DEUTSCHE
BUNDESBANK
EUROSYSTEM

MIDAS versus mixed-frequency VAR: nowcasting GDP in the euro area

Vladimir Kuzin
(DIW)

Massimiliano Marcellino
(EUI, Università Bocconi and CEPR)

Christian Schumacher
(Deutsche Bundesbank)

Discussion Paper
Series 1: Economic Studies
No 07/2009

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0
Telex within Germany 41227, telex from abroad 414431

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

ISBN 978-3–86558–508–0 (Printversion)
ISBN 978-3–86558–509–7 (Internetversion)

Abstract

This paper compares the mixed-data sampling (MIDAS) and mixed-frequency VAR (MF-VAR) approaches to model specification in the presence of mixed-frequency data, e.g., monthly and quarterly series. MIDAS leads to parsimonious models based on exponential lag polynomials for the coefficients, whereas MF-VAR does not restrict the dynamics and therefore can suffer from the curse of dimensionality. But if the restrictions imposed by MIDAS are too stringent, the MF-VAR can perform better. Hence, it is difficult to rank MIDAS and MF-VAR a priori, and their relative ranking is better evaluated empirically. In this paper, we compare their performance in a relevant case for policy making, i.e., nowcasting and forecasting quarterly GDP growth in the euro area, on a monthly basis and using a set of 20 monthly indicators. It turns out that the two approaches are more complementary than substitutes, since MF-VAR tends to perform better for longer horizons, whereas MIDAS for shorter horizons.

Keywords: nowcasting, mixed-frequency data, mixed-frequency VAR, MIDAS

JEL-Classification: E37, C53

Non-technical summary

Decision-making in fiscal or monetary policy is usually based on a large amount of macroeconomic information. Policy makers often face the problem of assessing the current state of the economy with incomplete statistical information, because important economic variables are released with considerable time lags and at low frequencies. For example, as a key indicator of real economic activity, GDP is published at quarterly frequency and with a considerable delay. Due to this limited availability of data, often business cycle indicators such as industrial production or surveys about business expectations might help monitoring the current state of the economy as well as forecasting. These business cycle indicators are published monthly and are available earlier than GDP. Thus, they could contain useful information about current and future GDP.

In the present paper, we discuss two econometric models capable of forecasting quarterly GDP based on monthly indicators, taking into account publication lags. The first approach is so-called MIDAS (mixed-data sampling). It is a single-equation approach, where quarterly GDP is explained by specifically weighted observations of monthly predictors. By taking into account autoregressive terms and lags of the indicators, MIDAS allows for a complicated dynamic relationship between the indicator and GDP. Distributed lags imply a parsimonious specification of the model. The second approach is a mixed-frequency vector-autoregressive model (MF-VAR). It specifies a monthly high-frequency VAR for GDP and the indicators, where monthly values of GDP are interpolated in a model-consistent way. The model is cast in state-space form, and interpolation of missing monthly obervations of GDP is carried out by means of Kalman filtering.

A theoretical comparison shows that both types of models exhibit specific relative advantages and disadvantages. For example, MIDAS exhibits a more parsimonious specification than MF-VAR, whereas the non-linear distributed lag function might be too rigid. MIDAS is a direct forecast approach and typically regarded as more robust to misspecification. However, if the high-frequency VAR model is close to the data-generating process, MF-VAR might be superior to MIDAS. Overall, the theroretical arguments indicate that an evaluation of the models with respect to their usefulness in regular forecasting exercises should be motivated by means of an empirical comparison.

In the present paper, we compare MF-VAR and MIDAS with respect to short-term forecasting GDP in the Euro Area. The dataset includes about twenty monthly business cycle indicators, from which the relevant predictors are chosen. The empirical results show that forecasts with both types of models have information content up to one quarter ahead. After sorting the best-performing models, we find both models among the best. Relative forecast comparisons based on selected indicators show that MIDAS performs well at short horizons, whereas MF-VAR outperforms at longer horizons.

Nicht-technische Zusammenfassung

Wirtschaftspolitische Entscheidungen werden üblicherweise auf Basis umfangreicher makroökonomischer Informationen getroffen. Dabei sind Entscheidungsträger oftmals mit dem Problem konfrontiert, dass die zur Verfügung stehenden statistischen Daten die gegenwärtige Lage einer Volkswirtschaft nur eingeschränkt abbilden können, da wichtige makroökonomische Variablen mit erheblichen Zeitverzögerungen oder nur in großen Zeitabständen veröffentlicht werden. So wird das BIP als eine Schlüsselvariable für die Wirtschaftsaktivität einer Volkswirtschaft lediglich vierteljährlich und mit einer erheblichen Zeitverzögerung publiziert. Aufgrund dieser eingeschränkten Informationslage werden oftmals höherfrequente Konjunkturindikatoren wie die Industrieproduktion oder Umfragedaten zu den Geschäftserwartungen der Unternehmen verwendet, um das BIP zu prognostizieren. Viele Konjunkturindikatoren werden monatlich berichtet, stehen zeitlich früher als das BIP zur Verfügung und könnten daher wertvolle Informationen über die aktuelle und zukünftige BIP-Entwicklung enthalten.

In dem vorliegenden Papier diskutieren wir zwei ökonometrische Modelle, die für Prognosen des vierteljährlichen BIP auf Basis monatlicher Indikatoren unter Berücksichtigung von Publikationsverzögerungen geeignet sind. Das erste Verfahren ist der sogenannte MIDAS-Ansatz (mixed-data sampling). Der Ansatz basiert auf einer Einzelgleichung, in der das vierteljährliche BIP durch speziell gewichtete Beobachtungen von monatlichen Indikatoren erklärt wird. Durch die Berücksichtigung von autoregressiven Termen und Verzögerungen der Indikatoren werden komplizierte dynamische Wechselwirkungen zwischen den Indikatoren und dem BIP zugelassen. Dabei werden die Koeffizienten der verzögerten Indikatoren durch nicht-lineare Verteilungsfunktionen spezifiziert, die eine sehr sparsame Parametrisierung zulassen. Der zweite Ansatz ist ein vektorautoregressives Modell auf Basis gemischt-frequenter Daten (MF-VAR). In diesem Ansatz wird ein monatliches VAR-Modell für das BIP und die Indikatoren spezifiziert, wobei monatliche Werte des BIP durch modellkonsistente Interpolation erzeugt werden. Hierzu wird das Modell in Zustandsraumform geschätzt, so dass die Interpolation der fehlenden monatlichen BIP-Beobachtungen mit dem Kalmanfilter erfolgen kann.

In einem theoretischen Vergleich zeigt sich, dass die beiden Modelltypen spezifische Vor- und Nachteile aufweisen. Beispielsweise ist der MIDAS-Ansatz sparsamer parametrisiert als das VAR-Modell, wenngleich die Wahl der (nicht-linearen) Funktionsform für die Koeffizientenmatrizen zu strikt sein könnte. Zudem wird MIDAS als robuster gegenüber Fehlspezifizierungen angesehen. Wenn das hochfrequente VAR-Modell jedoch dem „wahren“ datengenerierenden Prozess recht nahe kommt, kann dies gegenüber MIDAS vorteilhaft sein. Die theoretischen Vor- und Nachteile deuten an, dass eine Beurteiling beider Verfahren in Hinblick auf ihre Verwendung für angewandte Konjunkturprognosen letztlich anhand einer empirischen Analyse erfolgen sollte.

In einer empirischen Anwendung werden MF-VAR und MIDAS für Kurzfristprognosen des BIP im Euroraum verwendet. Als Datensatz dienen etwa zwanzig monatliche Konjunkturindikatoren, aus denen relevante Prediktoren ausgewählt werden. In den empirischen Ergebnissen zeigt sich, dass Prognosen mit beiden Modellklassen einen Informationsgehalt für das nächste Quartal haben. Werden alle Modelle gemäß ihrer Prognoseleistung sortiert, finden sich beide Modellklassen unter den am besten prognostizierenden Modellen. Relative Prognosevergleiche mit ausgewählten Indikatoren zeigen zudem, dass MIDAS bei kurzen Prognosehorizonten besser abschneidet, während MF-VAR bei längeren Horizonten dominiert.

Contents

MIDAS versus mixed-frequency VAR: Nowcasting GDP in the Euro Area[†]

1 Introduction

The development of econometric models based on mixed frequency data has attracted considerable attention recently. In particular, the mixed-data sampling (MIDAS) approach proposed by Ghysels, Santa-Clara and Valkanov (2004) and Ghysels, Sinko and Valkanov (2007) has proven useful for different forecasting purposes. MIDAS can be regarded as time-series regressions that allow the regressand and regressors to be sampled at different frequencies, where distributed lag polynomials are used to ensure parsimonious specifications. Whereas MIDAS has been initially used for financial applications, e.g. Ghysels, Santa-Clara and Valkanov (2006), it has been employed to forecast macroeconomic time series, in particular quarterly GDP with monthly indicators, in recent applications by Clements and Galvão (2008, 2009), Marcellino and Schumacher (2008a).

In this paper, we compare the MIDAS approach to a mixed-frequency VAR (MF-VAR) model as proposed by Zadrozny (1988), Mittnik and Zadrozny (2005) and Mariano and Murasawa (2007). The MF-VAR is a VAR operating at the highest sampling frequency of the time series to be included in the model. Low-frequency variables are interpolated according to their stock-flow nature implying specific time aggregation schemes. The high-frequency VAR together with the time aggregation restriction can be cast in state-space form and estimated by maximum likelihood. In this framework, the Kalman filter can tackle missing values at the end of the sample, and take into account the mixed-frequency nature of the data.

Compared to single-equation MIDAS, MF-VAR is a system approach that jointly explains indicators and predictant without imposing a-priori restrictions on the dynamics. This can be an advantage when few variables are modelled, their dynamics is limited, and the VAR provides a good approximation to the data generating process

[†]Correspondence: Kuzin: DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany, e-mail: vkuzin@diw.de; Marcellino: European University Institute, Department of Economics, via della Piazzuola 43, 50133 Florence, Italy, e-mail: massimiliano.marcellino@eui.eu; Schumacher: Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany, e-mail: christian.schumacher@bundesbank.de. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank. We are grateful to the seminar participants at University of Bielefeld, in particular Harry Haupt, for helpful comments and discussions. 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. Further codes were kindly provided by Arthur Sinko from www.unc.edu/~sinko/midas.zip.

(DGP). Otherwise, MIDAS can represent a more robust forecasting device. In addition, due to its single equation specification, a direct forecasting approach is preferable for MIDAS, while an iterated scheme is a more natural choice for the MF-VAR since it is cast in state-space form and iterated forecasts are directly provided by the Kalman filter. For a discussion of direct versus iterated forecasting see, e.g., Marcellino, Stock and Watson (2006) and Chevillon and Hendry (2005).

It is difficult to rank the MIDAS and MF-VAR approaches based purely on theoretical considerations since, as mentioned, their relative merits depend on the DGP, see also Marcellino and Schumacher (2008b). Therefore, their performance is better assessed in specific economic applications, and in this paper we focus on nowcasting and forecasting quarterly euro area GDP growth using a set of monthly indicators, a relevant issue also from the economic policy perspective.

In our application, we compare various specifications of MIDAS and MF-VAR models with single indicators, as well as combinations of these models. In addition, we take into account the different availability of monthly indicators that emerges from different statistical publication lags. The nowcast and forecast comparison is based on the relative mean-squared errors (MSE) at different horizons, and the analysis is conducted recursively, in a pseudo real-time way.

Our main finding is that in the case of euro area GDP growth, the two approaches are more complementary than substitutes, since MF-VAR tends to perform better for longer horizons, whereas MIDAS for shorter horizons.

The paper proceeds as follows. Section 2 provides a description of the MIDAS and MF-VAR approaches, as well as a discussion of their relative advantages. Section 3 presents the empirical results on nowcasting and forecasting quarterly euro area GDP growth with a set of monthly indicators. Section 4 summarizes our main findings and concludes.

2 Nowcasting quarterly GDP with ragged-edge data

In this paper we focus on 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 growth y_{t_m} is observed only in months $t_m = 3, 6, 9, \ldots, T_m^y$ with $T_m^y = 3T_q^y$. The aim is to nowcast or 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 quarters. In April, for example, Euro Area 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 e.g. Giannone et al. (2008).

In this Section we assume, for the sake of exposition, that the information set for now- and forecasting includes one stationary monthly indicator x_{t_m} in addition to the available observations of GDP. The time index t_m denotes a monthly sampling frequency of x_{t_m} for $t_m = 1, 2, 3, \ldots, T_m^x$, where T_m^x is the final month for which an observation is available. Usually, T_m^x is larger than $T_m^y = 3T_q^y$, as monthly observations for many relevant macroeconomic indicators are earlier available than GDP observations. The forecast for GDP is denoted as $y_{T_m^y+h_m|T_m^x}$, as we condition the forecast on information available in month T_m^x, which also includes GDP observations up to T_q^y in addition to the indicator observations up to T_m^x with $T_m^x \geq T_m^y = 3T_q^y$.

2.1 The MIDAS approach

To forecast quarterly GDP using monthly indicators, we rely on the mixed-data sampling (MIDAS) approach as proposed by Ghysels and Valkanov (2006), Ghysels et al. (2007), and Clements and Galvão (2008). The MIDAS regression approach is a direct forecasting tool. The dynamics of the indicators and joint dynamics between GDP and the indicators are not explicitly modelled. Rather, MIDAS directly relates future GDP to current and lagged indicators, thus yielding different forecasting models for each forecast horizon, see e.g. Marcellino, Stock and Watson (2006) as well as Chevillon and Hendry (2005) for detailed discussions of this issue in the single-frequency case.

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

$$y_{t_q+h_q} = y_{t_m+h_m} = \beta_0 + \beta_1 b(L_m, \boldsymbol{\theta}) x_{t_m+w}^{(3)} + \varepsilon_{t_m+h_m}, \tag{1}$$

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

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

with the monthly lag operator L_m defined as $L_m x_{t_m} = x_{t_m-1}$. In the MIDAS approach, quarterly GDP $y_{t_q+h_q}$ is directly related to the indicator $x_{t_m+w}^{(3)}$ and its lags, where $x_{t_m}^{(3)}$ is skip sampled from the monthly observations of x_{t_m} in the following way. The superscript three indicates that every third observation starting from the t_m-th one is included in the regressor $x_{t_m}^{(3)}$, thus $x_{t_m}^{(3)} = x_{t_m} \,\forall\, t_m = \ldots, T_m^x - 6, T_m^x - 3, T_m^x$. Lags of the monthly factors are treated accordingly, e.g. the k-th lag $x_{t_m-k}^{(3)} = x_{t_m-k} \,\forall\, t_m = \ldots, T_m^x - k - 6, T_m^x - k - 3, T_m^x - k$. In the time index of $x_{t_m+w}^{(3)}$, w is equal to 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 indicators as we can allow for large K to approximate the impulse response function of GDP to the indicators. 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 $x^{(3)}_{t_m-k}$, yielding coefficients $\widehat{\theta}_1$, $\widehat{\theta}_2$, $\widehat{\beta}_0$ and $\widehat{\beta}_1$. The forecast is given by

$$y_{T^y_m+h_m|T^x_m} = \widehat{\beta}_0 + \widehat{\beta}_1 b(L_m, \widehat{\boldsymbol{\theta}}) x_{T^x_m}. \tag{3}$$

Note that MIDAS is h-dependent, and thus has to be reestimated for multi-step forecasts. The same holds when new statistical information becomes available. For example, each month, new observations for the indicator are released, whereas GDP observations are released only once a quarter. Thus, also w changes from month to month, which also makes re-estimation 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

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

The autoregressive coefficient λ is not estimated unrestrictedly to rule out discontinuities of the impulse response function of $x^{(3)}_{t_m}$ 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’.

2.2 The mixed-frequency VAR

In contrast to the MIDAS approach and in line with a conventional VAR model based on single-frequency data, the MF-VAR model specifies the joint dynamics of monthly GDP, which is obtained from quarterly GDP by time disaggregation, and the monthly indicator. Following the notation of Mariano and Murasawa (2003, 2007), the disaggregation of quarterly GDP growth y_{t_m} into unobserved month-on-month GDP growth $y^*_{t_m}$ is based on the aggregation relation

$$y_{t_m} = \frac{1}{3}y^*_{t_m} + \frac{2}{3}y^*_{t_m-1} + y^*_{t_m-2} + \frac{2}{3}y^*_{t_m-3} + \frac{1}{3}y^*_{t_m-4}, \tag{5}$$

which holds for $t_m = 3, 6, 9, \ldots, T^y_m$, because GDP is observed only every third month of each quarter. The aggregation assumption represents the flow nature of GDP and allows for a linear state-space representation, see Mariano and Murasawa (2003) or Giannone et al. (2008). The latent month-on-month GDP growth $y^*_{t_m}$ and the corresponding monthly indicator x_{t_m} are then assumed to follow a bivariate VAR(p) process

$$\mathbf{\Phi}(L_m) \begin{pmatrix} y^*_{t_m} - \mu^*_y \\ x_{t_m} - \mu_x \end{pmatrix} = \mathbf{u}_{t_m}, \tag{6}$$

with $\mathbf{\Phi}(L_m) = \sum_{i=1}^{p} \mathbf{\Phi}_i L^i_m$ and $\mathbf{u}_{t_m} \sim \mathrm{N}(\mathbf{0}, \mathbf{\Sigma})$.

State-space representation To obtain the state-space representation of the MF-VAR, we define the state vector

$$\mathbf{s}_{t_m} = \begin{pmatrix} \mathbf{z}_{t_m} \\ \vdots \\ \mathbf{z}_{t_m-4} \end{pmatrix}, \quad \mathbf{z}_{t_m} = \begin{pmatrix} y^*_{t_m} - \mu^*_y \\ x_{t_m} - \mu_x \end{pmatrix} \tag{7}$$

consisting of demeaned monthly GDP growth with mean μ^*_y, and the monthly indicator demeaned with μ_x, as well as their lags. Transforming (6) into companion form and combining the latter with the aggregation constraint (5), we obtain the corresponding state-space form as

$$\mathbf{s}_{t_m+1} = \mathbf{A}\mathbf{s}_{t_m} + \mathbf{B}\mathbf{v}_{t_m}, \tag{8}$$

$$\begin{pmatrix} y_{t_m} - \mu_y \\ x_{t_m} - \mu_x \end{pmatrix} = \mathbf{C}\mathbf{s}_{t_m}, \tag{9}$$

where $\mathbf{v}_{t_m} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_2)$, and $\mu_y = 3\mu^*_y$ holds. Our experience shows that the mean parameters μ_y and μ_x are often quite difficult to estimate in the state-space framework. For this reason, we work with demeaned series for estimation. The system matrices are

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}, \quad \mathbf{A}_1 = \begin{bmatrix} \mathbf{\Phi}_1 & \dots & \mathbf{\Phi}_p & \mathbf{0}_{2\times 2(5-p)} \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{I}_8 & \mathbf{0}_{8\times 2} \end{bmatrix}, \tag{10}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{\Sigma}^{1/2} \\ \mathbf{0}_{8\times 2} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{H}_0 & \dots & \mathbf{H}_4 \end{bmatrix}, \tag{11}$$

where matrix $\mathbf{C}$ contains the lag polynomial $\mathbf{H}(L_m) = \sum_{i=0}^{4} \mathbf{H}_i L_m^i$ that is defined as

$$\mathbf{H}(L_m) = \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L_m + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} L_m^2 + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L_m^3 + \begin{bmatrix} 1/3 & 0 \\ 0 & 0 \end{bmatrix} L_m^4, \tag{12}$$

according to the aggregation constraint (5). For notational convenience, we consider only $p \leq 4$ for $\mathbf{A}$ and $\mathbf{B}$, however, the representation for $p > 4$ can be derived in a straightforward manner by modifying the state vector and system matrices accordingly.

Missing observations and estimation The state-space model consisting of (8) and (9) can be estimated with maximum likelihood techniques or the expectation-maximization (EM) algorithm, where we have to take into account missing observations due to publication lags and the low-frequency nature of GDP. We follow Mariano and Murasawa (2003, 2007) and first replace all missing values with zeros, where the missing values are assumed to be realizations of some iid standard normal random variable. Second, the signal equation (9) is also modified accordingly: for the first two months of each quarter, the upper row of $\mathbf{C}$ is set to zero and a standard normal error term is added, for details see Mariano and Murasawa (2003, 2007). Then, the EM algorithm is employed for parameter estimation.

Forecasting After estimation, the forecasting of GDP growth is done by means of the Kalman smoother. The application of the Kalman smoother ensures that all timely observations from the monthly indicator are taken into account. Whereas quarterly GDP is available up to $T_m^y = 3T_q^y$, we have monthly indicator observations up to T_m^x with difference in publication lag of $w = T_m^x - T_m^y$. Although GDP for a particular quarter is not available, the smoother considers the monthly indicator observations of the current quarter. Thus, both the MF-VAR and the MIDAS approach can consider timely within-quarter observations for nowcasting. For months without indicator observations, the Kalman smoother operates equally as the Kalman filter, as no updating step can be carried out. As the smoother is applied iteratively, we obtain iterative multi-step forecasts for the MF-VAR model, according to the definitions from Chevillon and Hendry (2005).

2.3 Discussion of MIDAS and MF-VAR

Both the MF-VAR and the MIDAS approaches can tackle the mixed-frequency nature of the data, and both can exploit timely indicator observations that are also available at higher frequency than GDP. However, in general, there are marked differences between the two methods:

- MIDAS is a single-equation approach whereas MF-VAR is a system approach that explains both GDP and the indicator. As such, misspecification in one equation can affect estimation and forecast accuracy of the other model equations. However, forecasts of the monthly indicators can be of interest by themselves.

- MIDAS has a sparse parameterization, whereas MF-VAR suffers more from the curse of dimensionality. For example, with MIDAS, adding a monthly variable to the predictors requires only 3 more coefficients (θ_1, θ_2, and β) to be estimated in the lag polynomial, whereas a VAR(p) with N variables requires N^2p coefficients of the VAR lag polynomial to be estimated. On the other hand, the MIDAS restrictions on the lag polynomial could be invalid, whereas the coefficients of VAR polynomials are estimated unrestrictedly.

- MIDAS is a direct multi-step forecast device, whereas MF-VAR provides iterative forecasts. Thus, the long-lasting discussion of the relative merits of direct versus iterative forecasting also applies here. Marcellino, Stock and Watson (2006) and Chevillon and Hendry (2005) are recent contributions, see Bhansali (2002) for a survey. The literature shows that there are arguments in favour of both approaches and, generally, the direct approach seems to dominate only in case of substantial misspecification.

- In Ghysels and Valkanov (2006) it is shown how the MIDAS can be regarded as an approximation to a general dynamic linear model, in their case a high-frequency VAR(1), where the low-frequency variable is a stock variable. Thus, in case the true high-frequency DGP behind the data is close to a VAR model, we can expect the MF-VAR to perform better than MIDAS, depending on the dimension and parsimony of the DGP. A more detailed discussion on these issues is provided in Marcellino and Schumacher (2008b).

This discussion suggests that we cannot expect one approach to be clearly superior than the other one for any DGP, and either approach could dominate in a specific empirical application. Therefore, the relative advanatges of MIDAS and MF-VAR should be evaluated empirically on a case-by-case basis, and in the next Section we focus on a policy-relevant case, i.e., nowcasting and forecasting quarterly GDP growth in the euro area, on a monthly basis.

3 Now- and forecasting Euro Area GDP with MIDAS and MF-VAR

The empirical comparison will be carried out in a recursive pseudo real-time context. In subsection 3.1, we describe the design of the exercise, the data used and the specification of the models. In the subsequent sections, we present and discuss the empirical results.

3.1 Design of the nowcast and forecast comparison exercise

Data The dataset contains Euro Area quarterly GDP from $1992Q1$ until $2008Q1$ and about 20 monthly indicators until $2008M06$. In particular, we consider industrial production by sector, survey on consumer sentiment, and business climate, raw material price indices, car registrations, interest rates, and monetary aggregates. 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, so that we cannot discuss the role of revisions on the relative forecasting accuracy here. However, we do not expect any major changes in the results from the use of real-time vintages, since the data revisions are typically small after 2000, see e.g. Marcellino and Musso (2008) for euro area GDP growth. Furthermore, many empirical findings such as Bernanke and Boivin (2003) and Schumacher and Breitung (2008) suggest that data revisions do not affect forecast accuracy considerably. However, we take into account another specific characteristic of multivariate data in real time, namely the different availability of variables due to publication lags. These differences in availablity of data lead to certain patterns of missing values at the end of every recursive sample, and recent papers find that accounting for this rather than using artificially balanced samples has a considerable impact on forecast acuracy, see Giannone et al. (2008), Schumacher and Breitung (2008), for example. In our paper, to consider the availability of the data at the end of each subsample, we follow Giannone et al. (2008), Marcellino and Schumacher (2008a), amongst others, and replicate the availability of data in pseudo real-time from a final vintage of data. When downloading the final data - the download date for the data used here was 11th July 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 July 2008, we observe interest rates until June 2008, thus there is only one missing value at the end of the sample, whereas industrial production is available up to April 2008, implying three 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 impose the same missing values pattern 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.

Nowcast and forecast design To evaluate the performance of the models, we carry out recursive estimation and nowcasting, 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 1999Q4 and 2008Q1. 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 1999Q4, we want to compute a nowcast in December 1999, one in November, and October, whereas the forecasts are computed from September 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 two quarters ahead. Thus, again for the initial evaluation quarter 1999Q4, we have six forecasts computed based on information available in April 1999 up to information available in September 1999. Overall, we have nine projections for each GDP observation of the evaluation period, depending on the monthly information available to make the projection.

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 1999Q4 in December 1999, 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 1999Q4 made in December 1999, we know from our full sample that at each period in time, we have one missing value for interest rates and three missing values of industrial production. These missing values are imposed also for the period December 1999, thus replicating the same pattern of data availability. We do this accordingly in every recursive subsample to determine the pseudo real-time final observation of each time series. To replicate the publication lags of GDP, we exploit the fact that in the Euro Area GDP of the previous quarter is available at the beginning of the third month of the next quarter. Note that we reestimate all forecast models recursively when new information becomes available, so that the estimated coefficients are allowed to change over time. For each evaluation period, we compute nine now- and forecasts depending on the available information. To compare the nowcasts with the realisations of GDP growth, the mean-squared error (MSE) is employed.

Lag length specification For estimating the MF-VAR model, a lag order determination is required. For this purpose, we apply the Bayesian information criterion (BIC) with a maximum lag order of $p = 4$ months. Experimenting with higher lag orders did not affect the main results, as the chosen lag lengths are usually very small with only one or two lags in most of the cases. Concerning the specifications of MIDAS and AR-MIDAS, we use a large variety of initial parameter specifications, and compute the residual sum of squares (RSS) from (1) and (4), respectively. 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 $\theta_1 < 0$. The maximum number of lags chosen for MIDAS is $K = 4$ months. Again, experimenting with higher lag orders did not affect the main results.

3.2 Empirical results

Individual models Below, we present a selection of well-performing models for different now- and forecast horizons h_m. The selection has been carried out with respect to relative MSE, defined as MSE of MIDAS or MF-VAR divided by the MSE of the benchmark forecast. In this application, the benchmark forecast is the in-sample mean of GDP growth recomputed every recursion. This benchmark outperforms a simple AR model of GDP growth, and, thus, was preferred in the present application. In Table 1 below, we show all models that have a relative MSE smaller than one for all $h_m = 1, \ldots, 6$. The ranking of models is chosen according toi their average performance over forecast horizons, defined as the mean of the relative MSE over $h_m = 1, \ldots, 6$ of each model. Models with smallest average MSE can be found in the upper part of the table. All the MIDAS and MF-VAR models clearly outperform the benchmark for the nowcast, but less so for the one-quarter ahead forecast. As most of the relative MSE are larger than one for $h_m = 8, 9$, there is little information content of the models for longer horizons, and the models should be regarded as short-term forecast models only.

Concerning the relative performance of MIDAS and MF-VAR, we cannot identify a clear winner from the results. Among the 24 best models shown in Table 1, there are both MF-VAR and MIDAS models with different indicators, in particular 17 MIDAS and 7 MF-VAR models.

Finally, it is interesting to note that the there is substantial agreement across methods on the best performing indicators, which are an index of industrial raw material prices and two survey variables, namely, the business confidence in industry and the business production expectations. However, among all the indicators in the table 1, we can find representatives of all important groups of predictors. In particular, also hard indicators like industrial production as well as financial indicators can be found in the ranking of best models.

Relative performance The selection above concentrates on the best-performing models only. To investigate the relative performance of MIDAS and MF-VAR further, we now follow Marcellino et al. (2006) and compare the relative performance of MIDAS and MF-VAR over the full set of indicators. For MIDAS, AR-MIDAS as well as MF-VAR, we compute the pairwise relative MSE of each model to the benchmark and average over all models within a class, see Table 2. On average, MIDAS and AR-MIDAS cannot do better than the benchmark for horizons larger than $h_m = 6$. MF-

Table 1: Forecasting performance for quarterly GDP growth of selected individual mixed-frequency models measured by MSE of the corresponding indicator relative to MSE of the benchmark

model	horizon h_m 1	2	3	4	5	6	7	8	9
hwwiind; midas	0.53	0.61	0.66	0.70	0.72	0.78	0.84	0.88	0.97
hwwiind; ar-midas	0.47	0.67	0.71	0.76	0.76	0.83	0.88	0.95	1.03
indconf; mf-var	0.71	0.68	0.67	0.69	0.75	0.73	0.86	0.86	0.91
prodexp; ar-midas	0.52	0.67	0.72	0.79	0.87	0.89	0.94	1.07	1.06
prodexp; mf-var	0.69	0.76	0.78	0.68	0.68	0.90	0.85	0.99	1.09
m1; ar-midas	0.62	0.74	0.75	0.74	0.76	0.87	0.85	0.95	1.01
indconf; ar-midas	0.54	0.76	0.79	0.82	0.89	0.82	0.91	1.07	1.06
prodexp; midas	0.59	0.65	0.74	0.84	0.89	0.95	1.00	1.02	1.16
ord-book; ar-midas	0.55	0.79	0.81	0.83	0.84	0.84	0.94	1.11	1.10
m1; mf-var	0.81	0.89	0.85	0.78	0.75	0.65	0.62	0.63	0.63
assstock; ar-midas	0.57	0.80	0.80	0.83	0.90	0.92	0.92	1.04	1.04
carpass; ar-midas	0.58	0.88	0.82	0.82	0.90	0.89	0.96	1.07	1.12
prcap; ar-midas	0.52	0.83	0.78	0.91	0.94	0.92	1.01	1.10	1.08
prcons; ar-midas	0.57	0.82	0.83	0.82	0.96	0.97	0.96	1.11	1.13
prcs; mf-var	0.98	0.90	0.75	0.77	0.83	0.87	0.93	0.96	0.97
hwwiind; mf-var	0.66	0.83	0.86	0.88	0.89	0.98	0.97	0.98	1.00
indconf; midas	0.71	0.79	0.83	0.91	0.95	0.97	1.01	1.13	1.14
m1; midas	0.88	0.90	0.87	0.86	0.88	0.86	0.81	0.87	0.90
loans; mf-var	0.87	0.85	0.84	0.95	0.89	0.85	0.96	0.97	0.98
prcs; midas	0.76	0.85	0.89	0.97	0.93	0.88	1.00	1.12	1.08
assstock; midas	0.75	0.85	0.85	0.91	0.97	0.97	1.00	1.01	1.09
een; midas	0.92	0.92	0.77	0.89	0.92	0.90	0.94	1.00	0.99
prcs; ar-midas	0.79	0.87	0.89	0.92	0.98	0.92	0.98	1.27	1.15
hwwi; mf-var	0.92	0.98	0.98	0.90	0.95	0.97	1.00	1.00	1.00

Note: We use the recursively estimated in-sample mean as benchmark forecast. Only models that outperform the benchmark for $h_m = 1, \ldots, 6$ are displayed in the table. The ordering of models is chosen according to the mean of relative MSE computed over $h_m = 1, \ldots, 6$. The first two columns in the table include the indicator name and model type (MIDAS, AR-MIDAS or MF-VAR). For the meaning of abbreviations of the particular indicators, see Appendix A. Details on the forecasting exercise are reported in Section 3.1.

Table 2: Average relative MSE performance for forecasting quarterly GDP growth of mixed-frequency model classes against benchmark

model class	horizon h_m 1	2	3	4	5	6	7	8	9
midas	0.82	0.90	0.90	0.99	1.01	0.98	1.01	1.08	1.09
ar-midas	0.62	0.83	0.82	0.87	0.96	0.96	1.02	1.14	1.14
mf-var	0.82	0.93	0.90	0.91	0.93	0.95	0.97	0.98	0.99

Note: The recursively estimated in-sample mean is used as benchmark forecast. The entries in the tables are obtained as follows: First, pairwise relative MSEs, defined as the MSE of a particular model relative to MSE of the benchmark, are calculated. Second, we take means over all models within a model class (MIDAS, AR-MIDAS or MF-VAR).

VAR provides an average relative MSE smaller than one up to horizon 9. This indicates that MF-VAR forecasts have information content for longer horizons than MIDAS, though the gains with respect to the benchmark are small. However, the AR-MIDAS models clearly outperform the MF-VAR approach for short nowcasting horizons, i.e. $h_m = 1, \ldots, 4$. This is due to the more flexible dynamic specification of MIDAS, which can be particularly helpful at short horizons.

Finally, to relate MIDAS and MF-VAR directly, we compute the relative MSE of MIDAS to MSE of MF-VAR. We then average over all these relative MSEs, see Table 3. The ranking is very similar to that emerging from Table 2. For short horizons up

Table 3: Relative performance: (AR-)MIDAS vs. MF-VAR

	horizon h_m								
	1	2	3	4	5	6	7	8	9
midas									
mean	1.02	0.99	1.02	1.10	1.09	1.05	1.04	1.12	1.11
median	0.97	0.99	1.02	1.08	1.07	1.03	1.01	1.06	1.07
ar-midas									
mean	0.76	0.89	0.91	0.96	1.04	1.02	1.05	1.20	1.27
median	0.74	0.88	0.89	0.95	1.03	1.01	1.04	1.14	1.15

Note: The entries in the table are average relative MSEs, where MF-VAR models serve as benchmark for MIDAS and AR-MIDAS. They are computed as follows: First, for each single indicator, the MSE of MIDAS and AR-MIDAS forecasts is respectively divided by the corresponding MSE of the corresponding MF-VAR model. Second, means and medians over all relative MSE (see Appendix A) are computed.

to $h_m = 4$, AR-MIDAS has an average relative MSE smaller than one, and thus tends to outperform MF-VAR. MIDAS without AR component is almost always worse than MF-VAR. For longer horizons MF-VAR clearly outperforms both MIDAS types.

Forecast combinations The availability of many indicators and the possible presence of model misspecification and parameter instability suggest that combining forecast from alternative models could yield sizeable gains, since these are the conditions when the advantages from forecast pooling are maximized, see e.g. the review by Timmermann (2006). Clements and Galvão (2008) consider combinations of MIDAS models. A more detailed evaluation of pooling in the presence of a large, mixed-frequency dataset is undertaken in Kuzin, Marcellino and Schumacher (2009). Here we focus on the small number of variables case. We provide results for the mean, the median, and the weighted mean of the models of a particular class, where combination weights are obtained from the inverse MSE of the previous four-quarter performance of a model.

Below, we provide the relative MSE of the combinations to the benchmark (Table 4), as well as the relative MSE of the combination of MIDAS and MIDAS-AR with respect to the combined MF-VARs (Table 5). To investigate the relative performance of the forecast combinations against the individual models, we compute the percentiles of the forecast combinations with respect to all MSEs of individual models within a corresponding class, see Table 6. The figures in Table 6 represent the percentage of single indicator models that outperform the combined forecast. All combinations do

Table 4: Relative MSE performance for forecasting quarterly GDP growth of model pooling within a given model class against benchmark

	horizon h_m								
midas	1	2	3	4	5	6	7	8	9
mean	0.62	0.74	0.75	0.82	0.89	0.87	0.91	1.00	1.05
weighted mean	0.60	0.70	0.69	0.78	0.79	0.80	0.84	0.91	1.01
median	0.65	0.80	0.81	0.85	0.95	0.91	0.97	1.04	1.04
ar-midas									
mean	0.54	0.75	0.72	0.79	0.87	0.84	0.95	1.08	1.09
weighted mean	0.54	0.74	0.73	0.79	0.82	0.82	0.91	1.00	1.05
median	0.55	0.81	0.79	0.83	0.90	0.89	0.95	1.07	1.05
mf-var									
mean	0.60	0.74	0.78	0.81	0.86	0.89	0.92	0.93	0.95
weighted mean	0.57	0.67	0.72	0.78	0.81	0.82	0.85	0.87	0.89
median	0.64	0.84	0.88	0.87	0.95	0.97	0.97	0.98	0.99

Note: The entries are obtained as follows: First, means, weighted averages based on past MSE performance and medians of all forecasts within a given class of models are computed. Second, the MSE of the combination is computed and finally divided by the MSE of the benchmark, the recursively in-sample sample mean.

Table 5: Relative MSE performance: Pooling of (AR-)MIDAS vs. pooling of MF-VAR

	horizon h_m								
midas	1	2	3	4	5	6	7	8	9
mean	1.03	1.00	0.96	1.01	1.04	0.98	0.99	1.07	1.10
weighted mean	1.04	1.05	0.96	1.00	0.98	0.97	1.00	1.05	1.13
median	1.01	0.96	0.92	0.98	1.01	0.94	1.00	1.06	1.04
ar-midas									
mean	0.90	1.00	0.93	0.96	1.00	0.95	1.02	1.15	1.15
weighted mean	0.94	1.12	1.02	1.01	1.01	1.00	1.07	1.15	1.17
median	0.86	0.96	0.89	0.95	0.95	0.92	0.98	1.09	1.06

Note: MF-VAR models serve as benchmark for MIDAS and AR-MIDAS. For further comments, see Tables 3 and 4.

well relative to the benchmark. Comparing Tables 2 and 4, we conclude that forecast combination is a useful method both in case of MIDAS and MF-VAR models, since the performance of forecast combinations relative to our benchmark is always better then the mean of all relative MSEs within a given class over all indicators. Again, AR-MIDAS seems to outperform MF-VAR at short forecast horizons, but its advantage

Table 6: Quantiles of MSEs of pooled (AR-)MIDAS and MF-VAR forecasts

	Horizon h_m								
midas	1	2	3	4	5	6	7	8	9
mean	0.09	0.12	0.16	0.06	0.11	0.18	0.10	0.14	0.42
weighted mean	0.07	0.09	0.11	0.05	0.04	0.08	0.06	0.08	0.28
median	0.11	0.21	0.22	0.09	0.29	0.33	0.24	0.47	0.38
ar-midas									
mean	0.17	0.13	0.11	0.10	0.15	0.15	0.26	0.42	0.48
weighted mean	0.15	0.12	0.12	0.10	0.09	0.07	0.10	0.12	0.32
median	0.21	0.38	0.28	0.27	0.28	0.28	0.27	0.24	0.34
mf-var									
mean	0.04	0.10	0.19	0.24	0.22	0.27	0.19	0.17	0.16
weighted mean	0.03	0.05	0.10	0.20	0.14	0.10	0.07	0.11	0.07
median	0.06	0.18	0.43	0.28	0.41	0.36	0.29	0.29	0.32

Note: We implement the pooling exercise as in Table 4 and then compute the quantiles of MSEs of pooled forecasts in the empirical distribution of all MSEs of individual indicators within a given class of models.

seems not so pronounced as in Table 2, where only individual models were compared. Moreover, Table 4 shows that pooling of MF-VARs performs very well at long forecast horizons ($h_m = 8, 9$), especially in case of pooling with weighted means, in contrast to only a small advantage resulting from Table 2.

The percentiles of the forecast combinations in Table 6 indicate that pooling is a useful alternative to individual models, since a lot of figures in Table 6 are clearly below 10%. However, the forecast combinations cannot outperform all of the individual models. For example, in the case of pooling with weighted means for AR-MIDAS at $h = 1$, there are 15% individual models within the AR-MIDAS class with smaller MSE than the combination. But it should be considered that with a large set of indicators, it is natural to find that some of them performing particularly well. In addition, the analysis of Banerjee and Marcellino (2005) clearly indicates that the best leading indicators for euro area GDP growth change over time, and the pooled forecast can protect from this instability.

4 Conclusions

This paper considers MIDAS and MF-VAR as alternative forecasting methods suitable for now- and forecasting with mixed-frequency data that is also subject to different publication lags.

Theoretical arguments indicate that we cannot expect one approach to be clearly superior than the other. For example, MIDAS is a direct multi-step forecast approach, whereas MF-VAR provides iterative forecasts. MIDAS is more parsimonious than MF-

VAR, but depends on certain distributed lag assumptions that might be too rigid. Thus, the relative performance of the two approaches will depend on the underlying unknown data generating process, and either MIDAS or the MF-VAR could dominate in a specific empirical application. Hence, we compare the alternative forecasting approaches empirically. In particular, we carry out a recursive comparsion exercise in terms of now- and forecasting quarterly Euro Area GDP with a set of about twenty monthly indicators.

The main results are the following.

1. If we look at the best-performing models, we find representatives of both MIDAS and MF-VAR classes of models, with different indicators. Thus, there seems to be no clear winner in terms of forecasting performance.

2. If we compare the models pairwise with the same indicator and compute the average MSE over the whole set of models, we find that MF-VAR outperforms MIDAS and AR-MIDAS at long forecast horizons, whereas AR-MIDAS can do better at short horizons up to three months.

3. When the single MIDAS and MF-VAR forecasts are combined, there are advantages with respect to most single indicator models. In addition, pooled MF-VAR forecasts are better at longer horizons, and pooled MIDAS forecasts at shorter horizons.

Overall, the MF-VAR seems to be a reasonable competitor to MIDAS in macroeconomic datasets such as the one chosen here. More generally, it can be useful to consider both classes of models for forecasting specific variables of interest, and pooling can provide additional advantages.

References

[1] Banerjee, A., Marcellino, M. and Masten, I. (2005), Leading indicators for Euro area inflation and GDP growth, Oxford Bulletin of Economics and Statistics, 67, 785-813.

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

[3] Bhansali, R. (2002), Multi-step forecasting, in: Clements, M., Hendry, D. (eds.), A Companion to Economic Forecasting, 206–221.

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

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

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

[7] Ghysels, E., Santa-Clara, P., Valkanov, R. (2004), The MIDAS touch: Mixed Data Sampling regression models, mimeo.

[8] Ghysels, E., Santa-Clara, P., Valkanov, R. (2006), Predicting volatility: Getting the most out of return data sampled at different frequencies, Journal of Econometrics 131, 59–95.

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

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

[11] 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.

[12] Kuzin, V., Marcellino, M., Schumacher, C. (2009), Pooling versus model selection for nowcasting with many predictors: An application to German GDP, Bundesbank Discussion Paper, forthcoming.

[13] Marcellino, M., Musso, A. (2008), Real time estimates of the euro area output gap: Reliability and forecasting performance, mimeo.

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

[15] Marcellino, M., Schumacher, C. (2008b), MIDAS and Dynamic Linear Models, mimeo.

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

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

[18] Mariano, R., Murasawa, Y. (2007), Constructing a Coincident Index of Business Cycles Without Assuming a One-Factor Model, Discussion Paper 2004-6, College of Economics, Osaka Prefecture University.

[19] Mittnik, S., Zadrozny, P. (2005), Forecasting German GDP at Monthly Frequency Using Monthly IFO Business Conditions Data, in: Sturm, J.-E., Wollmershäuser, T. (eds.), Ifo Survey Data in Business Cycle and Monetary Policy Analysis, Springer-Verlag, 19-48.

[20] Schumacher, C., Breitung, J. (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.

[21] Timmermann, A. (2006), Forecast Combinations, in: Elliot, G., Granger, C., Timmermann, A. (eds.), Handbook of Economic Forecasting, Vol 1, 135-196.

[22] Zadrozny, P. (1988), Gaussian-Likelihood of countinuous-time ARMAX models when dta are strocks and flows at different frequencies, Econometric Theory 4, 108-124.

A Euro Area dataset

This appendix describes the time series for the Euro Area economy used in the forecasting exercise. The whole data set for Euro Area contains 23 monthly time series over the sample period from $1992M1$ until $2008M6$. The time series cover broadly the following groups of data: industry statistics, surveys, financial data (interest rates, exchange rates, money stocks), and miscellaneous indicators, such as raw material price indices and car registrations. A complete list of variables is provided below, together with abbreviations used in the description of results in the main text.

The source of the time series is the databases of the Bundesbank and the ECB. Original sources are the European Commission, the ECB, and the HWWI. Natural logarithms were taken for all time series except interest rates and the surveys. Stationarity was obtained by appropriately differencing the time series. All of the time series taken from the above sources are already seasonally adjusted, where this was necessary.

A.1 Industrial production

prind - roduction: total
prcap - production: capital goods industry
print - production: intermediate goods industry

prcons - production: consumer goods industry
prcs - production: construction sector

A.2 Surveys

indconf - business confidence industry
prodexp - business production expectations
ordbook - business order books
assstock - assessment of stocks of finished goods
consconf - consumer confidence

A.3 Interest rates, exchange rates, money stocks

is3m - oney market rate, 3 months EURIBOR
il10 - yields on 10 year government bonds (GDP weights)
zdiff103 - yield spread: bond yields with 10 years minus 3 months EURIBOR
m1 - monetary aggregate M1
m3 - monetary aggregate M3
loans - loans
een - nominal effective exchange rate of the euro against the currencies of. the EER-22 group
eer - real effective exchange rate of the euro against the currencies of. the EER-22 group (on basis of consumer price index)

A.4 Raw material prices, car registrations

hwwi - HWWI raw material price index
hwwiind - HWWI raw material price index: industrial raw materials
hwwienerg - HWWI raw material price index: energy industrial raw materials
carcomm - car registrations: new commercial
carpass - car registrations: new passenger cars

The following Discussion Papers have been published since 2008:

Series 1: Economic Studies

01	2008	Can capacity constraints explain asymmetries of the business cycle?	Malte Knüppel
02	2008	Communication, decision-making and the optimal degree of transparency of monetary policy committees	Anke Weber
03	2008	The impact of thin-capitalization rules on multinationals' financing and investment decisions	Buettner, Overesch Schreiber, Wamser
04	2008	Comparing the DSGE model with the factor model: an out-of-sample forecasting experiment	Mu-Chun Wang
05	2008	Financial markets and the current account – emerging Europe versus emerging Asia	Sabine Herrmann Adalbert Winkler
06	2008	The German sub-national government bond market: evolution, yields and liquidity	Alexander Schulz Guntram B. Wolff
07	2008	Integration of financial markets and national price levels: the role of exchange rate volatility	Mathias Hoffmann Peter Tillmann
08	2008	Business cycle evidence on firm entry	Vivien Lewis
09	2008	Panel estimation of state dependent adjustment when the target is unobserved	Ulf von Kalckreuth
10	2008	Nonlinear oil price dynamics – a tale of heterogeneous speculators?	Stefan Reitz Ulf Slopek
11	2008	Financing constraints, firm level adjustment of capital and aggregate implications	Ulf von Kalckreuth

12	2008	Sovereign bond market integration: the euro, trading platforms and globalization	Alexander Schulz Guntram B. Wolff
13	2008	Great moderation at the firm level? Unconditional versus conditional output volatility	Claudia M. Buch Jörg Döpke Kerstin Stahn
14	2008	How informative are macroeconomic risk forecasts? An examination of the Bank of England's inflation forecasts	Malte Knüppel Guido Schultefrankenfeld
15	2008	Foreign (in)direct investment and corporate taxation	Georg Wamser
16	2008	The global dimension of inflation – evidence from factor-augmented Phillips curves	Sandra Eickmeier Katharina Moll
17	2008	Global business cycles: convergence or decoupling?	M. Ayhan Kose Christopher Otrok, Ewar Prasad
18	2008	Restrictive immigration policy in Germany: pains and gains foregone?	Gabriel Felbermayr Wido Geis Wilhelm Kohler
19	2008	International portfolios, capital accumulation and foreign assets dynamics	Nicolas Coeurdacier Robert Kollmann Philippe Martin
20	2008	Financial globalization and monetary policy	Michael B. Devereux Alan Sutherland
21	2008	Banking globalization, monetary transmission and the lending channel	Nicola Cetorelli Linda S. Goldberg
22	2008	Financial exchange rates and international currency exposures	Philip R. Lane Jay C. Shambaugh

23	2008	Financial integration, specialization and systemic risk	F. Fecht, H. P. Grüner P. Hartmann
24	2008	Sectoral differences in wage freezes and wage cuts: evidence from a new firm survey	Daniel Radowski Holger Bonin
25	2008	Liquidity and the dynamic pattern of price adjustment: a global view	Ansgar Belke Walter Orth, Ralph Setzer
26	2008	Employment protection and temporary work agencies	Florian Baumann Mario Mechtel, Nikolai Stähler
27	2008	International financial markets' influence on the welfare performance of alternative exchange rate regimes	Mathias Hoffmann
28	2008	Does regional redistribution spur growth?	M. Koetter, M. Wedow
29	2008	International financial competitiveness and incentives to foreign direct investment	Axel Jochem
30	2008	The price of liquidity: bank characteristics and market conditions	Falko Fecht Kjell G. Nyborg, Jörg Rocholl
01	2009	Spillover effects of minimum wages in a two-sector search model	Christoph Moser Nikolai Stähler
02	2009	Who is afraid of political risk? Multinational firms and their choice of capital structure	Iris Kesternich Monika Schnitzer
03	2009	Pooling versus model selection for nowcasting with many predictors: an application to German GDP	Vladimir Kuzin Massimiliano Marcellino Christian Schumacher

Series 2: Banking and Financial Studies

01	2008	Analyzing the interest rate risk of banks using time series of accounting-based data: evidence from Germany	O. Entrop, C. Memmel M. Wilkens, A. Zeisler
02	2008	Bank mergers and the dynamics of deposit interest rates	Ben R. Craig Valeriya Dinger
03	2008	Monetary policy and bank distress: an integrated micro-macro approach	F. de Graeve T. Kick, M. Koetter
04	2008	Estimating asset correlations from stock prices or default rates – which method is superior?	K. Düllmann J. Küll, M. Kunisch
05	2008	Rollover risk in commercial paper markets and firms' debt maturity choice	Felix Thierfelder
06	2008	The success of bank mergers revisited – an assessment based on a matching strategy	Andreas Behr Frank Heid
07	2008	Which interest rate scenario is the worst one for a bank? Evidence from a tracking bank approach for German savings and cooperative banks	Christoph Memmel
08	2008	Market conditions, default risk and credit spreads	Dragon Yongjun Tang Hong Yan
09	2008	The pricing of correlated default risk: evidence from the credit derivatives market	Nikola Tarashev Haibin Zhu
10	2008	Determinants of European banks' engagement in loan securitization	Christina E. Bannier Dennis N. Hänsel
11	2008	Interaction of market and credit risk: an analysis of inter-risk correlation and risk aggregation	Klaus Böcker Martin Hillebrand

12	2008	A value at risk analysis of credit default swaps	B. Raunig, M. Scheicher
13	2008	Systemic bank risk in Brazil: an assessment of correlated market, credit, sovereign and inter-bank risk in an environment with stochastic volatilities and correlations	Theodore M. Barnhill, Jr. Marcos Rietti Souto
14	2008	Regulatory capital for market and credit risk interaction: is current regulation always conservative?	T. Breuer, M. Jandačka K. Rheinberger, M. Summer
15	2008	The implications of latent technology regimes for competition and efficiency in banking	Michael Koetter Tigran Poghosyan
16	2008	The impact of downward rating momentum on credit portfolio risk	André Güttler Peter Raupach
17	2008	Stress testing of real credit portfolios	F. Mager, C. Schmieder
18	2008	Real estate markets and bank distress	M. Koetter, T. Poghosyan
19	2008	Stochastic frontier analysis by means of maximum likelihood and the method of moments	Andreas Behr Sebastian Tente
20	2008	Sturm und Drang in money market funds: when money market funds cease to be narrow	Stehpan Jank Michael Wedow
01	2009	Dominating estimators for the global minimum variance portfolio	Gabriel Frahm Christoph Memmel
02	2009	Stress testing German banks in a downturn in the automobile industry	Klaus Düllmann Martin Erdelmeier
03	2009	The effects of privatization and consolidation on bank productivity: comparative evidence from Italy and Germany	E. Fiorentino A. De Vincenzo, F. Heid A. Karmann, M. Koetter

04	2009	Shocks at large banks and banking sector distress: the Banking Granular Residual	Sven Blank, Claudia M. Buch Katja Neugebauer
05	2009	Why do savings banks transform sight deposits into illiquid assets less intensively than the regulation allows?	Dorothee Holl Andrea Schertler

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Among others under certain conditions visiting researchers have access to a wide range of data in the Bundesbank. They include micro data on firms and banks not available in the public. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a PhD and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

60431 Frankfurt
GERMANY