DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 26/2014

MIDAS and bridge equations

Christian Schumacher

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

Tel +49 69 9566-0

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 978–3–95729–066–3 (Printversion)
ISBN 978–3–95729–067–0 (Internetversion)

Non-technical summary

Research Question

In many policy institutions, nowcasts of quarterly GDP growth are regularly used to inform decision makers about the current state of the economy. A workhorse model for nowcasting is the bridge equation, which explains GDP growth by time-aggregated business cycle indicators. In the recent academic literature, another single-equation approach for nowcasting called mixed-data sampling, or MIDAS in brief, has received a lot of attention. So far, MIDAS and bridge equation approaches have been mostly discussed separately. Against this background, the paper addresses the following questions: What are the differences between MIDAS and bridge equations? Do they matter for nowcasting GDP growth in practice?

Contribution

The paper analytically discusses three features of the approaches: 1) how they tackle multi-step ahead nowcasting (direct versus iterative), 2) how the weights of high-frequency observations in the nowcasts differ (estimated versus time aggregation weights), and 3) how current-quarter observations of the business cycle indicator are considered in the equation. Given these results, the paper derives intermediate models between MIDAS and bridge equations that help to isolate the differences between them. The paper also compares the different approaches in an empirical nowcast exercise for Euro area GDP growth.

Results

The differences between the models are: 1) MIDAS is a direct multi-step nowcasting tool, whereas bridge equations are based on iterated forecasts; 2) MIDAS equations employ empirical weighting of high-frequency predictor observations with estimated functional lag polynomials, whereas the weights of indicator observations in bridge equations are partly fixed stemming from time aggregation. 3) MIDAS equations can consider current-quarter of high-frequency indicators, whereas bridge equations typically do not. In the empirical exercise for nowcasting Euro area growth, there is not a single dominating approach. MIDAS or bridge equations outperform each other depending on the indicators chosen and the evaluation sample. On the other hand, pooling of many single nowcasts yields stable results over time.

Nichttechnische Zusammenfassung

Fragestellung

In vielen Organisationen werden Kurzfristprognosen des vierteljährlichen BIP dazu verwendet, Entscheidungsträger über die aktuelle Lage der Volkswirtschaft zu informieren. Eine Standardmethode zu diesem Zweck ist der Brückengleichungsansatz, in dem das BIP durch zeitaggregierte Konjunkturindikatoren erklärt wird. In der jüngeren wissenschaftlichen Literatur wird der Mixed-Data Sampling Ansatz, oder kurz MIDAS, intensiv diskutiert. Bisher wurden beide Verfahren eher unabhänging voneinander betrachtet. Daher geht das Papier den folgenden Fragen nach: Was sind die Unterschiede zwischen dem MIDAS-Ansatz und Brückengleichungen? Welche Rolle spielen diese Unterschiede in der laufenden Anwendung?

Beitrag

Das Papier vergleicht analytisch, wie in den Ansätzen 1) Mehrschrittprognosen erstellt werden (direkt oder iterativ), wie 2) die Gewichte der hochfrequenten Indikatorwerte berechnet werden (geschätzt oder mit festen Zeitaggregationsgewichten), und 3) ob vorlaufende Beobachtungen aus dem laufenden Quartal in der Prognosegleichung berücksichtigt werden. Auf Grundlage dieser Ergebnisse werden alternative Modellvarianten abgeleitet um die Rolle der Unterschiede herauszuarbeiten. Die verschiedenen Ansätze werden ferner in einer empirischen Anwendung für die Kurzfristprognose des BIP im Euroraum miteinander verglichen.

Ergebnisse

Es zeigen sich folgende Unterschiede: 1) MIDAS impliziert direkte Mehrschrittprognosen, während Prognosen mit Brückengleichungen iterativ erstellt werden. 2) Im MIDAS-Ansatz werden die hochfrequenten Indikatorbeobachtungen mit geschätzten Lag-Polynomen gewichtet, während die Gewichtung in Brückengleichungen zum Teil auf festen Gewichten aus der Zeitaggregation der Indikatoren beruht. 3) MIDAS-Gleichungen können frühzeitig im Quartal verfügbare Indikatorbeobachtungen berücksichtigen, während dies bei Brückengleichungen üblicherweise nicht der Fall ist. In der empirischen Anwendung zur Kurzfristprognose des BIP im Euroraum zeigt sich, dass kein Ansatz eindeutig überlegen ist. Je nach verwendetem Konjunkturindikator und Evaluierungszeitraum zeigen im Prognosevergleich mal MIDAS, mal die Brückengleichungen genauere Prognosen. Prognosekombinationen über mehrere Modelle zeigen hingegen stabile Ergebnisse im Zeitablauf.

BUNDESBANK DISCUSSION PAPER NO 26/2014

MIDAS and bridge equations*

Christian Schumacher
Deutsche Bundesbank

Abstract

This paper compares two single-equation approaches from the recent nowcast literature: Mixed-data sampling (MIDAS) regressions and bridge equations. Both approach are used to nowcast a low-frequency variable such as quarterly GDP growth by higher-frequency business cycle indicators. Three differences between the approaches are discussed: 1) MIDAS is a direct multi-step nowcasting tool, whereas bridge equations are based on iterated forecasts; 2) MIDAS equations employ empirical weighting of high-frequency predictor observations with functional lag polynomials, whereas the weights of indicator observations in bridge equations are partly fixed stemming from time aggregation. 3) MIDAS equations can consider current-quarter leads of high-frequency indicators in the regression, whereas bridge equations typically do not. However, the conditioning set for nowcasting includes the most recent indicator observations in both approaches. To discuss the differences between the approaches in isolation, intermediate specifications between MIDAS and bridge equations are provided. The alternative models are compared in an empirical application to nowcasting GDP growth in the Euro area given a large set of business cycle indicators.

Keywords: Mixed-data sampling (MIDAS), bridge equations, GDP nowcasting

JEL classification: C51, C53, E37.

---

*Contact address: Deutsche Bundesbank, Economic Research Centre, Wilhelm-Epstein-Str. 14, 60431 Frankfurt, Germany. E-Mail: christian.schumacher@bundesbank.de. I thank Mark Kerßenfischer for excellent research assistance and Katja Sonderhof for kindly providing the data. Helpful comments were received by Knut Are Aastveit, Claudia Foroni, Klemens Hauzenberger, Nicolas Pinkwart, Oke Röhe, and seminar participants at the Deutsche Bundesbank.

1 Introduction

In policy institutions such as Central Banks, nowcasting GDP growth is an important task to inform decision makers about the current state of the economy. Nowcast models typically consider specific data irregularities: Whereas GDP is sampled at quarterly frequency and with a considerable delay only, many business cycle indicators are available at higher frequency and more timely, for example, monthly industrial production or high-frequency financial data. Policy analysts want to exploit this data for nowcasting in the most efficient way without a loss of information. Thus, methods for nowcasting should be able to tackle these data irregularities. This paper compares two single-equation approaches for nowcasting: 1) Mixed-data sampling (MIDAS) regressions and 2) bridge equations.

In MIDAS regressions, the observations of the low-frequency variable are directly related to lagged high-frequency observations of the predictors without time aggregation. If the differences in sampling frequencies are huge, functional lag polynomials are employed in order to keep the number of parameters to be estimated small. In this case, non-linear least squares (NLS) is used for parameter estimation as outlined in Ghysels, Sinko, and Valkanov (2007). Whereas MIDAS has been initially used for financial applications, for example in Ghysels, Santa-Clara, and Valkanov (2006), it has been recently employed as a macroeconomic forecast tool for quarterly GDP in many applications, starting with Clements and Galvão (2008), Clements and Galvão (2009). Recent contributions are Drechsel and Scheufele (2012a), Andreou, Ghysels, and Kourtellos (2013), Kuzin, Marcellino, and Schumacher (2011), Ferrara, Marsilli, and Ortega (2014), and Foroni, Marcellino, and Schumacher (2014), Duarte (2014), amongst others. Recent surveys include Armesto, Engemann, and Owyang (2010) and Andreou, Ghysels, and Kourtellos (2011).

Bridge equations are also dynamic, but explain the low-frequency variable by low-frequency lags of a predictor variable. For GDP nowcasting, quarterly values of the predictor are used on the right-hand side and are typically obtained from time aggregation of the high-frequency observations of the predictor if available. The bridge equations can be estimated by ordinary least squares (OLS). To make nowcasts, the predictors are themselves predicted using an additional high-frequency model, for example an autoregressive (AR) model. The high-frequency forecasts from this model are aggregated over time to the quarterly frequency and plugged into the bridge equation. Due to the simple estimation method and their transparency, bridge equations are widely used in policy organizations, in particular, central banks. Applications in the literature include Ingenito and Trehan (1996), Baffigi, Golinelli, and Parigi (2004), Golinelli and Parigi (2007), Diron (2008), Hahn and Skudelny (2008), Rünstler, Barhoumi, Benk, Cristadoro, Den Reijer, Jakaitiene, Jelonek, Rua, Ruth, and Van Nieuwenhuyze (2009), Bulligan, Golinelli, and Parigi (2010), Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011), Bulligan, Marcellino, and Venditti (2014), Camacho, Perez-Quiros, and Poncela (2014), Foroni and Marcellino (2013), and Foroni and Marcellino (2014), amongst others. Applications of bridge equations for nowcasting in Central Banks are documented in ECB (2008), Bundesbank (2013), and Bell, Co, Stone, and Wallis (2014) from the Bank of England.

In this paper, the relationship between MIDAS and bridge equations as nowcast tools is investigated in detail. In the literature, a few comparisons of the two approaches can be found, for example, Foroni and Marcellino (2013). This paper expands on this literature by providing analytical results to explain the differences between MIDAS and bridge equations. This is possible, because MIDAS and bridge equations do both belong to the class of distributed-lag models extended to mixed-frequency data. Three conceptual differences between the two model classes are established: 1) In the applications cited above, MIDAS is a direct multi-step forecasting tool, whereas bridge equations are mostly based on iterated multi-step forecasts from an additional high-frequency model, see Bhansali (2002) for a discussion of direct versus iterative forecasting. 2) MIDAS employs empirical weighting of high-frequency predictor observations often based on functional lag polynomials, whereas bridge equations are partly based on fixed weights stemming from statistical time aggregation rules. The different weighting schemes also imply different estimation methods, namely, OLS for bridge equations and NLS for MIDAS equations due to non-linear functional lag polynomials. 3) Finally, MIDAS can consider current-quarter observations of the high-frequency indicator in the mixed-frequency equation, whereas the bridge equation typically contains only contemporaneous or lagged observations of the indicator.

To assess the role of each of these differences, an intermediate model between MIDAS and bridge equations, called iterative MIDAS (MIDAS-IT), is derived. This approach differs from the bridge equation only by a different weighting scheme of the high-frequency observations on the right-hand side, and from standard MIDAS by an iterative solution of the model for nowcasting. Further model variants arise by different assumptions regarding leading terms of the indicators. Highlighting the differences between the approaches might help a practitioner to make modelling decisions in a class of regression-based models for nowcasting with mixed-frequency data that have been discussed mostly in isolation in the recent literature.

In an empirical exercise for nowcasting Euro area GDP, the alternative approaches are compared on recent data including the Great Recession. The predictor set comprises a large number of monthly indicators. MIDAS and bridge equations with single indicators are evaluated with respect to their out-of-sample nowcast accuracy. It turns out that the performance of the approaches varies over time, in particular during and after the recent financial crisis. On the other hand, pooling nowcasts from equations with different indicators provides stable results. The pooling results are more robust than single models, but the relative performance of MIDAS and bridge equations shows no clear winner. In the periods investigated here, however, the pooled mixed-frequency models are clearly better than naive benchmarks.

The paper proceeds as follows: Section 2 provides the analytical comparison of MIDAS and bridge equations, and discusses alternative models that link the two core approaches. In Section 3, the results of the empirical nowcast exercise are discussed. Section 3 also contains sensitivity checks. Section 4 concludes.

2 MIDAS and bridge equations for nowcasting

The focus in this paper is on quarterly GDP growth, which is denoted as y_t, where t is the quarterly time index $t = 1, 2, \ldots, T_y$ with T_y as the final quarter for which GDP is available. The aim is to nowcast or forecast GDP for period T_y+h, yielding a value for y_{T_y+h} with horizon $h = 1, \ldots, H$ quarters.

In this context, nowcasting means that in a particular calendar month, GDP for the current quarter is not observed. 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 $h = 2$. Typically, the GDP figure for the first quarter is published in mid-May. Thus, if a decision maker requests an estimate of current, namely second, quarter GDP in April, the horizon has to be set sufficiently large in order to provide the appropriate figures. Further information and details on nowcasting procedures can be found in the survey by Banbura, Giannone, and Reichlin (2011).

In this paper it is assumed for simplicity, 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 for monthly observations is defined as a fraction of the low-frequency quarter according to $t = 1 - 2/3, 1 - 1/3, 1, 2 - 2/3, \ldots, T_x - 1/3, T_x$ as in Clements and Galvão (2008). Usually, $T_x \geq T_y$ holds, as monthly observations for many relevant macroeconomic indicators are available earlier than GDP observations for the current quarter, where T_x is the final month for which the indicator is available. The now- or forecast for GDP is denoted as the conditional expectation $y_{T_y+h|T_x}$, as the nowcast is conditioned on information available in month T_x, which also includes GDP observations up to T_y.

2.1 MIDAS regressions

The mixed-data sampling (MIDAS) approach as proposed by Ghysels et al. (2007) (2007) and Clements and Galvão (2008) is a direct nowcast tool. The dynamics of the indicator is not explicitly modelled. Instead, MIDAS directly relates future GDP to current and lagged indicator observations, thus yielding different forecasting models for each horizon, see Marcellino, Stock, and Watson (2006) as well as Chevillon and Hendry (2005) for detailed discussions of this issue in the single-frequency case.

The MIDAS equation for GDP growth y_{t+h} in period $t+h$ with forecast horizon h quarters is

$$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B(L^{1/3};\theta)x^m_{t+w} + \varepsilon_{t+h}, \tag{1}$$

where $w = T_x - T_y$ denotes the lead of the high-frequency indicator as in Andreou et al. (2013). For nowcasting, specifying the right-hand side in terms of period-$t+w$ observations helps to condition the nowcast on the current-quarter indicator information, which is early available in contrast to GDP (Clements and Galvão (2008)). The MIDAS equation contains an autoregressive term λy_t. The term $B(L^{1/3};\theta)$ is a lag polynomial

$$B(L^{1/3};\theta) = \sum_{k=0}^{K} b(k;\theta)L^{k/3}, \tag{2}$$

where the high-frequency (monthly) lag operator is defined as $x^m_{t-1/3} = L^{1/3}x^m_t$. In the MIDAS literature, typically functional lag polynomials are chosen for $B(L^{1/3};\theta)$ to avoid parameter proliferation for long high-frequency lags K. A popular functional form of the polynomial is the exponential Almon lag

$$b(k;\theta) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=0}^{K}\exp(\theta_1 j + \theta_2 j^2)} \tag{3}$$

with parameters $\theta = [\theta_1, \theta_2]$ defined in Lütkepohl (1981). For given θ, the lag function provides a parsimonious way to consider a large number of K monthly lags of the indicators. The resulting functional form is typically unimodal, and can be hump shaped, declining or flat, as discussed in Ghysels et al. (2007). The weights $b(k;\theta)$ sum to one by construction, and β_1 is a regression coefficient that relates GDP to the weighted sum of high-frequency observations of the indicator.

The MIDAS parameters are estimated for each forecast horizon $h = 1,\ldots,H$ by NLS, and the direct forecast is given by the conditional expectation

$$\widehat{y}_{T_y+h|T_x} = \widehat{\beta}_0 + \widehat{\lambda}y_{T_y} + \widehat{\beta}_1 B(L^{1/3};\widehat{\theta})x^m_{T_x}, \tag{4}$$

where $T_x = T_y + w$ such that the most recent observations of the indicator are included in the conditioning set of the projection. For example, if one wants to nowcast second quarter GDP, industrial production is available for April and GDP for the first quarter, then the lead is $w = 1/3$. The specification of (1) does imply the projection (4), where the April observation is part of the conditioning set.

In the literature, alternative lag functions have been used in MIDAS regressions. For example, weights based on the beta distribution have been proposed in Ghysels et al. (2007). For mixing quarterly and monthly data, Foroni et al. (2014) discuss unrestricted lag polynomials if the lag order K is low. In this case, the polynomial is

$$\beta_1 B(L^{1/3};\theta) = b(L^{1/3};\theta) = \sum_{k=0}^{K} b(k;\theta)L^{k/3} = \sum_{k=0}^{K} b_k L^{k/3}. \tag{5}$$

This variant of MIDAS with unrestricted lag polynomials will be abbreviated by U-MIDAS from here on. Note that U-MIDAS implies the estimation of $K+1$ polynomial parameters, whereas the functional lag (3) requires three parameters to be estimated. Thus, standard MIDAS with functional lag polynomials are more parsimonious, whereas U-MIDAS is more flexible. Other weighting functions include non-exponential Almon as in Drechsel and Scheufele (2012a) or penalized changes in weights, see Breitung, Elengikal, and Roling (2013). An overview is provided in the on-line appendix of the paper Andreou et al. (2013).

2.2 Bridge equation

Following the recent papers of Foroni and Marcellino (2013) or Bulligan et al. (2014), a bridge equation with a single indicator can be defined as

$$y_t = \beta_0 + \lambda y_{t-1} + \beta(L)x_t^q + \varepsilon_t, \tag{6}$$

where y_t is again GDP growth in quarter t. The bridge equation contains a constant and an autoregressive term. The predictor x_t^q on the right-hand side is a quarterly indicator available for $t = 1, 2, 3, \dots, T_y$, the same periods as GDP. The lag polynomial $\beta(L)$ of order p is defined as $\beta(L) = \sum_{i=0}^{p} \beta_{i+1} L^i$ with $Ly_t = y_{t-1}$ and $Lx_t^q = x_{t-1}^q$.

The predictor x_t^q on the right-hand side is a quarterly indicator aggregated from the monthly indicator x_t^m over time to the quarterly frequency. Following the time aggregation literature such as Chow and Lin (1971), the mapping from the high-frequency indicator observations to the aggregated low-frequency observations is formalized through the aggregator function $\omega(L^{1/3})$ in the lag operator $L^{1/3}$ by $x_t^q = \omega(L^{1/3})x_t^m = \sum_{j=0}^{r} \omega_j L^{j/3} x_t^m$. The form of the aggregator function $\omega(L^{1/3})$ depends on the stock-flow nature of the indicator. For example, if x_t^m is a month-on-month growth rate for industrial production, the aggregated quarter-on-quarter growth rate is then defined by an aggregator function $\omega(L^{1/3})$ of order $r = 4$ according to

$$x_t^q = \omega(L^{1/3})x_t^m = x_t^m + 2x_{t-1/3}^m + 3x_{t-2/3}^m + 2x_{t-1}^m + x_{t-1-1/3}^m, \tag{7}$$

and holds for quarterly periods t only. If the variable is a stationary flow variable, the rule is $x_t^q = x_t^m + x_{t-1/3}^m + x_{t-2/3}^m$. The aggregate of a stock variable is simply $x_t^q = x_t^m$. Other aggregation rules are discussed in the appendix of Stock and Watson (2002).

Here, the lag polynomial applies to high-frequency lags of the indicator denoted by $x_{t-1/3}^q = L^{1/3}x_t^m$. Note that at the end of the sample, the filter (7) can be applied to complete quarters only. Thus, the practitioner has to wait until the three monthly observations of the indicator corresponding to the calendar quarter are available in order to apply the aggregator function (7) properly. Note that for the estimation of the bridge equation only sample periods up to T_y are used, in particular, there are no leads on the right-hand side of the bridge equation (6).

The bridge equation can be estimated by OLS, yielding parameter estimates $\widehat{\beta}_0$, $\widehat{\lambda}$, and $\widehat{\beta}(L)$. The goal is to obtain a forecast $y_{T_y+h|T_x}$ given information up to period T_x, which is just

$$\widehat{y}_{T_y+h|T_x} = \widehat{\beta}_0 + \widehat{\lambda} y_{T_y+h-1|T_x} + \widehat{\beta}(L)\widehat{x}_{T_y+h|T_x}^q. \tag{8}$$

Due to the AR(1) term, the forecast equation has to be solved forward to obtain multi-step forecasts, starting with $\widehat{\lambda} y_{T_y}$ for $h = 1$. The key ingredient is the forecast of the time-aggregated indicator $\widehat{x}_{T_y+h|T_x}^q$.

Forecasting the predictor is itself a two-step procedure including a forecast step for the monthly predictor, and a time aggregation step to obtain the quarterly projection. In the literature, typically very simple univariate models are chosen to predict the high-frequency indicator. For example, a simple AR forecast equation for x_t^m such as

$$x_t^m = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^m + \epsilon_t^m \tag{9}$$

can be used for $t = 1, 1/3, \ldots, T_x - 1/3, T_x$, where $\alpha(L^{1/3}) = \sum_{j=1}^{q} \alpha_j L^{j/3}$ is a q-order polynomial in the monthly lag operator $L^{1/3}$. The equation can be estimated by OLS and be iteratively solved to obtain monthly forecasts $\widehat{x}_{T_y+h|T_x}^m$. Note that (9) accounts for all monthly observations available up to period T_x, and can thus condition on more timely information than in the bridge equation (6).

Given the monthly forecast, the user has to aggregate the indicator forecast over time by $\widehat{x}_{T_y+h|T_x}^q = \omega(L^{1/3})\widehat{x}_{T_y+h|T_x}^m$. This forecast is plugged into the equation (8), yielding the GDP growth forecast. To illustrate the bridge nowcast, consider again the example discussed above for MIDAS. The aim is to nowcast second quarter GDP. Industrial production is available for April and GDP for the first quarter. The estimation of the AR model (9) as well as the indicator forecast consider the latest indicator information up to April. The iterative solution of (9) provides values for industrial production for May and June. The indicator projections for May and June together with the observations for April and earlier enter the time aggregation rule, for example (7). The bridge equation can be re-estimated on quarterly data, in this case up to the first quarter. Finally plugging in the time-aggregated indicator projection into the equation (8) yields the GDP nowcast. Note that recursive solutions are necessary for forecasting more than one period ahead due to the AR term in the bridge equation.

2.3 Differences between MIDAS and bridge equations

The two approaches differ in several ways from each other. Fortunately, they can both be regarded as extensions of distributed lag models to mixed-frequency data. This framework allows to isolate and highlight the key differences between the approaches.

2.3.1 Direct versus iterative multi-step forecasting

MIDAS is a direct multi-step forecast device, in the sense that the left-hand side of the equation (1), y_{t+h}, directly refers to the period $t+h$, whereas the right-hand side predictors x_{t+w}^{m} refer to period $t+w$ or earlier periods due to lags. This specification allows for a projection for horizon $h \geq 1$ in one step based on a single equation without any iterative model solution as outlined above for the bridge approach. To obtain projections for each forecast horizon $h = 1, \ldots, H$, the left-hand side variable y_{t+h} in the MIDAS equation has to be re-specified and the MIDAS equation to be estimated for each h.

The bridge equation, on the other hand, implies an iterative multi-step forecast in the sense that the model (9) of the high-frequency indicator is solved forward to produce the indicator forecasts over all horizons, which are then aggregated over time and plugged into the bridge equation (8). The bridge equation specifies y_t as a function of x_t^q and lags without considering the forecast horizon as in MIDAS. In the literature, there is a long-lasting discussion of the relative advantages of direct and iterative multi-step forecasting. Marcellino et al. (2006) and Chevillon and Hendry (2005) are recent contributions, and Bhansali (2002) provides an early survey. The literature shows that there are arguments in favour of both approaches. Generally, the direct approach is advantageous in case of mis-specification. If the model used for iterative forecasts is specified correctly, it should outperform the direct approach.

2.3.2 Differences in the functional form of the polynomials

If one disregards the differences due to direct versus iterative multi-step forecasting, it can be seen that MIDAS and bridge equations also differ with respect to the polynomials used and how the predictor observations enter the right-hand side of the equation. To make the polynomials comparable, assume the same order of the polynomials in both equations. This implies the restriction $K = 3p + r$, where the order of the MIDAS polynomial is K, and the order of the polynomial in the bridge equation is $3p + r$ coming from the convolution of $\beta(L)\omega(L^{1/3})$. To get the same number of weights on the right-hand sides of both equations, one can simply fix K, and specify $p = (K - r)/3$.

The functional lag polynomials (3) in MIDAS imply empirically estimated weights $\beta_1 B(L^{1/3}; \theta)$ for the high-frequency observations of the indicator x_t^m from (1). In contrast, the predictor enters the bridge equation (6) at quarterly frequency. But given the time aggregation scheme $\omega(L^{1/3})$ as in (7), the right-hand side of the bridge equation can also be written as a function of the high-frequency observations according to

$$\beta(L)x_t^q = \beta(L)\omega(L^{1/3})x_t^m. \tag{10}$$

Thus, the high-frequency observations are weighted by $\beta(L)\omega(L^{1/3})$ in the bridge equation, partly based on fixed time-aggregation weights $\omega(L^{1/3})$ and estimated weights $\beta(L)$. Hence, the bridge weights are not fully estimated by data compared to MIDAS. However, functional polynomials in MIDAS also impose restrictions on the weights, and the fit of MIDAS depends on the appropriateness of the chosen functional form. Thus, it is a-priori unclear whether bridge or MIDAS polynomials are supported better by the data.

The two polynomials differ also with respect to the number of parameters to be estimated: In the MIDAS example (3), we one can specify the functional polynomial by three parameters θ_1, θ_2 and β_1 to provide empirical weights for the monthly lags of the indicator. In the bridge equation, $p + 1 = (K - r)/3 + 1$ parameters have to be estimated, which is a function of K. Thus, for higher lag orders, the bridge equation approach might suffer from parameter proliferation compared to the parsimonious MIDAS (1) with weights (3).

The estimation of MIDAS and bridge equations also differs due to the different lag polynomials. Non-linear functional lag polynomials in MIDAS make NLS estimation necessary, which works iteratively and might depend on the choice of starting values. Coefficients in the bridge equation are simply estimated by OLS.

If the MIDAS regression is based on unrestricted lag polynomials, namely the U-MIDAS case with weights (5), one can immediately see that U-MIDAS actually nests the bridge equation. The U-MIDAS polynomial $\sum_{k=0}^{K} b_k L^{k/3}$ in (5) is freely estimated, whereas the bridge polynomial $\beta(L)\omega(L^{1/3})$ is restricted. Thus, U-MIDAS is more general than the bridge equation. However, this generality of U-MIDAS comes at the cost of having to estimate $K+1$ parameters in U-MIDAS, whereas the bridge approach requires estimating $(K-r)/3+1$ parameters. Generally, much lesser parameters have to be estimated in the bridge equation since $(K-r)/3+1 < K+1$ holds for any K.

Note that if K, the order of the lag polynomial, is equal to the discrepancy in sampling frequencies – in the case discussed here equal to three for quarterly and monthly data – MIDAS also nests the bridge equation approach. This case has been extensively discussed in Andreou, Ghysels, and Kourtellos (2010). They show that using fixed and equal time-aggregation weights leads to an omitted variable bias compared to MIDAS. If $K=3$, the predictor is a stationary flow variable, and AR terms are disregarded, the bridge equation is equivalent to the equal-weight model in Andreou et al. (2010). Note that the bridge equations used in practice, for example in Baffigi et al. (2004), typically contain more lags and differ with respect to the time-aggregation scheme of the indicator, see Foroni and Marcellino (2014).

2.3.3 Use of end-of-sample high-frequency data

As outlined in Clements and Galvão (2008), Clements and Galvão (2009), and Andreou et al. (2013), MIDAS can consider leads of the high-frequency indicator through the flexible direct approach in (1). Thus, when nowcasting in real-time, a newly available observation of the indicator can be used to re-estimate the MIDAS equation's parameters and to update the projection by conditioning on current-quarter indicator values. In the bridge equation approach, the most recent information can be used to update the indicator forecast model (9). When forecasting the indicator, the high-frequency model also enables the user to condition the projection on the latest indicator observation from period T_x. In this respect, both MIDAS and bridge equations use the same full-information conditioning set with observations up to T_x including current-quarter high-frequency observations.

However, the MIDAS and bridge equation themselves differ with respect to the way how current-quarter information is considered on the right-hand side. MIDAS will be re-estimated based on every incoming indicator observations. The bridge equation (6) is a low-frequency (quarterly) equation, and thus can only be updated, if the indicator data is fully available for the most recent low-frequency period. In the case discussed here, this means that the bridge equation can only be re-estimated once a quarter. But in the end, it is an empirical question, whether the indicator leads considered in the MIDAS equation help to fit the equation better than the specification in the bridge equation. Note again that this argument affects only the estimation of parameters, the conditioning set in the projection is the same in both approaches as outlined in the paragraph above.

2.3.4 Summary of comparison

The arguments discussed above do not uniquely favour MIDAS or bridge equations. It is not clear whether the more parsimonious form of the MIDAS polynomial will dominate the weighting in the bridge equation or whether the direct MIDAS approach is favourable to the iterative solution in the bridge equation approach. In the end it is an empirical question, which approach will do better.

Nonetheless, it might be interesting to see which of the differences discussed in this Section matter for nowcasting most. In the next Section, some of the differing modelling elements in MIDAS and bridge equations will be switched off and on in order to isolate the differences discussed so far.

2.4 A model in between MIDAS and bridge: Iterative MIDAS (MIDAS-IT)

In particular, intermediate models can be derived that contain aspects of MIDAS and bridge equations, but in a different way as the core models that have been discussed in the previous Sections. Consider first the discussion about direct and iterative forecasting. We can introduce iterative forecasting to the MIDAS approach, if the MIDAS equation (1) is neglected and specified in terms of GDP growth in period t rather than in period $t+h$ as in the direct approach. One obtains the equation

$$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x^m_{t+w} + \varepsilon_t. \tag{11}$$

This MIDAS equation can be estimated again with NLS. The conditional expectation for GDP growth for period $T_y + h$ given information up to T_x is

$$y_{T_y+h|T_x} = \widehat{\beta}_0 + \widehat{\lambda} y_{T_y+h-1|T_x} + \widehat{\beta}_1 B(L^{1/3};\widehat{\theta})x^m_{T_y+h+w|T_x}. \tag{12}$$

It requires a projection $x^m_{T_y+h+w|T_x}$ of the indicator. To facilitate comparability to the bridge equation approach, it is assumed that the simple AR model in (9) provides the iterative indicator projection. Equation (11) it is only estimated once for all horizons $h = 1, \ldots, H$. From here on, the model consisting of (11) and the AR model is called iterative MIDAS (MIDAS-IT) with leads.[1] Taking another perspective, it might be regarded as a bridge equation with functional high-frequency lag polynomials for the high-frequency indicator. Compared to the standard MIDAS approach in the literature (1), MIDAS-IT helps to assess whether

[1]The model with AR predictors in the case when K equals the discrepancy in sampling frequencies (equal to three for quarterly-monthly data) and without leads, has been discussed in Andreou et al. (2010).

the iterative approach can improve over the direct multi-step approach used in the standard MIDAS approach (1). The differences also stem from the use of the AR model to obtain the indicator forecast, which can be neglected in the direct MIDAS approach. Compared to the bridge equation, MIDAS-IT uses the same AR model for the indicator nowcasts. Thus, there are only two differences that can lead to different nowcast performances: 1) MIDAS-IT and bridge equations differ due to the differences in the lag polynomials $\widehat{\beta}_1 B(L^{1/3};\widehat{\theta})$ and $\widehat{\beta}(L)\omega(L^{1/3})$, and 2) also with respect to the information used for estimating the equation for the low-frequency variable. MIDAS-IT has x^m_{t+w} on the right-hand side including the lead, whereas the bridge equation contains x^m_t on the right-hand side without a lead.

To isolate the difference stemming from the lead of the indicator in the MIDAS equation, one can formulate another model that contains no leads in the MIDAS equations such that only x^m_t enters the right-hand side rather than x^m_{t+w}:

$$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x^m_t + \varepsilon_t. \tag{13}$$

This approach is called MIDAS-IT without leads. To nowcast with this approach, again the AR model is used to predict the indicator. Comparing MIDAS-IT with and without leads thus helps to answer the question, whether leads of the indicator should be used for estimating the equation of the low-frequency variable. The conditioning set for forecasting is the same in both approaches, as the AR model helps to make projections for the indicator. As an alternative, the forecaster can rely on U-MIDAS polynomials $\sum_{k=0}^{K} b_k L^{k/3}$ in both approaches. MIDAS-IT without leads and the bridge equation differ only by the polynomials $\widehat{\beta}_1 B(L^{1/3};\widehat{\theta})$ and $\widehat{\beta}(L)\omega(L^{1/3})$. If MIDAS-IT without leads turns out to be better, one can conclude that the functional lag weighting is better suited for nowcasting than the weighting implied by the bridge equation, which is partly based on a low-frequency polynomial and the time aggregation function.

To wrap-up the discussion in this Section, the full set of models and their features are summarized in Table 1.

3 Empirical application

3.1 Data

The dataset contains Euro area quarterly GDP growth from 1999Q1 until 2013Q4 and about 70 monthly indicators until 2013M12. The monthly indicators cover industrial production by sector, surveys on consumer sentiment and business climate including the Purchasing Managers Index (PMI), international and trade data, and

Table 1: Overview of MIDAS and bridge equations used in the empirical exercise

	equations for y_t and indicator x_t^m	direct or iterative	lag weights	leads in equation
A. MIDAS	$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B(L^{1/3};\theta)x_{t+w}^m + \varepsilon_{t+h}$	direct	functional or unrestricted poly	yes
B. MIDAS-IT, leads	$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x_{t+w}^m + \varepsilon_t$ $x_t^m = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^m + \epsilon_t^m$	iterative	functional or unrestricted poly	yes
C. MIDAS-IT, no leads	$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x_t^m + \varepsilon_t$ $x_t^m = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^m + \epsilon_t^m$	iterative	functional or unrestricted poly	no
D. Bridge	$y_t = \beta_0 + \lambda y_{t-1} + \beta(L)x_t^q + \varepsilon_t$ $x_t^m = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^m + \epsilon_t^m$	iterative	unrestricted poly and time aggregation	no

Note: The model abbreviations and model details are explained in Section 2 of the text.

financial data. Note that the beginning of the overall sample is restricted by the financial data for the Euro area and the PMI. Due to the attention the PMI has received recently, for example in Lahiri and Monokroussos (2013), it was decided to include this indicator in the dataset. A complete list of the variables used can be found in Appendix A.

The dataset is a final dataset and not a real-time dataset, so that the role of revisions on the relative forecasting accuracy cannot be discussed here. Partly, this is due to the fact that large monthly real-time data with sufficiently long samples are not available yet for the Euro area. Furthermore, it is not clear that major changes can be expected from the use of real-time vintages as in Bernanke and Boivin (2003), Schumacher and Breitung (2008), and Clements (2014). However, another important characteristic of multivariate data in real time is taken into account, namely the different availability of variables due to publication lags. These differences in availability of data lead to certain patterns of missing values at the end of every sub-sample, and recent papers find that accounting for this rather than using artificially balanced samples has a considerable impact on forecast accuracy, see Giannone, Reichlin, and Small (2008), for example. To consider the availability of the data at the end of each subsample, the nowcast exercise in this paper follows Giannone et al. (2008) and Foroni and Marcellino (2014), amongst others, where the availability of data is replicated from a final vintage of data in pseudo real-time.

3.2 Design of the empirical nowcast exercise

To evaluate the performance of the models, an empirical pseudo-real time exercise with rolling estimation and nowcasting will be carried out. Each rolling estimation sample contains 96 months of data. The evaluation sample is between 2008Q1 and 2013Q4, providing 6 years for comparison including the Great Recession. In the recent nowcast literature, it has been observed that the Great Recession has led to a strong decline in predictability in European countries, see for example Drechsel and Scheufele (2012b) and Kuzin, Marcellino, and Schumacher (2013). Foroni et al. (2014) and Foroni and Marcellino (2014) also report strong differences in predictability before and during the Great Recession in the Euro area. Below, the main interest lies in the period after the Great Recession. For this reason, the evaluation sample will be split in two subsamples, namely, 2008 to 2009 (Great Recession) and 2010 to 2013 (the aftermath of the Great Recession) in line with the papers cited above. For each period in the evaluation samples, nowcasts depending on different monthly information sets are computed. For example, for the initial evaluation quarter 2008Q1, nowcasts are computed in March 2008, one in February, and January. Every month, the nowcast is computed at the end of the month given the data available at that point in time. Note that the nowcast made at the end of the final month of the nowcast quarter might still be informative to decision makers, as the GDP figure for the corresponding reference quarter will be released about six weeks later. In the Tables below, nowcast error statistics will be provided for nowcasts made in each month of the reference quarters.

The estimation sample depends on the information available at each period in time when computing the nowcasts. Assume again the aim is to nowcast GDP for $2008Q1$ in March 2008, then the time series observations available at that period in time have to be identified. For this purpose, the ragged-edge structure from the end of the full sample of data is exploited, as discussed in the previous subsection. To replicate the publication lags of GDP, one can exploit the fact that in the Euro area GDP of the previous quarter is available in the middle of the second month of the next quarter. All nowcast models are re-estimated on the data in each subsample, such that the estimated coefficients are allowed to change over time. The maximum lag order in MIDAS is equal to $K = 12$ months. The maximum lag order of the indicator polynomial in the bridge equation is equal to $p = 4$. The AR model for the indicator has a maximum lag order of $q = 12$. The lag order used each subsample is determined by the Bayesian Information Criterion (BIC).

For each evaluation period, three nowcasts are computed. To compare the nowcasts with the realisations of GDP growth, the mean-squared error (MSE) is employed. Below, also relative MSEs will be reported. These will relate the performance of MIDAS to the performance of bridge equations, but also to the performance of an AR benchmark model. The AR model can have a lag order up to four, which is specified using the BIC.

3.3 Performance of single MIDAS and bridge equations

In Table 2, the nowcast performance of the MIDAS and bridge equations and their MIDAS-IT variants are compared to the AR benchmark. Given the large number of indicators and model groups, the Table contains summary statistics for the distribution of relative MSEs from all the models within a model class with the different predictors as in Stock and Watson (2012) or Foroni et al. (2014). For example, given the standard MIDAS model class, first the ratio of the MSE of MIDAS based on a single indicator is computed and divided by the MSE of the AR benchmark. The same is done for all MIDAS regressions with the other predictors, yielding a distribution of relative MSEs across predictors for this model class. The same procedure is applied to the bridge equations and the MIDAS-IT variants. In each case, values of the relative MSE smaller than one indicate a superior performance of MIDAS or the bridge equations compared to the benchmark. In Table 2, the median as well as the 10th and 90th percentiles of the relative MSE distributions are reported for each model class.

In the upper part of Table 2, the results for the evaluation period 2008 to

Table 2: Percentiles of relative MSEs from MIDAS and bridge equations relative to the AR benchmark

A. 2008-2009

model	percentiles	3rd	2nd	1st
MIDAS	0.10	0.49	0.52	0.44
	0.50	0.68	0.70	0.65
	0.90	0.91	1.02	0.86
MIDAS-IT with leads	0.10	0.61	0.61	0.48
	0.50	0.69	0.71	0.67
	0.90	0.91	0.95	0.85
MIDAS-IT, no leads	0.10	0.49	0.62	0.44
	0.50	0.67	0.69	0.66
	0.90	0.91	0.95	0.98
bridge	0.10	0.28	0.40	0.29
	0.50	0.68	0.76	0.69
	0.90	1.07	1.06	1.29

B. 2010-2013

model	percentiles	3rd	2nd	1st
MIDAS	0.10	0.96	0.93	1.11
	0.50	1.25	1.16	1.75
	0.90	2.52	2.36	4.86
MIDAS-IT with leads	0.10	0.82	0.86	0.86
	0.50	1.06	1.19	1.03
	0.90	1.83	1.96	2.10
MIDAS-IT, no leads	0.10	0.80	0.93	0.87
	0.50	1.15	1.19	1.20
	0.90	1.88	1.90	2.26
bridge	0.10	0.87	0.96	0.98
	0.50	1.75	1.57	1.69
	0.90	3.19	2.69	3.11

Note: The model abbreviations and model details are explained in Section 2 of the text. The third, fourth and fifth columns refer to the nowcasts computed in the 3rd, 2nd and 1st month of the respective nowcast period.

2009 are shown for the three monthly nowcasts. In the majority of cases for MIDAS and bridge equations, the percentiles are below one, indicating a superior performance than the benchmark for a large part of the model distributions. The 10th percentile of the bridge equations yields the lowest relative MSE compared to any other method, but the 90th percentile indicates a worse performance of some bridge equations compared to the benchmark. Standard MIDAS, MIDAS-IT with and without leads all perform similarly compared to the benchmark. In the lower part of Table 2, the results for the evaluation period 2010 to 2013 are shown. In all the model classes, the 90th percentiles and the medians over all horizons are greater than one, indicating that the majority of models within each class performs worse than the AR benchmark. However, the best 10% of the models can often provide a better performance than the benchmarks. For example, for MIDAS-IT with and without leads for the nowcast made in the 3rd month, the 10th percentile is about 0.8, whereas standard MIDAS and the bridge equation approach perform bit worse with 10th percentiles of 0.96 and 0.87, respectively.

A comparison betwen the upper and lower part of the Table 2 indicates a deline in relative predictability between Great Recession period from 2008 to 2009 and the evaluation period 2010 to 2013. Note also that the performance of the AR benchmark also changes over time such that the relative MSE results should not be confused with indicators of absolute nowcast accuracy. Actually, the performance of standard benchmarks during the Great recession was very bad as documented in Drechsel and Scheufele (2012b) for Germany, and for other countries by Kuzin et al. (2013). In the period after 2009, the AR model has performed much better. As an example, the absolute MSE for the AR nowcast made in the 3rd month in the quarter was 2.26 over the period 2008 to 2009, whereas it was 0.10 the evaluation period 2010 to 2013. The moderate relative decline in the performance of MIDAS and bridge equation indicate that their absolute MSEs have also declined, but to a lesser extent as the benchmark. Thus, the decline in relative MSEs should not be confused with an absolute decline in predictability.

To highlight the differences between MIDAS and bridge equations and their variants discussed in Section 2.3, in Table 3 summary statistics of relative MSEs between MIDAS and bridge equation model pairs are shown. The relative MSEs are technically computed in the same way as for the previous Table, but below the following model pairs are compared: 1) standard MIDAS relative to MIDAS-IT with leads to assess the relevance of iterative versus direct multi-step nowcasting; 2) MIDAS-IT with leads relative to MIDAS-IT without leads to assess the relevance of leads in the equation; 3) MIDAS-IT without leads relative to the bridge to assess the relevance of the weighting scheme in the bridge equation.

The results can be summarized as follows. For most of the model pairs, the 10th percentile is smaller than one for all horizons, whereas the 90th percentiles are greater than one. The medians vary around one without a clear tendency across

Table 3: Percentiles of relative MSEs from different combinations of MIDAS and bridge equations

A. 2008-2009

model	percentiles	3rd	2nd	1st
	0.10	0.74	0.77	0.66
MIDAS relative to MIDAS-IT with leads	0.50	0.93	0.97	0.99
	0.90	1.16	1.24	1.37
	0.10	0.89	0.92	0.74
MIDAS-IT with leads relative to MIDAS-IT, no leads	0.50	1.05	1.00	1.00
	0.90	1.49	1.14	1.41
	0.10	0.62	0.62	0.54
MIDAS-IT, no leads, relative to bridge	0.50	1.02	0.98	1.02
	0.90	2.20	1.79	2.09

B. 2010-2013

model	percentiles	3rd	2nd	1st
	0.10	0.75	0.67	1.04
MIDAS relative to MIDAS-IT with leads	0.50	1.09	1.01	1.69
	0.90	1.95	2.10	2.95
	0.10	0.70	0.84	0.64
MIDAS-IT with leads relative to MIDAS-IT, no leads	0.50	0.96	1.00	0.91
	0.90	1.32	1.21	1.27
	0.10	0.42	0.45	0.41
MIDAS-IT, no leads, relative to bridge	0.50	0.77	0.81	0.75
	0.90	1.17	1.16	1.15

Note: The model abbreviations and model details are explained in Section 2 of the text. The third, fourth and fifth columns refer to the nowcasts computed in the 3rd, 2nd and 1st month of the respective nowcast period.

horizons and model pairs under comparison. The only exception is MIDAS-IT without leads relative to the bridge equation in the evaluation period 2010-2013 in the lower part of the Table, where the median for all horizons is smaller than one. In that period, the MIDAS weighting scheme seems to dominate the time aggregation scheme of the bridge approach. But in general there are no clear-cut differences in the relative MSE distributions, and there is no superior performance of one of the model classes compared to the others. As all the 10th percentiles of the relative MSE pairs are smaller than one and all the 90th percentiles are greater than one, there seem to be competitive models and indicators in each model class. Thus, the relative accuracy of MIDAS and bridge equations might change on a case-by-case basis depending on the indicator and sample period chosen.

To discuss this issue further, I depart from the assessment of the whole distribution of relative MSEs and rather focus on the best-performing models. For this purpose, all the models are ranked according to their MSE performance relative to the AR benchmark. Table 4 shows the 20 best-performing of all the models for the nowcast computed in the third month of the quarter. The results are again presented for the two evaluation periods 2008 to 2009 and 2010 to 2013. In the Table, each model is characterized by the model class it belongs to and the predictor used for nowcasting in the model.

Regarding the performance of single nowcast models, one can see quite different rankings. For example, industrial production in manufacturing (ip_manu) nowcasts best with the bridge equation (rank 5 in 2008-2009), followed by MIDAS-IT without leads (rank 9 in 2008-2009), and standard MIDAS (rank 20). In the second sample period the relative ranking is the same. As another example, the survey headline index PMI manufacturing (pmi_man_head) yields the best nowcast results from MIDAS-IT without leads (rank 11 in 2008-2009), followed by standard MIDAS (rank 12), and the bridge approach (rank 20). In the latter evaluation period, only MIDAS-IT with leads (rank 7 in 2010-2013) based on the PMI manufacturing indicator makes it to the top 20 ranking. To sum up, the rankings of the different MIDAS and bridge models are depending on the indicators chosen, confirming the results from the previous Tables.

Taking a broader view on the results, there is again a striking difference in nowcastability between the two sample periods with respect to the benchmark. In the earlier period 2008 to 2009, the worst model in the ranking has a relative MSE of 0.38, whereas the worst MSE in the period 2010 to 2013 is equal to 0.75. In the evaluation period 2008 to 2009, there are 11 bridge equations in the list of the best 20 models. In the period until 2013, there are only 5 bridge equations among the top 20. In both periods, MIDAS-IT without leads is also in the top 20 ranking, in particular, six times in each evaluation period. The standard MIDAS as well as MIDAS-IT with leads can be found less often in the rankings. The general finding is that both rankings look very different in many respects depending on size of the

Table 4: Rankings of MIDAS and bridge equations based on relative MSEs

A. 2008-2009

rank	model	indicator	relative MSE
1	bridge	sur ind prod exp	0.14
2	bridge	ip x constr en	0.19
3	bridge	ip	0.20
4	bridge	sur ind emp exp	0.20
5	bridge	ip manu	0.21
6	MIDAS-IT, no leads	ip x constr en	0.26
7	bridge	ip x constr	0.26
8	MIDAS-IT, no leads	sur cons n12m	0.26
9	MIDAS-IT, no leads	ip manu	0.26
10	bridge	ip mig intermediate	0.27
11	MIDAS-IT, no leads	pmi man head	0.28
12	MIDAS	pmi man head	0.28
13	bridge	sur ind conf	0.28
14	bridge	ip mig cap goods	0.29
15	bridge	pmi man head	0.32
16	MIDAS-IT, no leads	sur ind prod exp	0.33
17	MIDAS	sur ind prod exp	0.34
18	MIDAS-IT, leads	ip x constr en	0.36
19	bridge	sur serv emp n3m	0.37
20	MIDAS	ip manu	0.38

B. 2010-2013

rank	model	indicator	relative MSE
1	bridge	ip manu bas met	0.45
2	MIDAS	ip mig intermediate	0.54
3	MIDAS-IT, no leads	ip mig cap goods	0.54
4	bridge	ip	0.54
5	MIDAS	ip manu plastic	0.58
6	bridge	ip manu	0.59
7	MIDAS-IT, leads	pmi man head	0.60
8	MIDAS	ip	0.60
9	MIDAS-IT, no leads	ip	0.63
10	MIDAS-IT, no leads	ip x constr en	0.65
11	MIDAS-IT, no leads	ip manu el eq	0.66
12	bridge	ip mig intermediate	0.67
13	MIDAS-IT, leads	ip manu plastic	0.67
14	MIDAS-IT, leads	sur ind prod trends	0.68
15	MIDAS-IT, no leads	ip manu	0.69
16	MIDAS-IT, no leads	ip manu plastic	0.69
17	bridge	ip x constr en	0.73
18	MIDAS-IT, leads	ip manu el eq	0.73
19	MIDAS-IT, leads	ip manu bas met	0.73
20	MIDAS	ip manu chem	0.75

Note: The model abbreviations from column 2 and the model details are explained in Section 2 of the text. The indicator abbreviations in column 3 are explained in the data appendix. The relative MSE in column 4 refers to the relative MSE of the MIDAS or bridge equation compared to the AR benchmark, where the nowcast is computed in the 3rd month of the respective quarter.

relative MSEs, composition of model classes, and the indicators used. Note again that the relative performance depends also on the AR benchmark performance.

3.4 Pooling MIDAS and bridge equations

A potential way to address nowcast instability in the single models as found in the results above is pooling of single nowcasts. The forecast literature such as Timmermann (2006), Hendry and Clements (2004), and Clark and McCracken (2010) has shown that combining forecasts from alternative models can be beneficial in the presence of misspecification and temporal instabilities. Related to the model classes discussed here, Clements and Galvão (2008), Clements and Galvão (2009) and Kuzin et al. (2013) consider combinations of MIDAS models, whereas bridge equations have been pooled in Hahn and Skudelny (2008) and Rünstler et al. (2009), amongst others. Below, nowcast pooling results for MIDAS and bridge equations are compared for three weighting schemes: the mean, the median, and the weighted mean of all the models of a particular class, where combination weights are obtained from the inverse MSE of the previous four-quarter performance of a model as in Kuzin et al. (2013). In Table 5, the relative MSEs compared to the AR benchmark of the nowcast combinations for the two subsamples are shown.

According to the results, all of the combinations perform well relative to the benchmark. This result is stable over all forecast horizons and evaluation samples. Also the absolute size of the relative MSEs is comparable between the two sample periods. Note that many individual models in the second evaluation period 2010 to 2013 perform worse than the benchmark as shown in Table 2. Compared to the single models, pooling nowcasts provides more stable results. Comparing the pooling results with the rankings from Table 4, it can be observed that pooling is not as good as the best-performing models, despite still performing clearly better than the benchmark.

With respect to the model classes, the results indicate a better performance of the combinations of bridge equation nowcasts than the variants of MIDAS in both evaluation samples. Within MIDAS, the MIDAS-IT without leads seems to perform slightly better than MIDAS-IT with leads.

With respect to pooling schemes, the results are similar. The performance of equal-weights pooling is comparable to MSE-weighted pooling, whereas the median performs slightly worse than the other weighting schemes.

3.5 Test results on the nowcast accuracy over time

To assess the nowcast performance over time further, the fluctuation test developed by Giacomini and Rossi (2010) will be applied. Their test for significant

Table 5: Pooling of MIDAS and bridge equations

A. 2008-2009

model	weighting scheme	3rd	2nd	1st
MIDAS	mean	0.61	0.67	0.60
MIDAS-IT, leads	mean	0.67	0.71	0.64
MIDAS-IT, no leads	mean	0.60	0.68	0.63
bridge	mean	0.57	0.66	0.58
MIDAS	median	0.65	0.67	0.63
MIDAS-IT, leads	median	0.68	0.67	0.66
MIDAS-IT, no leads	median	0.62	0.66	0.64
bridge	median	0.66	0.73	0.63
MIDAS	MSE-weighted mean	0.59	0.63	0.57
MIDAS-IT, leads	MSE-weighted mean	0.67	0.70	0.63
MIDAS-IT, no leads	MSE-weighted mean	0.59	0.67	0.58
bridge	MSE-weighted mean	0.52	0.64	0.51

B. 2010-2013

model	weighting scheme	3rd	2nd	1st
MIDAS	mean	0.72	0.72	0.86
MIDAS-IT, leads	mean	0.76	0.81	0.78
MIDAS-IT, no leads	mean	0.75	0.81	0.76
bridge	mean	0.52	0.58	0.55
MIDAS	median	0.81	0.87	0.85
MIDAS-IT, leads	median	0.91	0.89	0.89
MIDAS-IT, no leads	median	0.88	0.92	0.84
bridge	median	0.55	0.68	0.64
MIDAS	MSE-weighted mean	0.72	0.80	0.73
MIDAS-IT, leads	MSE-weighted mean	0.77	0.90	0.76
MIDAS-IT, no leads	MSE-weighted mean	0.77	0.89	0.64
bridge	MSE-weighted mean	0.50	0.58	0.65

Note: The model abbreviations and model details are explained in Section 2 of the text. The third, fourth and fifth columns refer to the nowcasts computed in the 3rd, 2nd and 1st month of the respective nowcast period.

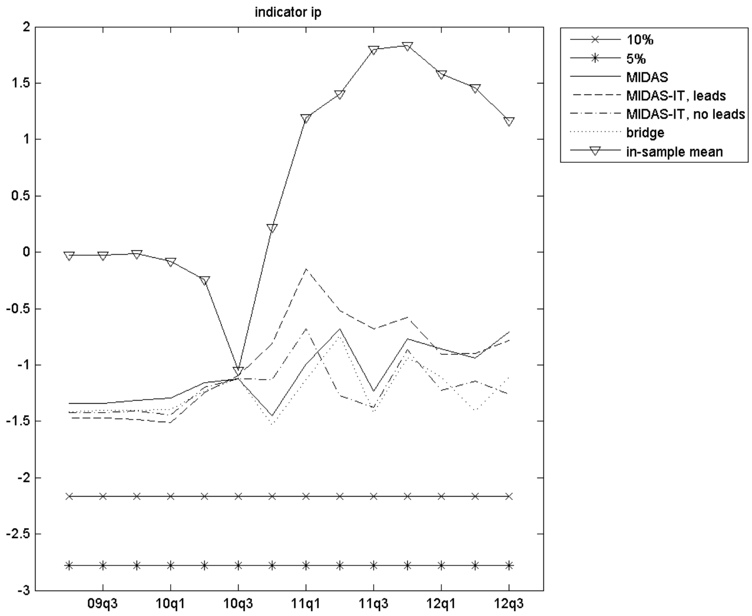
differences in the relative nowcast performances between two models is based on a measure of local relative nowcast loss, which is computed by the sequence of out-of-sample loss differences over centered rolling windows. The test is implemented by scaling the loss differences using a heteroscedasticity and autocorrelation consistent (HAC) estimator of the variance of the loss differential in each evaluation period and plotting the sample path of this statistic together with critical values reported in Giacomini and Rossi (2010). Below, the local loss differential is equal to the squared nowcast errors of the MIDAS or bridge equation nowcast minus the squared nowcast errors of the AR benchmark. Thus, negative values indicate a superior performance of MIDAS or bridge equations over the benchmark. The statistics are computed over rolling and centered windows with window size of eleven quarters, which is about one third of the overall evaluation sample size as in Rossi and Sekhposyan (2010). If the critical values are crossed by the local performance measure, the corresponding model significantly outperforms the benchmark at the respective period in time. Below, the one-sided test of Giacomini and Rossi (2010) is applied, where the null hypothesis states that the MIDAS or bridge nowcasts are worse than the benchmark.

In Figure 1, MIDAS and bridge equation nowcasts based on industrial production (panel A) and the PMI (panel B) are compared to the AR benchmark and the in-sample mean benchmark. In Figure 2, the performance of nowcast combinations across all indicators are shown. If the statistics exceed the 5 and 10% critical values, the benchmark is significantly outperformed in the respective evaluation period. In the Figures below, the test refers to the nowcast made at the end of the final month in the quarter.

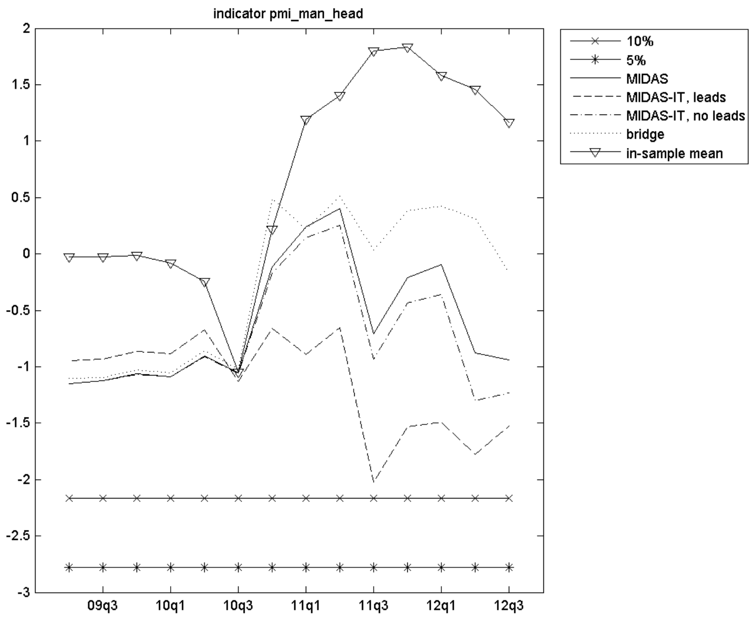
As in the application by Rossi and Sekhposyan (2010), the results show that it is generally difficult to outperform the AR benchmark, as the test statistics are not significant in most of the evaluation periods. In Figure 1, the nowcasts based on industrial production and the PMI are not significantly better than the benchmark in all the evaluation periods shown. PMI performs slightly better than IP in a number of periods, which is partly even worse than the AR nowcasts indicated by positive values of the statistic. The pooled nowcasts in Figure 2 are not significantly better than the benchmark in the early periods, but have a significantly better performance in the years 2010 to 2012. The performance over time of MIDAS and bridge equations is generally very similar. There are some periods where bridge equations do better, for example, pooling in the first part of the sample or MIDAS-IT with leads based on the PMI in the final part of the sample. But there is generally no systematic advantage of one of the model classes discussed. The nowcast accuracy is affected more by the indicators used or whether nowcasts are pooled or not.

Figure 1: Fluctuation test statistics for MIDAS and bridge equations with different indicators

A. Industrial production

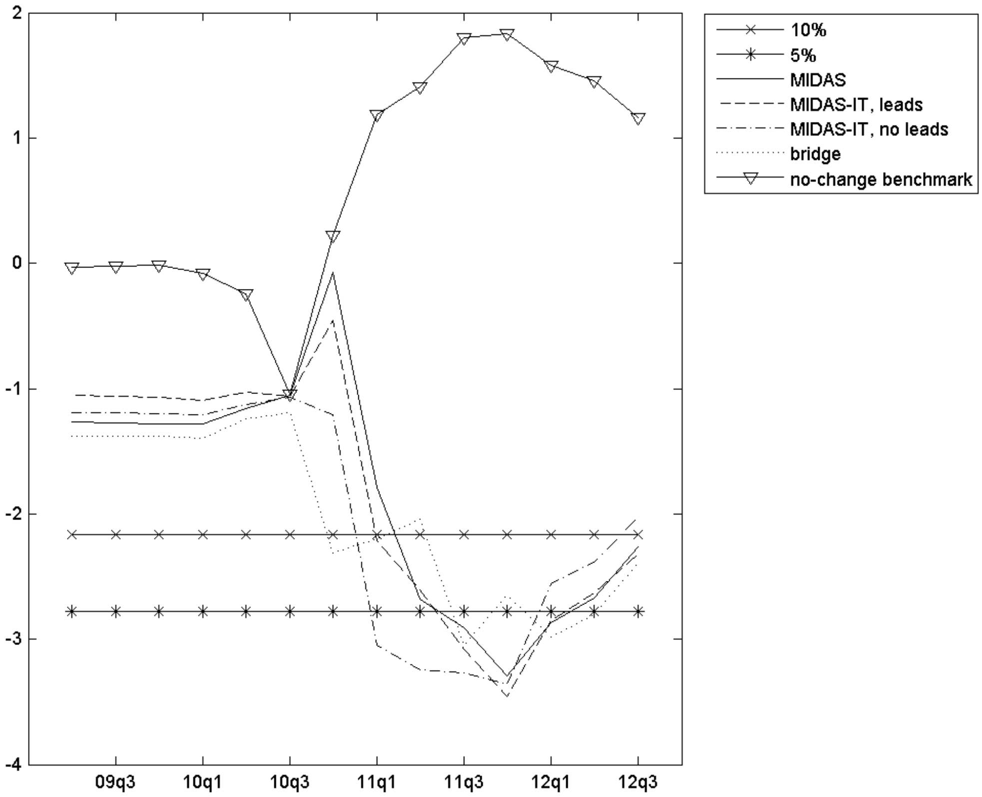


B. PMI



Note: The model abbreviations and model details are explained in Section 2 of the text.

Figure 2: Fluctuation test statistics for pools of MIDAS and bridge equations



Note: The model abbreviations and model details are explained in Section 2 of the text.

3.6 Further results to check robustness

In order to investigate the sensitivity of the results, additional empirical exercises were carried out. In particular, MIDAS regressions with unrestricted linear lag polynomials (U-MIDAS) were tried out. In addition to the rolling exercises with constant estimation sample size, also recursive exercises were carried out with an increasing sample size over time. Below, I report a summary of the additional findings only to save space. The results of the robustness checks can be provided upon request to interested readers.

The additional results can be summarized as follows. Single U-MIDAS regressions turned out to perform worse in the majority of cases compared to the exponential Almon lag MIDAS regressions. The relatively short estimation samples seemed to play a role here. Only in some cases for certain indicators, U-MIDAS did better. When pooling the U-MIDAS nowcasts, however, the results were quite close to the combinations of exponential Almon MIDAS regressions. The pooling performance compared to the benchmarks and the comparison between the variants of U-MIDAS and bridge equation was similar to the exponential Almon MIDAS case. Thus, the results from the previous Sections could be confirmed when using U-MIDAS regressions. In the recursive nowcast exercise, both MIDAS and bridge equations performed slightly worse compared to the rolling exercise.

4 Summary and conclusions

In the recent literature and in policy institutions, MIDAS and bridge equations have been widely used for nowcasting quarterly GDP growth based on monthly business cycle indicators. This paper compares MIDAS and bridge equations in a distributed-lag framework identifies three major differences: 1) MIDAS is a direct multi-step nowcast tool, whereas bridge equations are based on iterative multi-step nowcasts of the indicator, 2) MIDAS is based on functional or unrestricted lag polynomials of the monthly indicators, whereas the bridge equations are partly based on time aggregation, and 3) MIDAS equations consider current-quarter observations of the indicators on the right-hand side, whereas bridge equations generally do not. Highlighting the differences between the approaches might help a practitioner to make modelling decisions between MIDAS and bridge equations that have been discussed mostly in isolation in the literature.

In an empirical exercise for nowcasting Euro area GDP growth, MIDAS and bridge equations with single indicators are evaluated with respect to their out-of-sample nowcast accuracy. The evaluation periods include the recent Great Recession and the years after. It turns out that the performance of both single-indicator MIDAS and bridge equations varies over time. Compared to benchmarks, and many models do well during the Great Recession and perform badly thereafter.

Rankings of the MIDAS and bridge equations differ quite a lot for different indicators. Thus, if the user wants to pick a single MIDAS or bridge equations models, it might be necessary to evaluate the nowcast models for each indicator extensively to find out the best specification, as the nowcast results across models and predictors show no clear-cut ranking. In contrast, pooling nowcasts obtained from single MIDAS or bridge equations provides stable results over the whole evaluation period and ensures against misspecification to a certain extent.

A Euro area dataset

This Appendix describes the time series for the Euro area economy used in the nowcast exercise. The whole data set for Euro area contains 71 monthly indicators over the sample period from $1999M1$ until $2013M12$. A complete list of variables is provided in Tables 6 and 7 below, together with abbreviations used in the description of results in the main text.

The sources of the time series are the databases of the Bundesbank and Data Insight. 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.

References

Andreou, E., E. Ghysels, and A. Kourtellos (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics 158*(2), 246–261.

Andreou, E., E. Ghysels, and A. Kourtellos (2011). Forecasting with mixed-frequency data. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of economic forecasting*, pp. 225–245. Oxford University Press.

Andreou, E., E. Ghysels, and A. Kourtellos (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics 31*(2), 240–251.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin, and G. Rünstler (2011). Short-term forecasts of euro area GDP growth. *The Econometrics Journal 14*(1), C25–C44.

Armesto, M. T., K. M. Engemann, and M. T. Owyang (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review 92*(Nov/Dec), 521–536.

Table 6: List of monthly monthly indicators, part 1

1	industry	total industry	ip
2	industry	total ex constr	ip_x_constr
3	industry	manu	ip_manu
4	industry	constr	ip_constr
5	industry	total ex constr and mig energy	ip_x_constr_en
6	industry	energy	ip_energy
7	industry	mig cap goods ind	ip_mig_cap_goods
8	industry	mig durable cons goods ind	ip_mig_dur
9	industry	mig energy	ip_mig_energy
10	industry	mig intermed goods ind	ip_mig_intermediate
11	industry	mig non-durable cons goods ind	ip_mig_non_dur
12	industry	manu basic metals	ip_manu_bas_met
13	industry	manu chemicals	ip_manu_chem
14	industry	manu elect equipment	ip_manu_el_eq
15	industry	manu machinery and equipment	ip_manu_mach_eq
16	industry	manu rubber and plastic prod	ip_manu_plastic
17	industry	ex motor vehicles and motorcycles	retail_x_vehicles
18	industry	new passenger car registrations	car_registrations
19	financial	DJ Euro Stoxx 50	euro_stoxx_50
20	financial	DJ Euro Stoxx Price Index	euro_stoxx_total
21	financial	Standard& Poors 500 Index USA	sp500_usa
22	financial	us, dow jones, industrial average	dow_jones
23	financial	10-year government bond yield	10_y_gov_bond
24	financial	3-month interest rate, euribor	euribor_3m
25	financial	index of notional stock - money m1	m1
26	financial	index of notional stock - money m2	m2
27	financial	index of notional stock - money m3	m3
28	labour	unemployment rate, total	unemployment rate
29	international	US PMI manu	us_pmi_manu
30	international	US unemployment rate	us_unemp
31	international	US IP total ex construction	us_ip_x_cons
32	international	US employment, civilian	us_empl
33	international	US retail trade	us_retail
34	international	world market prices, raw materials, total	ind_raw_mat
35	international	world market prices raw materials, ex energy	ind_raw_mat_x_en
36	international	world market prices, crude oil, usd, hwwa	oil
37	international	brent crude, 1 month fwd, usd/bbl	brent_fut

Table 7: List of monthly monthly indicators, part 2

38	surveys	industrial confidence	sur_ind_conf
39	surveys	production trend recent months	sur_ind_prod_trends
40	surveys	assessment of order-book levels	sur_ind_order
41	surveys	assessment of export order-book levels	sur_ind_exp_order
42	surveys	assessment of stock of finished products	sur_ind_stocks
43	surveys	production expectations for the months ahead	sur_ind_prod_exp
44	surveys	selling price expectations for the months ahead	sur_ind_price_exp
45	surveys	employment expectations for the months ahead	sur_ind_emp_exp
46	surveys	consumer confidence	sur_cons_conf
47	surveys	general economic situation last 12 months	sur_cons_p12m
48	surveys	general econ situation next 12 months	sur_cons_n12m
49	surveys	price trends over last 12 months	sur_cons_pr_p12m
50	surveys	price trends next 12 m	sur_cons_pr_n12m
51	surveys	unemployment expectations over next 12 m	sur_cons_unemp_n12m
52	surveys	major purchases at present	sur_cons_purch_pres
53	surveys	major purchases over next 12 m	sur_cons_purch_n12m
54	surveys	construction confidence	sur_const_conf
55	surveys	trend of activity compared with preced. months	sur_const_act_trend
56	surveys	assessment of order books	sur_const_orders
57	surveys	employment next 3m	sur_const_emp_n3m
58	surveys	selling price expectations for the months ahead	sur_const_pr_fut
59	surveys	retail confidence	sur_retail_conf
60	surveys	present business situation	sur_retail_pres_sit
61	surveys	assessment of stocks	sur_retail_stocks
62	surveys	expected business situation	sur_retail_exp_sit
63	surveys	employment expectations	sur_retail_emp_exp
64	surveys	services confidence	sur_serv_conf
65	surveys	business situation past 3 m	sur_serv_busi_p3m
66	surveys	demand past 3 m	sur_serv_dem_p3m
67	surveys	demand over next 3 m	sur_serv_dem_n3m
68	surveys	employment past 3 m	sur_serv_emp_p3m
69	surveys	employment next 3 m	sur_serv_emp_n3m
70	surveys	PMI Manufacturing Headline Index	pmi_man_head
71	surveys	PMI Services Headline Index	pmi_serv_head

Baffigi, A., R. Golinelli, and G. Parigi (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting 20*(3), 447–460.

Banbura, M., D. Giannone, and L. Reichlin (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of economic forecasting*, pp. 193–224. Oxford University Press.

Bell, V., L. W. Co, S. Stone, and G. Wallis (2014). Nowcasting UK GDP growth. *Bank of England Quarterly Bulletin 54*(1), 58–68.

Bernanke, B. S. and J. Boivin (2003). Monetary policy in a data-rich environment. *Journal of Monetary Economics 50*(3), 525–546.

Bhansali, R. J. (2002). Multi-Step Forecasting. In M. P. Clements and D. F. Hendry (Eds.), *A Companion to Economic Forecasting*, pp. 206–221. Wiley-Blackwell.

Breitung, J., S. Elengikal, and C. Roling (2013). Forecasting inflation rates using daily data: A nonparametric MIDAS approach. University of Bonn, Mimeo.

Bulligan, G., R. Golinelli, and G. Parigi (2010). Forecasting monthly industrial production in real-time: from single equations to factor-based models. *Empirical Economics 39*(2), 303–336.

Bulligan, G., M. Marcellino, and F. Venditti (2014). Forecasting economic activity with targeted predictors. *International Journal of Forecasting*, forthcoming.

Bundesbank (2013). Forecasting models in short-term business cycle analysis - a workshop report. *Deutsche Bundesbank Monthly Report September 2013*, 69–83.

Camacho, M., G. Perez-Quiros, and P. Poncela (2014). Short-term forecasting for empirical economists. a survey of the recently proposed algorithms. *Foundations and Trends in Econometrics 6*, 101–161.

Chevillon, G. and D. F. Hendry (2005). Non-parametric direct multi-step estimation for forecasting economic processes. *International Journal of Forecasting 21*(2), 201–218.

Chow, G. C. and A.-L. Lin (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 372–375.

Clark, T. E. and M. W. McCracken (2010). Averaging forecasts from VARs with uncertain instabilities. *Journal of Applied Econometrics 25*(1), 5–29.

Clements, M. P. (2014). Real-time factor model forecasting and the effects of instability. University of Reading, Mimeo.

Clements, M. P. and A. B. Galvão (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. *Journal of Business & Economic Statistics 26*(4), 546–554.

Clements, M. P. and A. B. Galvão (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics 24*(7), 1187–1206.

Diron, M. (2008). Short-term forecasts of euro area real GDP growth: an assessment of real-time performance based on vintage data. *Journal of Forecasting 27*(5), 371–390.

Drechsel, K. and R. Scheufele (2012a). Bottom-up or Direct? Forecasting German GDP in a Data-rich Environment. Swiss National Bank Working Papers, No 2012-16.

Drechsel, K. and R. Scheufele (2012b). The performance of short-term forecasts of the German economy before and during the 2008/2009 recession. *International Journal of Forecasting 28*(2), 428–445.

Duarte, C. (2014). Autoregressive Augmentation of MIDAS Regressions. Bank of Portugal Working Papers 2014/1.

ECB (2008). Short-term forecasts of economic activity in the euro area. *ECB Monthly Bulletin April 2008*, 69–74.

Ferrara, L., C. Marsilli, and J.-P. Ortega (2014). Forecasting growth during the great recession: is financial volatility the missing ingredient? *Economic Modelling 36*, 44–50.

Foroni, C. and M. Marcellino (2013). A survey of econometric methods for mixed-frequency data. Norges Bank Working Paper 2013/06.

Foroni, C. and M. Marcellino (2014). A comparison of mixed frequency approaches for nowcasting euro area macroeconomic aggregates. *International Journal of Forecasting 30*(3), 554–568.

Foroni, C., M. Marcellino, and C. Schumacher (2014). U-MIDAS: MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society A*, forthcoming.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics 131*(1), 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Giacomini, R. and B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics 25*(4), 595–620.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Golinelli, R. and G. Parigi (2007). The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries. *Journal of Forecasting 26*(2), 77–94.

Hahn, E. and F. Skudelny (2008). Early estimates of euro area real GDP growth: a bottom up approach from the production side. ECB Working Paper 975.

Hendry, D. F. and M. P. Clements (2004). Pooling of forecasts. *The Econometrics Journal 7*(1), 1–31.

Ingenito, R. and B. Trehan (1996). Using monthly data to predict quarterly output. *Federak Reserve Bank of San Francisco Economic Review 3*, 3–11.

Kuzin, V., M. Marcellino, and C. Schumacher (2011). MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the Euro Area. *International Journal of Forecasting 27*(2), 529–542.

Kuzin, V., M. Marcellino, and C. Schumacher (2013). Pooling versus model selection for nowcasting GDP with many predictors: Empirical evidence for six industrialized countries. *Journal of Applied Econometrics 28*(3), 392–411.

Lahiri, K. and G. Monokroussos (2013). Nowcasting US GDP: The role of ISM business surveys. *International Journal of Forecasting 29*(4), 644–658.

Lütkepohl, H. (1981). A model for non-negative and non-positive distributed lag functions. *Journal of Econometrics 16*(2), 211–219.

Marcellino, M., J. H. Stock, and M. W. Watson (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics 135*(1), 499–526.

Rossi, B. and T. Sekhposyan (2010). Have economic models forecasting performance for US output growth and inflation changed over time, and when? *International Journal of Forecasting 26*(4), 808–835.

Rünstler, G., K. Barhoumi, S. Benk, R. Cristadoro, A. Den Reijer, A. Jakaitiene, P. Jelonek, A. Rua, K. Ruth, and C. Van Nieuwenhuyze (2009). Short-term forecasting of GDP using large datasets: a pseudo real-time forecast evaluation exercise. *Journal of Forecasting 28*(7), 595–611.

Schumacher, C. and 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*(3), 386–398.

Stock, J. H. and M. W. Watson (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Stock, J. H. and M. W. Watson (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics 30*(4), 481–493.

Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting 1*, 135–196.