International Journal of Forecasting 32 (2016) 257–270

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

A comparison of MIDAS and bridge equations ☆

Christian Schumacher

Deutsche Bundesbank, Economic Research Centre, Wilhelm-Epstein-Str. 14, 60431 Frankfurt, Germany

ARTICLE INFO

Keywords:
Mixed-data sampling
Bridge equations
Nowcasting

ABSTRACT

This paper compares two single-equation approaches from the recent nowcasting literature: mixed-data sampling (MIDAS) regressions and bridge equations. Both approaches are suitable for nowcasting low-frequency variables such as the quarterly GDP using higher-frequency business cycle indicators. Three differences between the approaches are identified: (1) MIDAS is a direct multi-step nowcasting tool, whereas bridge equations provide iterated forecasts; (2) the weighting of high-frequency predictor observations in MIDAS is based on functional lag polynomials, whereas the bridge equation weights are fixed partly by time aggregation; (3) for parameter estimation, the MIDAS equations consider current-quarter leads of high-frequency indicators, whereas bridge equations typically do not. To assist in discussing 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.

1. Introduction

In policy institutions such as central banks, nowcasting GDP growth is an important way of informing decision makers about the current state of the economy. Nowcasting models typically consider specific data irregularities: whereas GDP is sampled at a quarterly frequency and only with a considerable delay, many business cycle indicators are available at higher frequencies and in a more timely fashion; for example, monthly industrial production or high-frequency financial data. Policy analysts want to exploit this data for nowcasting in the most efficient way possible 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 related directly 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 ensure that the number of parameters to be estimated remains small. In this case, non-linear least squares (NLS) is used for parameter estimation, as outlined by Ghysels, Sinko, and Valkanov (2007). If the difference in sampling frequencies between the explained low-frequency variable and the high-frequency predictors is not too large (for example, given quarterly and monthly data), unrestricted linear polynomials have been considered in the literature as well, by Foroni, Marcellino, and Schumacher (2015). These polynomials can be estimated by ordinary least squares (OLS). Whereas MIDAS has been used initially for financial applications, by Ghysels, Santa-Clara, and Valkanov (2005, 2006) for example, it has been employed recently in many applications as a macroeconomic forecasting tool for quarterly GDP, starting with Clements and Galvão (2008, 2009). Recent contributions include those of Andreou, Ghysels, and Kourtellos (2013), Duarte

☆ This paper represents the author's personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.

E-mail address: christian.schumacher@bundesbank.de.

http://dx.doi.org/10.1016/j.ijforecast.2015.07.004

(2014), Drechsel and Scheufele (2012a), Ferrara, Marsilli, and Ortega (2014), Foroni et al. (2015), and Kuzin, Marcellino, and Schumacher (2011), amongst others. Recent surveys include those by Andreou, Ghysels, and Kourtellos (2011) and Armesto, Engemann, and Owyang (2010).

Bridge equations are also dynamic, but the variables on both sides of the equation are low-frequency variables. In particular, for nowcasting the quarterly GDP, the explanatory variables on the right-hand side of the equation are quarterly lags of the predictor. These quarterly observations are typically obtained from time aggregation of the high-frequency observations of the predictor. 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, such as an autoregressive (AR) model. The high-frequency forecasts from this model are then aggregated over time to the quarterly frequency and plugged into the bridge equation. Due to this simple estimation method and their transparency, bridge equations are used widely in policy organizations, and central banks in particular. Applications in the literature include those by Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011), Baffigi, Golinelli, and Parigi (2004), Bulligan, Golinelli, and Parigi (2010), Bulligan, Marcellino, and Venditti (2015), Camacho, Perez-Quiros, and Poncela (2014), Diron (2008), Foroni and Marcellino (2013), Foroni and Marcellino (2014), Golinelli and Parigi (2007), Hahn and Skudelny (2008), Ingenito and Trehan (1996), and Rünstler et al. (2009), amongst others. Applications of bridge equations to nowcasting in central banks are documented by 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 nowcasting tools is investigated in detail. In the literature, a few comparisons of the two approaches can be found, see for example Foroni and Marcellino (2013). This paper expands on this body of literature by providing analytical results to explain the differences between MIDAS and bridge equations. This is possible because MIDAS and bridge equations 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 an empirical weighting of high-frequency predictor observations, often based on functional lag polynomials, whereas bridge equations are based partly on fixed weights stemming from statistical time-aggregation rules. The different weighting schemes also imply different estimation methods, namely OLS for bridge equations and unrestricted MIDAS polynomials, but NLS for MIDAS equations based on non-linear functional lag polynomials. (3) Finally, MIDAS can consider current-quarter observations of the high-frequency indicator in the mixed-frequency equation, whereas bridge equations typically contain only contemporaneous or lagged observations of the indicator.

To assess the influences 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 in its use of a different weighting scheme for the high-frequency observations on the right-hand side, and from standard MIDAS in its iterative solution of the model for nowcasting. Further model variants arise from different assumptions regarding leading terms of the indicators. Highlighting the differences between the approaches could help practitioners in making 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.

To illustrate the differences between MIDAS and bridge equations, the two are compared in an empirical nowcasting exercise for Euro area GDP, where the evaluation period covers the years following the Great Recession. The predictor set comprises a large number of monthly indicators. Different specifications of MIDAS and bridge equations with single indicators are evaluated.

The paper proceeds as follows: Section 2 describes the MIDAS and bridge equations and how they can be used for nowcasting. Section 3 provides the analytical comparison of MIDAS and bridge equations, and discusses alternative models that link the two core approaches. In Section 4, the results of the empirical nowcasting exercise are discussed. Section 5 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 $t = 1, 2, \dots, T_y$, with T_y being the final quarter for which GDP data are available. The aim is to nowcast or forecast the GDP for period $T_y + h$, yielding a value for y_{T_y+h} with horizon $h = 1, \dots, H$ 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 available only with a delay of two quarters. In April, for example, the Euro area GDP is only available for the fourth quarter of the previous year, and a nowcast for the 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 the current, namely second, quarter GDP in April, the horizon has to be set sufficiently large. Further information and details on nowcasting procedures can be found in the survey by Banbura, Giannone, and Reichlin (2011).

In this paper, for simplicity, it is assumed that the information set for now- and forecasting includes one stationary monthly indicator x_t^M in addition to the available GDP observations. The time index for monthly observations is defined as a fraction of the low-frequency quarter according to $t = 1 - 2/3, 1 - 1/3, 1, 2 - 2/3, \dots, T_x - 1/3, T_x$, where T_x is the final month for which the indicator is available, as per Clements and Galvão (2008) and Ghysels et al. (2007). 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. We define the number of months earlier than GDP that the indicator is available by $w = T_x - T_y$. For simplicity, we might call w just the lead of the high-frequency indicator, as per Andreou et al. (2013). The now- or forecast for GDP is denoted as the conditional expectation $y_{T_y+h|T_x}$, which is conditional on information available in month T_x, including all monthly indicator observations until T_x and the GDP observations up to T_y.

2.1. MIDAS regressions

The mixed-data sampling (MIDAS) approach, proposed by Clements and Galvão (2008) and Ghysels et al. (2007), is a direct nowcasting tool, where the dynamics of the indicator are not modelled explicitly. The MIDAS equation for GDP growth y_{t+h} in period $t+h$ with a forecast horizon of 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 the lead of the high-frequency indicator $w = T_x - T_y$ indicates that we have more observations of the indicator than of GDP if $w > 0$, following 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 available early relative to GDP.

The MIDAS equation (Eq. (1)) contains a constant β_0 and an autoregressive term λy_t. The choice of one lag was made in order to keep the model's exposition simple. In general, it is easy to consider more autoregressive lags, as per Andreou et al. (2013) and Duarte (2014), and different specifications will be compared in the empirical application.

The effect of the monthly indicator x^M_{t+w} on y_{t+h} is determined by the high-frequency lag polynomial $\beta_1 B(L^{1/3};\theta)$, consisting of the regression coefficient β_1 and the polynomial $B(L^{1/3};\theta)$, defined as

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

which is a function of the high-frequency (monthly) lag operator $L^{1/3}x^M_t = x^M_{t-1/3}$. In the MIDAS literature, functional lag polynomials are chosen typically for $B(L^{1/3};\theta)$, to avoid parameter proliferation for long high-frequency lags K, as proposed by Ghysels et al. (2007). A popular functional form of the polynomial is the exponential Almon lag

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

with q shape parameters. One popular specification is $q = 2$ with $\theta = [\theta_1, \theta_2]$, defined by Lütkepohl (1981) and adopted for MIDAS by Ghysels et al. (2007). For a given θ, the lag function provides a parsimonious way of considering a large number of K monthly lags of the indicators. With $q = 2$, the resulting functional form is typically unimodal, and can be hump-shaped, declining or flat, as was discussed by Ghysels et al. (2007). For forecasting, Andreou et al. (2013) propose $q = 1$, which implies declining weights. 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 in the current quarter are included in the conditioning set of the projection. For example, if one wants to nowcast the second quarter GDP, and industrial production is available for April and GDP for the first quarter, then the lead is $w = 1/3$. The specification of Eq. (1) does imply the projection in Eq. (4), where the April observation is part of the conditioning set.

The MIDAS equation in Eq. (1) should be regarded as a stylized example from the large set of possible MIDAS specifications that have been discussed in the literature to date. For example, there is a large variety of alternative lag polynomials in addition to the exponential Almon in Eq. (3).

For mixing quarterly and monthly data, Foroni et al. (2015) discuss the unrestricted lag polynomial, defined as

$$B(L^{1/3}) = \sum_{k=0}^{K} b_k L^{k/3}, \tag{5}$$

which replaces $\beta_1 B(L^{1/3};\theta)$ in the basic MIDAS specification in Eq. (1). The parameters b_k for $k = 0, 1, \ldots, K$ can be estimated by OLS. Thus, this variant of MIDAS with unrestricted lag polynomials, which will be abbreviated as U-MIDAS hereafter, necessitates the estimation of $K + 1$ polynomial parameters, whereas the functional lag in Eq. (3) requires three parameters to be estimated. Hence, standard MIDAS equations with functional lag polynomials are more parsimonious, whereas U-MIDAS is more flexible.

Another weighting scheme, known as multiplicative MIDAS, or M-MIDAS in brief, relates MIDAS to the time aggregation literature. Following Bai, Ghysels, and Wright (2013) and Chen and Ghysels (2011), the M-MIDAS equation is defined as

$$y_{t+h} = \beta_0 + \lambda y_t + \beta(L)x^Q_{t+1}(\theta) + \varepsilon_{t+h}, \tag{6}$$

with the quarterly polynomial $\beta(L) = \sum_{i=0}^{p}\beta_{i+1}L^i$. The lag operator in $\beta(L)$ is the quarterly lag operator defined by $Lx^Q_t(\theta) = x^Q_{t-1}(\theta)$. The variable $x^Q_t(\theta)$ is a quarterly aggregated quantity derived from the monthly indicator observations by

$$x^Q_t(\theta) = \sum_{k=0}^{2} b(k;\theta)L^{k/3}x^M_t, \quad \text{for } t = 1, 2, \ldots, T_q, \tag{7}$$

$$x^Q_t(\theta) = \sum_{k=0}^{w-1} b((3-w)+k;\theta)L^{k/3}x^M_{t+w}, \quad \text{for } t = T_{q+1}. \tag{8}$$

Time aggregation in Eq. (7) provides a quarterly indicator value from the three monthly values in a quarter for

$t = 1, 2, \ldots, T_q$, see Andreou et al. (2013). The form of the time-aggregation function $b(k;\theta)$ can be either functional, as in Eq. (3) for example, or unrestricted, as in Eq. (5). M-MIDAS assumes that the intra-quarter monthly response of GDP to a change in the indicator is equal across quarters, scaled by the quarterly lag coefficients in $\beta(L)$. It was originally suggested by Chen and Ghysels (2011) for handling seasonal patterns; in their case, the intra-daily seasonality of volatility patterns. In the case of the monthly and quarterly data discussed here, M-MIDAS implies a seasonal pattern within the quarter. The time aggregation in Eq. (8) applies to the monthly indicator values that are available in quarter $T_q + 1$. Here, we assume that the monthly indicator leads the GDP by up to three months, $0 < w \leq 3$. Longer leads can be accounted for in a similar way. The shape parameters in θ, the quarterly polynomial parameters in $\beta(L)$, and the rest of the parameters can be estimated jointly by NLS if a functional polynomial has been chosen.

Other weighting functions include the beta distribution proposed by Ghysels et al. (2007), the non-exponential Almon lag of Drechsel and Scheufele (2012a), and penalized changes in weights, see Breitung, Elengikal, and Roling (2013). An overview is provided in the MIDAS toolbox manual (Ghysels, 2014).

2.2. Bridge equation

Following the recent papers by Bulligan et al. (2015) or Foroni and Marcellino (2013), a bridge equation with a single indicator can be defined as

$$y_t = \beta_0 + \lambda y_{t-1} + \beta(L)x_t^Q + \varepsilon_t, \tag{9}$$

where y_t is again the 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 that is available for $t = 1, 2, \ldots, T_y$, the same periods as the GDP. The quarterly 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 that is aggregated over time from the monthly indicator x_t^M to the quarterly frequency. Following the time aggregation literature (e.g., Chow & Lin, 1971), the mapping from the high-frequency indicator observations to the aggregated low-frequency observations is formalized through the deterministic aggregator function $\omega(L^{1/3})$ in the lag operator $L^{1/3}$ by

$$x_t^Q = \omega(L^{1/3})x_t^M = \sum_{j=0}^{r} \omega_j L^{j/3} x_t^M. \tag{10}$$

The exact form of the deterministic 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

$$\begin{aligned} x_t^Q &= \omega(L^{1/3})x_t^M = x_t^M + 2x_{t-1/3}^M + 3x_{t-2/3}^M \\ &\quad + 2x_{t-1}^M + x_{t-4/3}^M, \end{aligned} \tag{11}$$

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).

Note that, at the end of the sample, the filter in Eq. (10) can be applied only to complete quarters. Thus, in order to apply the aggregator function in Eq. (10) properly, the practitioner has to wait until all three monthly observations of the indicator corresponding to the calendar quarter are available. 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 in Eq. (9).

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

$$\widehat{y}_{T_y+h|T_x} = \widehat{\beta}_0 + \widehat{\lambda}\widehat{y}_{T_y+h-1|T_x} + \widehat{\beta}(L)\widehat{x}_{T_y+h|T_x}^Q. \tag{12}$$

Due to the AR(1) term, the forecast equation has to be solved forward in order 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, involving a forecast step for the monthly predictor, and a time aggregation step to obtain the quarterly projection. In the literature, it is typical for very simple univariate models to be chosen to predict the high-frequency indicator. For example, a simple AR forecast equation for x_t^M, such as

$$x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M, \tag{13}$$

can be used for $t = 1/3, 2/3, \ldots, T_x - 1/3, T_x$, where $\alpha(L^{1/3}) = \sum_{j=1}^{r} \alpha_j L^{(j-1)/3}$ is a polynomial in the monthly lag operator $L^{1/3}$, such that x_t^M enters Eq. (13) with a maximum lag length of $r/3$ quarters. The equation can be estimated by OLS and solved iteratively to obtain monthly forecasts $\widehat{x}_{T_y+h|T_x}^M$. Note that Eq. (13) accounts for all monthly observations that are available up to period T_x, and can thus condition on more timely information than the bridge equation in Eq. (9).

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 Eq. (12), yielding the GDP growth forecast. To illustrate the bridge nowcast, consider again the example discussed above for MIDAS. The aim is to nowcast the second quarter GDP. Industrial production is available for April and GDP for the first quarter. The bridge equation can be estimated on quarterly data only, in this case using data up to the first quarter. Both the estimation of the AR model in Eq. (13) and the indicator forecast consider the latest indicator information up to April. The iterative solution of Eq. (13) provides values for industrial production for May and June. The indicator projections for May and June enter the time-aggregation rule Eq. (10) together with the observations for April and earlier. Finally, plugging the time-aggregated indicator projection into Eq. (12) yields the GDP nowcast. Note that recursive solutions are necessary in order to forecast more than one period ahead, due to the AR term in the bridge equation.

3. Differences between MIDAS and bridge equations, and an intermediate model

The two approaches differ from each other in several ways. Fortunately, both of the models can be regarded as extensions of distributed lag models to mixed-frequency data, which allows us to isolate and highlight their main differences. For this purpose, note that all of the MIDAS approaches described above are functions of the monthly indicator observations x_t^M. We can also write the bridge equation as a function of x_t^M by using the time-aggregation scheme $\omega(L^{1/3})$ defined in Eq. (10) according to $\beta(L)x_t^Q = \beta(L)\omega(L^{1/3})x_t^M$. Now, we can write both MIDAS and bridge equations as functions of x_t^M and compare them:

$$\text{MIDAS: } y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B(L^{1/3};\theta)x_{t+w}^M + \varepsilon_{t+h}, \quad (14)$$

$$\text{bridge: } y_t = \beta_0 + \lambda y_{t-1} + \beta(L)\omega(L^{1/3})x_t^M + \varepsilon_t, \quad (15)$$

$$\text{bridge indicator: } x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M. \quad (16)$$

The MIDAS equation is still based on the basic exponential Almon polynomial, but can be modified easily to the U- and M-MIDAS polynomials. Below, we identify three major differences between MIDAS and bridge approaches.

3.1. Direct versus iterative multi-step forecasting

MIDAS is a direct multi-step forecasting device, in the sense that the left-hand side of Eq. (14), y_{t+h}, refers directly to the period $t+h$, whereas, due to lags, the right-hand side predictors x_{t+w}^M refer to period $t+w$ or earlier periods. This specification allows for the production of a projection for horizon $h \geq 1$ in one step, based on a single equation without any iterative model solution, as was 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 respecified and the MIDAS equation estimated again for each h.

The bridge equation, on the other hand, implies an iterative multi-step forecast, in the sense that the model of the high-frequency indicator (Eq. (16)) is solved forward to produce indicator forecasts over all horizons, which are plugged into the bridge equation (Eq. (15)). The bridge equation specifies y_t as a function of x_t^Q and lags, without considering the forecast horizon, as per MIDAS. Thus, it is estimated only once.

There is a long-running discussion in the literature of the relative advantages of direct and iterative multi-step forecasting. Chevillon and Hendry (2005) and Marcellino, Stock, and Watson (2006) made recent contributions, and Bhansali (2002) provided an early survey. The literature shows that there are arguments in favor of both approaches. In general, though, the direct approach is most advantageous in the case of misspecification. If the model used for iterative forecasts is specified correctly, it should outperform the direct approach.

3.2. Use of end-of-sample monthly data

The MIDAS equation in Eq. (14) considers w leads of the high-frequency indicator on the right-hand side. 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 in Eq. (16). 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, MIDAS and bridge equations both use the same full-information conditioning set with observations up to T_x, including current-quarter monthly observations.

However, as the bridge equation in Eq. (15) is a low-frequency (quarterly) equation, it can only be updated if the indicator data for the most recent low-frequency period are fully available, which implies that the bridge equation can only be re-estimated once a quarter. Thus, the lead of the indicator w is not considered as in the MIDAS approach. Note again that this argument affects only the estimation of parameters in the MIDAS and bridge equations; the conditioning set in the nowcasts is the same as that outlined in the previous paragraph for both approaches.

3.3. Differences in the functional form of the polynomials

Another key difference between the MIDAS equation in Eq. (14) and the bridge equation in Eq. (15) relates to the lag polynomials of the indicator. To make Eqs. (14) and (15) comparable with respect to the polynomials, assume $w = 0$ and $h = 0$ in the MIDAS equation (Eq. (14)). Apart from the polynomials, this makes Eq. (14) equal to the bridge equation (Eq. (15)). Furthermore, in order for the two polynomials to be comparable, we have to assume the same order of the polynomials in both the MIDAS and bridge equations. Let's compare the shape and parametric parsimony of the bridge polynomial with those of the different MIDAS polynomials:

1. **U-MIDAS**: If the MIDAS regression is based on unrestricted lag polynomials (Eq. (5)), one can see immediately that U-MIDAS actually nests the bridge equation. The U-MIDAS parameters in the polynomial $\sum_{k=0}^{K} b_k L^{k/3}$ are estimated freely, whereas the bridge polynomial $\beta(L)\omega(L^{1/3})$ is restricted by the deterministic time-aggregation function $\omega(L^{1/3})$. Thus, the U-MIDAS polynomial allows for a better approximation to any true polynomial. However, this generality of U-MIDAS comes at the cost of having to estimate $K+1$ parameters, whereas the bridge approach requires the estimation of only $p+1$ parameters.[1]
2. **M-MIDAS**: If the multiplicative MIDAS is chosen, one can still see directly that the multiplicative polynomial $\beta(L)x_t^Q(\theta)$ nests the bridge equation polynomial $\beta(L)x_t^Q$, because the time aggregation is based on empirical weights, where the shape is determined by θ,

[1] For example, for $p = 3$ quarterly bridge lags and a monthly time aggregation lag order of $r = 4$, we have $K = 16 = 3(p+1) + r$ parameters to be estimated for MIDAS, assuming it to have polynomial orders the same as those defined above.

whereas the bridge polynomial is restricted more by the fixed time-aggregation rule $\omega(L^{1/3})$. A necessary condition in order for this to hold is that the time-aggregation polynomials in M-MIDAS and the bridge equation approach must have the same length.[2] M-MIDAS requires $(p+1)+2$ parameters to be estimated if an exponential Almon lag function with two shape parameters is used for time aggregation in Eqs. (7) and (8). If we are using unrestricted polynomials, we have $(p+1)+3$ parameters. On the other hand, the bridge approach requires the estimation of only $p+1$ lag parameters, and is therefore slightly more parsimonious.

3. **Exponential Almon**: The exponential Almon lag polynomial (Eq. (3)) in MIDAS implies empirically estimated weights $\beta_1 B(L^{1/3};\theta)$. The bridge polynomial $\beta(L)\omega(L^{1/3})x_t^M$ is based partly on fixed time aggregation. Hence, the bridge weights are not estimated freely by the data, unlike with MIDAS. However, the functional form of the Exponential Almon lags with up to K lags does not generally nest the bridge equation polynomial. Only if the order of the lag polynomial K is equal to the discrepancy in sampling frequencies minus one does MIDAS with exponential Almon lags also nest the bridge equation approach.[3] In general, though, it is not clear whether or not the MIDAS weights in Eq. (3) are more flexible, because the functional form of the polynomial with a small number of parameters is only an approximation to any true polynomial with positive weights, as was discussed by Lütkepohl (1981). The exponential Almon lag polynomial requires the estimation of three parameters, namely β_1, θ_1, and θ_2. The bridge equation polynomial requires the estimation of $p+1$ parameters and depends on the maximum lag order of the polynomial. Thus, MIDAS is more parsimonious than the bridge approach for increasing numbers of lags. However, in macro applications with quarterly and monthly data, this advantage might not be huge.[4]

Summarizing the arguments, we can see that MIDAS polynomials are generally more flexible than bridge polynomials, as we can derive conditions under which the empirically estimated polynomials in MIDAS nest the bridge polynomial based on fixed time aggregation. The generality of the MIDAS approach is also documented by Bai et al. (2013), who show how closely MIDAS polynomial weights can approximate the Kalman filter weights of a mixed-frequency state space model. What are the implications of the more flexible form of the MIDAS polynomials compared to fixed time-aggregation weights in the bridge approach? The literature on time aggregation is relatively straightforward, and argues that time aggregation with fixed weights typically leads to a loss of information, see Marcellino (1999), or more recently Hassler (2014) and Hassler and Tsai (2013).

The distinction between flexible and fixed weights for time aggregation also has implications for parameter estimation. Andreou et al. (2010) study the asymptotic properties of the MIDAS NLS estimator based on exponential Almon lags with two shape parameters and the OLS estimator for a model with an equal-weight time-aggregation lag function. If the sample sizes of the high- and low-frequency variables go to infinity, in our case $T_y, T_x \to \infty$, they show that the OLS estimator based on the equal-weights polynomial is always less efficient than the MIDAS NLS estimator based on the exponential Almon polynomial. If the high-frequency regressor follows an AR(1) process, they also show that the equal-weights OLS estimator is both asymptotically biased and inefficient.

In practice, however, it is not clear whether MIDAS or the bridge polynomials will do best in an out-of-sample forecast environment with small samples and an uncertain DGP. If the bridge time-aggregation function $\omega(L^{1/3})$ and the lag polynomial are close to the true polynomial in the data, fixing the weights might avoid estimation uncertainty. In small samples, the cost of the greater flexibility of some MIDAS polynomials is the larger number of parameters that require estimation. In the discussion of the polynomials above, it turned out that the U-MIDAS and M-MIDAS polynomials require the estimation of more parameters than the bridge equation, whereas the exponential Almon lag polynomial is more parsimonious.

Note also that the estimation methods for MIDAS and bridge equations differ with respect to the functional forms of the lag polynomials. Non-linear functional lag polynomials in MIDAS necessitate NLS estimation, which works iteratively and might depend on the choice of starting values. The coefficients in the bridge equation and U-MIDAS are simply estimated by OLS.

3.4. A model in between MIDAS and bridge: iterative MIDAS

For practical purposes, it might be interesting to determine which of the differences discussed in the previous section matter the most for nowcasting. To check this, some of the differing modelling elements in MIDAS and bridge equations will be switched off and on in turn in order to isolate their effects. This provides us with intermediate models that contain selected aspects of MIDAS and bridge equations.

The first model is iterative MIDAS, or MIDAS-IT. It consists of the following equations:

$$\text{MIDAS-IT:}\ y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x_{t+w}^M + \varepsilon_t, \quad (17)$$

$$\text{MIDAS-IT indicator:}\ x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M. \quad (18)$$

On the left-hand side, we have GDP growth, which is t-dated, whereas the regressor on the right-hand side,

[2] The lag order in the M-MIDAS time-aggregation polynomial Eq. (7) above is two, namely the frequency mismatch minus one, in line with the literature, see Andreou et al. (2013). In practice, the time aggregation polynomial of a flow variable used in bridge equations has the same order; however, the famous month-on-month to quarter-on-quarter time-aggregation function Eq. (11) has an order of four. In this case, M-MIDAS can nest the bridge approach only if the polynomial orders have been adjusted.

[3] This case has been discussed extensively by Andreou, Ghysels, and Kourtellos (2010). They show that the use of fixed equal time-aggregation weights can lead to an omitted variable bias compared to MIDAS. More details on the consequences of equal-weight time aggregation for parameter estimation are provided below.

[4] For example, for a quarterly polynomial of order four, $p=4$, we have to estimate $5=p+1$ parameters in the bridge equation, compared to three in MIDAS based on exponential Almon lags.

x^M_{t+w}, contains current-quarter information up to period $t+w$, including the lead. The MIDAS equation is based on the exponential Almon lag, but can easily be extended to other polynomials, such as the U-MIDAS and M-MIDAS polynomials as discussed before. The nowcast of GDP growth for period $T_y + h$ given information up to time T_x is

$$y_{T_y+h|T_x} = \widehat{\beta}_0 + \widehat{\lambda} y_{T_y+h-1|T_x} + \widehat{\beta}_1 B(L^{1/3}; \widehat{\theta}) x^M_{T_y+h+w|T_x}. \quad (19)$$

This requires a projection $x^M_{T_y+h+w|T_x}$ of the indicator, which is provided by the iterative forward solution of the AR model in Eq. (18). From here on, the model consisting of Eq. (17) and the AR model is called MIDAS-IT with leads. As an alternative, I also consider the MIDAS-IT model with $w = 0$ in Eq. (17), which will be referred to below as MIDAS-IT without leads.[5]

A comparison of the MIDAS-IT variants to the core MIDAS and bridge approaches helps to isolate the roles of the key model features:

1. Standard MIDAS is a direct multi-step device, whereas MIDAS-IT is iterative. In the regression of MIDAS-IT with leads, the leading monthly information is used in the same way as in MIDAS, and the polynomial is also the same. Thus, if we compare the nowcasting performances of the standard MIDAS and MIDAS-IT with leads, any differences can stem only from the difference between the direct and iterative nowcast procedures for multi-step nowcasting.
2. The bridge equation differs from MIDAS-IT without leads only in the choice of the polynomial. The indicator models are the same in both models, and there are no leads considered on the right-hand side. Thus, MIDAS-IT without leads and bridge equations differ only in the lag polynomials.
3. Finally, comparing MIDAS-IT with and without leads helps to answer the question of whether leads of the indicator should be used for estimating the MIDAS equation. The conditioning set for nowcasting is the same in both approaches, as the AR model provides the projections for the indicator.

By comparing the intermediate approaches with each other and with the basic approaches, we can assess how much each of the different model features discussed in previous sections matters for nowcasting. To finish off the discussion in this section, the full set of models and their features are summarized in Table 1.

4. Empirical application

4.1. Data

The dataset includes Euro area quarterly GDP growth from 1999Q1 until 2014Q4 and 107 monthly indicators until 2015M4. The monthly indicators cover industrial production by sector, surveys on consumer sentiment and business climate including the Purchasing Managers Index (PMI), international data, prices, and financial data. A complete list of the variables can be found in Appendix A. To make the MIDAS and bridge equation nowcasts comparable, both will be estimated and applied to out-of-sample nowcasting using the same indicator taken from the monthly dataset.

The dataset is a final dataset, not a real-time dataset, meaning that the influence of revisions on the relative forecasting accuracy cannot be discussed here. This is due in part to the fact that no large monthly real-time datasets with sufficiently long samples are available for the Euro area yet. Furthermore, it is not clear that any major changes can be expected from the use of real-time vintages, as per Bernanke and Boivin (2003), Clements (2015), and Schumacher and Breitung (2008). However, another important characteristic of multivariate data in real time is taken into account, namely the differences in the availabilities of variables due to publication lags. These differences in data availability lead to certain patterns of missing values at the end of each sub-sample, and recent papers find that accounting for this rather than using artificially balanced samples has a considerable impact on the forecast accuracy, see for example Giannone, Reichlin, and Small (2008). To consider the availability of the data at the end of each subsample, the nowcast exercise in this paper follows Foroni and Marcellino (2014) and Giannone et al. (2008), amongst others, and replicates the data availability from a final vintage of data in pseudo real-time.

4.2. Design of the empirical nowcast exercise

To evaluate the models' performances, 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 2010Q1 and 2014Q4, providing five years for comparison after the Great Recession. The recent nowcasting literature has observed that the Great Recession between 2008 and 2009 coincided with a strong decline in the predictability of GDP in European countries, see for example Drechsel and Scheufele (2012b) and Kuzin, Marcellino, and Schumacher (2013). Foroni and Marcellino (2014) and Foroni et al. (2015) also report strong differences in predictability in the Euro area before and during the Great Recession. Below, our main interest lies in the period following the Great Recession, between 2010Q1 to 2014Q4.

For each period in the evaluation sample, nowcasts are computed at the end of the third month, given the data available at that point in time. This nowcast, made at the end of the nowcast quarter, is informative for decision makers because the official GDP figure for the corresponding reference quarter will not be released until about six weeks later. Results for only one horizon are presented, in order to save space. The main results regarding the differences between the MIDAS and bridge equations do not differ much for the other horizons, and are thus left out; see also Andreou et al. (2013) and Kuzin et al. (2013).

[5] The MIDAS-IT model without AR terms and with $w = 0$ in the case when K equals the discrepancy in sampling frequencies minus one (equal to two for quarterly-monthly data) has been discussed by Andreou et al. (2010). Note that Andreou et al. (2010) also consider an AR(1) model for the high-frequency indicator.

Table 1
Overview of the MIDAS and bridge equation approaches.

	Equations for y_t and the indicator x_t^M	Direct or iterative multi-step nowcasting	Lag polynomials	Lead w 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	Unrestricted, multiplicative, or functional	Yes
B. MIDAS-IT, leads	$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x_{t+w}^M + \varepsilon_t$ $x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M$	Iterative	Unrestricted, multiplicative, or functional	Yes
C. MIDAS-IT, no leads	$y_t = \beta_0 + \lambda y_{t-1} + \beta_1 B(L^{1/3};\theta)x_t^M + \varepsilon_t$ $x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M$	Iterative	Unrestricted, multiplicative, or functional	No
D. Bridge	$y_t = \beta_0 + \lambda y_{t-1} + \beta(L)x_t^Q + \varepsilon_t$ $x_t^M = \alpha_0 + \alpha(L^{1/3})x_{t-1/3}^M + \varepsilon_t^M$ $x_t^Q = \omega(L^{1/3})x_t^M$	Iterative	Unrestricted quarterly polynomial and time aggregation function	No

Note: The model abbreviations and model details are explained in Sections 2 (MIDAS and bridge) and 3.4 (MIDAS-IT) of the text.

At each period in pseudo real-time, the estimation samples for MIDAS and the bridge equations depend on the information available. For example, due to their shorter publication lags, financial and survey indicators are available earlier than industry statistics. These differences in data availability lead to the patterns of missing observations found in the multivariate data – the so-called ragged edge – which can be replicated from the full sample of retrieved data, as was discussed in Section 4.1. To replicate the publication lags of the GDP, one can exploit the fact that, in the Euro area, the GDP of a given quarter becomes available in the middle of the second month of the next quarter.

4.3. Specification of MIDAS and bridge equations

Concerning the specification of MIDAS and bridge equations, either models with fixed specifications over time are used, or model specification is carried out recursively using information criteria (IC).

In the class of MIDAS models, the exponential Almon and U-MIDAS polynomials will be compared. With respect to exponential Almon lags, two specifications with $q = 1$, as per Andreou et al. (2013), and $q = 2$, as per Clements and Galvão (2009) and Ghysels et al. (2007), are considered. NLS estimation of the exponential Almon lag parameters $\theta = [\theta_1, \theta_2]$ in MIDAS is carried out under the restriction $\theta_2 \leq 0$, such that the polynomial is stable for $K \to \infty$ as per Ghysels et al. (2007). In the case of one parameter in the polynomial, $q = 1$, the restriction $\theta_1 \leq 0$ is imposed.

Autoregressive (AR) terms are specified in the same way for MIDAS and bridge equations. When using IC model selection, we consider up to four lags. The lag order used in each subsample is determined by the Bayesian Information Criterion (BIC). As alternatives, parsimonious model specifications without an AR term and with only one AR lag are considered, following Clements and Galvão (2008) and Drechsel and Scheufele (2012a), respectively.

The lag order of the indicator K in MIDAS based on exponential Almon lags is equal to $K = 12$. When using U-MIDAS polynomials and IC model selection, the maximum number of lags is also $K = 12$. As a parsimonious alternative, we also consider $K = 0$, leading to a nowcast based only on the latest monthly indicator observation, see Eq. (4). The maximum lag length of the indicator polynomial in the bridge equation is equal to $p = 4$ quarters, and a sparse alternative specification considers only $p = 0$. The AR model for the indicator has a maximum lag order of $r = 12$.

To compare the end-of-quarter nowcasts with the realizations of GDP growth, the mean squared error (MSE) is employed. In particular, relative MSEs will be reported that relate the MSEs of MIDAS to the MSEs of the bridge equations, but also to the MSEs of an AR benchmark model. The AR model can have up to four lags, specified using the BIC.

4.4. Results: MIDAS and bridge equations

In Table 2, the nowcast performances of the U-MIDAS and bridge equations are compared to that of the AR benchmark in Panel A. In Panel B, the MIDAS equations are compared to bridge equations pairwise. Useful models should outperform both the AR benchmark and the competing models. Panels A and B of Table 3 contain the MIDAS results based on the exponential Almon lag relative to the AR model and to the bridge equations, respectively.

Given the large numbers of indicators and model groups, each of the tables contains summary statistics for the distribution of relative MSEs from all of the models within a model class with different predictors, as per Foroni et al. (2015) and Stock and Watson (2012). For example, for the direct MIDAS model class, the MSE of MIDAS based on a single indicator is computed first and divided by the MSE of the AR benchmark. The same is done for all of the 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 that are smaller than one indicate that the performance of MIDAS or the bridge equations is superior to that of the benchmark. In the tables below, the medians and 10th and 90th percentiles of the relative MSE distributions are reported for each model class. When comparing MIDAS and bridge equations in Panel B, the relative MSE is computed as follows. First, the MSE of MIDAS based on a single indicator is divided by the MSE of the bridge equation based on the same indicator. Second, the same is done for all the other predictors, and we end up with a distribution of relative MSEs. The results are as follows.

Table 2
Percentiles of relative MSEs from U-MIDAS and bridge equations.

Model	Percentiles	Specification	(a)	(b)	(c)	(d)
		AR lags	1	–	IC	1
		K	0	0	0	IC
		p	0	0	0	IC
A. U-MIDAS and bridge relative to AR model						
	0.10		0.93	0.96	0.97	1.00
U-MIDAS	0.50		1.19	1.64	1.53	2.13
	0.90		2.08	3.04	4.32	3.62
	0.10		0.86	0.89	0.89	0.99
U-MIDAS-IT with leads	0.50		1.01	1.37	1.03	1.89
	0.90		1.44	2.27	3.22	3.84
	0.10		0.87	0.94	0.93	0.98
U-MIDAS-IT, no leads	0.50		1.03	1.38	1.10	2.10
	0.90		1.59	2.55	4.33	3.97
	0.10		0.87	0.87	0.93	1.00
Bridge	0.50		1.16	1.48	1.47	1.58
	0.90		2.14	2.98	4.44	2.98
B. U-MIDAS relative to the bridge model						
	0.10		0.74	0.78	0.85	0.75
U-MIDAS	0.50		1.05	1.01	1.64	1.23
	0.90		1.41	1.68	3.41	2.32
	0.10		0.54	0.54	0.81	0.70
U-MIDAS-IT with leads	0.50		0.86	0.86	1.36	1.13
	0.90		1.31	1.44	2.70	2.00
	0.10		0.64	0.63	0.86	0.78
U-MIDAS-IT, no leads	0.50		0.96	0.97	1.43	1.21
	0.90		1.18	1.15	3.12	2.13

Note: The model specifications with respect to the AR terms apply to U-MIDAS and bridge equations in the same way, and '–' denotes no AR lags. K refers to the monthly lags of the indicator in U-MIDAS, where $k = 0, \ldots, K$, see Eq. (5). p refers to the quarterly lags of the time-aggregated indicator in the bridge equation with $i = 0, \ldots, p$, see Eq. (9). The model abbreviations and model details are explained in Sections 2 (MIDAS and bridge) and 3.4 (MIDAS-IT) of the text.

1. Panel A of Table 2 shows the results of U-MIDAS and bridge equations relative to the AR benchmark. In all cases, the 50th and 90th percentiles are greater than one, implying that the majority of the models in each class perform worse than the AR benchmark. If we look at the 10th percentiles, we see that the best 10% of the models can often provide better performances than the benchmark. For example, in specification (a) with one AR lag and $K = 0$, the 10th percentile for U-MIDAS-IT with leads is about 0.86, followed closely by U-MIDAS-IT without leads and the bridge equation approach; whereas the direct U-MIDAS approach performs slightly worse, with a 10th percentile equal to 0.93. Among specifications (a)–(d), the sparse specifications with one AR lag and $K = p = 0$ in column (a) perform better than the IC-based model selection of AR and the indicator lags from columns (c) and (d). This result holds for both U-MIDAS and bridge equations.
2. Panel B of Table 2 shows the results of U-MIDAS relative to bridge equations. The 10th percentiles are smaller than one, whereas the 90th percentiles are greater than one, indicating that neither U-MIDAS nor bridge equations outperforms the other for all indicators. Instead, both model classes seem to have models that perform very well. When we look at the median results, the results depend on the specification. When choosing the model specifications using information criteria, U-MIDAS performs worse than the bridge approach, and the direct U-MIDAS approach also yields a 50th percentile that is greater than one for all specifications. On the other hand, the sparse specifications (a) and (b) of U-MIDAS-IT outperform the corresponding bridge equations, as is indicated by a 50th percentile of 0.86 for U-MIDAS-IT with leads. Note that only the sparse specifications (a) and (b) yield good models compared to the benchmark in Panel A. We can conclude that, overall, the U-MIDAS-IT equations perform better in this comparison than either the AR benchmark or the bridge equation. Among the U-MIDAS-IT models, the U-MIDAS-IT with leads specification performs slightly better than U-MIDAS-IT without leads.
3. Table 3 contains the MIDAS results based on the exponential Almon lag. With respect to the AR model in Panel A, the 50th and 90th percentiles are greater than one. The best 10% of MIDAS equations often outperform the benchmark, whereas the best 10% of the bridge equations cannot all perform better than the benchmarks. Of specifications (a)–(f), the sparse specifications with one AR lag or none do better than IC-based model selection, while the exponential Almon polynomial with only one shape parameter, $q = 1$ (from Andreou et al., 2013) in specifications (c), (d), and (f) tends to outperform the specification with $q = 2$.
4. Panel B of Table 3 shows the results of MIDAS based on the exponential Almon lag relative to bridge equations. Both model classes seem to include models that perform very well, as is indicated by fact that the 10th percentiles are almost always smaller than one, whereas

Table 3
Percentiles of relative MSEs from MIDAS based on exponential Almon lags and bridge equations.

Model	Percentiles	Specification						
			(a)	(b)	(c)	(d)	(e)	(f)
		AR lags	1	–	1	–	IC	IC
		q	2	2	1	1	2	1
		p	IC	IC	IC	IC	IC	IC
A. MIDAS and bridge relative to AR model								
	0.10		1.03	0.90	0.94	0.87	1.06	0.92
MIDAS	0.50		1.44	1.86	1.11	1.68	2.14	1.88
	0.90		2.57	4.08	2.11	3.42	4.43	4.43
	0.10		0.90	0.81	0.81	0.78	0.95	0.91
MIDAS-IT with leads	0.50		1.44	1.83	1.17	1.70	2.13	1.68
	0.90		2.60	3.52	1.94	3.18	4.17	4.10
	0.10		1.01	0.83	0.87	0.78	1.04	0.97
MIDAS-IT, no leads	0.50		1.35	2.01	1.10	1.70	1.96	1.77
	0.90		2.69	3.85	2.20	3.18	4.01	4.01
	0.10		1.00	1.01	1.00	1.01	1.00	1.00
Bridge	0.50		1.58	1.80	1.58	1.80	1.73	1.73
	0.90		2.98	4.30	2.98	4.30	4.02	4.02
B. MIDAS relative to bridge model								
	0.10		0.60	0.65	0.51	0.64	0.73	0.65
MIDAS	0.50		0.93	0.94	0.82	0.92	1.11	0.99
	0.90		1.47	1.60	1.33	1.27	2.48	2.18
	0.10		0.60	0.62	0.53	0.52	0.70	0.52
MIDAS-IT with leads	0.50		0.89	0.91	0.76	0.81	1.05	0.99
	0.90		1.45	1.62	1.14	1.28	2.12	2.29
	0.10		0.64	0.68	0.53	0.61	0.70	0.61
MIDAS-IT, no leads	0.50		0.89	0.94	0.77	0.90	1.00	0.96
	0.90		1.39	1.57	1.09	1.26	2.05	1.83

Note: The model specifications with respect to the AR terms apply to MIDAS and bridge equations in the same way, and '–' denotes no AR lags. p refers to the quarterly lags of the time-aggregated indicator in the bridge equation, with $i = 0, \ldots, p$, see Eq. (9). The exponential Almon lag polynomial in MIDAS is specified by the monthly indicator lag order with $K = 12$ and polynomial order q, see Eqs. (2) and (3), respectively. The model abbreviations and model details are explained in Sections 2 (MIDAS and bridge) and 3.4 (MIDAS-IT) of the text.

the 90th percentiles are always greater than one. However, the median results indicate that almost all of the MIDAS model variants outperform the bridge equation for the majority of model pairs, as the 90th percentile entries are smaller than one. There are no major differences among the MIDAS models.

Given these empirical findings, we can summarize the following results with respect to the differences between MIDAS and bridge equations identified in the conceptual comparison in Section 3:

1. The iterative multi-step approach of MIDAS-IT tends to perform better than the direct MIDAS.
2. There are no systematic differences depending on whether leads are considered in MIDAS-IT or not.
3. With respect to the polynomials, it turns out that very parsimonious specifications are necessary in order to outperform the benchmark. In this regard, MIDAS-IT with unrestricted or exponential Almon lag polynomials tends to outperform the bridge polynomials in most cases but not all.

4.5. Results: groups of indicators

So far, the nowcast results have referred to the entire distribution of models across the 107 indicators used in the nowcast equations. We now check whether there are any systematic differences in the relative performances of MIDAS and bridge equations between variable groups: industry, surveys and financial (see Appendix A for the variables collected in each of these groups). Table 4 contains the percentiles of the relative MSEs of MIDAS relative to bridge equations, computed in the same way as in Panel B in the previous tables. Panel A of Table 4 contains group-wise results for U-MIDAS, specification (a) from Table 2. Panel B of Table 4 contains group-wise results for MIDAS based on the exponential Almon lag, specification (c) from Table 3.

The results show that the majority of the relative MSEs of the iterative MIDAS variants U-MIDAS-IT and MIDAS-IT, each with leads, are smaller than one for all variable groups, indicating that iterative MIDAS equations mostly perform better than the bridge equation approach. The iterative MIDAS variants with and without leads do only slightly worse. The direct MIDAS specifications U-MIDAS and MIDAS both have 50th percentiles that are greater than one for the industry indicators. Overall, the results give the impression that the group-wise results are in line with those based on the full set of indicators from the previous tables, as there are no major differences between the indicator groups.

4.6. Results: pooling MIDAS and bridge equations

The forecasting literature, such as the papers by Clark and McCracken (2010), Hendry and Clements (2004), and

Table 4
Percentiles of relative MSEs from MIDAS relative to bridge equations for different indicator groups.

Model	Percentiles	Indicator group		
		Industry	Surveys	Financial
A. U-MIDAS relative to bridge model				
	0.10	0.48	0.80	0.81
U-MIDAS	0.50	1.10	1.05	1.08
	0.90	1.75	1.23	1.36
	0.10	0.30	0.64	0.59
U-MIDAS-IT with leads	0.50	0.94	0.92	0.82
	0.90	1.49	1.43	1.04
	0.10	0.31	0.79	0.64
U-MIDAS-IT, no leads	0.50	0.90	1.05	0.94
	0.90	1.17	1.19	1.00
B. MIDAS based on exponential Almon relative to bridge model				
	0.10	0.49	0.52	0.52
MIDAS	0.50	1.07	0.70	0.91
	0.90	1.86	0.99	1.61
	0.10	0.35	0.55	0.55
MIDAS-IT with leads	0.50	0.81	0.71	0.79
	0.90	1.43	1.09	1.16
	0.10	0.61	0.52	0.59
MIDAS-IT, no leads	0.50	0.87	0.70	0.89
	0.90	1.71	0.99	1.20

Note: The results in Panel A refer to the U-MIDAS specification (a) from Table 2. The results in Panel B refer to the exponential-Almon MIDAS specification (c) from Table 3. The model abbreviations and model details are explained in Sections 2 (MIDAS and bridge) and 3.4 (MIDAS-IT) of the text.

Timmermann (2006), 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, 2009) and Kuzin et al. (2013) considered combinations of MIDAS models, whereas Hahn and Skudelny (2008) and Rünstler et al. (2009), among others, pooled bridge equations. Below, we compare nowcast pooling results for MIDAS and bridge equations for three weighting schemes: the mean, the median, and the weighted mean of all 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, as per Kuzin et al. (2013). The model specifications used in the model combinations are based on a wide range of specification parameters, as per Clark and McCracken (2010): AR lags of up to four, $q = 1, 2, K = 0, 3, 6, \ldots, 12$, and $p = 0, \ldots, 4$. Table 5 shows the relative MSEs of the nowcast pools compared to the AR benchmark.

According to the results in the table, all of the combinations perform well relative to the AR benchmark, as the relative MSEs are clearly smaller than one. This good performance holds across all of the pooling schemes, with no major differences. With respect to the pooled model classes, the results indicate neither a clear advantage of either the bridge or MIDAS equation nowcasts, nor big differences within the MIDAS model class.

5. Summary and conclusions

MIDAS and bridge equations have each been used widely for nowcasting quarterly GDP growth based on monthly business cycle indicators both in the recent literature and in policy institutions. This paper compares the

Table 5
Relative MSEs of pooled MIDAS and bridge equations relative to the AR model.

Model	Weighting scheme		
	Mean	Median	MSE
A. U-MIDAS and bridge			
U-MIDAS	0.70	0.65	0.70
U-MIDAS-IT with leads	0.68	0.71	0.68
U-MIDAS-IT, no leads	0.68	0.68	0.63
Bridge	0.65	0.74	0.63
B. MIDAS based on exponential Almon and bridge			
MIDAS	0.70	0.70	0.68
MIDAS-IT with leads	0.72	0.72	0.67
MIDAS-IT, no leads	0.71	0.73	0.65
Bridge	0.65	0.74	0.63

Note: The MSE weights are based on rolling four-quarter windows computed in pseudo real-time. The model abbreviations and model details are explained in Sections 2 (MIDAS and bridge) and 3.4 (MIDAS-IT) of the text. The model specifications used for pooling within each model class are described in Section 4.6.

two approaches in a distributed-lag framework and identifies three conceptual differences: (1) MIDAS is a direct multi-step nowcasting 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. To isolate these differences, an iterative multi-step MIDAS approach is derived, building mainly on MIDAS polynomials and the iterative indicator model from the bridge equation approach.

In an empirical exercise for nowcasting Euro area GDP growth, the variants of MIDAS and bridge equations with single indicators are evaluated with respect to their out-of-sample nowcast accuracies. The Euro area indicators used for nowcasting include over a hundred indicators from industry statistics, surveys, and financial data. The evaluation period covers the period since the end of the Great Recession. It turns out that both single-indicator MIDAS and bridge equation models for many indicators over this period find the simple AR model tough to beat. Among the MIDAS and bridge equations, the most parsimonious specifications, with only a few AR and indicator lags, perform best. Based on these specifications, MIDAS, and iterative MIDAS in particular, tends to outperform bridge equations for the majority of the predictors. Pooled nowcasts obtained from single MIDAS or bridge equations perform similarly well compared to the benchmark. Of course, the results depend on the particular dataset and sample chosen. Other applications would require a careful check as to which model specification performs best. The set of alternative models discussed here, including the novel iterative MIDAS approach, offers a wide range of options to guide a practitioner's modeling decisions for this purpose.

Acknowledgments

I thank the handling editor, Dick van Dijk, and two anonymous referees for very helpful comments. Further useful comments were received from Claudia Foroni, Klemens Hauzenberger, Nicolas Pinkwart, Oke Röhe, and seminar participants at the Deutsche Bundesbank, KOF Zurich, and University of Cologne. I also thank Mark Kerßenfischer for excellent research assistance, as well as Katja Sonderhof and Laura Wichert for kindly providing the data. This paper represents the author's personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank or its staff.

Appendix 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 the Euro area contains 107 monthly indicators over the sample period from 1999M1 to 2015M4. A complete list of the variables is provided in Table A.1.

The sources of the time series are the databases of the Bundesbank and Data Insight. Natural logarithms were taken for all time series except for interest rates and the surveys. Stationarity was obtained by differencing the time series appropriately. All of the time series taken from these sources are already seasonally adjusted, wherever necessary.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.ijforecast.2015.07.004.

Table A.1
List of monthly indicators.

1	Industry	IP, total industry
2	Industry	IP, total ex constr
3	Industry	IP, manu
4	Industry	IP, constr
5	Industry	IP, constr build
6	Industry	IP, constr civ eng
7	Industry	IP, total ex constr and mig energy
8	Industry	IP, energy
9	Industry	IP, mig cap goods ind
10	Industry	IP, mig durable cons goods ind
11	Industry	IP, mig energy
12	Industry	IP, mig intermed goods ind
13	Industry	IP, mig non-durable cons goods ind
14	Industry	Retail, manu basic metals
15	Industry	Retail, manu chemicals
16	Industry	Retail, manu elect equipment
17	Industry	New passenger car registrations
18	Labour	Unemployment rate, total
19	Surveys	ESI, industrial confidence
20	Surveys	ESI, production trend recent months
21	Surveys	ESI, assessment of order-book levels
22	Surveys	ESI, assessment of export order-book levels
23	Surveys	ESI, assessment of stock of finished products
24	Surveys	ESI, production expectations for the months ahead
25	Surveys	ESI, selling price expectations for the months ahead
26	Surveys	ESI, employment expectations for the months ahead
27	Surveys	ESI, consumer confidence
28	Surveys	ESI, general economic situation last 12 months
29	Surveys	ESI, general econ situation next 12 months
30	Surveys	ESI, price trends over last 12 months
31	Surveys	ESI, price trends next 12 m
32	Surveys	ESI, unemployment expectations over next 12 m
33	Surveys	ESI, major purchases at present
34	Surveys	ESI, major purchases over next 12 m
35	Surveys	ESI, construction confidence
36	Surveys	ESI, trend of activity compared with preced. months

(continued on next page)

Table A.1 (*continued*)

37	Surveys	ESI, assessment of order books
38	Surveys	ESI, employment next 3 m
39	Surveys	ESI, selling price expectations for the months ahead
40	Surveys	ESI, retail confidence
41	Surveys	ESI, present business situation
42	Surveys	ESI, assessment of stocks
43	Surveys	ESI, expected business situation
44	Surveys	ESI, employment expectations
45	Surveys	ESI, services confidence
46	Surveys	ESI, business situation past 3 m
47	Surveys	ESI, demand past 3 m
48	Surveys	ESI, demand over next 3 m
49	Surveys	ESI, employment past 3 m
50	Surveys	ESI, employment next 3 m
51	Surveys	PMI, manufacturing headline index
52	Surveys	PMI, services headline index
53	Surveys	PMI, composite headline index
54	Surveys	PMI, composite output index
55	Surveys	PMI, composite new orders index
56	Surveys	PMI, composite employment index
57	Prices	PPI, total industry ex constr
58	Prices	PPI, mig energy
59	Prices	PPI, mig intermediate goods
60	Prices	PPI, mig non-dur cons goods
61	Prices	HICP, overall index
62	Prices	HICP, all-items ex energy and unpr food
63	International	Intra euro area export
64	International	Extra euro area export
65	International	Intra euro area import
66	International	Extra euro area import
67	International	US, PMI manufacturing index
68	International	US, unemployment rate
69	International	US IP, total ex construction
70	International	US, employment, civilian
71	International	US, retail trade
72	Financial	HWWI raw material price index
73	Financial	HWWI raw material price index, wo energy
74	Financial	HWWI raw material price index, crude oil
75	Financial	Gold price, US Dollar, fine ounce
76	Financial	Brent future
77	Financial	Coal
78	Financial	Copper
79	Financial	ECB nominal effective exch rate
80	Financial	ECB real effective exch rate cpi deflated
81	Financial	ECB real effective exch rate producer price defl.
82	Financial	Euro/US dollar exchange rate
83	Financial	Euro/British Pound exchange rate
84	Financial	Euro/Yen exchange rate
85	Financial	Euro Stoxx 50 Index
86	Financial	Euro Stoxx 50 Volatility Index
87	Financial	Standard & Poors 500 Index USA
88	Financial	Dow Jones Index US
89	Financial	Interest rate, loans
90	Financial	Interest rate, housing loans
91	Financial	Spreads corporate AAA and government bond maturities 1–3 years
92	Financial	Spreads corporate AAA and government bond maturities 7–10 years
93	Financial	Spreads corporate BBB and government bond maturities 1–3 years
94	Financial	Spreads corporate BBB and government bond maturities 7–10 years
95	Financial	Spreads corporate AAA and BBB ML maturities 1–3 years
96	Financial	Spreads corporate AAA and BBB ML maturities 7–10 years
97	Financial	Eonia
98	Financial	1-month interest rate, euribor
99	Financial	3-month interest rate, euribor
100	Financial	6-month interest rate, euribor
101	Financial	1-week interest rate, euribor
102	Financial	10-year government bond yield
103	Financial	Spread Euribor 1 year 1 month
104	Financial	Spread 10 year 3 month
105	Financial	Money m1
106	Financial	Money m2
107	Financial	Money m3

References

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

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

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

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

Armesto, M. T., Engemann, K. M., & Owyang, M. T. (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St. Louis Review*, *92*(6), 521–536.

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

Bai, J., Ghysels, E., & Wright, J. H. (2013). State space models and MIDAS regressions. *Econometric Reviews*, *32*(7), 779–813.

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

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

Bernanke, B. S., & Boivin, J. (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, & D. F. Hendry (Eds.), *A companion to economic forecasting* (pp. 206–221). Wiley-Blackwell.

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

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

Bulligan, G., Marcellino, M., & Venditti, F. (2015). Forecasting economic activity with targeted predictors. *International Journal of Forecasting*, *31*, 188–206.

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

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

Chen, X., & Ghysels, E. (2011). News-good or bad—and its impact on volatility predictions over multiple horizons. *Review of Financial Studies*, *24*(1), 46–81.

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

Chow, G. C., & Lin, A.-L. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, *53*, 372–375.

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

Clements, M. P. (2015). Real-time factor model forecasting and the effects of instability. *Computational Statistics and Data Analysis*, http://dx.doi.org/10.1016/j.csda.2015.01.011. forthcoming.

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

Clements, M. P., & Galvão, A. B. (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., & Scheufele, R. (2012a). Bottom-up or Direct? Forecasting German GDP in a data-rich environment. Swiss National Bank working papers, No. 2012-16.

Drechsel, K., & Scheufele, R. (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 (pp. 69–74).

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

Foroni, C., & Marcellino, M. (2013). A survey of econometric methods for mixed-frequency data. Norges Bank working paper 2013/06.

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

Foroni, C., Marcellino, M., & Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society, Series A*, *178*, 57–82.

Ghysels, E. (2014). Matlab toolbox for mixed sampling frequency data analysis using MIDAS regression models. University of North Carolina, Chapel Hill, Mimeo.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, *76*(3), 509–548.

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*(1), 59–95.

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

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

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

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

Hassler, U. (2014). Persistence under temporal aggregation and differencing. *Economics Letters*, *124*(2), 318–322.

Hassler, U., & Tsai, H. (2013). Asymptotic behavior of temporal aggregates in the frequency domain. *Journal of Time Series Econometrics*, *5*(1), 47–60.

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

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

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

Kuzin, V., Marcellino, M., & Schumacher, C. (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.

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

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business and Economic Statistics*, *17*(1), 129–136.

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

Rünstler, G., Barhoumi, K., Benk, S., Cristadoro, R., Den Reijer, A., Jakaitiene, A., Jelonek, P., Rua, A., Ruth, K., & Van Nieuwenhuyze, C. (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., & 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*(3), 386–398.

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

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

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

Christian Schumacher is Senior Economist at the Economics Research Center of the Deutsche Bundesbank. He is coordinator of the research group "Forecasting in Macroeconomics and Finance". In the Eurosystem, he is Bundesbank representative in the Working Group on Econometric Modelling (WGEM). His research interests include forecasting, nowcasting, dynamic factor models with large data sets, and forecast pooling.