A New Business Cycle Barometer for Germany: Construction and Evaluation of the Nowcast Accuracy

Konstantin A. Kholodilin* and Stefan Kooths**

November 18, 2008

Abstract

The objective of this paper is to construct an indicator of the short-term growth rates of real German GDP and evaluate its out-of-sample forecast accuracy. This indicator is based on a three-stage procedure. In the first step, common factors are extracted from about 4000 monthly time series. In the second step, the Factor-MIDAS methodology is used to forecast the 5 quarterly production-side components of GDP. In the final step, the forecast of the growth rates of GDP is obtained as a weighted average of forecasts of its components. Out-of-sample performance of the indicator is compared to several alternative models: quarterly aggregated factor forecasts, VAR forecasts as well as forecast from various naïve models. It appears that the indicator provides a reasonably high forecast accuracy.

Keywords: Mixed data sampling; nowcasting; factor forecasts; German business conditions

JEL classification: C32; C52; C53

1 Introduction

The objective of this paper is to construct an indicator of the short-term growth rates of real German GDP and evaluate its out-of-sample forecast accuracy. As our experience related to the predecessor of this indicator produced by the German research institute DIW Berlin shows, such a measure

*DIW Berlin, Morenstraße 58, 10117 Berlin, Germany, e-mail: kkholodilin@diw.de

**DIW Berlin, Morenstraße 58, 10117 Berlin, Germany, e-mail: skooths@diw.de

has a large media impact, especially around the turning points of the business cycle. Both politicians, businesses, and general public need such an indicator as a reliable anchor when making their everyday economic decisions.

In fact, the barometer is designed to nowcast the economic growth, that is, to forecast the current quarter. This is a meaningful exercise, given that the official German statistics publishes its flash estimates only 45 days after the corresponding quarter is over. In contrast, our barometer appears already in the beginning of the first month of the quarter and hence almost 4.5 months earlier than the official estimate.

All in all, we want our barometer to satisfy three criteria: 1) timeliness - it should be an early estimate; 2) reliability - it should be an accurate enough estimate; and 3) interpretability - the sources of the macroeconomic shocks leading to the short-term fluctuations of GDP should be easy to interpret.

There is a vast recent literature on the short-term forecasting and nowcasting of German GDP using different approaches and different data. Whereas Hinze (2003), Langmantel (1999), Mittnik and Zadrozny (2004) use the already existing leading indicators (such as Ifo business climate or business expectations indices) to forecast the growth rate of real GDP, Dreger and Schumacher (2004), Kholodilin and Siliverstovs (2006), Schumacher and Breitung (2006), and Marcellino and Schumacher (2008) extract common unobserved factors and use them in the GDP growth predictions. However, in all the cases the direct GDP forecasting approach is adopted.

The idea of our approach is to forecast the production-side components of GDP and then aggregate these forecasts to obtain prediction of growth rates of the overall GDP. It is hoped that such an algorithm would not only allow to reduce the forecast errors but also increase the interpretability of results through the decomposition of the overall GDP growth rate into the contributions of different sectors.

As an estimation technique a novel (see, e.g., Marcellino and Schumacher (2008)) factor-MIDAS (**MI**ixed **DA**ta **S**ampling) methodology to make forecasts was used. The out-of-sample forecasting performance of this method is compared to a wide range of alternative models under various specifications.

The paper is structured as follows. In section 2, the data used in this study as well as the transformations they have undergone are shortly described. Section 3 explains the algorithm employed in forecasting of the overall GDP growth rates. In section 4, the forecast accuracy of the proposed indicator are compared to that of the alternative indicators. Finally, section 5 concludes.

2 Data

This study uses the largest available set of German economic data covering 2530 time series. To the best of our knowledge, nobody before has used such a large data set. Although many time series begin already in 1991, it was decided to cut the first several years because they were thought not representative due to the effects of the re-unification shock. Therefore, the data set employed here begins in January 1995 and ends in October 2008 (the latest available data at the time of writing the paper). All data come from the statistical database of Bundesbank and where downloaded from its web site (http://www.bundesbank.de/statistik/statistik_zeitreihen.php).

The original database includes about 7000 variables. Most of them are neither seasonally nor working-day adjusted. Therefore, the first transformation was to seasonally and working-day adjust the unadjusted time series. Simultaneously, the outliers were removed. The resulting time series were logged, provided they are strictly positive. Next, the time series were tested for unit root using the standard augmented Dickey-Fuller test. The time series, for which the null of unit root was not rejected, were differenced.

3 Estimation and forecasting algorithm

The new indicator is constructed using the following three-step procedure. Firstly, common unobserved factors are extracted from the multitude of the observed monthly time series. The extraction of the common factors allows taking advantage of the valuable information contained in the numerous available individual variables and reducing it to a small number of driving forces. Secondly, the growth rates of the five production-side components of the real German GDP are forecast using the factor-MIDAS model. Thirdly, these five forecasts are used to compute the forecast of the overall growth rate of the real GDP. The latter is simply a weighted average of the former, where the weights are shares of the value added within each production-side component in the total value added of the previous year. Below each of these steps is considered in more detail.

3.1 Step 1: Extraction of the common factors

There exist multiple methodologies used to construct common factor(s): both static and dynamics. We opt here for the static factor model suggested by Stock and Watson (1999) with the number of factors selected using the criterion of Bai and Ng (2002). As literature, e.g., Dreger and Schumacher

(2005), shows, this model is not significantly worse in terms of the out-of-sample forecasting performance compared to much more sophisticated and computationally intensive methods.

3.2 Step 2: Forecasting of the GDP components

Forecasts of the sectors of economy are conducted using the Factor-MIDAS methodology suggested by Marcellino and Schumacher (2008). MIDAS methodology allows to regress lower-frequency l.h.s. variables (e.g., quarterly GDP) on the higher-frequency r.h.s. variables (e.g., monthly industrial production, interest rates, leading indicators). It thus permits the use of all currently available information, whereas aggregating the monthly variables to the quarterly ones may imply incurring information losses. Factor-MIDAS uses the monthly unobserved common factor(s) as r.h.s. variables.

In this paper, we forecast the real growth rates of GDP, y_{t_q}, and its components, $y^c_{t_q}$, where t_q denotes the quarterly time index $t_q = 1, 2, 3, \ldots, T^y_q$ with T^y_q as the final for which the GDP components are available. The growth rate of GDP and its components can also be expressed at the monthly frequency by setting $y^c_{t_m} = y^c_{t_q}$ for all $t_m = 3t_q$, where t_m is the monthly time index. The objective is to nowcast GDP components h_q quarters, or $h_m = 3h_q$ months, ahead, $y^c_{T^y_m+h_m}$. Nowcasting implies that in each particular calender month, GDP and its components are not observed for the current quarter. As a rule, the first official (flash) estimate of GDP is published 1.5 months after the end of the reference quarter, whereas that of the GDP components is published only 2 months after the end of the reference quarter. Thus, in the first month of each quarter, t_q, the data on GDP and its components are only available for $t_q - 2$. In the second month of the quarter, information set includes GDP in period $t_q - 1$, while the latest figures on GDP component still come from the period $t_q - 2$. Finally, in the last month of the quarter, data on both GDP and its components for the period $t_q - 1$ become available. This implies that, in case of nowcasting, in the first month of the quarter, $h_q = 2$; in the second month, $h_q = 1$ for GDP and $h_q = 2$ for GDP components; whereas in the third month, $h_q = 1$ both for GDP and its components.

Assume for the sake of simplicity of exposition that there is only one stationary monthly indicator, x_{t_m}, on which the forecasts of the GDP components are based. The time index, t_m, denotes a monthly sampling frequency of x_{t_m} for $t_m = 1, 2, 3, \ldots, T^x_m$, where T^x_m is the final month, for which an observation of the indicator is available. Normally, $T^x_m > T^y_m = 3T^y_q$, since monthly macroeconomic variables have smaller publication lag than quarterly GDP.

Then the unrestricted factor-MIDAS model of the growth rate of c-th component of GDP is defined as:

$$y^c_{t_q} = \alpha_0 + \sum_{i=0}^{I} \beta_i x_{t_m - \omega - h_m - i} + \varepsilon_{t_m} \tag{1}$$

where $\omega = T^x_m - T^y_m$; x_t is the independent variable, in this case it is the monthly current and lagged values of the unobserved common factor(s).

The restricted factor-MIDAS model is formulated as:

$$y^c_{t_q} = \alpha_0 + \gamma \sum_{i=0}^{I} \delta_i x_{t_m - \omega - h_m - i} \tag{2}$$

where δ_i is a parameter at the i-th Alman lag defined as:

$$\delta_i = \frac{exp(\theta_1 i + \theta_2 i^2)}{\sum_{i=0}^{I} exp(\theta_1 i + \theta_2 i^2)} \tag{3}$$

Different specifications of the basic model are considered in order to explore all the possibilities of improving the forecast accuracy and to check the robustness of our results (see Wohlrabe and Robinzonov (2008) on the necessity of a thorough robustness check of the forecasts). Below are listed several specification issues, which can be relevant for the forecast accuracy.

1. It is possible to use the same factors based on all the information available in forecasting all GDP components. However, better results may be achieved when each component forecast is based on the factors extracted from the variables, which are somehow important for this particular component of GDP. In this latter case, the pre-selection of the data sets for the factor extraction can be made departing from different criteria. For example, it might be better to extract factors only from the variables that are leading the GDP component, as in den Reijer (2005), or that deliver higher forecast accuracy in the forecasting models including on the l.h.s the component and on the r.h.s. each observed variable separately, as in Siliverstovs and Kholodilin (2006).

2. Further issue relates to the fact that the data on the observed variables in the last few periods are missing, that is, different variables become available with different publication lags. This phenomenon is known as the ragged-edge problem. It can be dealt with in various ways. One

can either extrapolate the individual observed variables using some kind of ARIMA model or vertically realign the data as in Altissimo et al. (2006).

3. The model can be estimated using either growing or rolling window. The rolling window can reflect the changing nature of the process and better incorporate the structural shifts. However, given no structural changes, having each time more observations, as under the growing window approach, might lead to more parameter stability and precision.

3.3 Step 3: Aggregation of sectoral forecasts

In the final step, the forecasts of the growth rates of the five major sectors of the economy — 1) production sector excluding construction; 2) construction; 3) wholesale and retail trade, hotels and restaurants, and transport; 4) financial, real estate, renting, and business services; and 5) public and private services — are aggregated by multiplying them with their respective shares in the total value added measured in the previous year.

$$E(y^{GDP}_{T_q+h_q}|\Omega_{T_m}) = \sum_{c=1}^{C} w^c_{T_a-1} E(y^c_{T_q+h_q}|\Omega_{T_m}) \tag{4}$$

where C is the number of the production-side components and each component's weight, w_c, is defined as:

$$w^c_{T_a-1} = \frac{Y^c_{T_a-1}}{\sum_{c=1}^{C} Y^c_{T_a-1}} \tag{5}$$

where Y_{T_a-1} is the nominal value added in c-th sector in year $T_a - 1$.

Forecasting the GDP components first and then aggregating them to the GDP facilitates the interpretation of the growth rate fluctuations. Thus, contribution of each component to the overall growth rates can be determined and hence the sources of the shock hitting economy can be identified. Moreover, as the literature shows (e.g., (see, cf., Hendry and Hubrich, 2006; Drechsel and Maurin, 2008, *inter alia*)) aggregating forecasts both across components and countries can lead to more accurate and reliable results.

4 Evaluation of forecast accuracy

The estimation period is 1992:q1-2001:q4, whereas the forecasting period is 2001:q1-2008:q3. The estimation was conducted recursively, i.e., the initial estimation sample was each time increased by one observation.

The set of evaluated models includes: both unrestricted and restricted Factor-MIDAS models; autoregressive model (no factors included in regression); distributed lag model with quarterly aggregated factors (the so-called bridge equation); ARDL (autoregressive distributed lag) model with quarterly aggregated factors; naive models (lagged value as a forecast and mean value of the past observations as a forecast).

The forecast based on the unrestricted factor-MIDAS model can be formulated as:

$$E(y^c_{T_q+h_q}|\Omega_{T_m}) = \widehat{\alpha_0} + \sum_{i=0}^{I} \widehat{\beta}_i x_{T_m+h_m-i} \tag{6}$$

where Ω_{T_m} is the information set available by the period T_m and including monthly data; x_t is the independent variable, in this case it is the monthly current and lagged values of the unobserved common factor(s).

The forecast corresponding to restricted factor-MIDAS model is:

$$E(y^c_{T_q+h_q}|\Omega_{T_m}) = \widehat{\alpha_0} + \widehat{\gamma} \sum_{i=0}^{I} \widehat{\delta}_i x_{T_m+h_m-i} \tag{7}$$

where

$$\widehat{\delta}_i = \frac{exp(\widehat{\theta}_1 i + \widehat{\theta}_2 i^2)}{\sum_{i=0}^{I} exp(\widehat{\theta}_1 i + \widehat{\theta}_2 i^2)} \tag{8}$$

The forecast obtained under the autoregressive distributed lag (ARDL) model is:

$$E(y^c_{T_q+h_q}|\Omega_{T_q}) = \widehat{\alpha_0} + \sum_{i=1}^{I} \widehat{\alpha}_i y^c_{T_q+h_q-i} + \sum_{i=0}^{I} \widehat{\beta}_i x_{T_q+h_q-i} \tag{9}$$

where Ω_{T_q} is the information set available by the period T_q and including only quarterly data; x_{t_q} are the values of independent variable (here: common factor(s)) aggregated to the quarterly frequency.

A special case of the ARDL model is the purely autoregressive, where all $\beta_i = 0$. The forecast equation corresponding to this model can thus be written as:

$$E(y^c_{T_q+h_q}|\Omega_{T_q}) = \widehat{\alpha_0} + \sum_{i=1}^{I} \widehat{\alpha}_i y^c_{T_q+h_q-i} \tag{10}$$

The first of the naïve models is the one, where the forecast value is represented by the unconditional mean of all the past values:

$$E(y^c_{T_q+h_q}|\Omega_{T_q}) = \frac{1}{T_q} \sum_{t=1}^{T_q} y^c_t \tag{11}$$

Under the last naïve model, the forecast is simply the last observed value of the variable:

$$E(y^c_{T_q+h_q}|\Omega_{T_q}) = y^c_{t_q} \tag{12}$$

The forecast accuracy of these competing models is compared using two different measures: root mean squared forecast error (RMSFE) and sign test.

Table 1 reports the absolute RMSFE and relative (to the last naïve model) RMSFE of the alternative models. The first half of the table presents the forecasting results obtained with 3 factors, whereas the second half of the table contains those with 1 factor. In addition, the table compares forecast accuracy across the factors based on different data sets: the full data set ("All") and a reduced data set, where the variables, which are too strongly correlated with each other — pairwise correlation exceeding 0.5 — were dropped ("Corr $\leq$ 0.5"). Dropping the highly correlated variables was done because some groups of variables (e.g., various banking variables contained in the Bundesbank database) are too numerous and hence can lead towards bias in the factors. As result, the reduced data set contains 544 variables.

The forecasts take into account the publication lags of both the quarterly and monthly data. Thus, the GDP data for reference quarter become available in the second month of the next quarter, data on the GDP components become available in the third months, data on the monthly factors become available in the next month, while the data on the quarterly aggregated (bridged) factors become available in the first month of the next quarter.

One can see that the naïve model is very difficult to bit. Only with a reduced number of individual variables the factor models improve upon this model, although the accuracy improvement of about 6% is quite modest.

5 Conclusion

The paper examined the nowcasting properties of a new business cycle barometer of the growth rates of real German GDP. This indicator produces reasonably accurate out-of-sample forecasts in comparison to alternative methods when the data set is pre-selected in order to avoid the biasedness of factors stemming from the peculiar structure of the database. The improvement of the forecast accuracy compared to the best naïve model is about 6%.

References

Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, and G. Veronese (2006). New EuroCOIN: Tracking economic growth in real time. CEPR Discussion Papers 5633, C.E.P.R. Discussion Papers.

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

den Reijer, A. H. J. (2005). Forecasting Dutch GDP using large scale factor models. Technical Report 028, Netherlands Central Bank, Research Department.

Drechsel, K. and L. Maurin (2008). Flow on conjunctural information and forecast of euro area economic activity. Working Paper Series 925, European Central Bank.

Dreger, C. and C. Schumacher (2004). Estimating large-scale factor models for economic activity in Germany: Do they outperform simpler models? *Jahrbücher für Nationalökonomie und Statistik 224*, 732–750.

Dreger, C. and C. Schumacher (2005). The out-of-sample performance of leading indicators for the German business cycle. *Journal of Business Cycle Measurement and Analysis 2*, 71–88.

Hendry, D. F. and K. Hubrich (2006). Forecasting economic aggregates by disaggregates. CEPR Discussion Papers 5485, C.E.P.R. Discussion Papers.

Hinze, J. (2003). Prognoseleistung von Frühindikatoren: Die Bedeutung von Frühindikatoren für Konjunkturprognosen – Eine Analyse für Deutschland. *HWWA Discussion Paper* (236).

Kholodilin, K. A. and B. Siliverstovs (2006). On the forecasting properties of the alternative leading indicators for the German GDP: Recent evidence. *Jahrbücher für Nationalökonomie und Statistik, 226*, 234–259.

Langmantel, E. (1999). Das Ifo Geschäftsklima als Indikator für die Prognose des Bruttoinlandsprodukts. *Ifo Schnelldienst 16/17*, 16–21.

Marcellino, M. and C. Schumacher (2008). Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP. Economics Working Papers ECO2008/16, European University Institute.

Mittnik, S. and P. Zadrozny (2004). Forecasting quarterly German GDP at monthly intervals using monthly Ifo business conditions data. Technical Report CESifo Working Paper No. 1203, CESifo GmbH.

Schumacher, C. and J. Breitung (2006). Real-time forecasting of GDP based on a large factor model with monthly and quarterly data. Discussion Paper Series 1: Economic Studies 2006,33, Deutsche Bundesbank, Research Centre.

Siliverstovs, B. and K. Kholodilin (2006). On selection of components for a diffusion index model: It's not the size, it's how you use it. DIW Berlin, Discussion paper N 598.

Stock, J. and M. Watson (1999). Forecasting inflation. *Journal of Monetary Economcs 44*, 293–335.

Wohlrabe, K. and N. Robinzonov (2008). Freedom of choice in macroeconomic forecasting: An illustration with German industrial production and linear models. ifo Working Paper Nr. 57.

Appendix

Table 1: Comparison of forecast accuracy, 2001:m1-2008:m9

	3 factors				1 factor			
	All N=2530		Corr. ≤ 0.5 N=544		All N=2530		Corr. ≤ 0.5 N=544	
	RMSFE	Relative RMSFE	RMSFE	Relative RMSFE	RMSFE	Relative RMSFE	RMSFE	Relative RMSFE
Previous value of GDP	0.627	1.185	0.627	1.185	0.627	1.185	0.627	1.185
Average of past values of GDP	0.529	1.000	0.529	1.000	0.529	1.000	0.529	1.000
Direct factor bridge, NLag=3	0.657	1.242	0.575	1.087	0.581	1.098	0.532	1.005
Direct factor bridge, NLag=1	0.539	1.018	0.493	0.932	0.575	1.087	0.535	1.011
Indirect factor bridge, NLag=3	0.630	1.190	0.647	1.224	0.563	1.065	0.545	1.031
Indirect factor bridge, NLag=1	0.546	1.031	0.508	0.960	0.592	1.119	0.546	1.031
Direct factor-MIDASU, NLag=3	0.582	1.100	0.776	1.466	0.579	1.094	0.565	1.068
Direct factor-MIDASU, NLag=1	0.547	1.034	0.494	0.934	0.565	1.067	0.532	1.005
Indirect factor-MIDASU, NLag=3	0.563	1.065	0.791	1.495	0.570	1.078	0.576	1.088
Indirect factor-MIDASU, NLag=1	0.550	1.040	0.496	0.937	0.578	1.093	0.548	1.036