European University Institute

DEPARTMENT OF ECONOMICS

EUI Working Papers

ECO 2009/13

DEPARTMENT OF ECONOMICS

Pooling versus Model Selection for Nowcasting with Many Predictors: An Application to German GDP

Vladimir Kuzin, Massimiliano Marcellino and Christian Schumacher

EUROPEAN UNIVERSITY INSTITUTE, FLORENCE
DEPARTMENT OF ECONOMICS

Pooling versus Model Selection for Nowcasting with Many Predictors: An Application to German GDP

VLADIMIR KUZIN, MASSIMILIANO MARCELLINO
and
CHRISTIAN SCHUMACHER

EUI Working Paper **ECO** 2009/13

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Pooling versus model selection for nowcasting with many predictors: An application to German GDP[1]

Vladimir Kuzin
DIW Berlin
vkuzin@diw.de

Massimiliano Marcellino
EUI, Università Bocconi and CEPR
massimiliano.marcellino@eui.eu

Christian Schumacher
Deutsche Bundesbank
christian.schumacher@bundesbank.de

first version: 4 February 2008, this version: 19 January 2009

[1]This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank. We are grateful to seminar and workshop participants at the Bundesbank, DIW Berlin, University of Basle, and the University of Frankfurt for helpful comments. Helpful comments were also provided by Heinz Herrmann, Sylvia Kaufmann, and Karl-Heinz Tödter. The codes for this paper were written in Matlab. Some functions were taken from the Econometrics Toolbox written by James P. LeSage from `www.spatial-econometrics.com`. Other codes were kindly provided by Mario Forni from `www.economia.unimore.it/forni_mario/matlab.htm`, Arthur Sinko from `www.unc.edu/~sinko/midas.zip`, and Gerhard Rünstler.

Abstract

This paper discusses pooling versus model selection for now- and forecasting in the presence of model uncertainty with large, unbalanced datasets. Empirically, unbalanced data is pervasive in economics and typically due to different sampling frequencies and publication delays. Two model classes suited in this context are factor models based on large datasets and mixed-data sampling (MIDAS) regressions with few predictors. The specification of these models requires several choices related to, amongst others, the factor estimation method and the number of factors, lag length and indicator selection. Thus, there are many sources of mis-specification when selecting a particular model, and an alternative could be pooling over a large set of models with different specifications. We evaluate the relative performance of pooling and model selection for now- and forecasting quarterly German GDP, a key macroeconomic indicator for the largest country in the euro area, with a large set of about one hundred monthly indicators. Our empirical findings provide strong support for pooling over many specifications rather than selecting a specific model.

JEL Classification Codes: E37, C53

Keywords: nowcasting, forecast combination, forecast pooling, model selection, mixed-frequency data, factor models, MIDAS

1 Introduction

Forecast models that can take into account unbalanced datasets have received substantial attention in the recent literature. In real time, the unbalancedness of datasets arises due to the different sampling frequencies and different publication delays of business cycle indicators. For example, Gross Domestic Product (GDP), a key indicator of macroeconomic activity, is typically published at quarterly frequency and has a considerable publication lag. As policy makers regularly request information on the current state of the economy in terms of GDP, there is a need to provide estimates of current GDP in order to support policy decisions. Following the discussion in Giannone et al. (2008), we call the necessary projection of current GDP the 'nowcast' in this paper. In the same way, other business cycle indicators, that might serve as predictors for GDP, are released in an asynchronous way and exhibit complicated patterns of missing values at the end of the sample, which leads to the so-called 'ragged-edge' problem of multivariate data in econometrics, see Wallis (1986). Another difficulty arises, because GDP is released on a quarterly basis, whereas many important predictors are sampled at monthly or higher frequencies. Therefore, now- and forecast models should be able to account for mixed-frequency and ragged-edge data.

In the recent forecast literature, two alternative modeling approaches that can take into account these data irregularities have been discussed: mixed-data sampling (MIDAS) regressions with a few indicators and large factor models. In the MIDAS approach, as introduced by Ghysels and Valkanov (2006) and Ghysels, Sinko and Valkanov (2007), a low-frequency variable is regressed on higher frequency variables using skip-sampling and restricted lag polynomials. Clements and Galvão (2008, 2009) introduced the MIDAS approach to macroeconomics, and presented empirical results for US quarterly GDP predicted by monthly indicators. Due to the skip-sampling and direct projection, MIDAS can tackle mixed-frequency data as well as differences in data availability at the end of the sample. Whereas MIDAS is mainly a forecast tool based on a few selected indicators, the usefulness of factor models based on large datasets as forecast devices has been widely discussed in the recent literature, see the seminal papers by Stock and Watson (2002) and Forni et al. (2005). If ragged-edge and mixed-frequency data is present, factor estimation methods that take into proper account these data irregularities are required. Two prominent methods from the recent literature are: the two-step estimator in a state-space framework by Doz et al. (2006) and Giannone et al. (2008), which can account for statistical publication lags in the indicator dataset by using the Kalman smoother; and the dynamic principal components estimator by Altissimo et al. (2006), which can also handle ragged edge datasets, and thereby extends the dynamic estimator by Forni et al. (2005) based on balanced data.

Within the MIDAS and factor model classes, the practitioner has to make a set of auxiliary decisions when applying them for forecasting. For example, proper indicator selection is crucial for MIDAS regressions. However, in a related framework with single-frequency data, Banerjee and Marcellino (2006) for the US and Banerjee et al. (2005)

for the Euro area have found that selecting variables in real time can be much more difficult than what suggested by ex-post evaluations. The factor forecast framework is also not immune to mis-specification issues, e.g., there is an ongoing discussion regarding the appropriate factor estimation method, see Boivin and Ng (2005), Stock and Watson (2006), D'Agostino and Giannone (2006), and Schumacher (2007). And proper handling of dynamics is a problem for both approaches, even more than usual due to the mixed sampling frequencies of the indicators. Therefore, it is very likely that even a careful selection process can result in a mis-specified model.

In the present paper, we propose nowcast pooling as a simple way of dealing with this substantial model uncertainty, exacerbated by the use of large unbalanced datasets. From a theoretical point of view, it is difficult to rank model specification and pooling in finite and irregular samples. In addition, their relative performance will depend on the assumptions on the data generating process. Therefore, we prefer to take an empirical approach. In particular, we evaluate the nowcast performance of pooling and single models for quarterly German GDP, a key variable for the largest country in the euro area. Specifically, first we investigate the performance of a large number of MIDAS and factor models with different specifications, that are held fixed in the recursive evaluation exercise. In other words, on an ex-post basis, we search for the best specifications. Second, in order to allow for data-driven specification, we consider real-time model selection based either on information criteria or on the past forecast performance of the individual models, following the discussion in Inoue and Kilian (2006). Finally, we discuss to what extent alternative pooling schemes can circumvent potential mis-specification of single models. We consider averaging with equal weights, the median as well as performance-based weights over full set of models. As the sample under consideration is relatively small, and simple forecast combinations have turned out to provide robust results in the literature, we do not account for more sophisticated pooling methods, see e.g. Clark and McCracken (2008).

It is well known that pooling of forecasts provides a robust tool in the presence of mis-specification and parameter instability, see for example Timmermann (2005) and Clements and Hendry (2004) for theoretical results, and Clark and McCracken (2008), Assenmacher-Wesche and Pesaran (2008) and Garratt et al. (2009) for recent empirical applications. However, these papers do not take into account the data unbalancedness, which is pervasive in economics due to publication delays of statistical data and different sampling frequencies. Instead, we focus on pooling MIDAS and factor models as econometric specifications that take into explicit account the data unbalancedness. Hence, our first original contribution to the literature is to assess pooling in a more realistic context and for models potentially more useful for empirical analysis.

Our second original contribution is to compare MIDAS regressions based on few selected indicators with factor models based on large datasets, thus relating the MIDAS literature from Clements and Galvão (2008, 2009) to the factor nowcast literature from

Giannone et al. (2008), Altissimo et al. (2006) and Marcellino and Schumacher (2008).[1]

Our main results can be summarised as follows. First, searching in the set of all possible models on an ex-post basis, it is possible to find MIDAS and factor specifications that outperform a simple benchmark, and MIDAS models with a few indicators tend to outperform factor models in this ex-post evaluation. Since the search described above is based on full sample results, it might be subject to the data-mining critique. Second, when selecting the forecasting models in real time based either on information criteria or on their past performance, it is much more difficult to beat the benchmark, with the exception of factor model selection based on past forecasting performance. Third, pooling the whole set of MIDAS and factor now- and forecasts clearly outperforms single models selected according to information-criteria or based on their past performance. In comparison with the best fixed specifications selected on an ex-post basis, pooling is better than 93-100% of all the single indicator forecasts, and of 86-100% of all the factor forecasts, depending on the horizon. Furthermore, in real time, pooling of factor models seems to outperform pooling of MIDAS models with few indicators.

In summary, the main finding of our paper is that there is considerable uncertainty with respect to the appropriate specification of the compilcated econometric tools needed to handle large and unbalanced datasets of macroeconomic variables. In this context, pooling of many specifications within and across the MIDAS and factor model classes is overall superior to selecting a single model.

The paper proceeds as follows. Section 2 provides an overview of the individual MIDAS regressions and factor models employed here, as well as the combination methods. Section 3 describes the design of the forecast comparison exercise. Section 4 presents and compares the empirical results for fixed, information criteria and past performance based specifications. Section 5, discusses pooling over the whole set of MIDAS and Factor-MIDAS specifications. Section 6 conducts a variety of robustness analyses. Section 7 summarizes and concludes.

2 Nowcasting quarterly GDP with ragged-edge data: MIDAS, factor models, and pooling

To forecast quarterly GDP using monthly indicators, we mainly rely on the mixed-data sampling (MIDAS) approach as proposed by Ghysels and Valkanov (2006), Ghysels et al. (2007), and Clements and Galvão (2008, 2009). MIDAS is a single-equation approach that allows a low-frequency variable like GDP to be explained by high-frequency regressors. In our application, we will consider different types of regressors: either a small number of business cycle indicators, following the work by Clements and Galvão (2008, 2009),

[1]Barhoumi et al. (2008) also consider forecasting with ragged-edge data, but do not consider MIDAS approaches and specification uncertainty as in the present paper, in particular, with respect to specification uncertainty of factor models.

or factors estimated from a large set of indicators, following Marcellino and Schumacher (2008). For both types of regressors, the MIDAS regression approach serves as a way to compute the projections. Below, in subsection 2.1, we first introduce the MIDAS regression, then discuss the choice of monthly predictors in subsection 2.2, in particular the different factor estimation approaches that can be applied to large sets of indicators. When discussing the alternative approaches, we will also address the different specifications that are necessary when applying the models in real time. Finally, the alternative pooling methods are described in subsection 2.3.

2.1 The MIDAS approach as a now- and forecasting tool

In our application, the predictand is quarterly GDP growth, which is denoted as y_{t_q} where t_q is the quarterly time index $t_q = 1, 2, 3, \ldots, T_q^y$ with T_q^y as the final quarter for which GDP is available. GDP growth can also be expressed at the monthly frequency by setting $y_{t_m} = y_{t_q} \forall t_m = 3t_q$ with t_m as the monthly time index. Thus, GDP y_{t_m} is observed only at months $t_m = 3, 6, 9, \ldots, T_m^y$ with $T_m^y = 3T_q^y$. The aim is to forecast GDP h_q quarters ahead, or $h_m = 3h_q$ months ahead, yielding a value for $y_{T_m^y + h_m}$.

Nowcasting means that in a particular calender month, we do not observe GDP for the current quarter. It can even be the case that GDP is only available with a delay of two periods. In April, for example, German GDP is only available for the fourth quarter of the previous year, and a nowcast for second quarter GDP requires $h_q = 2$. Thus, if a decision maker requests an estimate of current quarter GDP, the forecast horizon has to be set sufficiently large in order to provide the appropriate figures. For further discussion on nowcasting, see Giannone et al. (2008).

To now- and forecast quarterly GDP growth, we can make use of a stationary monthly predictor z_{t_m}. For simplicity, we assume that there is only one predictor, and generalise this case later on to more than one indicators or factors. The time index t_m denotes a monthly period, and observations of z_{t_m} are available for $t_m = 1, 2, 3, \ldots, T_m^z$, where T_m^z is the final month for which an observation is available. Usually, T_m^z is larger than $T_m^y = 3T_q^y$, as monthly observations for many relevant macroeconomic indicators, in particular financial or survey data, are earlier available than GDP observations. The forecast for GDP is denoted as $y_{T_m^y + h_m | T_m^z}$, as we condition the forecast on information available in month T_m^z, which also includes GDP observations up to T_q^y in addition to the indicator observations up to T_m^z with $T_m^z \geq T_m^y = 3T_q^y$. Thus, the indicator is available $w_{zy} = T_m^z - T_m^y$ months ahead of GDP.

Basic MIDAS The forecast model for forecast horizon h_q quarters with $h_q = h_m/3$ is

$$y_{t_q + h_q} = y_{t_m + h_m} = \beta_0 + \beta_1 b(L_m, \boldsymbol{\theta}) z^{(3)}_{t_m + w_{zy}} + \varepsilon_{t_m + h_m}, \quad (1)$$

where $w_{zy} = T_m^z - T_m^y$ and the polynomial $b(L_m, \boldsymbol{\theta})$ is the exponential Almon lag

$$b(L_m, \boldsymbol{\theta}) = \sum_{k=0}^{K} c(k, \boldsymbol{\theta}) L_m^k, \quad c(k, \boldsymbol{\theta}) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=0}^{K} \exp(\theta_1 k + \theta_2 k^2)}, \tag{2}$$

with the monthly lag operator L_m defined as $L_m z_{t_m} = z_{t_m - 1}$. In the MIDAS approach, quarterly GDP $y_{t_q+h_q}$ is directly related to the indicator $z_{t_m+j}^{(3)}$ and its lags, where $z_{t_m}^{(3)}$ is a skip-sampled version of the monthly z_{t_m}. The superscript three indicates that every third observation starting from the t_m-th one is included in the regressor $z_{t_m}^{(3)}$, thus $z_{t_m}^{(3)} = z_{t_m} \,\forall\, t_m = \ldots, T_m^z - 6, T_m^z - 3, T_m^z$. Lags of the monthly factors are treated accordingly, e.g. the k-th lag $z_{t_m-k}^{(3)} = z_{t_m-k} \,\forall\, t_m = \ldots, T_m^z - k - 6, T_m^z - k - 3, T_m^z - k$. In the regression, the variable w_{zy} denotes the number of monthly periods, the monthly indicator is earlier available than GDP. Thus, we take into account that a monthly indicator is typically available within the quarter for which no GDP figure is available, see Clements and Galvão (2008, 2009).

For given $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$, the exponential lag function $b(L_m, \boldsymbol{\theta})$ provides a parsimonious way to consider monthly lags of the factors as we can allow for large K to approximate the impulse response function of GDP from the factors. The longer the lead-lag relationship in the data is, the less MIDAS suffers from sampling uncertainty compared with the estimation of unrestricted lags, where the number of coefficients increases with the lag length.

The MIDAS model can be estimated using nonlinear least squares (NLS) in a regression of y_{t_m} onto $z_{t_m+w_{zy}-h_m}^{(3)}$ and lags, yielding coefficients $\widehat{\theta}_1$, $\widehat{\theta}_2$, $\widehat{\beta}_0$ and $\widehat{\beta}_1$. The forecast is given by

$$y_{T_m^y+h_m|T_m^z} = \widehat{\beta}_0 + \widehat{\beta}_1 b(L_m, \widehat{\boldsymbol{\theta}}) z_{T_m^z}. \tag{3}$$

According to this forecast equation, the MIDAS approach is a direct forecasting tool, as it relates future GDP to current and lagged indicators, see Marcellino, Stock and Watson (2006) as well as Chevillon and Hendry (2005) for detailed discussions of this issue in the single-frequency case. MIDAS is horizon-dependent, and thus has to be reestimated for multi-step forecasts for all h_m. The same holds for the case new statistical information becomes available. For example, each month, new observations for the indicator is released, whereas GDP observations are released only once in a quarter. Thus, also w_{zy} changes from month to month, which also makes a new regression necessary.

Autoregressive MIDAS As an extension to the basic MIDAS approach, Clements and Galvão (2008) consider autoregressive dynamics in the MIDAS approach. In particular, they propose the model

$$y_{t_m+h_m} = \beta_0 + \lambda y_{t_m} + \beta_1 b(L_m, \boldsymbol{\theta})(1 - \lambda L_m^3) z_{t_m+w}^{(3)} + \varepsilon_{t_m+h_m}. \tag{4}$$

The autoregressive coefficient λ is not estimated unrestrictedly to rule out discontinuities of the impulse response function of $z_{t_m}^{(3)}$ on $y_{t_m+h_m}$, see the discussion in Ghysels et al. (2007), pp. 60. The restriction on the coefficients is a common-factor restriction to ensure a smooth impulse response function, see Clements and Galvão (2008). The AR coefficient λ can be estimated together with the other coefficients by NLS. As an AR model is often supposed to be an appropriate benchmark specification for GDP, the extension of MIDAS might give additional insights in which direction the other MIDAS approaches considered so far might be improved. Henceforth, we denote this approach as 'AR-MIDAS', whereas we denote MIDAS without AR terms just as 'MIDAS'.

Multiple MIDAS regression MIDAS regressions can easily be extended to the multiple predictor case. Assume we have M predictors z_{i,t_m} for $i = 1, \ldots, M$. The corresponding MIDAS equation is

$$y_{t_q+h_q} = y_{t_m+h_m} = \beta_0 + \sum_{i=1}^{M} \beta_{1,i} b_i(L_m, \boldsymbol{\theta}_i) z_{i,t_m+w_{zy}}^{(3)} + \varepsilon_{t_m+h_m}, \tag{5}$$

where the coefficients $\beta_{1,i}$ and b_i differ with respect to the different indicators chosen. In particular, each indicator can have a different impulse response function through $\boldsymbol{\theta}_i = \{\theta_{1,i}, \theta_{2,i}\}$ that determine the polynomial b_i.

2.2 The MIDAS predictors

In our empirical application, we have available a large set of monthly predictors, collected in the N-dimensional vector $\mathbf{X}_{t_m} = (x_{1,t_m}, \ldots, x_{N,t_m})'$ for months $t_m = 1, 2, 3, \ldots, T_m$. Here T_m is the latest observation available in the entire set of monthly time series. However, due to publication lags, some elements at the end of the sample can be missing for certain predictors, thus rendering an unbalanced sample. We will distinguish two types of MIDAS regressors: 1) single indicators selected from the a large set of indicators; 2) factors estimated from $\mathbf{X}_{t_m}$. Thus, regarding factor now- and forecasting, we follow the Factor-MIDAS approach of Marcellino and Schumacher (2008), where factors are estimated in the first step, and these factors are plugged into a MIDAS regression for computing the forecasts.

2.2.1 MIDAS forecasting with a single indicator

In our application, we will now- and forecast with a large range of MIDAS models, where in each model GDP is explained by a single indicator, $z_{t_m} \in \mathbf{X}_{t_m}$. Thus, we end up with N single-indicator MIDAS regressions and N single-indicator MIDAS with autoregressive terms. As we will see, some of these simple models will perform very well. However, in order to check the robustness of the results with respect to this specification choice, we will perform a sensitivity analysis later on and use more than one predictor in MIDAS.

In real-time, when a practitioner aims at minimising forecast error loss, the question is how to specify the MIDAS with respect to variable selection, the choice of the AR term, as well as the maximum length of the lag polynomial. We will focus on the variable selection issue in our application below, as well as on the choice of the AR term.

2.2.2 MIDAS forecasting with factors

We want to model $\mathbf{X}_{t_m}$ using a factor specification, and particularly assume that the monthly observations have a factor structure according to

$$\mathbf{X}_{t_m} = \mathbf{\Lambda F}_{t_m} + \boldsymbol{\xi}_{t_m}, \tag{6}$$

where the r-dimensional factor vector is denoted as $\mathbf{F}_{t_m} = (f'_{1,t_m}, \ldots, f'_{r,t_m})'$. The factors times the $(N \times r)$ loadings matrix $\mathbf{\Lambda}$ represent the common components of each variable. The idiosyncratic components $\boldsymbol{\xi}_{t_m}$ are that part of $\mathbf{X}_{t_m}$ not explained by the factors. Under the assumption that the $(T_m \times N)$ data matrix $\mathbf{X}$ is balanced, various ways to estimate the factors have been provided in the literature. For example, two of the most widely used approaches are based on principal components analysis (PCA) as in Stock and Watson (2002) or dynamic PCA according to Forni et al. (2005). Note that, according to (6), all the factor models to be discussed below will work at the higher monthly frequency, thus factor estimates are available for all monthly periods $t_m = 1, 2, \ldots, T_m$. Below, we compare two ways of estimating the factors in the presence of ragged-edge data. In the empirical application, we will employ both models to account for model uncertainty.

Vertical realignment of data and dynamic principal components factors A very convenient way to solve the ragged-edge problem is provided by Altissimo et al. (2006) for estimating the New Eurocoin indicator. They propose to realign each time series in the sample in order to obtain a balanced dataset. Assume that variable i is released with k_i months of publication lag. Thus, given a dataset in period $T_m^{x_i}$, the final observation available of this time series is for period $T_m^{x_i} - k_i$. The realignment proposed by Altissimo et al. (2006) is then simply

$$\widetilde{x}_{i,T_m} = x_{i,T_m - k_i} \tag{7}$$

for $t_m = k_i + 1, \ldots, T_m^{x_i}$. Applying this procedure to all the time series, and harmonising at the beginning of the sample, yields a balanced data set $\widetilde{\mathbf{X}}_{t_m}$ for $t_m = \max(\{k_i\}_{i=1}^{N}) + 1, \ldots, T_m^{x_i}$.

Given this monthly data, Altissimo et al. (2006) propose dynamic PCA to estimate the factors. As the dataset is balanced, the two-step estimation techniques by Forni et al. (2005) directly apply. In our applications below, we will denote the combination of vertical realignment and dynamic principal components factors as 'VA-DPCA'. Details on how the estimation is carried out, can be found in the appendix B.

The vertical realignment solution to the ragged-edge problem is easy to use. A disadvantage is that the availability of data determines dynamic cross-correlations between variables. Furthermore, statistical release dates for data are not the same over time, for example, due to major revisions. In this case, dynamic correlations within the data change and factors can change over time. The same holds if factors are reestimated at a higher frequency than the frequency of the factor model. This is a very common scenario, for example, if a monthly factor model is reestimated several times within a month when new monthly observations are released. If this the case, the realignment of the data changes the correlation structure all the time. On the other hand, dynamic PCA as in Forni et al. (2005) exploits the dynamic cross-correlations in the frequency domain and might be in principle able to account for these changes in realignments of the data.

Estimation of a large parametric factor model in state-space form The factor estimation approach followed by Doz et al. (2006) is based on a complete representation of the large factor model in state-space form. The complete model consists of a factor representation of the large vector of monthly time series and an explicit VAR structure is assumed to hold for the factors. The full state-space model has the form

$$\mathbf{X}_{t_m} = \mathbf{\Lambda}\mathbf{F}_{t_m} + \boldsymbol{\xi}_{t_m}, \tag{8}$$

$$\mathbf{\Psi}(L_m)\mathbf{F}_{t_m} = \mathbf{B}\boldsymbol{\eta}_{t_m}. \tag{9}$$

Equation (8) is the static factor representation of $\mathbf{X}_{t_m}$ as above in (6). Equation (9) specifies a VAR of the factors with lag polynomial $\mathbf{\Psi}(L_m) = \sum_{i=1}^{p} \mathbf{\Psi}_i L_m^i$. The q-dimensional vector $\boldsymbol{\eta}_{t_m}$ contains the orthogonal dynamic shocks that drive the r factors, where the matrix $\mathbf{B}$ is $(r \times q)$-dimensional. The model is already in state space form, since the factors $\mathbf{F}_{t_m}$ are the states. If the dimension of $\mathbf{X}_{t_m}$ is small, the model can be estimated using iterative maximum likelihood (ML). In order to account for large datasets, Doz et al. (2006) propose quasi-ML to estimate the factors, as iterative ML is infeasible in this framework. For a given number of factors r and dynamic shocks q, the estimation proceeds in the following steps:

1. Estimate $\widehat{\mathbf{F}}_{t_m}$ using PCA as an initial estimate. Here, estimation is based on the balanced part of the data. We can obtain this by removing as many values at the end of the sample as long the dataset is unbalanced. The sample size employed for the initial estimation of the factors is then $t_m = 1, \ldots, \min(\{T_m^{x_i}\}_{i=1}^N)$.

2. Estimate $\widehat{\mathbf{\Lambda}}$ by regressing $\mathbf{X}_{t_m}$ on the estimated factors $\widehat{\mathbf{F}}_{t_m}$. The covariance of the idiosyncratic components $\widehat{\boldsymbol{\xi}}_{t_m} = \mathbf{X}_{t_m} - \widehat{\mathbf{\Lambda}}\widehat{\mathbf{F}}_{t_m}$, denoted as $\widehat{\mathbf{\Sigma}}_{\boldsymbol{\xi}}$, is also estimated.

3. Estimate a VAR(p) on the factors $\widehat{\mathbf{F}}_{t_m}$ yielding $\widehat{\mathbf{\Psi}}(L)$ and the residual covariance of $\widehat{\boldsymbol{\varsigma}}_{t_m} = \widehat{\mathbf{\Psi}}(L_m)\widehat{\mathbf{F}}_{t_m}$, denoted as $\widehat{\mathbf{\Sigma}}_{\boldsymbol{\varsigma}}$.

4. To obtain an estimate for $\mathbf{B}$, given the number of dynamic shocks q, apply an eigenvalue decomposition of $\widehat{\mathbf{\Sigma}}_{\boldsymbol{\varsigma}}$. Let $\mathbf{M}$ be the $(r \times q)$-dimensional matrix of the eigen-

vectors corresponding to the q largest eigenvalues, and let the $(q \times q)$-dimensional matrix $\mathbf{P}$ contain the largest eigenvalues on the main diagonal and zero otherwise. Then, the estimate of $\mathbf{B}$ is $\widehat{\mathbf{B}} = \mathbf{M} \times \mathbf{P}^{-1/2}$. The coefficients and auxiliary parameters of the system of equations (8) and (9) is fully specified numerically. The model is cast into state-space form.

5. The Kalman filter or smoother then yield new estimates of the monthly factors. The dataset used for Kalman smoother estimation is now the unbalanced dataset for $t_m = 1, \ldots, T_m$, and T_m is the latest observation available in the entire set of monthly time series

If missing values at the end of the sample are present, as in our setup, the Kalman filter also yields optimal estimates and forecasts for these values conditional on the model structure and properties of the shocks. Thus, it is well suited to tackle ragged-edge problems as in the present context. Nonetheless, one has to keep in mind that in this case the coefficients in system matrices have to be estimated from a balanced sub-sample of data, as in step 1 a fully balanced dataset is needed for PCA initialisation. However, although the system matrices are estimated on balanced data in the first step, the factor estimation based on the Kalman filter applies to the unbalanced data and can tackle ragged-edge problems. The solution is to estimate coefficients outside the state-space model and avoid estimating a large number of coefficients by iterative ML. In the applications below, we will denote the state-space model Kalman filter estimator of the factors as 'KFS-PCA'.

Specification uncertainty The factor approach requires many decisions concerning the specification by the practitioner, starting with the choice of the factor estimation method. In the description of the methods above, we have already provided a few pros and cons. Hence, there might be proponents of either dynamic PCA with vertical realignment of the data or the state-space approach. Indeed, there is an exhaustive literature concerning the relative advantages of factor estimation methods. For example, Marcellino and Schumacher (2008) find only minor differences between alternative estimation methods for factor models in the presence of ragged-edge data. For balanced datasets, there is a long debate on the choice between dynamic or static PCA, see for example Forni et. al (2003), Boivin and Ng (2005), Stock and Watson (2006), D'Agostino, and Giannone (2006), and Schumacher (2007). In the empirical literature on factor forecasting, there is also considerable uncertainty on how to choose the number of factors. For example, the application of information criteria sometimes leads to inferior model specifications in terms of forecast accuracy, see Bernanke and Boivin (2003), footnote 7, Giannone, Reichlin, and Sala (2005), footnote 8, and Schumacher (2007). Thus, when applying factor models for forecasting, there are many decisions that can lead to mis-specification. Below, we will discuss the relevance of the estimation method as well as the number of factors on the now- and forecast accuracy with mixed-frequency and ragged-edge data. In addition to the factor-specific specification issues, decisions concerning the MIDAS regression have to be made.

2.3 Nowcast pooling over many specifications of models

All in all, we have the following groups of individual models: MIDAS and autoregressive MIDAS with single indicators, MIDAS and autoregressive MIDAS with factors estimated by two alternative methods. Below, we will compare many different fixed specifications of these models. In addition to the fixed specifications, we consider model selection based on information criteria and on the past forecasting performance. As a third approach to now- and forecasting, we evaluate alternative ways of pooling.

We pool over alternative specifications of the individual models, following the recent literature by Clark and McCracken (2008), Assenmacher-Wesche and Pesaran (2008) and Garratt et al. (2009), for example. Concerning the relevant model set of pooling, we pool three groups and all the differently specified models therein:

- all models from the single-indicator MIDAS group,
- all models from Factor-MIDAS, and,
- the whole set of single-indicator MIDAS models and Factor-MIDAS.

Therefore, we can assess, first, to what extent nowcast pooling helps within a class of models; second, whether combining the forecasts from single indicator models is better than combining the indicators by means of factors; and, third, whether there are any additional gains from pooling over the forecast models and the indicators together.

Pooling of all the models in a given class and across classes takes into account model uncertainty in its widest sense given the set of models in this exercise. However, when combining across classes, we have to account for the different number of models within each model class. For example, there are substantially more single-indicator MIDAS forecasts than factor models, as the variable selection in MIDAS implies more specifications than the different numbers of factors in the factor approach. To avoid that the size of a group has an effect on the combination of nowcasts, we pool the models in two steps: we first pool the forecasts within a model class (e.g. within single-indicator MIDAS), and then across model classes.

Concerning the weighting schemes, we rely on relatively simple ones only. As the sample under consideration is relatively small, and simple forecast combinations have turned out to provide robust results in the literature, we do not account for more sophisticated pooling methods. The potential presence of model mis-specification and parameter instability suggests that already simple combinations from alternative MIDAS regressions and factor models could yield sizeable gains, see also Clark and McCracken (2008) in this regard. In our application, we use the following weighting schemes:

- equal-weight averaging,
- the median, and
- weighted averaging based on the past performance.

The merits of simple equal-weights pooling or the median are widely known in case structural breaks occur, for example, see Timmermann (2005). However, it might also be beneficial to exploit potential systematic patterns in the past performance of a particular model. For this purpose, we evaluate the past performance of a particular model in terms of mean-squared error (MSE), where we employ a moving window over the previous four quarters. We do this for of all models to be combined in our application and normalise these MSEs to sum to one. The combination weight of a model is finally the inverse of its standardised MSE, see Stock and Watson (2006), p. 522, for a similar weighting scheme. Of course, the forecast weights will be updated for every new recursion in our exercise.

Note that the combinations of MIDAS regressions with single indicators can be regarded as an extension of a particular forecast combination by Stock and Watson (2006), where forecasts from distributed lag models with single-indicators are pooled. We extend their work to the case with mixed-frequency and ragged-edge data. However, the novel aspect of the application carried out here is the combination over different model classes, whereas most of the existing literature on forecasting with mixed-frequency and ragged-edge data, such as Giannone et al. (2008) and Marcellino and Schumacher (2008), is mainly concerned with individual models.

3 Design of the nowcast and forecast comparison exercise

In this section we describe: first, the data used; second, the design of the exercise; finally, the specification of the models.

3.1 Data and replication of the ragged edge

The dataset contains German quarterly GDP growth from 1992Q1 until 2007Q4 and 111 monthly indicators until 2008M2. The monthly indicators include industrial production by sector, incoming orders, turnover, survey on consumer sentiment and business climate, construction, financial time series, raw material price indices, as well as car registrations. More information about the data can be found in appendix A.

The dataset is a final dataset. It is not a real-time dataset and does not contain vintages of data, as they are not available for Germany for such a broad coverage of time series. Furthermore, in Schumacher and Breitung (2008), a considerably smaller real-time dataset for Germany is used, but the results indicate that data revisions do not affect the forecast accuracy considerably. Similar results have been found by Bernanke and Boivin (2003) for the US in a similar context. Thus, we cannot discuss the role of revisions on the relative forecasting accuracy here. However, we take into account that GDP and the monthly indicators are subject to different publication lags, and these lead to certain patterns of missing values at the end of every recursive sample. To consider the availability of the data at the end of the sample due to different publication lags, we follow Giannone et al. (2008) and Banbura and Rünstler (2007) and replicate the availability from the final vintage of data that is available. When downloading the data - the download date for the data used here was 7th March 2008 -, we observe the data availability pattern in terms of the missing values at the end of the data sample. For example, at the beginning of March 2008, we observe interest rates until February 2008, thus there is only one missing value at the end of the sample, whereas industrial production is available up to January 2008, implying two missing values. For each time series, we store the missing values at the end of the sample. Under the assumption that these patterns of data availability remain stable over time, we can impose the same missing values at each point in time of the recursive experiment. Thus, we shift the missing values back in time to mimic the availability of information as in real time.

3.2 Nowcast and forecast design

To evaluate the performance of the models, we estimate and nowcast recursively, where the full sample is split into an evaluation sample and an estimation sample, which is recursively expanded over time. The evaluation sample is between $2000Q1$ and $2007Q4$. For each of these quarters, we want to compute nowcasts and forecasts depending on different monthly information sets. For example, for the initial evaluation quarter $2000Q1$, we want to compute a nowcast in March 2000, one in February, and January, whereas the forecasts are computed from December 1999 backwards in time accordingly. Thus, we have three nowcasts computed at the beginning of each of the intra-quarter months. Concerning the forecasts, we present results up to one quarters ahead. Thus, again for the initial evaluation quarter $2000Q1$, we have three forecasts computed based on information available in October 1999 up to information available in December 1999. Overall, we have six projections for each GDP growth observation of the evaluation period, depending on the information available to make the projection. Note that we have also results for forecast horizons longer than one quarter ahead. However, in line with similar findings by Giannone et al. (2008) for the US, these forecasts generally turned out to be uninformative and will not be reported below.

The estimation sample depends on the information available at each period in time when computing the now- and forecasts. Assume again we want to nowcast GDP for $2000Q1$ in March 2000, then we have to identify the time series observations available at that period in time. For this purpose, we exploit the ragged-edge structure from the end of the full sample of data, as discussed in the previous subsection. For example, for the nowcast GDP for 2000$Q1$ made in March 2000, we know from our full sample that at each period in time, we have one missing value for interest rates and two missing values of industrial production. These missing values are imposed also for the period March 2000, thus replicating the same ragged-edge pattern of data availability. We do this accordingly in every recursive subsample to determine the pseudo real-time observation of each time series. The first observation for each time series is the same for all recursions, namely 1992$M1$. This implies the recursive design with increasing information over time available for estimating the MIDAS regressions and factor models. To replicate the publication lags of GDP, we exploit the fact that GDP of the previous quarter is available for now- and forecasting at the beginning of the third month of the next quarter. Note that we reestimate the factors and forecast equations every recursion when new information becomes available, so factor weights and forecast model coefficients are allowed to change over time.

For each evaluation period, we compute six now- and forecasts depending on the available information in the respective months. To compare the nowcasts with the realisations of GDP growth, we use the mean-squared error (MSE). In our tables, we provide relative MSE, where the MSE of a particular forecast model is divided by the in-sample mean of GDP growth. A relative MSE smaller than one indicates that the forecast of a model for the chosen now- and forecast horizon is to some extent informative for current and future GDP, as the in-sample mean has turned out to be a tough competitor, see Giannone et al. (2008).

3.3 Specification of MIDAS and factor models

To specify the now- and forecast models in the applications below, we follow three approaches: fixed specification over recursions, recursive specification by information criteria, and recursive specification by past performance.

The range of auxiliary parameters to choose the fixed specifications from is set as follows: In the factor model framework, we compute now- and forecasts for all possible combinations of r and q and evaluate them with a maximum of $r = 6$ static factors. Given r, we consider all possible combinations of r and the number of dynamic factors with $q \leq r$. The maximum lag order for MIDAS was set to six, $K = 6$. The empirical estimation results show, that longer lags typically play no role, so the choice of K is not restrictive. Estimation of single-indicator MIDAS is carried out with all combinations of indicators and with and without AR terms, so we end up with 222 models used for now- and forecasting. Regarding the factor models, we have 42 different specifications with different r and q for the state-space factor model and the dynamic PCA approach each. Additionally, we have the two different Factor-MIDAS projections with and without AR terms, so we end up with 168 models.

The information criteria chosen for model selection are the following: We determine the number of static and dynamic factors, r and q, respectively, using information criteria from Bai and Ng (2002), in particular their criterion IC_{p2}, and Bai and Ng (2007) with $m = 1.0$ following the Monte Carlo results in Bai and Ng (2007). The maximum number of factors is the same as in the fixed case above. For estimating the state-space factor model, a lag order determination is required to specify the factor VAR. For this purpose, we apply the Bayesian information criterion (BIC) with a maximum lag order of $p = 6$ months. The chosen lag lengths are usually very small with only one or two lags in most of the cases. For single-indicator MIDAS, the selection of variables as well as the AR terms is carried out using the Bayesian information criterion (BIC). For a motivation of the use of BIC in the MIDAS context, see Galvão (2007), p. 14. To compute the BIC, we have to take into account the exponential lag polynomial determined by $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$, and the number of coefficients in MIDAS is set to two in case no AR term is incorporated and three otherwise, see equation (4).

To specify the models by inspecting their past performance, we refer to the MSE computed over the previous four quarters for each model, in line with the weighting scheme for pooling in subsection 2.3. The MSEs are computed recursively for the entire set of models, then the best-performing one is chosen within a class. Thus, model specifications can change over time regarding variable selection and the number of factors as well as the AR terms.

Concerning the NLS estimation of MIDAS equations, we use a large variety of initial parameter specifications, and compute the residual sum of squares (RSS). The parameter set with the smallest RSS then serves as the initial parameter set for NLS estimation. The parameters of the exponential lag function are restricted to $\theta_1 < 2/5$ and $\theta_1 < 0$. To specify the dynamic PCA estimator of the factors following Forni et al. (2005), we use the frequency-domain auxiliary parameters $M = 24$ and $H = 60$ for estimating the spectral density, see appendix B for details.

4 Now- and forecasts from single models

In the first subsection we compute forecasts over the entire range of indicators in MIDAS regressions, and over specifications with and without AR terms. During the recursive application, we hold the respective specifications fixed. When nowcasting with factor models, we consider all combinations of dynamic and static factors. For both types of models, we obtain a large set of results that helps to identify the best-performing models and specifications within and across the model classes ex-post.

In the second subsection, we consider sequential (ex-ante) specification by information criteria. Specifically, we apply information criteria for model and variable selection to the MIDAS and Factor-MIDAS models estimated over recursive subsamples. In the same subsection, we evaluate specification based on the past performance. Specifically, we use the forecast performance in terms of MSE over the past four periods in order to select the best-performing specification within the group of Factor- and single-indicator MIDAS.

This procedure, as well as selection by information criteria, relies on in-sample information only.

When using fixed specifications over all recursions, a comparison of the best models within each category of models and a comparison across groups allows for an assessment of the potential forecast accuracy in case a practitioner knew the right specification in real time. Thus, searching ex post for the right specification is to some extent data mining. Instead, the use of information criteria and selection based on the past performance comes closer to the specification problems in a real-time context, and shows to what extent the results based on fixed specifications can be matched under more realistic conditions.

4.1 Fixed specifications

Now- and forecast results for the factor models and single-indicator MIDAS based on fixed specifications can be found in table 1. The table shows relative MSEs to the naive benchmark, which is the in-sample mean of GDP growth. The now- and forecasts are shown for monthly horizons $h_m = 1, \dots, 6$, where horizons one to three belong to the nowcast. Horizon $h_m = 1$ is a nowcast made in the third month of the respective quarter, whereas horizon $h_m = 2$ is the nowcast made in the second month of the current quarter. Thus, similar to standard forecast comparisons, increasing horizons correspond to less information available for now- and forecasting, and we expect an increasing MSE for increasing horizons h_m. In the table, MSE results are shown for selected MIDAS single-indicator models and factor models. To find the best-performing models in terms of MSE, we chose those with a relative MSE smaller than one for $h_m = 1, 2, 3$. To order the models, we use the average of the MSE over $h_m = 1, 2, 3$.

In panel A of table 1, we find results concerning single-indicator MIDAS. We see that there are 20 models that have a relative MSE smaller than one up to $h_m = 3$. Regarding forecasts ($h_m = 4, 5, 6$), only half of the models can consistently outperform the naive benchmark, and in most of the cases only to a small extent. We do not report results for $h_m > 6$, as the forecasts are almost always uninformative compared to the benchmark. Among the top-performing models, surveys on business expectations play a big role, whereas industry statistics like incoming orders or turnover as well as interest rates play only a minor role. Concerning the MIDAS projections, both regressions with and without AR terms can be found among the best-performing models. Panel B of table 1 provides results for Factor-MIDAS. Here only 5 models yield relative MSEs consistently smaller than one for $h_m = 1, 2, 3$. Regarding forecasts ($h_m = 4, 5, 6$), the factor models in most of the cases perform worse than the benchmark. Concerning the specifications, models with only one factor ($r = q = 1$) do best, and we find both MIDAS projections with and without AR terms in the ranking.

According to the results in table 1, factor models tend to perform worse than the best-performing single-indicator MIDAS models. However, in terms of the size of the MSE, the overall best-performing single-indicator model (survey: bus. exp., wholesale trade)

Table 1: Now- and forecast results for single-indicator MIDAS and factor models, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
A. Single-indicator MIDAS							
survey: bus. exp., wholesale trade	MIDAS	0.72	0.67	0.78	0.80	0.67	0.87
survey: bus. exp., consumer goods prod.	MIDAS	0.78	0.67	0.75	0.89	0.92	0.96
survey: bus. conditions, wholesale trade	MIDAS	0.70	0.79	0.90	0.88	1.12	1.19
survey: bus. exp., consumer goods prod.	AR-MIDAS	0.93	0.71	0.80	0.91	0.94	0.97
stocks finished goods, consumer goods prod.	MIDAS	0.82	0.80	0.82	1.01	1.01	1.07
survey: bus. exp., retail trade	AR-MIDAS	0.79	0.79	0.87	1.16	1.17	0.91
survey: bus. exp., retail trade	MIDAS	0.79	0.77	0.91	1.23	1.22	0.98
survey bus. exp., wholesale trade	AR-MIDAS	0.97	0.72	0.84	0.82	0.68	1.12
survey consumer sentiment (GfK)	AR-MIDAS	0.74	0.84	0.95	0.93	1.12	1.24
stocks finished goods, consumer goods prod.	AR-MIDAS	0.87	0.84	0.86	1.01	1.00	1.11
survey: bus. cond., wholesale trade	AR-MIDAS	0.81	0.85	0.95	0.92	1.06	1.26
long-term interest rate (1-2 years mat.)	MIDAS	0.86	0.90	0.87	0.89	0.85	0.86
turnover (abroad), intermediate goods prod.	MIDAS	0.89	0.83	0.92	0.83	0.89	1.04
production, intermediate goods prod.	MIDAS	0.82	0.91	0.95	0.96	0.99	1.03
survey: bus. exp., non-dur. cons. goods prod.	MIDAS	0.95	0.88	0.85	1.04	0.94	1.52
long-term interest rate (5-6 years mat.)	MIDAS	0.86	0.92	0.94	0.98	1.02	0.93
survey: bus. cond., investm. goods prod.	MIDAS	0.90	0.93	0.93	1.06	1.12	1.11
orders (domestic), intermediate goods prod.	MIDAS	0.88	0.92	0.99	1.23	1.39	1.04
short-term employed	AR-MIDAS	0.95	0.95	0.96	0.98	0.92	1.04
turnover (abroad), mechanical engineering	AR-MIDAS	0.94	0.97	0.97	0.95	0.95	1.28
B. Large factor models							
VA-DPCA, $r = 1$, $q = 1$	AR-MIDAS	0.77	0.66	0.84	1.00	0.97	1.09
VA-DPCA, $r = 1$, $q = 1$	MIDAS	0.69	0.76	0.96	0.96	1.02	1.08
KFS-PCA, $r = 1$, $q = 1$	MIDAS	0.73	0.89	0.85	1.09	1.07	0.89
VA-DPCA, $r = 2$, $q = 2$	AR-MIDAS	0.85	0.77	0.93	1.08	1.03	1.06
KFS-PCA, $r = 1$, $q = 1$	AR-MIDAS	0.85	0.90	0.82	1.12	1.08	1.02

Note: The entries in the table are relative MSEs relative to the in-sample mean, where the mean is recomputed every subsample. The model abbreviations in the first column are: VA-DPCA refers to the vertical realignment and dynamic PCA used in Altissimo et al. (2006), and KFS-PCA is the Kalman smoother of state-space factors according to Doz et al. (2006). The projection MIDAS-basic is the projection from Ghysels and Valkanov (2006), and AR-MIDAS is the basic MIDAS regression with an autoregressive term as proposed by Clements and Galvão (2007).

and the best-performing factor model (VA-DPCA, $r = 1$, $q = 1$) seem to work similarly well for the nowcast, as the ranking of top models is changing over horizons.

The results obtained so far are based on ex-post forecast MSEs only. Taking the results literally, the potential user of these methods could make use of the best-performing specifications. However, it is unclear whether the same results can be obtained in real-time also, when no a-priori knowledge about the best specifications is available to the practitioner. We consider this issue in the next subsection.

4.2 Information-criteria model selection and specification based on past performance

The first question we address in this subsection is whether we can find the best-performing specifications with in-sample information only. In particular, can we find the best-performing indicator variables for MIDAS and the optimal number of factors without resorting on the ex-post forecast errors? The second question we ask is whether it is better to use model specification based on information criteria or on the past forecasting performance.

To address both questions, we will now compare the performance of fixed specifications to time-varying specifications, where we use only information from the recursive subsamples to determine the model specifications. In table 2, we report the relative MSEs of the models specified using information criteria and the past MSE performance, as described in subsection 3. In panel A of the table, we present the results based on information-criteria

Table 2: Now- and forecast results for single-indicator MIDAS and Factor-MIDAS, information criteria model selection, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
A. Information criteria model selection							
single-indicator MIDAS/AR-MIDAS	BIC	0.96	1.07	1.50	1.04	1.70	1.01
VA-DPCA, MIDAS	Bai, Ng (2002, 2007)	1.09	1.00	0.95	1.19	1.35	1.08
VA-DPCA, AR-MIDAS	Bai, Ng (2002, 2007)	1.17	0.84	0.83	1.28	1.05	0.77
KFS-PCA, MIDAS	Bai, Ng (2002, 2007)	1.29	1.66	0.83	1.59	1.04	0.73
KFS-PCA, AR-MIDAS	Bai, Ng (2002, 2007)	1.48	1.53	0.88	1.26	1.22	1.07
B. Model and variable selection by past MSE performance							
single-indicator MIDAS	MSE	0.86	1.26	0.99	1.20	1.05	1.24
large factor models	MSE	0.89	0.84	0.85	0.93	0.91	0.66

Note: See table 1.

model selection. When BIC is employed for selecting the predictor in MIDAS as well as the AR terms, we find only for $h_m = 1$ a relative MSE smaller than one. For all the other horizons, the now- and forecasts are uninformative. Regarding the factor models, the information criteria also select specifications that perform worse than the benchmark for almost all the horizons with only a few exceptions. Panel B of table 2 contains the results with model specification based on the past performance of the models in terms of MSE. For single-indicator MIDAS, where both AR terms as well as variable selection is done by BIC recursively, there is again only for $h_m = 1$ a relative MSE smaller than one. The factor models, however, where the number of factors as well as AR terms are specified using the past MSE, yield a good performance compared with the benchmark. For all horizons, the time-varying specifications yield relative MSEs smaller than one. Note that the factor model performance is for some of the horizons even better than the fixed specifications from the table 1. Therefore, the past performance seems to contain some information that can - in contrast to the fixed specifications over time - be exploited for now- and forecasting with factors.

If we compare the overall results from table 2 based on information criteria and performance-based model selection to the results with fixed specifications in table 1, the general impression is, that forecasting is much more difficult when the model specifications are unknown in pseudo real-time, as the relative MSEs in table 2 are generally larger than those in table 1. In particular, the information criteria applied to model selection lead to clearly inferior results. For example, without knowing the preferable predictor for MIDAS or the correct number of factors a priori, it is difficult to specify these forecast models properly, and it is not possible to achieve the optimistic now- and forecast results from table 1. In this context, however, factor model specification based on the past performance can still outperform the benchmark.

5 Nowcast pooling

After discussing the individual models' performance, we now assess nowcast pooling. As for information criteria and selection based on the past performance, pooling is only to a small extent subject to the data-mining critique, as only in-sample information is used to specify the weights.

In table 3, we present now- and forecast results of the alternative pooling schemes described in section 2.3. The first three rows in the table contain the results when all the single-indicator MIDAS now- and forecasts are combined using equally weighted mean, MSE-based mean as well as the median. The results indicate an information content for both the nowcast and the forecast one quarter ahead, as the relative MSEs are smaller than one in many cases.[2] Concerning the pooling methods, the median tends to perform worse than the unweighted mean, and both are outperformed by the MSE-based weighted mean. Compared with figures based on model selection from table 2, the results are now

[2]Note that results for larger horizons $h_m > 6$ are generally still uninformative with few exceptions.

Table 3: Now- and forecast results for nowcast pooling, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
single-indicator MIDAS	equal-weight mean	0.90	0.93	0.95	0.94	0.97	1.00
single-indicator MIDAS	MSE-weighted mean	0.84	0.89	0.92	0.88	0.92	0.95
single-indicator MIDAS	median	0.95	0.97	0.98	1.00	1.01	1.04
large factor models	equal-weight mean	0.88	0.91	0.83	1.00	0.89	0.64
large factor models	MSE-weighted mean	0.87	0.88	0.80	0.82	0.88	0.61
large factor models	median	0.89	0.84	0.85	0.93	0.91	0.66
all models	equal-weight mean	0.76	0.81	0.83	0.94	0.88	0.75
all models	MSE-weighted mean	0.79	0.84	0.82	0.81	0.84	0.67
all models	median	0.79	0.81	0.85	0.94	0.91	0.78

Note: See table 1.

clearly better, indicating advantages of pooling over model selection. Rows four to six contain results from pooling all the factor models that differ with respect to the number of factors and the AR term in the Factor-MIDAS projection. The results are again better than those based on model selection from table 2, and the MSE-based weighted mean outperforms the other weighting schemes for most of the horizons, although the differences are smaller than in the case of single-indicator MIDAS. Comparing the levels of relative MSEs between factor models and single-indicator MIDAS, we find a slightly better performance of the factor combinations.

The final three rows contain now- and forecast combinations of all the models under consideration. Here, the ranking of the different pooling methods is less clear. The interesting result is that the combination of all forecast models provides overall smaller relative MSEs than the combinations of factor and single-indicator MIDAS alone. Thus, taking into account model uncertainty to a wider extent than just pooling within a model class seems to improve the forecasting performance. Furthermore, pooling over all models almost entirely outperforms the individual models chosen by information criteria or the past performance in table 2.

But what about the performance compared to the fixed specifications in table 1? Is nowcast pooling also competitive to the ex-post best-performing models? A direct comparison of tables 1 and 3 suggests that even pooling cannot perform as well as the ex-post best performing models, though the differences are often small.

In order to analyze this issue in more details, we investigate the relationship between the groups of individual models and the forecast combinations. For this purpose, we present percentiles of the relative MSEs from the alternative nowcast pools for each horizon. The percentiles provide an indication of how the pooling MSE values compare to those of the individual models. Table 4 contains the results. The entries in the table can

Table 4: Percentiles of the MSEs from now- and forecast pooling in the cumulative distribution of individual models

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
A. Pooling vs single-indicator MIDAS							
MIDAS models	equal-weight mean	0.13	0.16	0.15	0.13	0.21	0.22
MIDAS models	MSE-weighted mean	0.08	0.11	0.10	0.07	0.11	0.13
MIDAS models	median	0.23	0.22	0.21	0.26	0.31	0.29
all models	equal-weight mean	0.03	0.04	0.02	0.13	0.07	0.02
all models	MSE-weighted mean	0.05	0.07	0.02	0.02	0.04	0.00
all models	median	0.05	0.04	0.03	0.13	0.10	0.02
B. Pooling vs individual large factor models							
large factor models	equal-weight mean	0.25	0.22	0.11	0.20	0.06	0.08
large factor models	MSE-weighted mean	0.24	0.17	0.06	0.00	0.06	0.06
large factor models	median	0.25	0.14	0.13	0.06	0.07	0.10
all models	equal-weight mean	0.07	0.12	0.11	0.09	0.06	0.21
all models	MSE-weighted mean	0.12	0.14	0.08	0.00	0.04	0.10
all models	median	0.11	0.12	0.13	0.08	0.07	0.24

Note: The entries in the table can be interpeted as follows. An entry x implies that the MSE of the combination of single-indicator MIDAS models is larger than $100 \cdot x$ percent of the MSEs of the individual MIDAS models, and accordingly smaller than $100 \cdot (1 - x)$ percent of the MSEs from the worse-performing models. Thus, the pool is in the $(100 \cdot x)$th percentile of the distribution of individual models. In the table, the model set used in the combination of now- and forecasts can be found in the first column of the table. The second column contains the weighting methods employed.

be interpreted as follows. In panel A of table 4, entry 0.13 for $h_m = 1$ implies that the MSE of the combination of single-indicator MIDAS models is larger than 13 percent of the MSEs of the individual MIDAS models, and accordingly smaller than 87 percent of the MSEs from the worse-performing models. Thus, the pool is in the 13th percentile of the distribution of individual models. The results in table 4 confirm that nowcast pooling is in almost all of the cases not the best-performing method. However, based on the MSE-weighted mean for all horizons reported, the pool is between the 7th and 13th percentile compared to the individual single-indicator MIDAS models (row 2). Combining factor models and single-indicator MIDAS reduces the relative MSE further, and the pooled forecast ends up in the 7th percentile and lower (row 5). The best combinations can outperform between 93 and 100 percent of the individual MIDAS models, depending on the forecast horizon.

Looking at the distribution of factor models in panel B, we find that pooling of the factor models only using the MSE-weighted mean is doing better than 76 to 94 percent of the individual factor models for $h_m = 1, 2, 3$. Regarding the forecast performance for

$h_m = 4, 5, 6$, combinations can outperform between 94 and 100 percent of the individual factor models. Thus, pooling improves the performance of the forecast one quarter ahead more than the nowcast. The forecast combinations of all models are in most of the cases better than the combinations of the factor models only (row 2 and 5 in panel B). Compared with the distribution of single-indicator MIDAS from panel A, we see generally smaller percentiles for the single-indicator MIDAS. This implies that the combinations do better than the majority of single-indicator MIDAS forecasts, whereas the individual factor models can be outperformed to a lesser extent. Thus, factor models that already exploit the large information set seem to be a tougher competitor to the combinations than the individual single-indicator MIDAS models.

Overall, the combinations seem to work well and leave most of the individual models behind. They are in most of the cases not the best-performing now- and forecast devices, as there are a few fixed specifications that can do better ex-post, but they cannot be identified in real-time.

6 Robustness of the results

In this section we briefly report results to evaluate the robustness of the findings we have obtained so far. In particular, we discuss nowcast results in a subsample to check the robustness of the results over time; we extend the number of indicators in the MIDAS regression; and we employ alternative information criteria for specifying Factor-MIDAS.

6.1 Subsample analysis

In order to check the stability of the results over time, we split the evaluation sample and provide results for the second, more recent period. There are two reasons for this. Banerjee et al. (2005) and Banerjee and Marcellino (2006) find that forecast models with single indicators often have a time-varying information content for future economic activity. Similar problems occur for the specification of factor models, in particular related to the number of static and dynamic factors, see Schumacher (2007). By splitting the sample, we can discuss how stable the rankings based on the chosen specifications are over time.

Another argument is based on the general finding for many industrialised countries that, complementary to the Great Moderation phenomenon, the forecast performance of many sophisticated forecast models has broken down, see D'Agostino et al. (2006) and Campbell (2007), for example. In particular, for very recent samples, outperforming naive forecasts has proven to be difficult.

Due to the relatively short sample size in the current exercise, as common in empirical work on euro area macro data (see e.g. Banbura and Rünstler (2007)), the evaluation sample for the stability check is $2004Q1$-$2007Q4$. For detailed results, the reader is referred to appendix C, whereas the main findings can be summarised as follows. First, the overall performance of the models has improved a little compared to the benchmark.

Second, the relative ranking of single-indicator MIDAS and factor models remains the same. With fixed specifications, there are a few single-indicator MIDAS forecasts that outperform the factor models. With model and variable selection based on information criteria, the single-indicator MIDAS models become mostly uninformative, and only past MSE performance helps to find models with some information content for future GDP.

Third, pooling over many specifications and models is again the most robust device for now- and forecasting, and it outperforms model selection based on information criteria and the past performance.

Fourth, the ranking of the single-indicator models has changed, in line with the evidence in the papers cited above. Thus, one might find ex post favourable evidence on a particular indicator but, due to changing information content, these relevant indicators cannot be detected in real time. In this regard, our results are in line with De Mol et al. (2008) and Banerjee and Marcellino (2006). In our framework and given the dataset used, we find evidence that nowcast pooling can circumvent these problems to a good extent.

Finally, the ranking of the factor models has changed to a smaller extent. This is not in contrast to the previous result, since the weight of each indicator in the estimated factors can change over time.

6.2 Double-indicator MIDAS

We now consider more than one indicator in MIDAS regressions. In particular, we include industrial production into MIDAS and add sequentially all the other indicators to the MIDAS regression. As industrial production is one of the key indicators for GDP and subject to investigation in many mixed-frequency studies, see for example Clements and Galvão (2008), it might be a natural candidate for extending the MIDAS regression. Thus, we end up with double-indicator MIDAS. Details can be found in appendix D.

In brief, we find two main results. First, comparing fixed specifications of single- and double-indicator MIDAS, we find that models with one indicator still represent the top 5 models in terms of relative MSE. Among the best models with information content up to three months, there about 25% double-indicator MIDAS models and 75% models with a single indicator.

Second, now- and forecast pooling over all model classes, including factor models and MIDAS with one and two predictors, provides very similar results as before. It again outperforms individual models specified by information criteria and past performance. Hence, the information in the cross section of data is already successfully exploited by the combination of the models with single predictors, and adding further regressors seem to add little information for future GDP.

6.3 BIC specification of Factor-MIDAS

The information criteria by Bai and Ng (2002, 2007) aim at minimising the overall idiosyncratic variance, given a certain penalty function that depends on the sample size. Thus, the number of factors chosen are not optimised with respect to forecasting GDP. In order to take into account the dynamic correlations between the factor estimates and GDP, we will follow the seminal work by Stock and Watson (2002), where BIC is employed for selecting the relevant factors in the projection with single-frequency data. In particular, we set the number of static factors equal to $r = 6$, and estimate factors with all possible combinations of $q \leq r$. We then compute now- and forecasts with Factor-MIDAS for each set of factors. The BIC is used recursively for choosing the Factor-MIDAS regression used for the now- and forecasts.

The main result is that the factor now- and forecasts based on BIC are informative only for the horizons one and three.[3] Compared to the results based on information criteria in table 2, the results are only better for horizon one. Thus, we confirm our main conclusions that information criteria tend to select model specifications with almost uninformative now- and forecast performance.

7 Conclusions

In this paper, we discuss nowcast pooling versus nowcasting with single models in the presence of model uncertainty, exacerbated by the presence of mixed frequency data with ragged edges. The nowcasts are based on MIDAS regressions with few indicators and Factor-MIDAS based on large datasets, and both models can tackle the 'ragged-edge' data as well as the different sampling frequencies of GDP and many business cycle indicators. Thus, the nowcasting perspective followed in this paper takes into account the publication lags of statistical data that decision makers face in their everyday business of assessing the current state of the economy.

To address model uncertainty in the set of nowcast models chosen, we compare the performance of many alternative specifications with respect to alternative factor estimation methods, number of factors, indicators selected for MIDAS, the role of autoregressive dynamics, and others. The different models are applied to a German post-unification dataset, containing of about one hundred monthly indicators. The now- and forecasts of the individual models and pooling are compared with respect to their predictive ability for German GDP growth.

In this framework, we discuss three main questions. First, searching in the set of all possible models under analysis, is it possible to find specifications that outperform a simple benchmark in terms of mean-squared error (MSE)? The answer is yes, perhaps not surprisingly given the extensive search in such a large model set. More interestingly,

[3] The exact relative MSEs are 0.72, 1.03, 0.85, 1.02, 1.07, and 0.99 for horizon one to six, respectively.

single indicator models tend to outperform factor models in this ex-post evaluation.

Second, since the search described above is based on full sample results, it might be subject to the data-mining critique. Therefore, it may not be suited for a real-time implementation, and the question arises whether we get similar gains with respect to the benchmark by selecting the forecasting models based either on information criteria or on their past performance. The answer is that it is much more difficult to beat the benchmark in this case, with the exception of Factor-MIDAS specifications based on past forecasting performance. In general, now- and forecasting based on information-criteria model selection performs clearly worse than the fixed specifications identified ex post.

Third, as a method to avoid the specification search, all the nowcasts and forecasts can be pooled together, using different weighting schemes. The question is then whether this approach yields additional gains with respect to the factor specification based on the past performance. The answer is yes, and this is particularly the case when all single-indicator and all Factor-MIDAS forecasts are combined together using inverse MSE weights. While in general the resulting pooled now- and forecasts still cannot outperform the very best single models based on fixed specifications, it is better than 93-100% of all the single indicator forecasts, and of 86-100% of all the Factor-MIDAS forecasts, depending on the horizon. Additionally, a subsample analysis of the results shows that the ranking of the best fixed specifications changes, in particular, with respect to variable selection in MIDAS, in line with other findings about the time-varying predictive power of single indicators.

To conclude, the results obtained in the present paper provide strong support in favour of pooling for nowcasting and short-term forecasting. Actually, with respect to previous studies, pooling seems to play an even more important role in a context characterized by a large set of models and mixed-frequency and ragged-edge indicators.

References

[1] Assenmacher-Wesche, K., H. Pesaran (2008), Forecasting the Swiss economy using VECX* models: an exercise in forecast combination across models and observation windows, National Institute Economic Review 203, 91-108.

[2] Altissimo, F., R. Cristadoro, M. Forni, M. Lippi, G. Veronese (2006), New Eurocoin: Tracking Economic Growth in Real Time, CEPR Working Paper 5633.

[3] Bai, J., S. Ng (2002), Determining the Number of Factors in Approximate Factor Models, Econometrica 70, 191-221.

[4] Bai, J., S. Ng (2007), Determining the Number of Primitive Shocks in Factor Models, Journal of Business and Economic Statistics 25, 52-60.

[5] Banbura, M., G. Rünstler (2007), A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP, ECB Working Paper 751.

[6] Banerjee, A., M. Marcellino (2006), Are there any reliable leading indicators for US inflation and GDP growth?, International Journal of Forecasting 22, 137-151.

[7] Banerjee, A., M. Marcellino, I. Masten (2005), Leading Indicators for Euro-area Inflation and GDP Growth, Oxford Bulletin of Economics and Statistics, 67, 785-813.

[8] Banerjee, A., M. Marcellino, I. Masten (2007), Forecasting macroeconomic variables using diffusion indexes in short samples with structural change, in Rapach, D., Wohar, M. (eds.), Forecasting in the Presence of Structural Breaks and Model Uncertainty, Elsevier, forthcoming.

[9] Barhoumi, K., S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, G. Rünstler, K. Ruth, C. Van Nieuwenhuyze (2008), Short-term forecasting of GDP using large monthly datasets: a pseudo real-time forecast evaluation exercise, ECB Occasional Paper 84.

[10] Bernanke, B., J. Boivin (2003), Monetary Policy in a Data-Rich Environment, Journal of Monetary Economics 50, 525-546.

[11] Boivin, J., S. Ng (2005), Understanding and Comparing Factor-Based Forecasts, International Journal of Central Banking 1, 117-151.

[12] Campbell, S. (2007), Macroeconomic Volatility, Predictability, and Uncertainty in the Great Moderation: Evidence From the Survey of Professional Forecasters, Journal of Business & Economic Statistics 25, 191-200.

[13] Chevillon, G., D. Hendry (2005), Non-parametric direct multi-step estimation for forecasting economic processes, International Journal of Forecasting 21, 201-218.

[14] Clark, T., M. McCracken (2008), Averaging Forecasts from VARs with Uncertain Instabilities, Journal of Applied Econometrics, forthcoming.

[15] Clements, M., A. Galvão (2008), Macroeconomic Forecasting With Mixed-Frequency Data: Forecasting Output Growth in the United States, Journal of Business & Economic Statistics 26, 546-554.

[16] Clements, M., A. Galvão (2009), Forecasting US output growth using Leading Indicators: An appraisal using MIDAS models, Journal of Applied Econometrics, forthcoming.

[17] Clements, M., D. Hendry (2004), Pooling of Forecasts, Econometrics Journal 7, 1-31.

[18] D'Agostino, A., D. Giannone, P. Surico (2006), (Un)Predictability and macroeconomic stability, ECB Working Paper 605.

[19] D'Agostino, A., D. Giannone (2006), Comparing alternative predictors based on large-panel factor models, Working Paper 680.

[20] De Mol, C., D. Giannone, L. Reichlin (2008), Forecasting Using a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?, Journal of Econometrics 146, 318-328.

[21] Den Reijer, A. (2005), Forecasting Dutch GDP using large scale factor models, DNB Working Paper 28.

[22] Doz, C., D. Giannone, L. Reichlin (2006), A two-step estimator for large approximate dynamic factor models based on Kalman filtering, ECARES, mimeo.

[23] Eickmeier, S., C. Ziegler (2007), How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach, Journal of Forecasting, forthcoming.

[24] Forni, M., M. Hallin, M. Lippi, L. Reichlin (2003), Do financial variables help forecasting inflation and real activity in the euro area? Journal of Monetary Economics 50, 1243 - 1255.

[25] Forni, M., M. Hallin, M. Lippi, L. Reichlin (2005), The Generalized Dynamic Factor Model: One-sided estimation and forecasting, Journal of the American Statistical Association 100, 830-840.

[26] Galvão, A. (2007), Changes in Predictive Ability with Mixed Frequency Data, Queen Mary University of London Working Paper 595.

[27] Garratt, A., E. Mise, G. Koop, S. Vahey (2009), Real-time Prediction with UK Monetary Aggregates in the Presence of Model Uncertainty, Journal of Business and Economic Statistics, forthcoming.

[28] Ghysels, G., R. Valkanov (2006), Linear Time Series Processes with Mixed Data Sampling and MIDAS Regression Models, mimeo.

[29] Ghysels, G., A. Sinko, R. Valkanov (2007), MIDAS Regressions: Further Results and New Directions, Econometric Reviews 26, 53-90.

[30] Giannone, D., L. Reichlin, L. Sala (2005), Monetary Policy in Real Time, in: M. Gertler, K. Rogoff (eds.), NBER Macroeconomic Annual 2004, No. 19, MIT Press, 161-224.

[31] Giannone, D., L. Reichlin, D. Small (2008), Nowcasting GDP and Inflation: The Real-Time Informational Content of Macroeconomic Data Releases, Journal of Monetary Economics 55, 665-676.

[32] Inoue, A., L. Kilian (2006), On the selection of forecasting models, Journal of Econometrics 130, 273-306.

[33] Kapetanios, G., V. Labhard, S. Price (2008), Forecast combination and the Bank of England's suite of statistical forecasting models, Economic Modelling 25, 772–792.

[34] Lütkepohl, H., R. Brüggemann (2006), A small monetary system for the euro area based on German data, Journal of Applied Econometrics, 21, 683-702.

[35] Marcellino, M. (2004), Forecast Pooling for Short Time Series of Macroeconomic Variables, Oxford Bulletin of Economic and Statistics 66, 91-112.

[36] Marcellino, M., G. Kapetanios (2006), A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions, CEPR WP 5620.

[37] Marcellino, M., C. Schumacher (2008), Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP, CEPR Working Paper 6708.

[38] Marcellino, J. Stock, M. Watson (2003), Macroeconomic forecasting in the Euro area: country specific versus Euro wide information, European Economic Review 47, 1-18.

[39] Marcellino, M., J. Stock, M. Watson (2006), A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series, Journal of Econometrics 135, 499-526.

[40] Mariano, R., Y. Murasawa (2003), A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series, Journal of Applied Econometrics 18, 427-443.

[41] Schneider, M., M. Spitzer (2004), Forecasting Austrian GDP using the generalized dynamic factor model, OeNB Working Paper 89.

[42] Schumacher, C. (2007), Forecasting German GDP using alternative factor models based on large datasets, Journal of Forecasting 26, 271-302.

[43] Schumacher, C., J. Breitung (2008), Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data, International Journal of Forecasting, 24, 368-398.

[44] Stock, J., M. Watson (2002), Macroeconomic Forecasting Using Diffusion Indexes, Journal of Business & Economic Statistics 20, 147-162.

[45] Stock, J., M. Watson (2004), Combination Forecasts of Output Growth in a Seven-Country Data Set, Journal of Forecasting 23, 405-430.

[46] Stock, J., M. Watson (2006), Forecasting with Many Predictors, in: Elliot, G., Granger, C., Timmermann, A. (eds.), Handbook of Economic Forecasting, Vol 1, 515-554.

[47] Wallis, K. (1986), Forecasting with an econometric model: The 'ragged edge' problem, Journal of Forecasting 5, 1-13.

[48] Watson, M. (2003), Macroeconomic Forecasting Using Many Predictors, in: Dewatripont, M., Hansen, L., Turnovsky, S. (eds.), Advances in Economics and Econometrics, Theory and Applications, Eight World Congress of the Econometric Society, Vol 3, 87-115.

A Monthly dataset

This appendix describes the time series for the German economy used in the forecasting exercise. The whole data set for Germany contains 111 monthly time series over the sample period from $1992M1$ until $2008M2$. The time series cover broadly the following groups of data: prices, labour market data, financial data (interest rates, stock market indices), industry statistics, construction statistics, surveys and miscellaneous indicators.

The source of the time series is the Bundesbank database. The download date of the dataset is 7th March 2008. In this dataset, there are differing missing values at the end of the sample. For example, whereas financial time series are available up to $2008M2$, industrial time series like production, orders and so on are only available up to $2008M1$. This leads to a ragged-edge structure at the end of the sample, which serves as a template to replicate the ragged edges in past pseudo real-time periods as described in the main text.

Natural logarithms were taken for all time series except interest rates. Stationarity was obtained by appropriately differencing the time series. Most of the time series taken from the above source are already seasonally adjusted. Remaining time series with seasonal fluctuations were adjusted using Census-X12 prior to the forecast simulations. Extreme outlier correction was done using a modification of the procedure proposed by Watson (2003). Large outliers are defined as observations that differ from the sample median by more than six times the sample interquartile range, see Watson (2003), p. 93. The identified observation is set equal to the respective outside boundary of the interquartile.

A.1 Prices

producer price index
producer price index without energy

consumer price index
consumer price index without energy
export prices
import prices
oil price Brent GB

A.2 Labour market

unemployed
unemployment rate
employees and self-employed
employees, short-term
productivity, per employee
productivity, per hour
wages and salaries per employee
wages and salaries per hour
vacancies

A.3 Interest rates, stock market indices

money market rate, overnight deposits
money market rate, 1 month deposits
money market rate, 3 months deposits
bond yields on public and non-public long term bonds with average maturity from 1 to 2 years
bond yields on public and non-public long term bonds with average maturity from 5 to 6 years
bond yields on public and non-public long term bonds with average maturity from 9 to 10 years
yield spread: bond yields with maturity from 1 to 2 years minus 3 months money market rate
yield spread: bond yields with maturity from 5 to 6 years minus 3 months money market rate
yield spread: bond yields with maturity from 9 to 10 years minus 3 months money market rate
CDAX share price index
DAX German share index
REX German bond index
exchange rate US dollar/Deutsche Mark
indicator of the German economy's price competitiveness against 19 industrial countries based on consumer prices
monetary aggregate M1
monetary aggregate M2
monetary aggregate M3

A.4 Manufacturing turnover, production and received orders

production: intermediate goods industry

production: capital goods industry

production: durable and non-durable consumer goods industry

production: mechanical engineering

production: electrical engineering

production: vehicle engineering

export turnover: intermediate goods industry

domestic turnover: intermediate goods industry

export turnover: capital goods industry

domestic turnover: capital goods industry

export turnover: durable and non-durable consumer goods industry

domestic turnover: durable and non-durable consumer goods industry

export turnover: mechanical engineering

domestic turnover: mechanical engineering

export turnover: electrical engineering industry

domestic turnover: electrical engineering industry

export turnover: vehicle engineering industry

domestic turnover: vehicle engineering industry

orders received by the intermediate goods industry from the domestic market

orders received by the intermediate goods industry from abroad

orders received by the capital goods industry from the domestic market

orders received by the capital goods industry from abroad

orders received by the consumer goods industry from the domestic market

orders received by the consumer goods industry from abroad

orders received by the mechanical engineering industry from the domestic market

orders received by the mechanical engineering industry from abroad

orders received by the electrical engineering industry from the domestic market

orders received by the electrical engineering industry from abroad

orders received by the vehicle engineering industry from the domestic market

orders received by the vehicle engineering industry from abroad

industrial production

A.5 Construction

orders received by the construction sector: building construction

orders received by the construction sector: civil engineering

orders received by the construction sector: residential building

orders received by the construction sector: non-residential building construction

man-hours worked in building construction

man-hours worked in civil engineering
man-hours worked in residential building
man-hours worked in industrial building
man-hours worked in public building
turnover: building construction
turnover: civil engineering
turnover: residential building
turnover: industrial building
turnover: public building
production in the construction sector

A.6 Surveys

ifo surveys: business situation: capital goods producers
ifo surveys: business situation: producers durable consumer goods
ifo surveys: business situation: producers non-durable consumer goods
ifo surveys: business situation: retail trade
ifo surveys: business situation: wholesale trade
ifo surveys: business expectations for the next six months: producers capital goods
ifo surveys: business expectations for next six months: producers durable consumer goods
ifo surveys: business expectations for next six months: producers non-durable consumer goods
ifo surveys: business expectations for next six months: retail trade
ifo surveys: business expectations for next six months: wholesale trade
ifo surveys: stocks of finished goods: producers of capital goods
ifo surveys: stocks of finished goods: producers of durable consumer goods
ifo surveys: stocks of finished goods: producers of non-durable consumer goods
GfK consumer surveys: income expectations
GfK consumer surveys: business cycle expectations
GfK consumer surveys: propensity to consume: consumer climate
GfK consumer surveys: price expectations
ZEW financial market survey: business cycle expectations

A.7 Miscellaneous indicators

current account: exports
current account: imports
current account: services import
current account: services export
current account: transfers from abroad
current account: transfers to foreign countries
HWWA raw material price index

HWWA raw material price index without energy
HWWA raw material price index: industrial raw materials
HWWA raw material price index: energy industrial raw materials
new car registrations
new car registrations by private owners
retail sales turnover

B The two-step factor estimator by Forni et al. (2005)

The two-step estimation technique based on dynamic principal components by Forni et al. (2005) proceeds in two steps: Estimation of the dynamic common components and idiosyncratic components as well as their covariances is carried out in a first step, and the static factors are estimated in a second step. Let T_{mb} denote the balanced sample size of monthly indicators obtained from realignment (7) applied to all the N time series $\widetilde{\mathbf{X}}_{t_m}$ for $t_m = 1, \ldots, T_{mb}$. :

1. Covariances of the common and idiosyncratic components: To estimate the q dynamic factors, Forni et al. (2005) propose dynamic principal component analysis in the frequency domain. Let $\widehat{\mathbf{\Gamma}}(k) = T_{mb}^{-1} \sum_{t_m=1}^{T_{mb}} \widetilde{\mathbf{X}}_{t_m} \widetilde{\mathbf{X}}'_{t_m-k}$ be the k-lag estimated autocovariance of the vector of time series. An estimator of spectral density of $\widetilde{\mathbf{X}}_{t_m}$ is then given by $\widehat{\mathbf{\Sigma}}(\theta_h) = \sum_{k=-M}^{M} w_k \widehat{\mathbf{\Gamma}}(k) e^{-ik\theta_h}$ at frequency $\theta_h = 2\pi h/(2H)$ for $h = 0, \ldots, 2H$, and with Bartlett lag weights $w_k = 1 - |k|/(M+1)$. For each frequency, compute the dynamic eigenvalues and eigenvectors of $\widehat{\mathbf{\Sigma}}(\theta_h)$, and denote $\mathbf{\Lambda}(\theta_h)$ as the $(q \times q)$ diagonal matrix with the largest q dynamic eigenvalues on the main diagonal, and the $(N \times q)$ matrix $\widehat{\mathbf{P}}(\theta_h) = (\widehat{\mathbf{P}}_1(\theta_h), \ldots, \widehat{\mathbf{P}}_q(\theta_h))$ of the corresponding eigenvectors, see Forni et al. (2003), p. 1253. The variance of the common components is then given by $\widehat{\mathbf{\Sigma}}_\chi(\theta_h) = \widehat{\mathbf{P}}(\theta_h)\mathbf{\Lambda}(\theta_h)\widehat{\mathbf{P}}^*(\theta_h)$, where a star denotes complex conjugates. The covariance of the idiosyncratic components can be obtained by $\widehat{\mathbf{\Sigma}}_\xi(\theta_h) = \widehat{\mathbf{\Sigma}}(\theta_h) - \widehat{\mathbf{\Sigma}}_\chi(\theta_h)$. Inverse discrete Fourier transform provides time-domain autocovariances of the common components $\widehat{\mathbf{\Gamma}}_\chi(k) = (2H+1)^{-1} \sum_{h=0}^{2H} \widehat{\mathbf{\Sigma}}_\chi(\theta_h) e^{ik\theta_h}$ for $k = -M, \ldots, M$. The autocovariance of the idiosyncratic component $\widehat{\mathbf{\Gamma}}_\xi(k)$ can be obtained accordingly.

2. The factors: The aim is to find the r linear combinations of the time series $\widehat{\mathbf{Z}}'_j \widetilde{\mathbf{X}}_{t_m}$ for $j = 1, \ldots, r$ that maximise the contemporaneous covariance explained by the common factors $\widehat{\mathbf{Z}}'_j \widehat{\mathbf{\Gamma}}_\chi(0) \widehat{\mathbf{Z}}_j$. As a restriction, Forni et al. (2005) impose the normalisation $\widehat{\mathbf{Z}}'_j \widehat{\mathbf{\Gamma}}_\xi(0) \widehat{\mathbf{Z}}_i = 1$ for $i = j$ and 0 for $i \neq j$.[4] This optimisation problem can be reformulated as a generalised eigenvalue problem $\widehat{\mathbf{\Gamma}}_\chi(0)\widehat{\mathbf{Z}}_j = \widehat{\mu}_j \widehat{\mathbf{\Gamma}}_\xi(0)\widehat{\mathbf{Z}}_j$, where $\widehat{\mu}_j$

[4] The off-diagonal elements of the covariance matrix of the idiosyncratic components are forced to be zero in order to improve the forecasting properties of the model, see Forni et al. (2005), p. 836.

denotes the j-th generalised eigenvalue and $\widehat{\mathbf{Z}}_j$ its $(N \times 1)$ corresponding eigenvector. The factors are obtained as

$$\widehat{\mathbf{F}}_{t_m} = \widehat{\mathbf{Z}}' \widetilde{\mathbf{X}}_{t_m}, \tag{10}$$

where $\widehat{\mathbf{Z}}_j = (\widehat{\mathbf{Z}}_1, \ldots, \widehat{\mathbf{Z}}_r)$ denotes the $(N \times r)$ matrix of the eigenvectors corresponding to the r largest eigenvalues.

Note that although the first step to obtain the covariance matrix of the common components is essentially dynamic, the final step of the estimation of the factors is finding a linear combination of contemporaneous variables. For the estimation of the factors, the auxiliary variables to be specified by the user are M, H, q and r. In the empirical comparison, we will concentrate on the specification of the number of factors q and r, as there is some disagreement in the literature concerning their choice, see Bernanke and Boivin (2003), footnote 7, Giannone, Reichlin, and Sala (2005), footnote 8, and Schumacher (2007).

C Subsample results

The evaluation sample is now between $2004Q1$ and $2007Q4$. The recursive simulation design is otherwise the same as in the main text.

Now- and forecast results based on fixed specifications for the factor models and single-indicator MIDAS can be found in table 5. The table shows relative MSEs to the naive benchmark, which is the in-sample mean of GDP. The now- and forecasts are again shown for monthly horizons $h_m = 1, \ldots, 6$. In the table, MSE results are shown for selected MIDAS single-indicator models and factor models. To find the best-performing models in terms of MSE, we chose those, that have a relative MSE smaller than one for $h_m = 1, 2, 3$. To order the models, we use the average of the MSE over $h_m = 1, 2, 3$. The general performance in the second subsample is better than in the while sample, as there are now models that outperform the benchmark. To save space, we only report only the top 20 of models of both classes. An important result is changed ranking of variables in single-indicator MIDAS. Now, the survey consumer sentiment (GfK) is doing best. Also, a few more specifications with interest rates appear in the top 20. Also compared with table 1 in the main text, some predictors do not make it to the top 20 now, such as production of intermediate goods or vacancies. Thus, there is some time variation in the ranking of best-performing models. Regarding the factor models, the ranking has also changed to some extent. However, the first models are still the ones with only one factor.

In table 6, the results based on information criteria and past performance are shown. As in the main text, information criteria do help little to identify specifications that perform well. Specification based on the past performance however now also works relatively well for selecting single-indicator MIDAS models.

Table 7 contains the pooling results for the subsample chosen. The results are again

Table 5: Subsample analysis: Now- and forecast results for single-indicator MIDAS and factor models, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
A. Single-indicator MIDAS							
survey consumer sentiment (GfK)	AR-MIDAS	0.35	0.48	0.67	0.58	0.61	0.85
survey consumer sentiment (GfK)	MIDAS	0.35	0.48	0.67	0.59	0.61	0.85
stocks finished goods, consumer goods prod.	MIDAS	0.58	0.56	0.58	0.87	0.91	1.02
stocks finished goods, consumer goods prod.	AR-MIDAS	0.58	0.58	0.60	0.84	0.92	1.09
survey: bus. exp., consumer goods prod.	MIDAS	0.67	0.64	0.69	0.75	0.83	0.83
survey: bus. cond., retail trade	AR-MIDAS	0.48	0.65	0.90	0.72	0.99	0.87
long-term interest rate (1-2 years mat.)	MIDAS	0.68	0.68	0.69	0.72	0.75	0.84
survey: bus. exp., consumer goods prod.	AR-MIDAS	0.73	0.65	0.70	0.75	0.83	0.83
long-term interest rate (5-6 years mat.)	MIDAS	0.67	0.69	0.75	0.81	0.90	0.95
long-term interest rate (1-2 years mat.)	AR-MIDAS	0.80	0.69	0.69	0.72	0.75	0.82
survey: bus. cond., consumer goods prod.	AR-MIDAS	0.64	0.67	0.88	0.89	0.95	0.96
German bond index REX	MIDAS	0.74	0.70	0.76	0.80	0.88	0.88
long-term interest rate (5-6 years mat.)	AR-MIDAS	0.77	0.69	0.75	0.81	0.90	0.93
survey: bus. cond., consumer goods prod.	MIDAS	0.62	0.70	0.89	0.92	0.96	0.95
export prices	MIDAS	0.69	0.87	0.74	0.79	0.81	0.79
turnover (abroad), intermediate goods prod.	MIDAS	0.78	0.75	0.80	0.86	0.88	0.96
orders (domestic), consumer goods prod.	AR-MIDAS	0.86	0.84	0.64	1.01	0.83	1.05
survey: bus. exp., retail trade	AR-MIDAS	0.84	0.83	0.69	0.85	0.90	0.93
long-term interest rate (9-10 years mat.)	MIDAS	0.74	0.77	0.85	0.96	0.94	1.08
long-term interest rate (9-10 years mat.)	AR-MIDAS	0.83	0.75	0.82	0.95	0.94	1.04
B. Large factor models							
VA-DPCA, $r = 1$, $q = 1$	MIDAS	0.62	0.69	0.79	0.78	0.84	0.92
VA-DPCA, $r = 1$, $q = 1$	AR-MIDAS	0.67	0.68	0.82	0.76	0.80	0.92
VA-DPCA, $r = 2$, $q = 1$	MIDAS	0.64	0.75	0.82	0.91	0.93	0.84
VA-DPCA, $r = 2$, $q = 1$	AR-MIDAS	0.72	0.69	0.85	0.87	0.92	0.81
KFS-PCA, $r = 1$, $q = 1$	MIDAS	0.68	0.79	0.79	0.84	0.84	0.82
KFS-PCA, $r = 2$, $q = 2$	MIDAS	0.61	0.98	0.70	0.83	0.82	0.76
VA-DPCA, $r = 6$, $q = 3$	MIDAS	0.95	0.65	0.71	1.06	1.31	0.82
KFS-PCA, $r = 1$, $q = 1$	AR-MIDAS	0.75	0.80	0.80	0.90	0.83	0.83
VA-DPCA, $r = 5$, $q = 4$	MIDAS	0.83	0.70	0.87	1.03	1.33	0.62
VA-DPCA, $r = 2$, $q = 2$	MIDAS	0.72	0.73	0.97	1.10	0.95	0.91
VA-DPCA, $r = 2$, $q = 2$	AR-MIDAS	0.85	0.69	0.93	1.05	0.91	0.95
VA-DPCA, $r = 3$, $q = 1$	AR-MIDAS	0.88	0.79	0.89	1.35	0.90	0.69
VA-DPCA, $r = 3$, $q = 1$	AR-MIDAS	0.76	0.90	0.91	1.33	1.02	0.80
VA-DPCA, $r = 5$, $q = 2$	MIDAS	0.76	0.92	0.90	1.04	0.97	0.55
VA-DPCA, $r = 6$, $q = 4$	AR-MIDAS	0.94	0.85	0.80	1.20	1.24	1.27
VA-DPCA, $r = 6$, $q = 4$	MIDAS	0.93	0.88	0.79	1.11	1.13	0.89
VA-DPCA, $r = 4$, $q = 2$	AR-MIDAS	0.87	0.77	0.97	1.06	1.08	0.76
VA-DPCA, $r = 6$, $q = 2$	MIDAS	0.90	0.98	0.73	1.20	0.99	0.98
VA-DPCA, $r = 4$, $q = 2$	MIDAS	0.73	0.95	0.97	1.14	1.02	0.89
VA-DPCA, $r = 3$, $q = 3$	AR-MIDAS	0.88	0.87	0.96	1.32	1.09	0.80

Note: See table 1 in the main text.

Table 6: Subsample analysis: Now- and forecast results for single-indicator MIDAS and Factor-MIDAS, information criteria model selection, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
A. Information criteria model selection							
single-indicator MIDAS/AR-MIDAS	BIC	0.96	1.36	1.11	1.10	1.38	1.13
VA-DPCA, MIDAS	Bai, Ng (2002, 2007)	1.02	1.02	0.94	1.44	1.17	1.07
VA-DPCA, AR-MIDAS	Bai, Ng (2002, 2007)	1.19	0.89	0.81	1.43	0.92	0.62
KFS-PCA, MIDAS	Bai, Ng (2002, 2007)	1.28	1.68	0.73	1.06	1.18	1.06
KFS-PCA, AR-MIDAS	Bai, Ng (2002, 2007)	1.40	1.36	0.87	0.94	1.27	1.35
B. Model and variable selection by past MSE performance							
single-indicator MIDAS	MSE	0.85	1.79	0.88	0.93	0.61	0.79
large factor models	MSE	0.83	0.79	0.83	0.96	0.86	0.73

Note: See table 1 in the main text.

Table 7: Subsample analysis: Now- and forecast results for nowcast pooling, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
single-indicator MIDAS	equal-weight mean	0.80	0.83	0.84	0.87	0.89	0.90
single-indicator MIDAS	MSE-weighted mean	0.72	0.80	0.80	0.82	0.86	0.87
single-indicator MIDAS	median	0.87	0.88	0.88	0.90	0.94	0.92
large factor models	equal-weight mean	0.81	0.84	0.81	0.98	0.87	0.73
large factor models	MSE-weighted mean	0.82	0.78	0.75	0.95	0.88	0.76
large factor models	median	0.83	0.79	0.83	0.96	0.86	0.73
all models	equal-weight mean	0.67	0.66	0.75	0.90	0.83	0.73
all models	MSE-weighted mean	0.72	0.73	0.72	0.85	0.81	0.76
all models	median	0.69	0.68	0.76	0.90	0.85	0.75

Note: See table 1 in the main text.

line with those from the main text. There are little differences between weighting schemes, with small advantages of MSE-weighted pooling over the equal-weight mean and the median. Pooling over all factor models and single-indicator MIDAS is doing best overall (lines 7-9). The pooling over all specifications and models also outperforms the model selection techniques as shown in table 6.

D Single- and double-indicator MIDAS

Table 8 contains results with both single- and double-indicator MIDAS. Double-indicator MIDAS includes industrial production and an additional predictor from the set of monthly time series in the MIDAS regression. Overall, we computed now- and forecasts with fixed specifications for 220 models, containing models with and without AR terms and all the remaining monthly indicators apart from industrial production. In lines 4, 12, 16, 20, 23,

Table 8: Now- and forecast results for double- and single-indicator MIDAS, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
survey: bus. exp., wholesale trade	MIDAS	0.72	0.67	0.78	0.80	0.67	0.87
survey: bus. exp., consumer goods prod.	MIDAS	0.78	0.67	0.75	0.89	0.92	0.96
survey: bus. conditions, wholesale trade	MIDAS	0.70	0.79	0.90	0.88	1.12	1.19
IP; survey: bus. exp., wholesale trade	MIDAS	0.68	0.94	0.79	0.68	0.92	1.31
survey: bus. exp., consumer goods prod.	AR-MIDAS	0.93	0.71	0.80	0.91	0.94	0.97
stocks finished goods, consumer goods prod.	MIDAS	0.82	0.80	0.82	1.01	1.01	1.07
survey: bus. exp., retail trade	AR-MIDAS	0.79	0.79	0.87	1.16	1.17	0.91
survey: bus. exp., retail trade	MIDAS	0.79	0.77	0.91	1.23	1.22	0.98
survey bus. exp., wholesale trade	AR-MIDAS	0.97	0.72	0.84	0.82	0.68	1.12
survey consumer sentiment (GfK)	AR-MIDAS	0.74	0.84	0.95	0.93	1.12	1.24
stocks finished goods, consumer goods prod.	AR-MIDAS	0.87	0.84	0.86	1.01	1.00	1.11
IP; survey: bus. exp., consumer goods prod.	MIDAS	0.77	0.92	0.89	0.88	0.95	1.02
survey: bus. cond., wholesale trade	AR-MIDAS	0.81	0.85	0.95	0.92	1.06	1.26
long-term interest rate (1-2 years mat.)	MIDAS	0.86	0.90	0.87	0.89	0.85	0.86
turnover (abroad), intermediate goods prod.	MIDAS	0.89	0.83	0.92	0.83	0.89	1.04
IP plus survey: bus. exp., wholesale trade	AR-MIDAS	0.83	0.96	0.86	0.67	0.95	1.44
production, intermediate goods prod.	MIDAS	0.82	0.91	0.95	0.96	0.99	1.03
survey: bus. exp., non-dur. cons. goods prod.	MIDAS	0.95	0.88	0.85	1.04	0.94	1.52
long-term interest rate (5-6 years mat.)	MIDAS	0.86	0.92	0.94	0.98	1.02	0.93
IP; turnover (abroad), intermediate goods prod.	MIDAS	0.90	0.95	0.91	0.85	0.85	1.10
survey: bus. cond., investm. goods prod.	MIDAS	0.90	0.93	0.93	1.06	1.12	1.11
orders (domestic), intermediate goods prod.	MIDAS	0.88	0.92	0.99	1.23	1.39	1.04
IP; survey: bus. exp., consumer goods prod.	AR-MIDAS	0.90	0.96	0.95	0.94	0.98	1.16
IP; stocks finished goods, consumer goods prod.	MIDAS	0.87	0.98	0.98	0.98	1.09	1.09
short-term employed	AR-MIDAS	0.95	0.95	0.96	0.98	0.92	1.04
turnover (abroad), mechanical engineering	AR-MIDAS	0.94	0.97	0.97	0.95	0.95	1.28

Note: Double-indicator MIDAS models with industrial production and another predictor can be identified in the table by the abbreviation ‘IP; ’. For further details, see table 1 in the main text.

and 24, we find double-indicator models with industrial production, whereas the other models are the single-indicators that have the same ranking as in table 3 in the main text. Thus, only six out of 26 models are double indicator models. Thus, MIDAS regressions with single indicators already do very well in now- and forecasting. Note that we have also tried to include survey business expectations as a second variable, a variable that performed well in the fixed specifications. However, the main results remained unchanged with this modification.

Table 9 contains the pooling results over the broader model set. Pooling over all

Table 9: Subsample analysis: Now- and forecast results for nowcast pooling with double- and single-indicator MIDAS, MSE relative to in-sample mean forecast of GDP

		nowcast current quarter			forecast 1 quarter		
	horizon h_m	1	2	3	4	5	6
double- and single-indicator MIDAS	equal-weight mean	0.87	0.94	0.95	0.93	0.96	1.00
double- and single-indicator MIDAS	MSE-weighted mean	0.83	0.92	0.91	0.88	0.92	0.96
double- and single-indicator MIDAS	median	0.90	0.96	0.97	0.97	1.00	1.03
all models	equal-weight mean	0.77	0.82	0.84	0.94	0.88	0.75
all models	MSE-weighted mean	0.79	0.85	0.82	0.81	0.84	0.67
all models	median	0.81	0.82	0.85	0.93	0.91	0.77

Note: Pooling over all models now contains the group of double-indicator MIDAS, in addition to single-indicator MIDAS and the factor models. For details on the models, see table 3 in the main text.

models again outperforms pooling over single- and double-indicator MIDAS. However, comparing the results with those obtained from using single-indicator MIDAS only from table 3, we can see that there are only small differences. Thus, extending the predictors in the MIDAS regressions does not seem to contribute much, and we can confirm our previous results in this sensitivity check.