Supplementary Appendix: Density forecasts with MIDAS models*

Knut Are Aastveit† Norges Bank

Claudia Foroni‡ European Central Bank

Francesco Ravazzolo§ Free University of Bozen/Bolzano

July 28, 2016

The appendix contains supplementary material for the paper "Density forecasts with MIDAS models". In the first section, we report the results of additional Monte Carlo simulations. In the next section, we provide additional results from our empirical exercise, using alternative measures of absolute and relative density forecast accuracy, as well as point forecast accuracy. In the last section, we document the results of a case study, where we investigate how weekly updates of financial data change the nowcasts of real output growth for the U.S. during 2008q3 and 2008q4.

JEL codes: C10, C53, E37

Keywords: Mixed Data Sampling, Density Forecasts, Bootstrapping, Nowcasting

*The views expressed are those of the authors and do not necessarily reflect those of the European Central Bank or Norges Bank.

†Knut-Are.Aastveit@norges-bank.no

‡Claudia.Foroni@ecb.europa.eu

§Francesco.Ravazzolo@unibz.it

A Additional results - Simulation exercise

In this section, we provide results from additional Monte Carlo simulations. First, we report results from robustness checks with respect to the data generating process of the high-frequency indicator and the estimation sample size, respectively. Then we report the results of Monte Carlo simulations where we assume that the data generating process follows a UMIDAS specification, instead of following a MIDAS specification as in Section 4 in the main text.

Results when the high-frequency indicator follows a heteroskedastic process

In the Monte Carlo simulations, reported in Section 4 in the main text, we assumed that the high-frequency regressor followed an autoregressive process. This is typically the case when the high-frequency regressor is a macroeconomic variable. Here, we will report results when the high-frequency regressor instead follows a heteroskedastic process. This is often the case for financial variables. More precisely, we will assume that the high-frequency regressor follows the heteroskedastic process: $x^m_{t_m} = \sqrt{(\sigma^m_{t_m})^2}\epsilon^m_{t_m}$, $(\sigma^m_{t_m})^2 = c_0 + c_1(x^m_{t_m-1})^2$, $c_0 = 0.25, c_1 = 0.85, \epsilon^m_{t_m} \sim N.i.i.d(0,1)$. We report results in Table A.1 for four different cases of the disturbance, u_t:

1. $u_t \sim N.i.i.d(0,1)$
2. $u_t = \varepsilon_t + 0.5\varepsilon_{t-1}$, where $\varepsilon_t \sim N.i.i.d(0,1)$
3. $u_t = \sqrt{h_t}\varepsilon_t$, $h_t = 0.1 + 0.3u^2_{t-1} + 0.6h_{t-1}$, where $\varepsilon_t \sim N.i.i.d(0,1)$
4. $u_t \sim N.i.i.d(-1,1)$ with probability 0.9 or $u_t \sim N.i.i.d(9,1)$ with probability 0.1.

Comparing the results in Table A.1 with the results in Table 1 in the main text, the specification for the exogenous high-frequency process does not seem to matter. That is, we obtain very similar results when the high-frequency regressor follows an AR process as when it is heteroskedastic. This is not very surprising as the high-frequency regressor is exogenous and NLS is a consistent and efficient estimator for the MIDAS model, see Andreou et al. (2010).

Table A.1. Coverage rates for MIDAS simulations with various DGPs and heteroskedastic high-frequency indicator

		90 percent coverage				70 percent coverage			
		Coverage	**SE**	**Lower**	**Upper**	**Coverage**	**SE**	**Lower**	**Upper**
	IID	88.83	0.15	5.64	5.53	68.91	0.21	15.49	15.60
DGP 1	**Normal**	88.41	0.15	5.87	5.72	67.97	0.22	16.00	16.03
	Block Wild	89.83	0.14	5.08	5.09	70.10	0.20	14.83	15.07
	IID	88.60	0.15	5.56	5.84	68.59	0.22	15.60	15.81
DGP 2	**Normal**	88.17	0.16	5.79	6.04	67.64	0.23	16.04	16.32
	Block Wild	90.09	0.14	4.97	4.94	70.59	0.23	14.70	14.71
	IID	88.90	0.14	5.38	5.72	68.83	0.21	15.62	15.55
DGP 3	**Normal**	88.54	0.14	5.61	5.85	68.06	0.21	16.01	15.93
	Block Wild	89.98	0.13	4.92	5.10	69.99	0.21	15.00	15.01
	IID	90.30	0.13	3.48	6.22	73.90	0.20	12.81	13.29
DGP 4	**Normal**	89.86	0.13	0.05	10.09	87.50	0.15	2.34	10.16
	Block Wild	90.65	0.13	3.22	6.14	74.10	0.20	12.64	13.26

Note: The table reports associated coverage rates for simulation exercises where the DGP follows a MIDAS model with various assumptions for the DGP disturbance. DGP 1 refers to the disturbance process $u_t \sim N.i.i.d(0,1)$, DGP 2 refers to the disturbance process $u_t = 0.5u_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim N.i.i.d(0,1)$, DGP 3 refers to the disturbance process $u_t = \sqrt{h_t}\varepsilon_t = \sqrt{0.1 + 0.3u_{t-1}^2 + 0.6\varepsilon_{t-1}\varepsilon_t}$, where $\varepsilon_t \sim N.i.i.d(0,1)$, and DGP 4 refers to a process where the disturbance is drawn from a mixture of two distributions, $u_t \sim N.i.i.d(-1,1)$ with probability 0.9 or $u_t \sim N.i.i.d(9,1)$ with probability 0.1. Results are provided for MIDAS models using two different bootstrapping methods, the residual-based bootstrap (IID) and the block wild bootstrap, as well as for constructing predictive densities by drawing from a normal distribution. We report nominal coverage rates, measured as the mean nominal coverage of the 1000 simulations, for the 70 percent and 90 percent fan charts, their associated standard errors and the mean proportions below and above the nominal coverage.

Results for various sample sizes

As a robustness check, we have experimented with increasing the estimation sample (T). Table A.2 reports results in the case of DGP 1 (from the main text) where we change the estimation sample T to either 50 or 300. When $T = 50$, the relative improvement from the block wild bootstrap compared with the other two approaches, in terms of actual coverage levels, are larger than in our baseline simulation exercise (where $T = 100$). In contrast, for the larger estimation sample, the three methods provide very similar and accurate actual coverage levels. This suggests that the block wild bootstrap has better small-sample properties than the residual-based bootstrap for the MIDAS model.

Finally, in the main text, we suggest following Davidson and MacKinnon (2006) and rescale

the residuals so that they have the correct variance by $\ddot{\varepsilon}_t \equiv \left(\frac{n}{n-k}\right)^{0.5} \widehat{\varepsilon}_t$. Results documented in Table A.2, suggest that rescaling the residuals provides modest improvements, in terms of more accurate density forecasts, for the block wild and residual-based bootstrap.

Table A.2. Robustness with respect to sample size. Coverage rates for DGP MIDAS normal case

		90 percent coverage				**70 percent coverage**			
		Coverage	**SE**	**Lower**	**Upper**	**Coverage**	**SE**	**Lower**	**Upper**
	IID	88.68	0.14	5.74	5.59	68.65	0.20	15.71	15.64
T=100	**Normal**	88.22	0.14	5.98	5.80	67.91	0.20	16.18	15.91
	Block Wild	89.66	0.13	5.19	5.15	69.93	0.20	15.14	14.94
	IID	88.09	0.14	5.99	5.93	68.21	0.19	15.90	15.89
T=50	**Normal**	87.31	0.15	6.46	6.22	66.79	0.20	16.62	16.59
	Block Wild	89.74	0.14	5.15	5.10	70.08	0.21	14.87	15.06
	IID	89.38	0.14	5.31	5.31	69.63	0.20	15.10	15.27
T=300	**Normal**	89.31	0.14	5.31	5.38	69.41	0.21	15.21	15.38
	Block Wild	89.80	0.13	5.12	5.08	69.80	0.21	15.10	15.09
	IID	88.28	0.14	5.75	5.97	68.22	0.20	15.75	16.03
No Rescale	**Normal**	88.57	0.14	5.58	5.85	68.32	0.20	15.72	15.96
	Block Wild	89.23	0.14	5.55	5.22	69.08	0.22	15.63	15.29

Note: The table reports associated coverage rates for simulation exercises with different estimation sample sizes T. where the DGP follows a MIDAS with the disturbance $u_t \sim N.i.i.d(0.1)$. The table reports results for $T = 50$ and $T = 300$ where $x^m_{t_m} = 0.9x^m_{t_m-1} + \epsilon^m_{t_m}$. where $\epsilon^m_{t_m} \sim N.i.i.d(0.1)$. Moreover, No Rescale, refers to simulations with $T = 100$, but where we do not rescale the residuals. Results are provided for two different bootstrapping methods, the residual-based bootstrap (IID) and the block wild bootstrap, as well as for constructing predictive densities by drawing from a normal distribution. We report nominal coverage rates, measured as the mean nominal coverage of the 1000 simulations, for the 70 percent and 90 percent fan charts, their associated standard errors and the mean proportions below and above the nominal coverage.

Simulation results when DGP is a UMIDAS model

In Section 4 in the main text, we report simulation results assuming that the DGP follows a MIDAS. Here, we instead assume that the DGP follows a UMIDAS model.

The Monte Carlo design is exactly as in Section 4.1 of the paper. The only difference is the way we generate the data, this time assuming that the true DGP follows a UMIDAS regression model. In particular, y_t is generated as follows:

$$y_t = \beta_0 + \beta_1 x_{t-1} + \beta_1^2 x_{t-2} + \beta_1^3 x_{t-3} + \beta_1^4 x_{t-4} + \beta_1^5 x_{t-5} + \beta_1^6 x_{t-6} + u_t, \qquad t = 1, 2, \ldots, T \quad \text{(A.1)}$$

where $\beta_0 = 0$ and $\beta_1 = 0.8$. This scheme is chosen because it mimics well the behavior of macroeconomic series, in which typically the most recent observations have more importance, and therefore the weights in the lags slowly decay.

Tables A.3 and A.4 report results when the high-frequency regressor follows an autocorrelated process and a heteroskedastic process, respectively. In both cases, we report results for four different cases of the disturbance, u_t:

1. $u_t \sim N.i.i.d(0, 1)$
2. $u_t = \varepsilon_t + 0.5\varepsilon_{t-1}$, where $\varepsilon_t \sim N.i.i.d(0, 1)$
3. $u_t = \sqrt{h_t}\varepsilon_t$, $h_t = 0.1 + 0.3u_{t-1}^2 + 0.6h_{t-1}$, where $\varepsilon_t \sim N.i.i.d(0, 1)$
4. $u_t \sim N.i.i.d(-1, 1)$ with probability 0.9 or $u_t \sim N.i.i.d(9, 1)$ with probability 0.1.

The simulation results for the UMIDAS are in general very similar to the simulation results for the MIDAS, reported in Section 4.2 in the main text. In all cases, the actual coverage levels from the block wild bootstrap are very close to the nominal coverage levels, and considerably closer to the nominal coverage levels than actual coverage levels provided by the two other methods for computing predictive densities.

Table A.3. Coverage rates for UMIDAS simulations with various DGPs and autocorrelated high-frequency indicator

		90 percent coverage				70 percent coverage			
		Coverage	**SE**	**Lower**	**Upper**	**Coverage**	**SE**	**Lower**	**Upper**
	IID	88.85	0.14	5.46	5.68	69.01	0.21	15.35	15.64
DGP 1	**Normal**	88.24	0.14	5.75	6.01	67.98	0.21	15.90	16.12
	Block Wild	89.82	0.14	5.33	4.85	70.06	0.21	15.18	14.76
	IID	88.65	0.15	5.57	5.79	68.61	0.23	15.59	15.80
DGP 2	**Normal**	88.03	0.16	5.88	6.09	67.57	0.23	16.08	16.34
	Block Wild	89.82	0.15	5.10	5.08	70.31	0.22	14.96	14.74
	IID	88.75	0.14	5.54	5.71	68.74	0.21	15.59	15.67
DGP 3	**Normal**	88.21	0.15	5.82	5.97	67.72	0.21	16.17	16.11
	Block Wild	89.63	0.13	5.16	5.21	70.05	0.22	14.92	15.03
	IID	90.44	0.14	4.03	5.54	72.43	0.21	13.73	13.84
DGP 4	**Normal**	90.23	0.13	0.03	9.75	88.29	0.15	1.89	9.82
	Block Wild	89.91	0.14	0.92	9.17	72.35	0.21	16.63	11.02

Note: The table reports associated coverage rates for simulation exercises where the DGP follows a UMIDAS model with various assumptions for the DGP disturbance. DGP 1 refers to the disturbance process $u_t \sim N.i.i.d(0,1)$, DGP 2 refers to the disturbance process $u_t = 0.5u_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim N.i.i.d(0,1)$, DGP 3 refers to the disturbance process $u_t = \sqrt{h_t}\varepsilon_t = \sqrt{0.1 + 0.3u_{t-1}^2 + 0.6\varepsilon_{t-1}}\varepsilon_t$, where $\varepsilon_t \sim N.i.i.d(0,1)$, and DGP 4 refers to a process where the disturbance is drawn from a mixture of two distributions, $u_t \sim N.i.i.d(-1,1)$ with probability 0.9 or $u_t \sim N.i.i.d(9,1)$ with probability 0.1. Results are provided for UMIDAS models using two different bootstrapping methods, the residual-based bootstrap (IID) and the block wild bootstrap, as well as for constructing predictive densities by drawing from a normal distribution. We report nominal coverage rates, measured as the mean nominal coverage of the 1000 simulations, for the 70 percent and 90 percent fan charts, their associated standard errors and the mean proportions below and above the nominal coverage.

Table A.4. Coverage rates for UMIDAS simulations with various DGPs and heteroskedastic high-frequency indicator

		90 percent coverage				70 percent coverage			
		Coverage	**SE**	**Lower**	**Upper**	**Coverage**	**SE**	**Lower**	**Upper**
	IID	88.88	0.13	5.45	5.66	68.86	0.20	15.53	15.61
DGP 1	**Normal**	88.23	0.14	5.82	5.95	67.75	0.21	16.07	16.19
	Block Wild	89.52	0.14	5.33	5.15	69.67	0.21	15.35	14.99
	IID	88.50	0.15	5.63	5.87	68.50	0.22	15.61	15.89
DGP 2	**Normal**	87.94	0.16	5.88	6.18	67.41	0.23	16.08	16.51
	Block Wild	89.87	0.15	5.08	5.05	70.28	0.23	14.96	14.76
	IID	89.06	0.15	5.57	5.38	69.23	0.20	15.49	15.28
DGP 3	**Normal**	88.53	0.15	5.75	5.71	68.28	0.21	15.96	15.76
	Block Wild	89.63	0.14	5.18	5.19	69.91	0.21	14.91	15.18
	IID	90.68	0.13	3.80	5.52	72.53	0.20	13.46	14.01
DGP 4	**Normal**	90.14	0.13	0.03	9.82	88.34	0.15	1.75	9.91
	Block Wild	90.65	0.13	3.79	5.55	72.48	0.20	13.44	14.08

Note: The table reports associated coverage rates for simulation exercises where the DGP follows a UMIDAS model with various assumptions for the DGP disturbance. DGP 1 refers to the disturbance process $u_t \sim N.i.i.d(0,1)$, DGP 2 refers to the disturbance process $u_t = 0.5u_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim N.i.i.d(0,1)$, DGP 3 refers to the disturbance process $u_t = \sqrt{h_t}\varepsilon_t = \sqrt{0.1 + 0.3u_{t-1}^2 + 0.6\varepsilon_{t-1}}\varepsilon_t$, where $\varepsilon_t \sim N.i.i.d(0,1)$, and DGP 4 refers to a process where the disturbance is drawn from a mixture of two distributions, $u_t \sim N.i.i.d(-1,1)$ with probability 0.9 or $u_t \sim N.i.i.d(9,1)$ with probability 0.1. Results are provided for UMIDAS models using two different bootstrapping methods, the residual-based bootstrap (IID) and the block wild bootstrap, as well as for constructing predictive densities by drawing from a normal distribution. We report nominal coverage rates, measured as the mean nominal coverage of the 1000 simulations, for the 70 percent and 90 percent fan charts, their associated standard errors and the mean proportions below and above the nominal coverage.

B Additional results - empirical exercise

In this section, we provide additional results from our empirical exercise, using alternative measures of absolute and relative density forecast accuracy, as well as point forecast accuracy.

Absolute accuracy

In Section 5 in the main text, we gauge calibration by testing jointly for uniformity and independence of the PITs, applying the test proposed by Berkowitz (2001). The Berkowitz test works with the inverse normal cumulative density function transformation of the PITs, which permits testing for normality instead of for uniformity. For one-step ahead forecasts, the null hypothesis is that the transformed PITs are iid N(0,1). The test statistic is χ^2, with three degrees of freedom. For longer horizons, we do not test for independence, and thus the null hypothesis is that the transformed PITs are identically standard normally distributed. The test statistics are then χ^2, with two degrees of freedom.

For multi-step ahead forecasts, the PITs are typically serially correlated and the Berkowitz test is less suitable. To account for serial correlation in the PITs, Knüppel (2015) have recently developed a raw-moments test. The raw-moments test is based on the standardized PITs instead of the inverse normal transforms. Despite of the autocorrelation of the PITs, the raw-moments tests rely on standard critical values. Although this test is more suitable for multi-step ahead forecasts, the test lacks power in small samples.

Table B.1 report the results of the Knüppel (2015) test. The results are broadly in line with those of the Berkowitz test, reported in Table 3 in the main text. The main difference is that the results for models with capacity utilization and the NFCI are somewhat more encouraging when we apply the Knüppel test than the Berkowitz test. In this case, the null hypothesis is rejected for the models with capacity utilization at the 5% significance level for the backcast for all four models and at horizon $h_m = 1$ and $h_m = 2$ for some of the models. For all other horizons and for all horizons using the NFCI, the null hypothesis cannot be rejected at the 5% significance level for any of the MIDAS models.

Table B.1. Absolute forecast accuracy for U.S. output growth. Evaluation sample: 2001q3:2015q2

h_m		*-1*	*1*	*2*	*3*	*4*	*5*	*6*
		Knüppel						
Employment	**UMIDAS**	0.737	0.915	0.690	0.590	0.586	0.521	0.763
	AR-UMIDAS	0.590	0.743	0.727	0.510	0.542	0.322	0.751
	MIDAS	0.977	0.914	0.870	0.746	0.666	0.512	0.365
	AR-MIDAS	0.957	0.798	0.829	0.439	0.294	0.238	0.519
	ADL	0.878	0.886	0.763	0.839	0.298	0.286	0.859
Industr. prod.	**UMIDAS**	0.161	0.096	0.188	0.237	0.236	0.235	0.097
	AR-UMIDAS	0.141	0.104	0.204	0.372	0.271	0.240	0.133
	MIDAS	0.214	0.190	0.181	0.409	0.239	0.266	0.377
	AR-MIDAS	0.157	0.180	0.193	0.445	0.472	0.247	0.336
	ADL	0.212	0.397	0.362	0.301	0.295	0.201	0.328
Capacity util.	**UMIDAS**	**0.013**	**0.020**	**0.031**	0.084	0.181	0.412	0.209
	AR-UMIDAS	**0.034**	**0.045**	0.091	0.131	0.277	0.308	0.220
	MIDAS	**0.012**	**0.015**	**0.028**	0.064	0.144	0.209	0.189
	AR-MIDAS	**0.021**	0.071	0.058	0.335	0.349	0.352	0.382
	ADL	**0.027**	0.107	0.132	0.122	0.325	0.290	0.217
CFNAI	**UMIDAS**	0.246	0.669	0.374	0.753	0.496	0.303	0.320
	AR-UMIDAS	0.122	0.268	0.571	0.806	0.504	0.395	0.206
	MIDAS	0.136	0.548	0.761	0.834	0.711	0.539	0.388
	AR-MIDAS	0.071	0.627	0.591	0.871	0.879	0.322	0.326
	ADL	**0.045**	0.274	0.557	0.479	0.271	0.269	0.327
NFCI	**UMIDAS**	0.061	0.090	0.098	0.121	0.115	0.134	0.268
	AR-UMIDAS	0.076	0.131	0.062	0.106	0.153	0.095	0.297
	MIDAS	0.134	0.091	0.099	0.111	0.171	0.098	0.075
	AR-MIDAS	0.113	0.163	0.141	0.153	0.107	0.099	0.093
	ADL	0.125	0.197	0.170	0.102	0.262	0.172	0.189
AR		0.331	0.413	0.278	0.330	0.306	0.238	0.207

Note: The table reports measures of the absolute density forecasting performance based on the PITs test by Knüppel (2015) of the four types of MIDAS models described in Section 2 in the main text (UMIDAS, AR-UMIDAS, MIDAS, AR-MIDAS), the ADL and AR models. Predictive densities are simulated using the block wild bootstrap approach for all models and the test is described in Section B. The results are computed on an evaluation sample from 2001q3 to 2015q2. The results are reported for model specifications using five different high-frequency indicators. Results for the corresponding quarterly ADL model are reported for each indicator and serve as a comparison in our tests. The table shows the results for the backcast, nowcast and 1-quarter ahead forecasts. Bold numbers indicate that the null of correct calibration predictive densities is rejected at a 5% significance level.

Relative accuracy

In the main text, we evaluated relative density forecast accuracy using the logarithmic score (LS). Here, 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+h}, which is simply a step function in y_{t+h}, and the empirical CDF that is associated with the model predictive density. Formally, the CRPS of a component density for a particular observation can be defined as:

$$CRPS_{t+h,t_m+w-h_m} = \int \left(F(z_{t+h}) - \mathbb{I}_{[y_{t+h},+\infty)}(z_{t+h})\right)^2 \mathrm{d}z_{t+h} \tag{B.2}$$

where $F(z_{t+h})$ is the CDF of the predictive density $f(z_{t+h}|I_{t_m+w-h_m})$ defined above, see Panagiotelis and Smith (2008) for more formal details and Ravazzolo and Vahey (2013) for an illustrative example. The integral is approximated using the draws from the forecast density. Smaller CRPS values imply higher model precision.

We report results for relative density forecast performance, measured in terms of CRPS, in Table B.2. The results are broadly in line with those reported in the main text, using LS as the measure of relative density forecast accuracy. Importantly, by comparing the forecasting performance of the various MIDAS models with the AR and ADL benchmarks, the table shows sizeable decreases in the CRPS using MIDAS models when monthly and weekly data on the current quarter are available. For each model, there is a clear and steady increase in forecasting performance as more information becomes 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 $h_m = 3$) and the forecasts produced at the quarter before the quarter to forecast (i.e., at horizon $h_m = 4, 5, 6$). However, the gains from using higher-frequency information compared to the quarterly ADL and AR benchmarks is somewhat smaller when using the CRPS than when using the LS.

Figure B.1 shows how the relative cumulative CRPS from the MIDAS and AR-MIDAS regressions have evolved over time. The figure describes relative forecasting performance using each of the four monthly indicators for the backcast and the three nowcasting 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

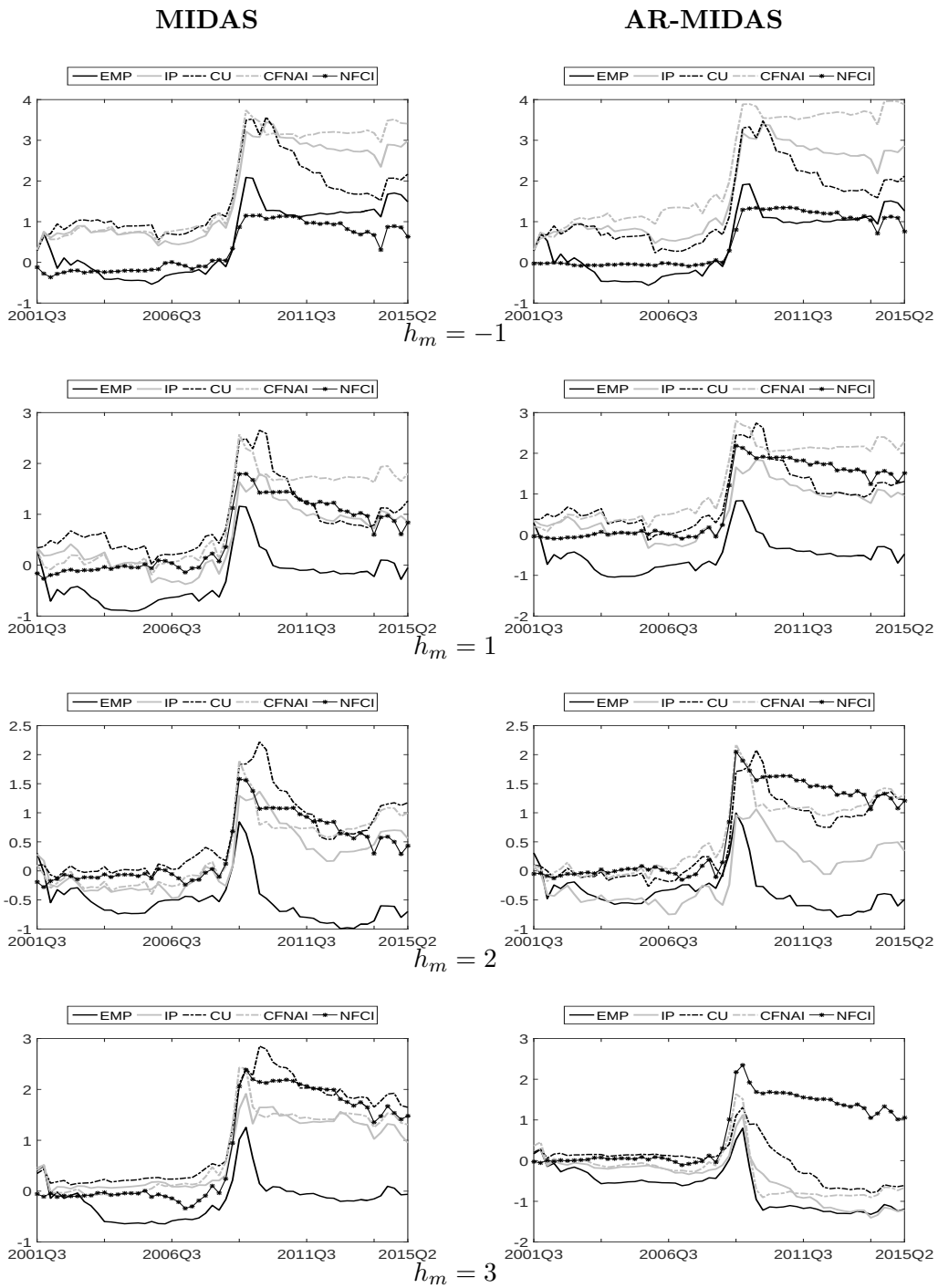
Table B.2. Relative forecast accuracy for U.S. output. Evaluation sample: 2001q3:2015q2

h_m		*-1*	*1*	*2*	*3*	*4*	*5*	*6*
				CRPS				
Employment	**UMIDAS**	**0.257**	**0.277***	**0.284****	**0.292****	**0.296**	**0.300**	**0.308**
	AR-UMIDAS	**0.261**	**0.279****	**0.283****	**0.290****	**0.298**	**0.298**	**0.308**
	MIDAS	**0.263**	**0.280***	**0.293***	**0.302**	**0.307**	**0.313**	**0.316**
	AR-MIDAS	**0.267**	**0.288***	**0.290****	0.322	0.325	0.321	0.330
	ADL	0.271	0.307	0.305	0.308	0.310	0.314	0.319
Industr. prod.	**UMIDAS**	0.262	**0.283**	0.277	0.275	0.283	0.311	0.304
	AR-UMIDAS	0.260	**0.282**	0.275	0.280	**0.281**	**0.301**	**0.304**
	MIDAS	**0.236**	**0.264**	**0.271**	0.284	**0.287**	0.326	0.328
	AR-MIDAS	**0.238**	**0.262**	0.275	0.322	0.309	0.318	0.350
	ADL	0.251	0.265	0.271	0.277	0.304	0.307	0.313
Capacity util.	**UMIDAS**	0.266	0.273	0.276	**0.273****	**0.282**	0.316	**0.308**
	AR-UMIDAS	0.272	0.266	0.277	**0.274***	**0.284**	**0.306**	**0.309**
	MIDAS	**0.251**	**0.257**	**0.260**	**0.271**	**0.287**	0.324	0.317
	AR-MIDAS	**0.252**	**0.256**	**0.259**	0.311	**0.304**	0.307	0.346
	ADL	0.259	0.265	0.270	0.289	0.306	0.306	0.314
CFNAI	**UMIDAS**	0.220	**0.249**	**0.254**	0.277	**0.281**	0.309	0.301
	AR-UMIDAS	0.220	**0.241**	**0.262**	0.274	**0.280**	0.306	0.301
	MIDAS	0.229	**0.247**	0.263	0.277	**0.285**	0.317	0.305
	AR-MIDAS	0.220	**0.238**	**0.257**	0.312	0.315	0.328	0.343
	ADL	0.217	0.254	0.262	0.269	0.305	0.301	0.298
NFCI	**UMIDAS**	0.304	0.299	0.321	0.344	**0.278**	**0.275**	0.357
	AR-UMIDAS	0.296	0.280	0.305	0.334	**0.285**	**0.275**	0.348
	MIDAS	0.278	0.264	0.273	**0.274**	**0.278**	**0.278**	**0.297**
	AR-MIDAS	0.276	**0.252**	**0.259**	0.282	**0.299**	**0.297**	**0.304**
	ADL	0.262	0.262	0.259	0.279	0.312	0.312	0.304
AR		0.290	0.279	0.281	0.301	0.299	0.298	0.298

Note: The table reports the relative forecasting performance, measured in terms of CPRS, of the four types of MIDAS models described in Section 2 in the main text (UMIDAS, AR-UMIDAS, MIDAS, AR-MIDAS). Predictive densities are simulated using the block wild bootstrap approach for all models and the evaluation criterion is described above. The results are computed on an evaluation sample from 2001q3 to 2015q2. The results are reported for model specifications using five different high-frequency (monthly) indicators. Results for the corresponding ADL model are reported for each indicator and serve as a comparison in our tests. Statistics for the standard AR model are also reported. The table shows the results for the backcast, nowcast and 1-quarter ahead forecasts. Bold numbers indicate that a model specification provides more accurate forecasts than the ADL benchmark. Differences in accuracy that are statistically different from zero at a 10% and 5% significance level are denoted by one and two asterisks, 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). We report p-values based on one-sided tests.

with the AR benchmark. Again, the results for the CRPS are in line with those reported for the LS in the main text. Specifically, for all indicators and for both the MIDAS and AR-MIDAS regressions, the relative cumulative CRPS increases sharply during the Great Recession. However this increase is then followed by a weak, but persistent, decrease during the recovery period. Similarly to the results for the LS, this may indicate that the high-frequency indicators are particularly informative about sharp decreases in real output growth.

Figure B.1. Relative Cumulative CRPS



Note: The figure shows the relative Cumulative CRPS $= \sum_{s=1}^{t}(CRPS_{s+h,s_m+w-h_m,l} - CRPS_{s+h,s_m+w-h_m,i})$, $t = 1, ..., T$, $l =$ 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 model. The different rows show the results for a selection of different forecasting horizons.

Point forecast accuracy

So far in the literature, MIDAS models have been evaluated based on their point forecasting performance. Studies, such as Clements and Galvão (2008), Andreou et al. (2013) and Foroni et al. (2015) found sizeable gains from the use of high-frequency information in terms of point forecast accuracy. In the main text, we documented also sizeable gains from the use of high-frequency information in terms of density forecast accuracy. However, since the sample periods, as well as some of the explanatory variables in the aforementioned studies differ from ours, we also report results for point forecasting. In this way, we can make a clear comparison of the gains of using higher frequency information in terms of point and density forecast accuracy.

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

$$MSPE_{t+h,t_m+w-h_m} = \frac{1}{t^*} \sum_{t=\underline{t}}^{\bar{t}} e^2_{t+h,t_m+w-h_m} \tag{B.3}$$

where $t^* = \bar{t} - \underline{t} + h$, $\bar{t}$ and $\underline{t}$ denote the beginning and end of the evaluation period, and e^2_{t+h,t_m+w-h_m} is the h-step ahead square prediction error conditional on the information set available at time $t_m + w - h_m$

Table B.3 reports real-time point forecasting performance, measured in terms of MSPE, for the different MIDAS models. As for the results obtained with density forecasting, there 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 becomes available. Compared to the quarterly-frequency ADL benchmark, the point nowcasting improvements from the various MIDAS models are broadly in line with the density nowcasting improvements, measured by LS, reported in the main text. Similar to findings in Andreou et al. (2013), we achieve large point forecasting gains, relative to the ADL model, by exploiting information from a large number of high-frequency financial time series, measured by the NFCI. However, note that the relative gains are partly driven by a relatively poor point forecasting performance from the ADL model.

Table B.3. Point forecast accuracy for U.S. output. Evaluation sample: 2001q3:2015q2

h_m		*-1*	*1*	*2*	*3*	*4*	*5*	*6*
				MSPE				
Employment	**UMIDAS**	**0.216**	**0.247***	**0.276****	**0.301***	**0.321**	**0.335**	**0.355**
	AR-UMIDAS	**0.219**	**0.250****	**0.274****	**0.303***	**0.324**	**0.336**	**0.351**
	MIDAS	**0.227**	**0.256***	**0.285****	**0.321**	**0.337**	**0.371**	0.372
	AR-MIDAS	**0.233**	**0.269***	**0.287****	0.384	0.397	0.395	0.406
	ADL	0.238	0.324	0.327	0.335	0.363	0.373	0.370
Industr. prod.	**UMIDAS**	0.256	0.302	0.267	**0.270**	**0.295**	0.381	0.355
	AR-UMIDAS	0.259	0.307	0.259	0.282	**0.297**	**0.350**	**0.348**
	MIDAS	**0.182**	**0.231**	**0.241**	0.284	**0.311**	0.395	0.406
	AR-MIDAS	**0.189**	**0.233**	**0.248**	0.373	**0.340**	0.392	0.465
	ADL	0.222	0.242	0.252	0.283	0.356	0.355	0.353
Capacity util.	**UMIDAS**	0.239	0.262	0.283	**0.262**	**0.287**	0.376	0.366
	AR-UMIDAS	0.249	0.247	0.280	**0.259**	**0.290**	0.356	0.372
	MIDAS	**0.205**	**0.214**	**0.224**	**0.252**	**0.294**	0.387	0.374
	AR-MIDAS	**0.199**	**0.215**	**0.225**	0.337	**0.330**	0.363	0.460
	ADL	0.219	0.236	0.246	0.299	0.352	0.354	0.349
CFNAI	**UMIDAS**	0.160	0.205	0.234	0.262	0.282	0.371	0.348
	AR-UMIDAS	0.160	**0.194**	0.247	0.261	**0.285**	0.362	0.341
	MIDAS	0.169	**0.197**	0.230	0.265	**0.291**	0.384	0.361
	AR-MIDAS	0.164	**0.190***	0.229	0.372	0.377	0.441	0.457
	ADL	0.153	0.202	0.222	0.251	0.365	0.355	0.335
NFCI	**UMIDAS**	0.423	0.331	0.373	0.527	**0.272**	**0.272**	0.551
	AR-UMIDAS	0.401	**0.286**	**0.336**	0.516	**0.280**	**0.269**	0.521
	MIDAS	**0.265**	**0.237**	**0.253**	**0.261**	**0.268**	**0.277**	**0.323**
	AR-MIDAS	**0.257**	**0.205**	**0.222**	**0.273**	0.322	0.328	**0.339**
	ADL	0.390	0.327	0.343	0.348	0.307	0.327	0.356
AR		0.302	0.279	0.293	0.350	0.349	0.349	0.349

Note: The table reports point forecasting performance, measured in terms MSPE, of the four types of MIDAS models described in Section 2 in the main text (UMIDAS, AR-UMIDAS, MIDAS, AR-MIDAS). Predictive densities are simulated using the block wild bootstrap approach for all models and the evaluation criterion is described above. The results are computed on an evaluation sample from 2001q3 to 2015q2. The results are reported for model specifications using five different high-frequency (monthly) indicators. Results for the corresponding ADL model are reported for each indicator and serve as a comparison in our tests. Statistics for the standard AR model are also reported. The table shows the results for the backcast, nowcast and 1-quarter ahead forecasts. Bold numbers indicate that a model specification provides more accurate forecasts than the ADL benchmark. Differences in accuracy that are statistically different from zero at a 10% and 5% significance level are denoted by one and two asterisks, 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). We report p-values based on one-sided tests.

C A case study: The importance of high-frequency financial data when forecasting during the Great Recession

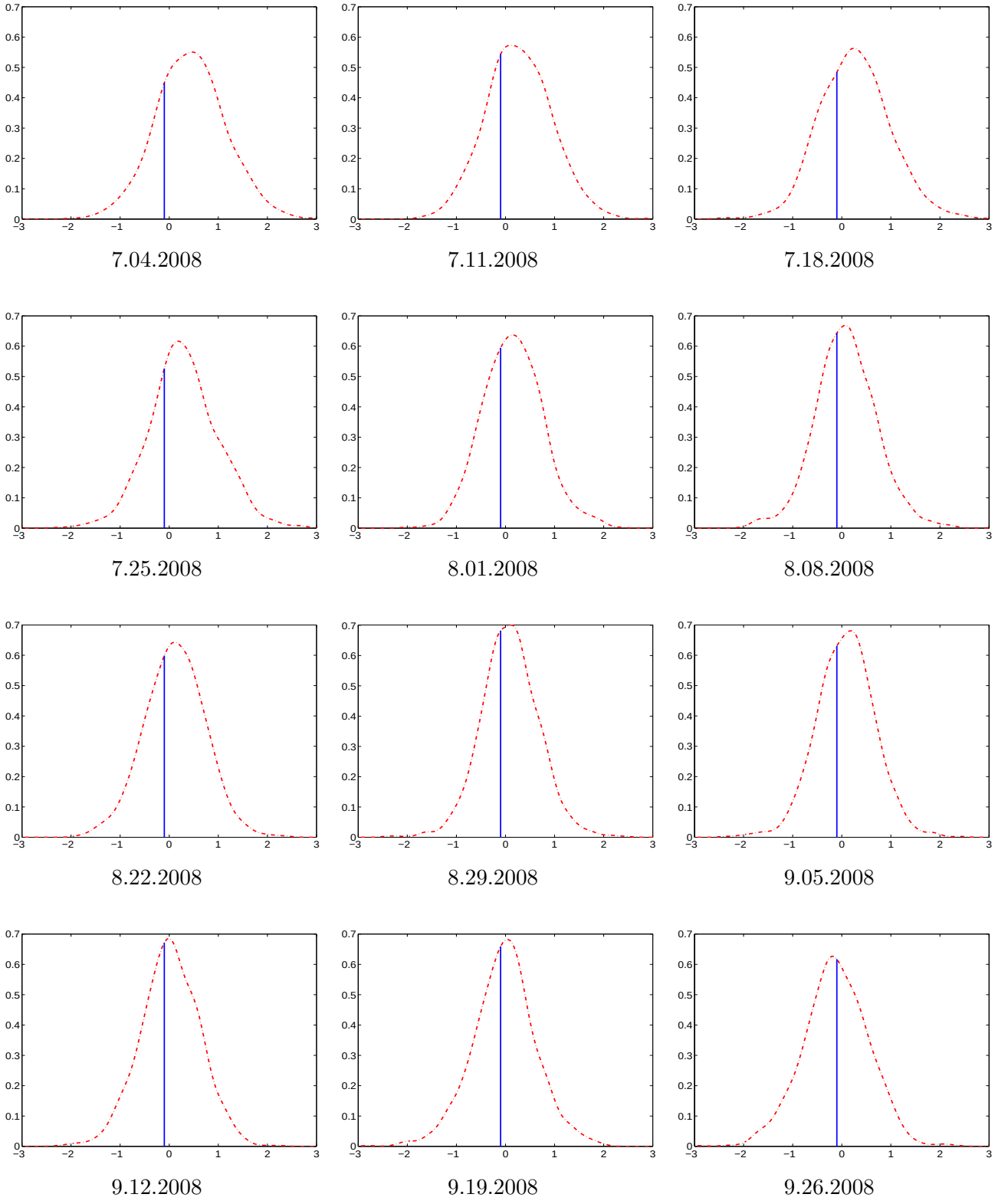
Theory suggests that the forward-looking nature of financial asset prices should contain information about the future state of the economy and therefore should be considered as highly relevant for macroeconomic forecasting. Andreou et al. (2013) have recently stressed that macroeconomic forecasters should use information from a large number high-frequency financial time series when forecasting macroeconomic variables. In Section 5.2 in the main text, we showed that MIDAS models which included weekly updates of the Chicago Fed's National Financial Condition Index (NFCI) provided useful information for nowcasts of real output growth for the U.S. Motivated by this, as a case study, we investigate how weekly updates of financial data change the nowcasts of real output growth for the U.S. during 2008q3 and 2008q4.

We specify a MIDAS regression with weekly values of the NFCI as the high-frequency indicator and compute density nowcasts for real output growth using our suggested block wild bootstrapping approach. Figures C.1 and C.2 report how the density nowcasts for 2008q3 and 2008q4 change as a result of the weekly updates of the NFCI.

We start by presenting the results for 2008q3. Figure C.1 shows that for almost every weekly update of the NFCI, the predictive densities shift to the left, attributing a larger probability mass to negative outcomes. The figure shows that the density nowcasts for real output growth in 2008q3 had already changed between the first and second week of July 2008. During the second week of July 2008, IndyMac announced they had failed to raise capital since their May quarterly report. Later the same week, citing liquidity concerns, the Federal Deposit Insurance Corporation put IndyMac Bank into conservatorship. On July 31, IndyMac Bancorp filed for bankruptcy. Interestingly, there are again visible changes in the density nowcast from our weekly MIDAS regression from July 25th to August 1st. While there were no noteworthy changes in the density nowcasts during August 2008, the predictive densities changed considerably during September 2008. The week of September 15 to September 19 was particularly turbulent, with Lehman Brothers Holding Inc. filing for bankruptcy and the equivalent of a bank run on the money market funds. By the end of September 2008, the density nowcasts from our weekly MIDAS regression attributes the majority of its predictive probability mass to negative outcomes for real output growth.

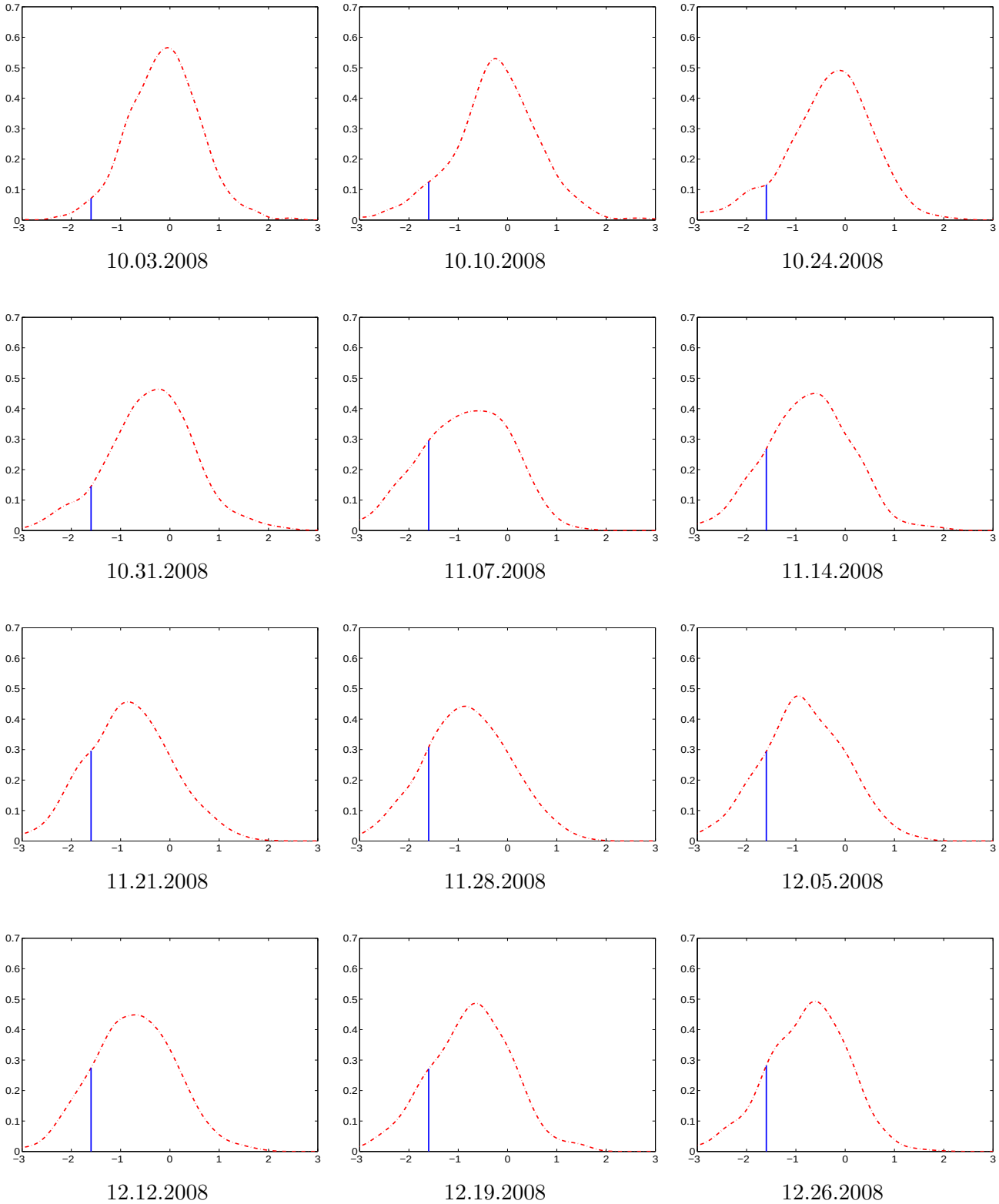
Turning to the density nowcasts for 2008q4, Figure C.2 reveals even more dramatic changes in the density nowcasts during 2008q4. During the first ten days of October, the Emergency Economic Stabilization Act, which implemented the Troubled Asset Relief Program (TARP), was signed into law, the Federal Reserve Bank (Fed) cut the federal funds rate by 50 basis points to 1.5 percent at an unscheduled emergency meeting and the TED spread reached its highest level. This is visible in Figure C.2 in the sense that the density nowcast after the second week of October has a considerably fatter lower tail than the density nowcast after the first week of October. In late October, the Fed cut the federal funds rate again by 50 basis points to 1 percent. During the rest of the quarter, it became more and more evident that the U.S. economy was going into a deep recession, and the financial turmoil had also spread to Europe and the rest of the world. Figure C.2 shows that during November and December, the weekly density nowcasts shift continuously to the left. By mid-December, when the Fed decided to establish a target range for the federal funds rate of 0 to 0.25 percent and formerly launched its first quantitative easing program, the density nowcast from our weekly model attributes only a small probability mass to positive outcomes of real GDP growth, i.e., indicating a distinct slowdown in output growth.

Figure C.1. Case study: Updates of weekly density nowcasts during 2008q3



Note: The figure shows the density forecasts produced by the MIDAS model based on the weekly NFCI indicator on twelve selected dates in 2008q3. The red dashed line represents the kernel of the computed density forecast and the blue line shows the probability given to realized actual GDP growth.

Figure C.2. Case study: Updates of weekly density nowcasts during 2008q4



Note: Density forecasts produced by the MIDAS model based on the weekly NFCI indicator on twelve selected dates in 2008q4. The red dashed line represents the kernel of the computed density forecast and the blue line shows the probability given to realized actual GDP growth.

References

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

Andreou, E., E. Ghysels, and A. Kourtellos (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.

Berkowitz, J. (2001). Testing density forecasts, with applications to risk management. *Journal of Business & Economic Statistics 19*(4), 465–474.

Clements, M. P. and A. B. Galvão (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics 26*, 546–554.

Davidson, R. and J. G. MacKinnon (2006). Bootstrap methods in econometrics. In *Palgrave Handbooks of Econometrics: Volume 1 Econometric Theory*, pp. 812–838. Basingstoke: Palgrave Macmillan.

Foroni, C., M. Marcellino, and C. Schumacher (2015). U-midas: Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society - Series A 29*(1), 57–82.

Knüppel, M. (2015). Evaluating the Calibration of Multi-Step-Ahead Density Forecasts Using Raw Moments. *Journal of Business & Economic Statistics 33*(2), 270–281.

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 18*(4), 367–381.