 and 12

	Fig. 11: IP growth, late release	Fig. 12: IP growth, early release
Unrestricted beta pdf with non-zero last lag	0	0
Restricted beta pdf with non-zero last lag	0	0
Unrestricted beta pdf with zero last lag	0	0
Restricted beta pdf with zero last lag	2.7%	2.0%
Exponential Almon lag	0	0
Non-exponential Almon lag, order 1	22.5%	24.0%
Non-exponential Almon lag, order 2	2.0%	2.2%
Non-exponential Almon lag, order 3	0.2%	0
Non-exponential Almon lag, order 4	3.7%	2.7%
U-MIDAS	68.9%	69.1%
Bridge	0	0
Average number GDP growth lags	1.5	1.5
Average number indicator lags	4.0	3.7

17 forecast horizons at 24 out-of-sample forecast periods (Q1 2008 to Q4 2013) imply $(17 \times 24 =)$ 408 separate optimal model estimations according to the rolling window model selection procedure outlined in Section 2. This figure presents the shares of the different weight/model schemes (cf. Section 2) in the 408 optimal models along with the average number of GDP growth and indicator lags.