

Density forecasts with MIDAS models*

Knut Are Aastveit[†]
Norges Bank

Claudia Foroni[‡]
Norges Bank

Francesco Ravazzolo[§]
Norges Bank and BI Norwegian Business School

October 23, 2014

Abstract

In this paper we derive a general parametric bootstrapping approach to compute density forecasts for various types of mixed-data sampling (MIDAS) regressions. We consider both classical and unrestricted MIDAS regressions with and without an autoregressive component. First, we compare the forecasting performance of the different MIDAS models in Monte Carlo simulation experiments. We find that the results in terms of point and density forecasts are coherent. Moreover, the results do not clearly indicate a superior performance of one of the models under scrutiny when the persistence of the low frequency variable is low. Some differences are instead more evident when the persistence is high, for which the AR-MIDAS and the AR-U-MIDAS produce better forecasts. Second, in an empirical exercise we evaluate density forecasts for quarterly US output growth, exploiting information from typical monthly series. We find that MIDAS models provide accurate and timely density forecasts.

JEL codes: C10, C53, E37

Keywords: Mixed Data Sampling, Density Forecasts, Nowcasting

*The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We would like to thank Todd Clark, Mike Clements, Eric Ghysels, Anne Sofie Jore, James Mitchell, Anders Rygh Swensen, Christian Schumacher and seminar and workshop participants at Norges Bank, the Macroeconomic and Econometric Conference in Birmingham, the 22nd Annual Symposium of the Society for Nonlinear Dynamics and Econometrics in New York, the IAAE 2014 Annual Conference in London, the 34th International Symposium on Forecasting in Rotterdam and the 68th European Meeting of the Econometric Society for useful comments.

[†]Knut-Are.Aastveit@norges-bank.no

[‡]Claudia.Foroni@norges-bank.no

[§]Francesco.Ravazzolo@norges-bank.no

1 Introduction

Economic time series are often available at different frequencies. As an example, while Gross Domestic Product (GDP) is released quarterly and with a substantial delay, many indicators are available at a monthly frequency and in a more timely manner. These timely indicators often contain important indications of the current state of the economy.

In order to obtain good nowcasts and short-term forecasts, it is therefore crucial to exploit the information from high-frequency data, see e.g. Evans (2005) and Banbura et al. (2011). Hence, methods that can deal with mixed-frequency data are becoming increasingly popular for nowcasting and short-term forecasting. Factor models and mixed-frequency VARs are examples of models that can deal with mixed-frequency data when cast in a state-space form with time aggregation scheme, see e.g. Mariano and Murasawa (2010) and Kuzin et al. (2011) for examples of mixed-frequency VARs, and Giannone et al. (2008) and Banbura and Modugno (2014) for examples of mixed-frequency dynamic factor models.¹ An alternative univariate approach is the mixed data-sampling (MIDAS) regression proposed by Ghysels et al. (2005, 2006) in the context of financial applications and extended to a macroeconomic context by Clements and Galvão (2008, 2009).

Policymakers and forecasters are increasingly interested in forecast metrics that require density forecasts of macroeconomic variables, as complete probability distributions over outcomes provide information helpful for making economic decisions, see e.g. Tay and Wallis (2000), Garratt et al. (2003), Gneiting (2011) and Clark (2011). Accordingly, several central banks, including the Bank of England, Norges Bank and Sveriges Riksbank have committed to publishing density or interval forecasts for macroeconomic aggregates in recent years. However, despite the flourishing theoretical and empirical literature on the use of mixed-frequency approaches, the focus has so far mainly been on point forecasts, Aastveit et al. (2014), Mazzi et al. (2013) and Carriero et al. (2014) being notable exceptions.²

In this paper, we extend the existent literature by using the MIDAS approach to obtain density forecasts. As a first contribution, we compute density forecasts from different MIDAS

¹See also Kuzin et al. (2013) and Aastveit et al. (2014) for applications that combine nowcasts from different models that use mixed-frequency data.

²Aastveit et al. (2014) study combinations of density nowcasts from a wide set of factor models, mixed-frequency VARs and bridge equation models, while the two latter studies focus on either combining density nowcasts from a small set of bridge equation models, Mazzi et al. (2013), or density nowcasts from models with time-varying parameters and volatility, Carriero et al. (2014).

model specifications. In particular, we consider the classical MIDAS model, introduced by Ghysels et al. (2004), and the extended version by Clements and Galvão (2008), which includes an autoregressive component. These models rely on the exponential lag polynomials to exploit high-frequency information while at the same time being parsimonious. Due to the non-linearity of the lag polynomials, MIDAS regressions are typically estimated by non-linear least squares (NLS). Furthermore, we also examine the unrestricted MIDAS regressions, U-MIDAS and AR-U-MIDAS, introduced by Forni et al. (2013). The U-MIDAS and AR-U-MIDAS approaches do not resort to functional distributed lag polynomials and can be estimated by OLS.

We propose a general stochastic simulation method based on parametric bootstrapping for computing predictive densities for both MIDAS and U-MIDAS regressions. Our method accounts for both parameter and shock uncertainty and extends the bootstrapping procedures in Berkowitz and Kilian (2000) and Clements and Taylor (2001) to the class of MIDAS models. We evaluate our bootstrapping approach in Monte Carlo simulations as well as in an empirical exercise applied to US real-time data.

First, we compare the forecasting performance of the different MIDAS models using Monte Carlo simulations. The basic design of the exercise is similar to that of Forni et al. (2013), where a high-frequency VAR(1) is specified. We consider three different residual assumptions: normal distributed residuals, Student's t-distributed residuals with 5 degrees of freedom, and residuals that have a time-varying variance modeled as a GARCH(1,1) process. We evaluate density forecasts according to two different criteria, the Logarithmic Score (LS) and the Continuous Rank Probability Score (CRPS). Point forecasts are evaluated based on root mean square forecasting errors. Our findings indicate that the forecasting results in terms of point and density forecasts are coherent. Moreover, the results do not clearly indicate a superior performance of one of the models under scrutiny, especially when the persistence of the low-frequency variable is low. Some differences are instead more evident when the persistence is high. In this case, the AR-U-MIDAS performs somewhat better, both in terms of point and density forecasts.

Second, in an empirical nowcasting exercise using US real-time data, we compute recursive density forecasts from the various MIDAS regressions for quarterly real output growth for the evaluation period 1985q2-2013q2. We consider 6 monthly explanatory variables. Following Clements and Galvão (2008), we use industrial production, employment (nonfarm payrolls) and

capacity utilization. In addition, we include the Purchasing Managers Index (PMI), the Chicago Fed National Activity Index (CFNAI) and the Philadelphia Fed Business Outlook Survey for general business activity (PFBOS). The monthly real-time vintages that we use correspond to the information available around the 20-25th of each month in the quarter, i.e., just after the release of the CFNAI.

Looking ahead to the results, we find that the use of within-quarter information on monthly indicators can result in marked improvements in density forecast accuracy, with reductions in CRPS and increases in LS compared to the benchmark quarterly-frequency AR and AR distributed-lag (ADL) models. In particular, models that use the CFNAI index substantially improve the forecast accuracy relative to the benchmark models. However, gains are smaller when focusing on more traditional macroeconomic series such as industrial production, employment (nonfarm payrolls) and capacity utilization. The results are robust to ending the evaluation sample in 2007q4, i.e. prior to the Great Recession.

The plan of the rest of our paper is as follows. Section 2 provides an overview of the different MIDAS models that we consider in our analysis. Section 3 describes our bootstrapping approach for computing density forecasts for MIDAS models and how these forecasts are evaluated. In Section 4, results from our Monte Carlo experiments are presented, while results from the empirical application are presented in Section 5. Finally, Section 6 concludes.

2 Mixed-data Sampling

To take into account mixed-frequency data, Ghysels et al. (2004) introduced the mixed-data sampling (MIDAS) approach, in which the dependent variable y_{t_q} , sampled at a lower frequency, is regressed on a distributed lag of x_{t_m} , which is sampled at a higher frequency.

In this section, we present the two main MIDAS approaches that have been suggested in the literature; the standard MIDAS regression introduced by Ghysels et al. (2004) and the unrestricted version as in Forni et al. (2013). For each of the two approaches, we present the basic model and their extension with an AR component.

In the rest of the paper, we define $t_q = 1, \dots, T_q$ as the basic (quarterly) time unit and $t_m = 1, \dots, T_m$ as the higher frequency (monthly) time unit. In every basic time unit we observe the high frequency unit $m = 3$ times (e.g. 3 months in a quarter).

2.1 The MIDAS model

As introduced by Ghysels et al. (2004), MIDAS regressions are reduced form regressions that involve processes sampled at different frequencies. Highly parsimonious distributed lag polynomials are used to link the higher-frequency variable to the low-frequency one. These polynomials are chosen in such a way to prevent the proliferation of parameters while at the same time being very flexible, allowing for various shapes, where only a few parameters need to be estimated.

The basic MIDAS model for a single explanatory variable, and h_q -step-ahead forecasting, with $h_q = h_m/m$, is given by:

$$y_{t_q+mh_q} = y_{t_m+h_m} = \beta_0 + \beta_1 b(L_m; \theta) x_{t_m+w}^{(m)} + \varepsilon_{t_m+h_m} \quad (1)$$

where $b(L^{1/m}; \theta) = \sum_{k=0}^K c(k; \theta) L_m^k$, and $L_m^x x_{t_m}^{(m)} = x_{t_m-x}^{(m)} \cdot x_{t_m+w}^{(m)}$ is skip-sampled from the high frequency indicator x_{t_m} .

Different choices for $c(k; \theta)$ have been proposed in the literature. The most common one is the exponential Almon lag, defined as:

$$c(k; \theta) = \frac{\exp(\theta_1 k + \dots + \theta_Q k^Q)}{\sum_{k=1}^K \exp(\theta_1 k + \dots + \theta_Q k^Q)} \quad (2)$$

The key advantage of this specification is that it allows for long lags with a limited number of parameters. It also allows for different values of the parameters to take various shapes. Therefore, estimating the parameters from the data automatically determines the shape of the weights and the number of lags to be included in the regression. The MIDAS model is estimated using nonlinear least squares (NLS) in a regression of y_t onto $x_{t-h}^{(m)}$.

Notice that the MIDAS model is h -dependent, and thus needs to be re-estimated for each forecast horizon. It is therefore a model which provides direct and not recursive forecasts. There is therefore no need to have forecasts of the independent variable, but the model is adjusted accordingly to include the very last observations of the $x^{(m)}$ available. The forecast is given by

$$\widehat{y}_{T_m^y+h_m|T_m^x} = \widehat{\beta}_0 + \widehat{\beta}_1 b(L_m; \widehat{\theta}) x_{T_m^x}^{(m)}. \quad (3)$$

2.2 The AR-MIDAS model

The main extension to the MIDAS model is the inclusion of an autoregressive term. However, this is not straightforward, since it would result in the creation of seasonal patterns in the explanatory variables, as highlighted in Clements and Galvão (2009).

To avoid the inconvenience of seasonal effects, Clements and Galvão (2009) suggest the introduction of the AR dynamics as a common factor:

$$y_{t_m} = \beta_0 + \lambda y_{t_m-h_m} + \beta_1 b(L_m; \theta) \left(1 - \lambda L_m^{h_m}\right) x_{t_m+w-h_m}^{(m)} + \varepsilon_{t_m}. \quad (4)$$

In this way, the response of y to $x^{(m)}$ remains non-seasonal.

To estimate the AR-MIDAS model, we follow the procedure described in Clements and Galvão (2009), for which we briefly highlight the steps. First, estimate the standard MIDAS (the basic model), take the residuals $\widehat{\varepsilon}_{t_m}$ and estimate an initial value for λ , say λ_0 , where $\widehat{\lambda}_0 = \left(\sum \widehat{\varepsilon}_{t_m+w-h_m}^2\right)^{-1} \sum \widehat{\varepsilon}_{t_m} \widehat{\varepsilon}_{t_m+w-h_m}$. Then construct $y_{t_m}^* = y_{t_m} - \widehat{\lambda}_0 y_{t_m-h_m}$ and $x_{t_m+w-h_m}^{*(m)} = x_{t_m+w-h_m}^{(m)} - \widehat{\lambda}_0 x_{t_m-(h_m-w)-h_m}^{(m)}$. The estimator $\widehat{\theta}_1$ is obtained by applying nonlinear least squares to:

$$y_{t_m}^* = \beta_0 + \beta_1 b(L_m; \theta) x_{t_m+w-h_m}^{*(m)} + \varepsilon_{t_m}. \quad (5)$$

A new value of λ , $\widehat{\lambda}_1$, is obtained from the residuals of this regression. Then a new step is run, using $\widehat{\lambda}_1$ and $\widehat{\theta}_1$ as the initial values. In this way, the procedure gets the estimates of $\widehat{\lambda}$ and $\widehat{\theta}$ which minimize the sum of squared residuals.

2.3 The Unrestricted MIDAS model

Foroni et al. (2013) study the performance of MIDAS regressions that do not resort to functional distributed lag polynomials, referred to as the U-MIDAS. The U-MIDAS model is based on a linear lag polynomial such that

$$\begin{aligned} c(L^m)\omega(L)y_{t_m} &= \delta_1(L)x_{1t_m-1} + \dots + \delta_N(L)x_{Nt_m-1} + \epsilon_{t_m}, \\ t &= 1, 2, 3, \dots \end{aligned} \quad (6)$$

where $c(L^m) = (1 - c_1L^m - \dots - c_cL^{mc})$, $\delta_j(L) = (\delta_{j,0} + \delta_{j,1}L + \dots + \delta_{j,v}L^v)$, $j = 1, \dots, N$.

Under the assumption that the lag orders c and v are large enough to make the error term ϵ_{t_m} uncorrelated, then all the parameters in the U-MIDAS model, equation (6), can be estimated by simple OLS, while the aggregation scheme $\omega(L)$ is assumed known. The lag order can be selected according to any information criteria.

The simple U-MIDAS model, which represents the unrestricted counterpart of the MIDAS model with one explanatory variable is, therefore, as follows:

$$\omega(L)y_{t_m} = \delta_1(L)x_{1t_m-1} + \epsilon_{t_m}, \quad (7)$$

with $\delta_1(L) = (\delta_{1,0} + \delta_{1,1}L + \dots + \delta_{1,\nu}L^\nu)$, and ν selected according to the BIC criterion.

The direct forecast in this case is constructed as:

$$\bar{y}_{T_m^x+h_m|T_m^x} = \bar{\delta}_1(L)x_{1T_m^x}, \quad (8)$$

where $\bar{\delta}_1(L)$ is obtained by projecting y_{t_m} on the last information available.

2.4 The Unrestricted MIDAS model with an AR component

In the U-MIDAS it is straightforward to add an autoregressive component to derive the AR-U-MIDAS (see Foroni et al. (2013) for further details). In contrast with the MIDAS with polynomial restrictions, the autoregressive term in this case, can be included easily without any common factor restriction as in Clements and Galvão (2009).

The h_m -step ahead forecasts given information in T_m^x are given in this case by:

$$\bar{y}_{T_m^x+h_m|T_m^x} = \bar{c}(L^k)y_{T_m^x} + \bar{\delta}_1(L)x_{1T_m^x} + \dots + \bar{\delta}_N(L)x_{NT_m^x}, \quad (9)$$

where the polynomials $\bar{c}(Z)$ and $\bar{\delta}_i(L)$ are obtained by projecting y_{t_m} on information dated $mt - h_m$ or earlier, for $t = 1, 2, \dots, T_m^x$.

3 Introducing density forecasts in MIDAS models

So far in the literature, MIDAS models have been evaluated based on their point forecasting performance. Policymakers and forecasters are increasingly interested in forecast metrics that require density forecasts of macroeconomic variables. Moreover, if the decision maker's loss function is not quadratic, then it no longer suffices to focus solely on the first moments of possible outcomes (point forecasts). To ensure appropriate decision making, the decision maker should be given suitable characterizations of forecast uncertainty. Density forecasts provide an estimate of the probability distribution of forecasts (see Gneiting (2011) for a detailed discussion of the difference between point forecasting and density forecasting).

In this section, we first derive a bootstrapping approach that can be used to compute density forecasts for the different MIDAS models, described in the previous section, and then we present how we evaluate the forecast densities.

3.1 Bootstrapping MIDAS density forecasts

Forecast uncertainty cannot be derived analytically in the case of non-linear models, as in the MIDAS and AR-MIDAS models.³ We propose a stochastic simulation method based on parametric bootstrapping to forecast densities accounting for both parameter and shock uncertainty and thereby extend the bootstrapping procedures in Berkowitz and Kilian (2000) and Clements and Taylor (2001) to the class of MIDAS models. Our method relies on the algorithm in Davison and Hinkley (1997) (section 7.2.4) for prediction in generalized linear models. Recall the general MIDAS model:

$$y_{t_m} = \beta_0 + \lambda y_{t_m - h_m} + \delta(\Theta) x_{t+w-h_m}^{(m)} + \varepsilon_{t_m}, \quad (10)$$

where $\delta(\Theta)$ is the polynomial which describes the impact of the high-frequency variable on the low-frequency one. Depending on the functional form, we have the MIDAS or U-MIDAS specification. When the coefficient $\lambda = 0$, we obtain the models with no autoregressive terms.

Then, we assume that:

³Note that the U-MIDAS and AR-U-MIDAS models are linear regression models. Assuming a normal distribution of the random shocks would yield future outcomes that have a normal distribution. However, in order to account for parameter uncertainty, we also apply our bootstrap procedure to these models.

(A1) ε_{t_m} are i.i.d. with $E(\varepsilon_{t_m}) = 0$, $E(\varepsilon_{t_m}^2) = \sigma^2$ with $\sigma^2 < \infty$, and $E(\varepsilon_{t_m}^{2(s+1)}) < \infty$ for $s \geq 3$.

(A2) $(\varepsilon_{1_m}, \varepsilon_{1_m}^2)$ satisfies Cramer's condition, i.e., for every $d > 0$, there exists δ such that $\sup_{|t| > d} |E \exp(it'(\varepsilon_{1_m}, \varepsilon_{1_m}^2))| \leq \exp(-\delta)$.

(A3) $x_{t_m+w-h_m}^{*(m)}$ are exogenous fixed variables.

(A4) $|\lambda| < 1$ for the AR-MIDAS and AR-U-MIDAS models.

The steps conducted in the bootstrap are as follows.

1. Estimate equation (10) using the adequate least square or nonlinear least square method, and obtain $\widehat{\beta}_0, \widehat{\lambda}, \widehat{\delta}(\Theta)$.
2. For $r = 1, \dots, R$, simulate $\widetilde{y}_{r,t_m} = \widehat{\beta}_0 + \widehat{\lambda}\widetilde{y}_{r,t_m-h_m} + \widehat{\delta}(\Theta) x_{t_m+w-h_m}^{(m)} + \widetilde{\varepsilon}_{r,t_m}$, where $\widetilde{\varepsilon}_{r,t_m}$ is resampled from $\widehat{\varepsilon}_{t_m}$.
3. Reestimate eq. (10) for each \widetilde{y}_{r,t_m} , and obtain $\widetilde{y}_{r,T_m^x+h_m|T_m^x}$, where the shock uncertainty is included by resampling from $\widehat{\varepsilon}_{t_m}$.

Following Davison and Hinkley (1997), we fix the value of y_T equal to the value of the original series.

In practice, R vectors of pseudo-random numbers are generated to replicate the same properties of the residuals of the model, via the bootstrapping technique. For each $r = 1, \dots, R$ replications, a new set of simulated data is generated, and based on that a new forecast $\widetilde{y}_{r,t_m+h_m}$ is obtained. The empirical distribution of $\{\widetilde{y}_{r,t_m+h_m}\}_{r=1}^R$ is then our density.

Given the assumptions (A1)-(A4), Davison and Hinkley (1997) discuss how the method is a generalization of the bootstrapping algorithm for linear models and Bose (1988) provides proofs of its convergence.⁴

⁴Bose (1988) focuses on linear AR models with imposed stationarity (see assumption (A4) above). For an extension accounting for a possible unit root, see Inoue and Kilian (2002). Moreover, relaxing assumption (A1) by allowing for conditional heteroskedasticity or serial correlation could be done by applying a wild bootstrap procedure, see, e.g., Goncalves and Kilian (2004).

3.2 Evaluating the density forecasts

As a conventional measure for point forecast accuracy, we use the Mean Squared Predictive Error (MSPE), which is defined as:

$$\frac{1}{T_m^S} \sum_{s=T_m+1}^{T_m^S} e_s^2, \quad (11)$$

where e_{t_m} is the forecast error, and T_m^S and T_m denote the starting period for the forecasts and the number of observations in the estimation sample.

Following Gneiting and Raftery (2007), Billio et al. (2013) and Clark and Ravazzolo (2014) for the evaluation of the density forecasts, we provide two metrics, the logarithmic score (LS) and the continuous ranked probability score (CRPS).

The LS is defined as the logarithm of the probability density function evaluated at the outturn of the forecast and is given by

$$\ln S_{i,h} = \frac{1}{T_m - h_m - T_m^S + 1} \sum_{t=T_m^S}^{T_m-h_m} \ln f_{t_m+h_m,t_m,i}(y_{t+h_m}), \quad (12)$$

where $\ln f_{t_m+h_m,t_m,i}(y_{t+h_m})$ denotes the density for $Y_{t_m+h_m}$ conditional on some information set available i at time t_m , and T_m^S and T_m are defined as above. The LS chooses the model that on average gives the higher probability to events that actually occurred. The LS is related to the Kullback Leibler Information Criterion (KLIC), see Mitchell and Hall (2005) and Hall and Mitchell (2007), which measures the distance between the true unknown density forecast and the one proposed.⁵ Since we do not know the true density, we cannot use the KLIC, but we can still compare different densities by searching for the maximum log score.

We also evaluate our forecasts based on a second metric, the Continuous Rank Probability Score (CRPS). The CRPS for a specific model measures the average absolute distance between the empirical cumulative distribution function (CDF) of $y_{t_m+h_m}$, which is simply a step function in $y_{t_m+h_m}$, and the empirical CDF that is associated with the model predictive density. Formally,

⁵See also Amisano and Giacomini (2007), Kascha and Ravazzolo (2010) and Jore et al. (2010).

the CRPS of a component density for a particular observation can be defined as:

$$CRPS = \int \left(F(z) - \mathbb{I}_{[y_{t+h_m}, +\infty)}(z) \right)^2 dz \quad (13)$$

where F is the CDF from the predictive density $f_{t_m+h_m, t_m, i}$ defined above, see Panagiotelis and Smith (2008) for more formal details and Ravazzolo and Vahey (2013) for an illustrative example. The expectation terms are approximated using the draws from the forecast density. Smaller CRPS values imply a higher precision of the model.

The CRPS circumvents some of the drawbacks of the LS, as the latter does not reward values from the predictive density that are close but not equal to the realizations (see, e.g., Gneiting and Raftery (2007)). In the rest of the paper we will analyze the results both in terms of CRPS and in terms of LS.

4 Monte Carlo experiments

In this section, we present Monte Carlo experiments, which focus on the performance of different MIDAS models, comparing them according to their point forecasting and density forecasting performance. We first describe the simulation design, and then we present the results.

4.1 Simulation design

We follow the simulation design in Foroni et al. (2013) and consider a bivariate high-frequency VAR as the Data Generating Process (DGP):

$$\begin{pmatrix} y_{t_m} \\ x_{t_m} \end{pmatrix} = \begin{pmatrix} \rho & \delta_l \\ \delta_h & \rho \end{pmatrix} \begin{pmatrix} y_{t_m-1} \\ x_{t_m-1} \end{pmatrix} + \begin{pmatrix} e_{y, t_m} \\ e_{x, t_m} \end{pmatrix}, \quad (14)$$

where y_{t_m} is the low frequency variable and x_{t_m} is the high-frequency variable. t_m is the high-frequency time index with $t_m = 1, \dots, (T + ES) \times k$ and T defines the size of the estimation sample expressed in the low-frequency unit (quarters in our example). For forecasting purposes, we generate an additional number of observations to define our evaluation sample, ES . The Monte Carlo simulations are set to replicate monthly and quarterly data, assuming that the quarterly variable is a monthly variable skip-sampled every third period and therefore available

only for $t_m = k, 2k, \dots, (T + ES) \times k$. In our case, k is equal to 3, as there are three monthly observations in a quarter.

In our first simulation design, the shocks e_{y,t_m} and e_{x,t_m} in the DGP are sampled independently from normal distributions with mean zero for all $t_m = 1, \dots, (T + ES) \times k$, and the variances are chosen such that the unconditional variance of y is equal to one and the explanatory power of x is high.⁶ For the persistence parameter, we choose different combinations of ρ and δ_l to mimic different persistence and dependence in the process, see values in tables. Moreover, we fix $\delta_h = 0$, so that only y depends on x and not vice versa.

In our second simulation design, the set-up is similar to the first one except that the shocks e_{y,t_m} and e_{x,t_m} are sampled independently from a Student's t-distribution with 5 degrees of freedom. The motivation behind this is to study the accuracy of our bootstrapping method when the error terms have fat tails.

Finally, in our third simulation design the set-up is again similar to the first simulation exercise, except that the shocks e_{y,t_m} now have a time-varying variance, modeled as a GARCH(1,1) process:

$$\begin{aligned} e_{y,t_m} &= \sigma_{y,t_m} u_{y,t_m}, \quad u_{y,t_m} \sim N(0, 1), \\ \sigma_{y,t_m}^2 &= a_0 + a_1 u_{y,t_m-1} + b_1 \sigma_{y,t_m-1}^2 \end{aligned}$$

with $a_0 > 0$, $a_1 \geq 0$, $b_1 \geq 0$ and $a_1 + b_1 \leq 1$. Since Assumption (A1) for our bootstrap method implies homoscedasticity, results from the third simulation exercise will illustrate the accuracy of our bootstrapping approach when this assumption is violated.

In both experiments, T is equal to 100 and ES is equal to 50 (both T and ES are indicated in quarters). We produce one-step ahead forecasts for y_{T+es} , where $es = 1 : ES$, using information for y up to $T + es - 1$ (equivalently to $3T + 3es - 3$ months). The monthly variable is available up to $3T + 3es - 1$ months. Our one-step ahead forecast corresponds therefore to a one-month ahead forecast.

In this Monte Carlo experiment, the number of Monte Carlo replications is fixed at 1000. For each replication we compute the point forecasts for the 4 MIDAS models described in Section 2 (MIDAS, AR-MIDAS, U-MIDAS, AR-U-MIDAS). The densities are constructed computing

⁶We have also tried different combinations in which the explanatory power of the x is low, and most of the variation in y is due to noise. The results are essentially the same.

1000 draws. The forecasts are then evaluated based on MSPE (point forecasts) and LS and CRPS (density forecasts).

4.2 DGP with normally distributed shocks: results

The results for the Monte Carlo forecast experiments where the DGP has normally distributed residuals, are summarized in Table 2. We compare the forecasting performance of the different MIDAS models based on their (out-of-sample) MSPEs, CRPS, and LS computed over the 50 periods of the evaluation sample. Moreover, we report the results for two standard quarterly benchmarks, an AR(1) process and a Random Walk (RW) process. The densities for the benchmark are computed using our bootstrap method without the high frequency components, therefore similar to Berkowitz and Kilian (2000) and Clements and Taylor (2001). In this way, we are also able to compare the performance of our mixed-frequency models relative to models that consider only quarterly information.

From Table 2, we observe several interesting results. First, when the persistence of the low-frequency variable is low ($\rho = 0.1, 0.5$), using high-frequency information substantially improves the forecasting performance relative to the AR and RW benchmarks. This is not very surprising as these benchmarks discard the useful information provided by the high-frequency variable. Moreover, by comparing the performance of the U-MIDAS vs. the AR-U-MIDAS and the MIDAS vs. AR-MIDAS, including an AR-component in the mixed frequency models does not substantially affect the forecasting performance. Obviously, results change when the persistence is high, i.e. setting $\rho = 0.9$ in our simulations. In this case, only the AR-MIDAS and the AR-U-MIDAS produce superior forecasts relative to the two benchmarks.

The results in terms of point and density forecasts are coherent: whenever we find a smaller MSPE, the models also obtain an improvement in the density forecasts. We interpret this as supporting evidence for our bootstrapping approach, which is then precise not only for the mean part of the distribution but more importantly also for the full distribution.

When $\rho = 0.9$, the results do not change in the case of the U-MIDAS, but some issues arise in the case of the AR-MIDAS. In a few of the bootstrapping replications, imposing the common factor restriction in the AR component of the AR-MIDAS model to avoid seasonality effects creates convergence issues and forecast performance deteriorates as a result.

Finally, ranking the mixed-frequency models is not straightforward, as results do not clearly indicate a superior performance to one of the models under scrutiny. However, we can highlight some features that emerge from the table. First of all, when the persistence of the model is low, there are no sizeable differences in performance across the models. Some differences are instead more evident when the persistence is high. The AR-U-MIDAS seems to perform slightly better, in terms of both point and density forecasts. This evidence is in line with the findings in Foroni et al. (2013) in the case of point forecasts.

It is worth mentioning that the equally good or even superior performance of the AR-U-MIDAS is obtained in a much shorter computational time and with no convergence issues, since the model and bootstrap replications are estimated by a simple OLS.

4.3 DGP with Student's t-distributed shocks: results

In this subsection, we present the results of the Monte Carlo forecast experiments for the second DGP, in which the residuals are drawn from a Student's t distribution. As in the previous simulation design, we compare the forecasting performance of the different MIDAS models based on their MSPEs, CRPS, and LS. The results of the mixed-frequency models, together with the two benchmarks mentioned in the previous subsection, are reported in Table 3.

In general, the results display the same features as in the previous DGP, with shocks generated by a normal distribution. First, we can confirm that when the persistence of the low frequency variable is low, using high-frequency information hugely improves forecasting performance relative to a quarterly benchmark. Moreover, the different mixed-frequency models perform in a very similar way, and ranking them is impossible because their forecasting ability is nearly the same. In contrast, when the persistence is high, it is the AR component which turns out to be necessary to have a forecasting performance able to compete with the benchmarks. Therefore, the models that incorporate an AR term clearly outperform their counterparts that do not.

The results in terms of point and density forecasts are also in this case coherent: whenever a model performs well in terms of point forecasts, it also performs well in terms of densities.

4.4 DGP with time-varying volatility shocks: results

As a final Monte Carlo experiment, we consider a DGP where the residuals of the low-frequency variable are drawn from a GARCH(1,1) distribution. The design of the experiment is similar to the previous subsections, except for the error distribution for the variable y . The results of the mixed-frequency models, together with the two benchmarks, are reported in Table 4.

Even in this case, the results share the same features as in the previous two cases, where the errors in the DGPs were distributed respectively as a Normal and as a Student's t . Once again, therefore, we confirm the importance of mixed-frequency information in both point and density forecasting, especially in the case in which the persistence of the low frequency variable is low.

5 An empirical application: nowcasting US output growth

In this section we provide an empirical application which focuses on the performance of different MIDAS models for real-time out-of-sample nowcasting of quarterly US real output growth using monthly information. By using real-time data, we take into account the publication lags of the different series and ensure that we are only using data that were available on the date of the forecast origin. We evaluate the performance of our models in terms of point and density forecasts.

5.1 The data and design of the empirical exercise

Our dataset includes quarterly US real output growth and six monthly explanatory indicators, which are recognized to be important predictors. Table 1 provides an overview of the data and some of their real-time characteristics. Following Clements and Galvão (2008), we consider industrial production, employment (nonfarm payrolls), and capacity utilization. The three monthly indicators are converted to monthly growth rates by taking the log-differences of the series. Quarterly real output growth is measured as the log-difference of the real GDP series. We obtained the real-time data vintages of the aforementioned variables from the Federal Reserve Bank of Philadelphia's Real-Time Dataset for Macroeconomists (RTDSM), described in Croushore and Stark (2001). Moreover, we include three additional indicators: the Purchasing Managers Index (PMI), the Chicago Fed National Activity Index (CFNAI) and the Philadel-

Table 1: Description of the data

Name	Description	Transf	First vintage	Timing	Publishing lag
EMP	Employment	$\Delta \ln$	1985m1	1st Friday of month	1 month
IP	Industrial Production	$\Delta \ln$	1985m1	15-18th of month	1 month
CU	Capacity Utilization	$\Delta \ln$	1985m1	15-18th of month	1 month
PMI	NAPM Purchasing Managers Index	<i>Level</i>	2002m1	1st business day of month	1 month
CFNAI	Chicago Fed National Activity Index	<i>Level</i>	2001m1	22-25th of month	1 month
PFBOS	Philly Fed's Business Outlook Survey	<i>Level</i>	Last vintage	3rd Thursday of month	Current month

Note: Variable names and variable descriptions are reported in Column 1 and 2. Column 3 reports the transformation that is used for each of the explanatory variables in various MIDAS regressions. The first available real-time vintage for each variable is reported in Column 4. Finally, Column 5 and Column 6 indicate the official dates of the publication and the lag with which the data are reported, respectively.

phia Fed Business Outlook Survey for general business activity (PFBOS). Monthly real-time vintages of PMI are available from the ALFRED (Archival Federal Reserve Economic Data) database maintained by the Federal Reserve Bank of St. Louis. The CFNAI is a coincident indicator of the overall economic activity and is computed as a weighted average of 85 existing monthly indicators of national economic activity. We include the CFNAI index, since Andreou et al. (2013) find it useful for improving short-term forecasting. Real-time vintages of monthly CFNAI are collected from the website of the Federal Reserve Bank of Chicago. Finally, the PFBOS series is collected from the website of the Federal Reserve of Philadelphia. We include the PFBOS since Giannone et al. (2008) find that in a dynamic factor model PFBOS is the data release that has the largest impact on the nowcast and its precision. Real-time vintages of the PFBOS are not available. However, this series only undergoes revisions due to changes in seasonal factors. Since these revisions are substantially smaller than revisions in e.g. national account data, we include this series in our empirical analysis.

We compute recursive point and density forecasts for US real GDP growth from the different MIDAS regressions. As benchmarks, we consider quarterly-frequency AR and autoregressive distributed lag (ADL) models.⁷ For the AR model we select the lag length according to the BIC criterion, where the maximum lag is fixed to 4. For the ADL model, we fix the lag of GDP growth equal to one and select the lags of the exogenous variables with a BIC criterion. Our estimation sample covers the period 1970m1 to 2013m6. The full recursive forecast evaluation period runs from 1985q1 to 2013q2. As robustness, we also provide results for an evaluation sample that ends in 2007q4 (before the Great Recession). We use monthly real-time vintages from 1985m2 to 2013m7 and compute forecasts using monthly vintages that correspond to the

⁷Notice that when the monthly exogenous variable is only available for the first month or the two first months of the quarter, the quarterly ADL model ignores this information.

available information between the 20th and 25th of the month.⁸ More precisely, this corresponds to the timing just after the release of the CFNAI. The quarterly US GDP is available as an advance estimate at the end of the first month after the end of the referring quarter. Accordingly, when computing the nowcasts at the second and third month of each quarter (e.g. February and March), GDP growth for the previous quarter (e.g. q4 in our example) is available. On the other hand, when we compute nowcasts in the first month of each quarter (e.g. January), the GDP value for the previous quarter is not available. Hence, the nowcasts are then two step ahead forecasts.

For each monthly indicator we specify 4 different MIDAS regressions (MIDAS, AR-MIDAS, UM-IDAS and AR-U-MIDAS). We compute a nowcast from each model specification for each monthly vintage, i.e. 3 forecasts per quarter (which correspond to month horizons $h_m = 1, 2, 3$). Due to the release delay of US real output, we also compute a backcast of the previous quarter for the first month of each quarter (we refer to this as $h_m = -1$). A key issue in this exercise is the choice of a benchmark for the “actual” measure of GDP. We follow Romer and Romer (2000) in using the second available estimate of GDP as the actual measure. For a discussion of alternative benchmark data vintages, see Stark and Croushore (2002)

The forecasting performance of the different models is assessed in terms of CRPS and LS for the density forecasts. To provide a rough gauge of whether the differences in forecast accuracy between the ADL model and the alternative MIDAS models are significant, we follow Clark and Ravazzolo (2014) and apply a Diebold and Mariano (1995) t -tests for equality of the average loss (with loss defined as LS, CRPS or RMSE).⁹ In the tables, differences in accuracy that are statistically different from zero are denoted by one, two or three asterisks, corresponding to significance levels of 10%, 5% and 1%, respectively. The underlying p -values are based on t -statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992).¹⁰ Since the AR-MIDAS and the AR-U-MIDAS model nests the ADL model, for the comparison of these models to the ADL, we report

⁸For the PMI and CFNAI, real-time vintage data exist only for parts of the evaluation period, from 2002m1 and 2001m1, respectively. For these, we use the first available real-time vintage and truncate these series backward recursively.

⁹Amisano and Giacomini (2007) extend the results of Giacomini and White (2006) for point forecasts to density forecasts, developing a set of weighted likelihood ratio tests. In our application to tests for density forecasts, we do not employ a weighting scheme (or put another way, we use equal weights).

¹⁰The issue of real-time data complicates any assessment of whether the resulting differences in forecast accuracy between models are significant, see Clark and McCracken (2009).

p -values based on one-sided tests, taking the ADL as the null and the AR-MIDAS/AR-U-MIDAS as the alternative. Because the ADL and the MIDAS/UMIDAS models are not nested, for these model comparisons we report p -values based on two-sided tests.

5.2 Results

Table 5 presents real-time density forecasting performance for the different MIDAS models for the evaluation sample 1985q2-2013q2. Density forecast performance is measured in terms of CRPS and LS. In addition, we provide results for the two quarterly benchmarks (an AR model and an ADL model).

The table reveals three interesting results. First, by comparing the forecasting performance from the various MIDAS models with the AR and ADL benchmarks, the table shows sizeable reductions in the CRPS and sizeable increases in the LS for the MIDAS models when monthly data are available on the current quarter. For each model, there is a clear and steady increase in forecasting performance as more information is available, i.e. the nowcasts produced in the third month of the quarter (at horizon $h_m = 1$) are better than the forecasts produced in the first month of the quarter (at horizon $h_m = 3$).¹¹ Compared to the quarterly-frequency ADL benchmark the density nowcasting improvements for the MIDAS models at horizon $h_m = 1$ are of the order of 15-20 percent in terms of LS and 10-15 percent in terms of CRPS. Results based on the t-tests indicate a significant improvement for several of the models compared to the ADL benchmark.¹² In particular, this is the case for model specifications using the CFNAI.

Second, a comparison of the forecasting performance of the 4 different MIDAS models for each of the 6 monthly indicators reveals very small differences in terms of forecasting performance. There is a small tendency for MIDAS and AR-MIDAS models to perform relatively better than the U-MIDAS and AR-U-MIDAS when applied to the CFNAI index and the PF-BOS. However, the differences are not significant. Thus, in terms of model comparison, we do not obtain clearly a superior approach. This is consistent with results in Foroni et al. (2013)

¹¹We obtained qualitatively similar results for MIDAS specifications using either a credit spread, the S&P 500, the Chicago Fed's National Financial Condition Index or the oil price. However, the relative nowcasting performance from these model specifications were somewhat worse than those from the model specifications reported here. Note that in model specifications that included financial variables, we used monthly averages of the financial variables as predictors. Thus, we did not consider the use of daily financial data as in Andreou et al. (2013).

¹²In most cases, the ADL and MIDAS models provide statistically superior forecasts, in terms of the t-type test, to the AR model.

based on point forecasts comparison.

Third, the CFNAI index is the single monthly indicator that improves the density forecasts the most both in terms of LS and CRPS. This holds for all the 4 MIDAS models. This is not very surprising as the CFNAI is a weighted average of 85 existing monthly indicators of national economic activity. However, the PFBOS, IP and CU also provide useful information that improves the forecasts.

As a robustness for our results, Table 6 reports results when the evaluation sample ends in 2007q4, i.e., prior to the Great Recession. The absolute performance of the individual models is, as expected, superior compared with the performance when including the crisis period. The table also shows that even when excluding the crisis period, the various MIDAS models outperform the AR benchmark and in most cases also the ADL benchmark.

Figures 1 and 2 show how the relative cumulative LS and CRPS from the MIDAS and AR-MIDAS regressions have evolved over time, relative to the AR benchmark model. The figures describe the relative forecasting performance using each of the 6 monthly indicators and for each of the 4 forecasting horizons. The measures are constructed so that an increase in the relative value measures a relative improvement in the forecasting performance of the MIDAS and AR-MIDAS regressions compared with the AR benchmark. The figures reveal three interesting results. First, for all models and indicators, except for the AR-MIDAS model with the PMI,¹³ there has been a steady increase in the relative cumulative LS and CRPS. Second, for both the MIDAS and the AR-MIDAS regressions the cumulative LS and CRPS increase the most for CFNAI. This also holds for all forecast horizons and for most of the recursive time period. Thus, the superior forecast performance from models that use the CFNAI is a robust finding. Third, for all indicators and for both the MIDAS and AR-MIDAS regressions, the relative cumulative LS and CRPS increase sharply during the Great Recession, but this increase is then followed by a weak, but persistent, decrease during the recovery period. This may indicate that the monthly indicators are particularly informative about sharp decreases in real output growth.

Finally, Table 7 reports real-time point forecasting performance for the different MIDAS models for the evaluation samples 1985q2-2013q2 and 1985q2-2007q4. We measure point forecast accuracy in terms of RMSE. As for the results obtained with density forecasting, there

¹³The AR-MIDAS with PMI seems to experience convergence issues similar to what we found in the Monte Carlo experiments in the case of high persistence.

are small differences between the different MIDAS specifications in terms of forecasting performance and there is a clear and steady increase in forecasting performance as more information is available. However, based on the t-test for equality of the average loss as well as the relative forecasting performance of the different MIDAS specifications compared to the benchmarks, the gains from using mixed-frequency information are smaller for point forecasting than for density forecasting. This indicates that the gains in density forecast accuracy associated with having more and more information on the quarter derive not only from better centering (i.e., better point forecasts) of the forecast densities but also from improvement of the density shape.

6 Conclusions

In this paper, we extend the existent literature by using the MIDAS approach to obtain density forecasts. First, we compute density forecasts from different MIDAS models, the classical MIDAS models and the unrestricted version, U-MIDAS. We derive a parametric bootstrap algorithm to compute densities valid for both classes of models when accounting for parameter uncertainty. To compare the different models, we run Monte Carlo simulations that provide evidence on the forecasting performance of the models in finite samples. We find that the results in terms of point and density forecasts are coherent, but do not clearly indicate a superior performance of one of the models under analysis. Finally, in our empirical applications, we compute density forecasts from the different mixed-frequency models for quarterly US output growth, considering six monthly variables: industrial production, employment, capacity utilization, the Purchasing Managers Index, the Chicago Fed National Activity Index and the Philadelphia Fed Business Outlook Survey for general business activity. The results indicate that including monthly indicators largely improves density forecasting performance.

Nowcasting real GDP growth is only one of many examples where our parametric bootstrapping method can be applied to compute predictive densities. The generic question we address is how density forecasts can be computed in a MIDAS model. There are many other macroeconomic series to which our method can be applied and the method is therefore of general interest beyond the specific application considered in the present article.

References

- Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business and Economic Statistics* 32(1), 48–68.
- Amisano, G. and R. Giacomini (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business and Economic Statistics* 25(2), 177–190.
- 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.
- Andrews, D. and J. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953–966.
- Banbura, M., D. Giannone, and L. Reichlin (2011). Nowcasting. In M. P. Clements and D. F. Hendry (Eds.), *The Oxford Handbook of Economic Forecasting*. Oxford Handbooks in Economics.
- Banbura, M. and M. Modugno (2014). Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29(1), 133–160.
- Berkowitz, J. and L. Kilian (2000). Recent developments in bootstrapping time series. *Econometric Reviews* 19(1), 1–48.
- Billio, M., R. Casarin, F. Ravazzolo, and H. K. van Dijk (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics* 177(2), 213–232.
- Bose, A. (1988). Edgeworth corrections by bootstrap in autoregressive. *Annals of Statistics* 16(4), 1709–1722.
- Carriero, A., T. E. Clark, and M. Marcellino (2014). Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility. *Journal of Applied Econometrics*, forthcoming.
- Clark, T. and F. Ravazzolo (2014). The macroeconomic forecasting performance of autoregressive models with alternative specifications of time-varying volatility. *Journal of Applied Econometrics* (forthcoming).

- Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics* 29(3), 327–341.
- Clark, T. E. and M. W. McCracken (2009). Tests of Equal Predictive Ability With Real-Time Data. *Journal of Business & Economic Statistics* 27(4), 441–454.
- Clements, M. P. and A. B. Galvão (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics* 26, 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.
- Clements, M. P. and N. Taylor (2001). Bootstrapping prediction intervals for autoregressive models. *International Journal of Forecasting* 17(2), 247–267.
- Croushore, D. and T. Stark (2001). A real-time data set for macroeconomists. *Journal of Econometrics* 105(1), 111–130.
- Davison, A. and D. Hinkley (1997). *Bootstrap Methods and their Applications*. Cambridge University Press.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13(3), 253–63.
- Evans, M. D. (2005). Where are we now? Real-time estimates of the macro economy. *International Journal of Central Banking* 1(2), 127–175.
- Froni, C., M. Marcellino, and C. Schumacher (2013). U-midas: Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society - Series A* (forthcoming).
- Garratt, A., K. Lee, M. H. Pesaran, and Y. Shin (2003). Forecast uncertainties in macroeconomic modeling: An application to the UK economy. *Journal of the American Statistical Association* 98(464), 829–38.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. CIRANO Working Papers 2004s-20.

- Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a risk-return trade-off after all. *Journal of Financial Economics* 76(3), 509–548.
- 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(1-2), 59–95.
- Giacomini, R. and H. White (2006). Tests of Conditional Predictive Ability. *Econometrica* 74(6), 1545–1578.
- 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.
- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association* 106(494), 746–762.
- Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102, 359–378.
- Goncalves, S. and L. Kilian (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics* 123(1), 89–120.
- Hall, S. G. and J. Mitchell (2007). Combining density forecasts. *International Journal of Forecasting* 23(1), 1–13.
- Inoue, A. and L. Kilian (2002). Bootstrapping autoregressive processes with possible unit roots inoue. *Econometrica*, 377–391.
- Jore, A. S., J. Mitchell, and S. P. Vahey (2010). Combining forecast densities from VARs with uncertain instabilities. *Journal of Applied Econometrics* 25(4), 621–634.
- Kascha, C. and F. Ravazzolo (2010). Combining inflation density forecasts. *Journal of Forecasting* 29(1-2), 231–250.
- 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, 04). Pooling versus model selection for nowcasting GDP with many predictors: empirical evidence for six industrialized countries. *Journal of Applied Econometrics* 28(3), 392–411.

- Mariano, R. S. and Y. Murasawa (2010). A coincident index, common factors, and monthly real GDP. *Oxford Bulletin of Economics and Statistics* 72(1), 27–46.
- Mazzi, G. L., J. Mitchell, and G. Montana (2013). Density nowcasts and model combination: nowcasting Euro-area GDP growth over the 2008-9 recession. *Oxford Bulletin of Economics and Statistics* (forthcoming).
- Mitchell, J. and S. G. Hall (2005). Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation. *Oxford Bulletin of Economics and Statistics* 67(s1), 995–1033.
- Panagiotelis, A. and M. Smith (2008). Bayesian density forecasting of intraday electricity prices using multivariate skew t distributions. *International Journal of Forecasting* 24(4), 710–727.
- Ravazzolo, F. and S. Vahey (2013). Forecast densities for economic aggregates from disaggregate ensembles. *Studies of Nonlinear Dynamics and Econometrics* (forthcoming).
- Romer, C. D. and D. H. Romer (2000). Federal reserve information and the behavior of interest rates. *American Economic Review* 90(3), 429–457.
- Stark, T. and D. Croushore (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics* 24(4), 507–531.
- Tay, A. and K. F. Wallis (2000). Density Forecasting: A Survey. *Journal of Forecasting* 19, 235–254.

Table 2: Forecast accuracy for the Monte Carlo experiment (DGP: normally distributed shocks)

	Specification			Specification			Specification					
	ρ	δ_h	δ_l	ρ	δ_h	δ_l	ρ	δ_h	δ_l			
	0.1	0	0.5	0.5	0	0.5	0.9	0	0.5			
		MSPE CRPS LS			MSPE CRPS LS			MSPE CRPS LS				
	umidas	0.104	0.182	-0.301	umidas	0.143	0.214	-0.472	umidas	0.633	0.454	-1.214
	ar-umidas	0.105	0.183	-0.302	ar-umidas	0.137	0.210	-0.448	ar-umidas	0.258	0.288	-0.753
	median	0.102	0.181	-0.292	median	0.137	0.210	-0.449	median	0.547	0.420	-1.139
Mixed	ar-midas	0.104	0.182	-0.302	ar-midas	0.149	0.219	-0.492	ar-midas	0.334	0.329	-0.886
Frequency	umidas	0.105	0.183	-0.303	umidas	0.146	0.216	-0.483	umidas	0.663	0.461	-1.239
	ar-umidas	0.106	0.183	-0.308	ar-umidas	0.139	0.211	-0.461	ar-umidas	0.262	0.289	-0.760
	mean	0.103	0.181	-0.296	mean	0.139	0.211	-0.464	mean	0.570	0.427	-1.157
	ar-midas	0.105	0.183	-0.304	ar-midas	0.152	0.220	-0.508	ar-midas	0.345	0.331	-0.895
	AR	0.997	0.566	-1.431	AR	0.905	0.539	-1.380	AR	0.301	0.310	-0.833
	RW	1.928	0.773	-1.743	RW	1.335	0.656	-1.574	RW	0.321	0.320	-0.863
Benchmarks	AR	1.008	0.568	-1.433	AR	0.913	0.540	-1.383	AR	0.305	0.312	-0.838
	RW	1.983	0.782	-1.754	RW	1.354	0.658	-1.580	RW	0.324	0.322	-0.867

Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on different criteria: MSPE for point forecasts, CRPS and LS for density forecasts. The criteria are described in Section 3.2. The results are computed as averages across the 1000 replications, with an evaluation sample equal to 50 low frequency periods (quarters). The different DGP specifications are reported in the header of the table. The results from two benchmarks (an AR process with lag length selected according to the BIC criterion and a Random Walk (RW)) are included for comparison.

Table 3: Forecast accuracy for the Monte Carlo experiment (DGP: Student's t-distributed shocks)

	Specification				Specification				Specification			
	ρ	δ_h	δ_l		ρ	δ_h	δ_l		ρ	δ_h	δ_l	
	0.1	0	0.5		0.5	0	0.5		0.9	0	0.5	
		MSPE CRPS LS				MSPE CRPS LS				MSPE CRPS LS		
	umidas	0.167	0.225	-0.564	umidas	0.229	0.267	-0.720	umidas	1.042	0.584	-1.473
	ar-umidas	0.167	0.225	-0.567	ar-umidas	0.221	0.262	-0.698	ar-umidas	0.422	0.363	-1.023
	median	0.164	0.223	-0.558	median	0.219	0.260	-0.705	median	0.882	0.534	-1.393
	ar-midas	0.167	0.225	-0.550	ar-midas	0.241	0.273	-0.752	ar-midas	0.553	0.418	-1.145
	umidas	0.176	0.227	-0.696	umidas	0.242	0.270	-0.821	umidas	1.117	0.593	-1.528
	ar-umidas	0.178	0.228	-0.696	ar-umidas	0.232	0.264	-0.808	ar-umidas	0.438	0.366	-1.095
	mean	0.174	0.225	-0.702	mean	0.231	0.263	-0.807	mean	0.948	0.546	-1.451
	ar-midas	0.177	0.227	-0.658	ar-midas	0.256	0.278	-0.859	ar-midas	0.571	0.420	-1.188
	AR	1.630	0.705	-1.684	AR	1.462	0.678	-1.626	AR	0.493	0.394	-1.085
	RW	3.191	1.003	-2.016	RW	2.178	0.834	-1.827	RW	0.524	0.408	-1.113
	AR	1.684	0.709	-1.765	AR	1.503	0.682	-1.663	AR	0.513	0.398	-1.125
	RW	3.297	1.006	-2.046	RW	2.247	0.839	-1.843	RW	0.547	0.412	-1.149

Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on different criteria: MSPE for point forecasts, CRPS and LS for density forecasts. The criteria are described in Section 3.2. The results are computed as averages across the 1000 replications, with an evaluation sample equal to 50 low frequency periods (quarters). The different DGP specifications are reported in the header of the table. The results from two benchmarks (an AR process with lag length selected according to the BIC criterion and a Random Walk (RW)) are included for comparison.

Table 4: Forecast accuracy for the Monte Carlo experiment (DGP: GARCH(1,1) distributed shocks)

	Specification				Specification				Specification			
	ρ	δ_h	δ_l		ρ	δ_h	δ_l		ρ	δ_h	δ_l	
	0.1	0	0.5		0.5	0	0.5		0.9	0	0.5	
		MSPE CRPS LS				MSPE CRPS LS				MSPE CRPS LS		
	umidas	0.034	0.104	0.241	umidas	0.046	0.121	0.086	umidas	0.216	0.266	-0.700
	ar-umidas	0.034	0.104	0.239	ar-umidas	0.043	0.118	0.115	ar-umidas	0.084	0.163	-0.216
	median	0.033	0.103	0.242	median	0.043	0.118	0.109	median	0.186	0.246	-0.617
	ar-midas	0.035	0.105	0.231	ar-midas	0.055	0.133	-0.002	ar-midas	0.144	0.214	-0.469
Mixed	umidas	0.034	0.104	0.225	umidas	0.047	0.122	0.062	umidas	0.230	0.271	-0.731
Frequency	ar-umidas	0.034	0.105	0.226	ar-umidas	0.044	0.118	0.085	ar-umidas	0.086	0.164	-0.241
	mean	0.034	0.104	0.226	mean	0.044	0.118	0.085	mean	0.199	0.251	-0.654
	ar-midas	0.035	0.106	0.221	ar-midas	0.060	0.138	-0.075	ar-midas	0.150	0.217	-0.492
	AR	0.936	0.547	-1.398	AR	0.813	0.510	-1.328	AR	0.125	0.199	-0.400
	RW	1.841	0.769	-1.737	RW	1.190	0.618	-1.521	RW	0.129	0.204	-0.415
Benchmarks	AR	0.949	0.550	-1.403	AR	0.826	0.513	-1.332	AR	0.127	0.201	-0.404
	mean	1.854	0.769	-1.739	mean	1.212	0.622	-1.525	mean	0.131	0.204	-0.418
	RW				RW				RW			

Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on different criteria: MSPE for point forecasts, CRPS and LS for density forecasts. The criteria are described in Section 3.2. The results are computed as averages across the 1000 replications, with an evaluation sample equal to 50 low frequency periods (quarters). The different DGP specifications are reported in the header of the table. The results from two benchmarks (an AR process with lag length selected according to the BIC criterion and a Random Walk (RW)) are included for comparison.

Table 5: Forecast accuracy for US output. Evaluation sample: 1985q2:2013q2

	<i>hm</i>	CRPS				Log Score			
		<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>
'EMP'	umidas	0.262	0.281	0.307	0.289	-0.683	-0.745*	-0.837	-0.871
	ar-umidas	0.261	0.281	0.309	0.293	-0.673	-0.726**	-0.838	-0.866
	midas	0.271	0.279	0.295	0.288	-0.739	-0.766	-0.802	-0.842
	ar-midas	0.270	0.281	0.285	0.313	-0.738	-0.767	-0.758**	-0.925
	ADL	0.249	0.279	0.284	0.287	-0.640	-0.833	-0.842	-0.859
'IP'	umidas	0.240	0.251	0.270	0.276	-0.619	-0.668*	-0.760	-0.821
	ar-umidas	0.237	0.249	0.270	0.277	-0.618	-0.668**	-0.763	-0.812
	midas	0.235	0.248	0.267	0.265	-0.610	-0.661**	-0.752	-0.782
	ar-midas	0.237	0.244	0.261	0.301	-0.619	-0.649**	-0.722	-0.869
	ADL	0.220	0.256	0.257	0.261	-0.566	-0.750	-0.744	-0.756
'CU'	umidas	0.247	0.250	0.278	0.277	-0.665	-0.676	-0.775	-0.813
	ar-umidas	0.242	0.248	0.274	0.275	-0.652	-0.663**	-0.755	-0.814
	midas	0.240	0.248	0.273	0.265	-0.646	-0.675	-0.764	-0.770
	ar-midas	0.242	0.246	0.269	0.302	-0.643	-0.66**	-0.748	-0.878
	ADL	0.235	0.259	0.263	0.268	-0.619	-0.752	-0.769	-0.784
'PMI'	umidas	0.271	0.285	0.295	0.309	-0.781	-0.797**	-0.824	-0.870
	ar-umidas	0.271	0.282*	0.295	0.317	-0.791	-0.787***	-0.817*	-0.902
	midas	0.250*	0.259***	0.272**	0.285	-0.738	-0.763***	-0.833	-0.880
	ar-midas	0.352	0.391	0.417	0.410	-1.056	-1.175	-1.203	-1.188
	ADL	0.271	0.304	0.300	0.305	-0.767	-0.883	-0.869	-0.883
'CFNAI'	umidas	0.213	0.221**	0.242	0.257	-0.519*	-0.551***	-0.623**	-0.734**
	ar-umidas	0.213	0.220**	0.238*	0.256	-0.532*	-0.545***	-0.598***	-0.749
	midas	0.210	0.215***	0.235	0.254	-0.507**	-0.524***	-0.586***	-0.715**
	ar-midas	0.208*	0.215***	0.236*	0.284	-0.521**	-0.543***	-0.587***	-0.825
	ADL	0.225	0.260	0.261	0.262	-0.617	-0.743	-0.748	-0.758
'PFBOS'	umidas	0.267	0.268	0.265	0.277	-0.720	-0.729***	-0.745**	-0.786
	ar-umidas	0.267	0.268	0.274	0.280	-0.723	-0.718**	-0.760*	-0.800
	midas	0.255	0.258	0.263	0.278	-0.684*	-0.703***	-0.719**	-0.795
	ar-midas	0.256**	0.260	0.261	0.276	-0.692**	-0.698***	-0.720***	-0.797
	ADL	0.270	0.269	0.284	0.287	-0.732	-0.726	-0.851	-0.852
AR		0.319	0.318	0.317	0.352	-0.993	-0.996	-0.995	-1.095

Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on CRPS and LS for evaluating density forecasts. The criteria are described in Section 3.2. The results are computed on an evaluation sample from 1985q2 to 2013q2. The results are reported for each of the monthly indicators included in the models. Results for the corresponding ADL model are reported for each indicator and serve as a comparison in our tests. Moreover, the results from an AR process with lag length selected according to the BIC criterion are included for comparison. The table shows the results for the backcast and nowcast. Bold numbers indicate that a MIDAS provide more accurate forecasts than the ADL and AR benchmarks. Differences in accuracy that are statistically different from zero are denoted by one, two or three stars corresponding to significance levels of 10%, 5% and 1%, respectively. The underlying p -values are based on t -statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992). Since the ar-umidas and ar-midas models nest the ADL model, we report p -values based on one-sided tests, taking the ADL as the null and the MIDAS models as the alternative. Because the umidas and midas models are not nested, for these model comparisons we report p -values based on two-sided tests.

Table 6: Forecast accuracy for US output. Evaluation Sample: 1985q2:2007q4

	<i>hm</i>	CRPS				Log Score			
		<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>
'EMP'	umidas	0.260	0.276	0.294	0.271	-0.678	-0.735	-0.794	-0.817
	ar-umidas	0.258	0.272	0.297	0.271	-0.668	-0.706**	-0.798	-0.817
	midas	0.261	0.268	0.280	0.268	-0.708	-0.733	-0.764	-0.807
	ar-midas	0.259	0.271	0.272	0.288	-0.708	-0.740	-0.730**	-0.863
	ADL	0.250	0.261	0.264	0.268	-0.647	-0.787	-0.791	-0.814
'IP'	umidas	0.224	0.229	0.248	0.280	-0.570	-0.604**	-0.693	-0.837
	ar-umidas	0.222	0.229	0.250	0.280	-0.577	-0.602***	-0.686	-0.823
	midas	0.220	0.229	0.246	0.264	-0.562	-0.601**	-0.685	-0.783
	ar-midas	0.220	0.226	0.243	0.276	-0.558	-0.585***	-0.658	-0.809
	ADL	0.212	0.247	0.246	0.249	-0.549	-0.732	-0.721	-0.733
'CU'	umidas	0.226	0.227	0.257	0.275	-0.589	-0.601**	-0.709	-0.807
	ar-umidas	0.224	0.226	0.254	0.271	-0.583	-0.589***	-0.693	-0.801
	midas	0.219	0.224*	0.252	0.258	-0.571	-0.595**	-0.693	-0.753
	ar-midas	0.223	0.224*	0.252	0.277	-0.587	-0.587**	-0.692	-0.816
	ADL	0.217	0.246	0.248	0.250	-0.567	-0.714	-0.728	-0.734
'PMI'	umidas	0.261	0.275	0.275	0.305	-0.755	-0.76**	-0.775**	-0.858
	ar-umidas	0.262	0.273*	0.278	0.315	-0.777	-0.755***	-0.77**	-0.889
	midas	0.24**	0.245***	0.251**	0.262**	-0.712	-0.729***	-0.790	-0.846
	ar-midas	0.303	0.340	0.365	0.353	-0.927	-1.019	-1.076	-1.046
	ADL	0.267	0.292	0.290	0.293	-0.758	-0.852	-0.839	-0.850
'CFNAI'	umidas	0.201	0.206**	0.229	0.245	-0.474**	-0.499***	-0.594**	-0.707**
	ar-umidas	0.202	0.206**	0.221*	0.246	-0.477**	-0.499***	-0.548***	-0.727
	midas	0.193	0.197***	0.219*	0.241	-0.454***	-0.451***	-0.539***	-0.689**
	ar-midas	0.193***	0.198***	0.219*	0.251	-0.461***	-0.474	-0.549***	-0.753
	ADL	0.221	0.252	0.251	0.253	-0.611	-0.723	-0.727	-0.739
'PFBOS'	umidas	0.261	0.262	0.263	0.275	-0.708	-0.711	-0.740**	-0.784
	ar-umidas	0.262	0.263	0.271	0.275	-0.709	-0.703	-0.761*	-0.796
	midas	0.252	0.254	0.262	0.273	-0.681	-0.695	-0.723**	-0.791
	ar-midas	0.252*	0.253*	0.257	0.260	-0.675*	-0.678**	-0.719***	-0.772**
	ADL	0.263	0.261	0.271	0.274	-0.712	-0.706	-0.830	-0.834
AR		0.301	0.300	0.301	0.345	-0.972	-0.975	-0.972	-1.107

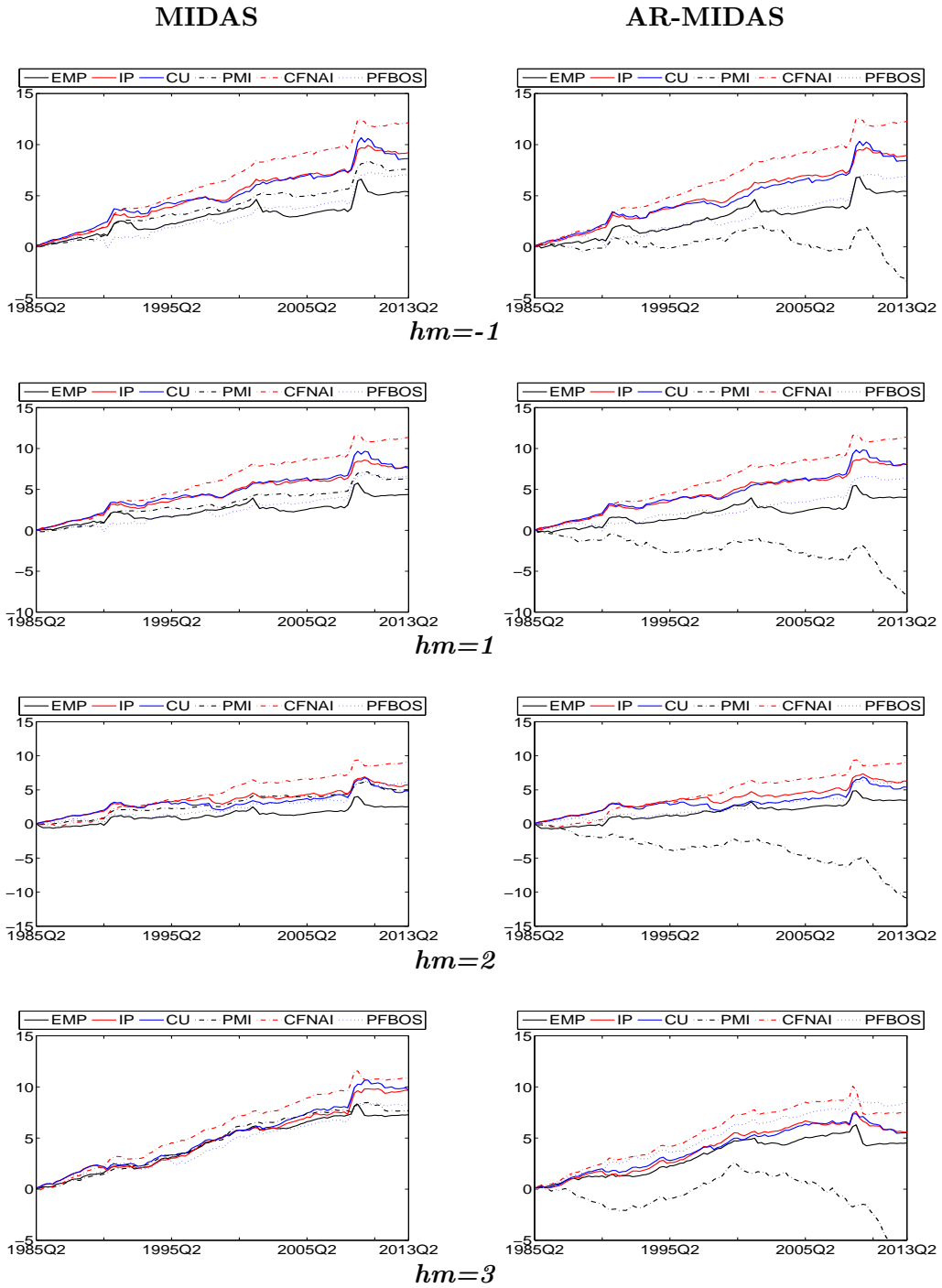
Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on CRPS and LS for evaluating density forecasts. The criteria are described in Section 3.2. The results are computed on an evaluation sample from 1985q2 to 2007q4. The results are reported for each of the monthly indicators included in the models. Results for the corresponding ADL model are reported for each indicator and serve as a comparison in our tests. Moreover, the results from an AR process with lag length selected according to the BIC criterion are included for comparison. The table shows the results for the backcast and nowcast. Bold numbers indicate that a MIDAS provide more accurate forecasts than the ADL and AR benchmarks. Differences in accuracy that are statistically different from zero are denoted by one, two or three stars corresponding to significance levels of 10%, 5% and 1%, respectively. The underlying p -values are based on t -statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992). Since the ar-umidas and ar-midas models nest the ADL model, we report p -values based on one-sided tests, taking the ADL as the null and the MIDAS models as the alternative. Because the umidas and midas models are not nested, for these model comparisons we report p -values based on two-sided tests.

Table 7: Forecast accuracy for US output. Point forecasts. Evaluation sample: 1985q2:2013q2 and 1985q2:2007q4

		1985 - 2013				1985 - 2007			
	<i>hm</i>	<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>-1</i>	<i>1</i>	<i>2</i>	<i>3</i>
'EMP'	umidas	0.458	0.489	0.543	0.498	0.452	0.475	0.511	0.442
	ar-umidas	0.457	0.494	0.547	0.505	0.452	0.477	0.521	0.446
	midas	0.478	0.490	0.524	0.495	0.459	0.470	0.493	0.441
	ar-midas	0.476	0.494	0.509	0.552	0.451	0.472	0.477	0.480
	ADL	0.432	0.474	0.490	0.494	0.435	0.432	0.441	0.439
'IP'	umidas	0.419	0.437	0.471	0.462	0.383	0.388	0.427	0.467
	ar-umidas	0.413	0.434	0.473	0.466	0.378	0.389	0.430	0.470
	midas	0.408	0.430	0.465	0.445	0.371	0.388	0.425	0.440
	ar-midas	0.414	0.429	0.457	0.535	0.370	0.385	0.416	0.470
	ADL	0.369	0.433	0.436	0.447	0.350	0.404	0.403	0.409
'CU'	umidas	0.430	0.445	0.494	0.470	0.387	0.389	0.449	0.465
	ar-umidas	0.421	0.436	0.484	0.466	0.381	0.384	0.440	0.460
	midas	0.420	0.440	0.483	0.447	0.375	0.383	0.440	0.434
	ar-midas	0.416	0.432	0.475	0.532	0.376	0.381	0.439	0.468
	ADL	0.399	0.438	0.443	0.457	0.363	0.403	0.406	0.409
'PMI'	umidas	0.466	0.500	0.520	0.558	0.441	0.482	0.482	0.540
	ar-umidas	0.467	0.495	0.521	0.579	0.447	0.479	0.488	0.569
	midas	0.418**	0.437**	0.450**	0.473*	0.386**	0.405**	0.393**	0.405**
	ar-midas	0.615	0.677	0.730	0.727	0.525	0.591	0.644	0.613
	ADL	0.473	0.530	0.522	0.532	0.461	0.503	0.495	0.503
'CFNAI'	umidas	0.375	0.395	0.436	0.432	0.342	0.358	0.399	0.403
	ar-umidas	0.379	0.392	0.426	0.432	0.347	0.355	0.386	0.402
	midas	0.362	0.379	0.420	0.430	0.322	0.336*	0.380	0.397
	ar-midas	0.361	0.378*	0.424	0.514	0.319**	0.335**	0.383	0.415
	ADL	0.373	0.437	0.442	0.440	0.358	0.419	0.420	0.415
'PFBOS'	umidas	0.467	0.469	0.456	0.473	0.460	0.445	0.459	0.444
	ar-umidas	0.468	0.471	0.476	0.477	0.458	0.469	0.474	0.451
	midas	0.439**	0.445*	0.454	0.471	0.439	0.448	0.455	0.427
	ar-midas	0.447**	0.455*	0.452	0.474	0.437*	0.447	0.449	0.402
	ADL	0.476	0.471	0.489	0.496	0.462	0.457	0.440	0.450
AR		0.535	0.529	0.533	0.564	0.458	0.453	0.461	0.523

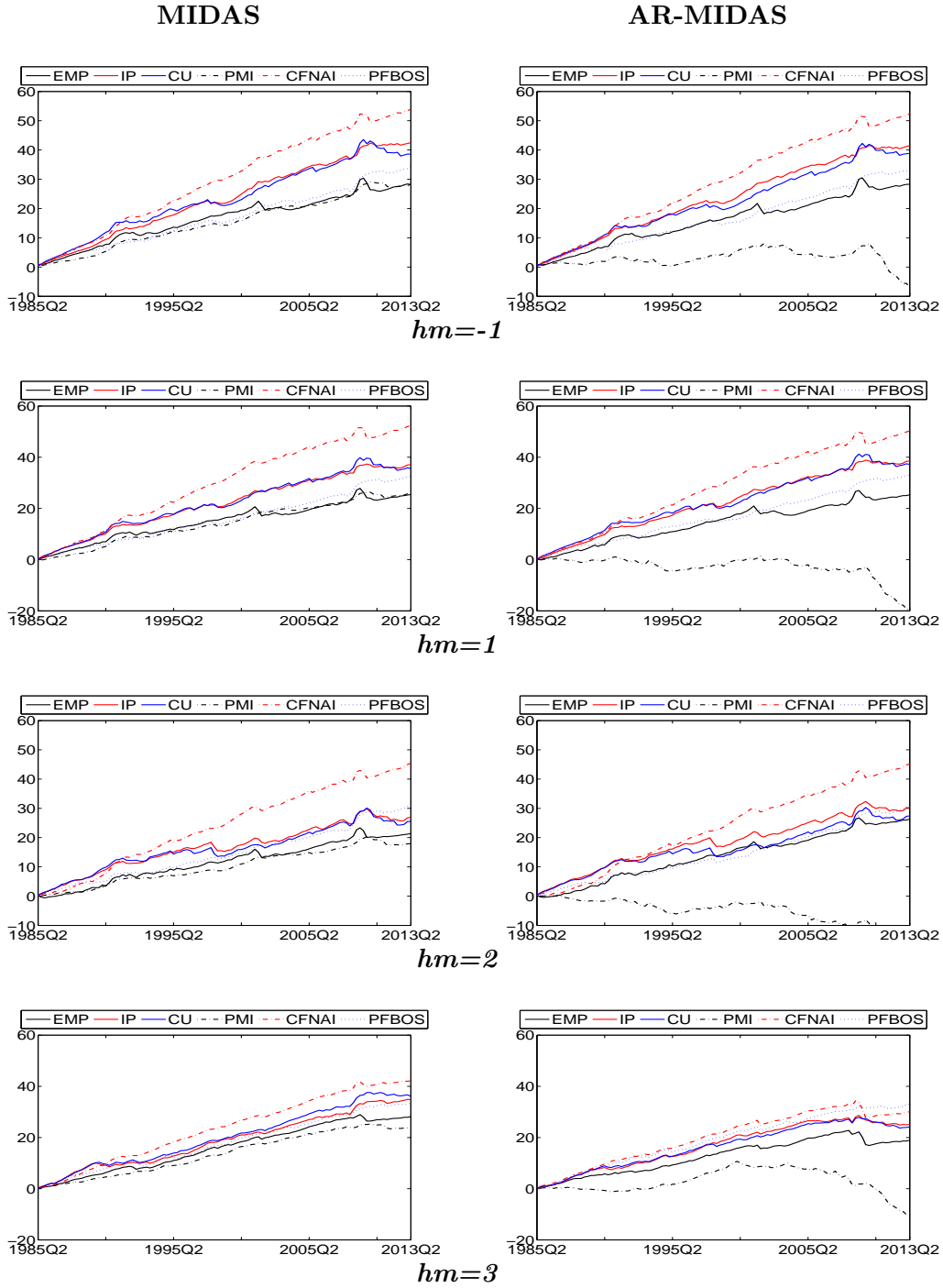
Note: The table compares the forecasting performance of the four types of MIDAS models described in Section 2 (umidas, ar-umidas, midas, ar-midas), based on RMSPE for evaluating point forecasts. The criterion is described in Section 3.2. The results are computed on an evaluation sample from 1985q2 to 2013q2 and on 1985q2 to 2007q4. The results are reported for each of the monthly indicators included in the models. Results for the corresponding ADL model are reported for each indicator and serve as a comparison in our tests. Moreover, the results from an AR process with lag length selected according to the BIC criterion are included for comparison. The table shows the results for the backcast and nowcast. Bold numbers indicate that a MIDAS provide more accurate forecasts than the ADL and AR benchmarks. Differences in accuracy that are statistically different from zero are denoted by one, two or three stars corresponding to significance levels of 10%, 5% and 1%, respectively. The underlying p -values are based on t -statistics computed with a serial correlation-robust variance, using the pre-whitened quadratic spectral estimator of Andrews and Monahan (1992). Since the ar-umidas and ar-midas models nest the ADL model, we report p -values based on one-sided tests, taking the ADL as the null and the MIDAS models as the alternative. Because the umidas and midas models are not nested, for these model comparisons we report p -values based on two-sided tests.

Figure 1: Relative Cumulative CRPS



Note: Relative Cumulative CRPS = $\sum_1^{t+h(m)} (CRPS_{t+h(m)}^l - CRPS_{t+h(m)}^i)$, $t = 1, \dots, T$, $l =$ AR benchmark model and $i =$ alternative information set. The first column shows the results for the MIDAS model and the second column the results for the AR-MIDAS models. The different rows show the results for different forecasting horizons.

Figure 2: Relative Cumulative Log Score



Note: Relative Cumulative Log Score = $\sum_1^{t+h(m)} (\ln f_{t_m+h_m, t_m, i}(y_{t+h_m}) - \ln f_{t_m+h_m, t_m, l}(y_{t+h_m}))$, $t = 1, \dots, T$, $l =$ AR benchmark model and $i =$ alternative information set. The first column shows the results for the MIDAS model and the second column the results for the AR-MIDAS models. The different rows show the results for different forecasting horizons.