Forecasting Low Frequency Macroeconomic Events with High Frequency Data*

Ana Beatriz Galvao
University of Warwick
ana.galvao@wbs.ac.uk

Michael T. Owyang
Federal Reserve Bank of St. Louis
owyang@stls.frb.org

September, 2020

Abstract

High-frequency financial and economic activity indicators are usually time aggregated before forecasts of low-frequency macroeconomic events, such as recessions, are computed. We propose a mixed-frequency modelling alternative that delivers high-frequency probability forecasts (including their confidence bands) for these low-frequency events. The new approach is compared with single-frequency alternatives using loss functions adequate to rare event forecasting. We provide evidence that: (i) weekly-sampled spread improves over monthly-sampled to predict NBER recessions, (ii) the predictive content of the spread and the Chicago Fed Financial Condition Index (NFCI) is supplementary to economic activity for one-year-ahead forecasts of contractions, and (iii) a weekly activity index can date the 2020 business cycle peak two months in advance using a mixed-frequency filtering.

keywords: mixed frequency models, recession, financial indicators, weekly activity index, event probability forecasting.

JEL Codes: C25, C53, E32

*This paper builds on a draft written and circulated by Michael Owyang in 2011. M. Owyang benefitted from conversations with Eric Ghysels, Jim Hamilton, and Mike McCracken. Kristie Engemann, Kate Vermann, Hannah Shell, and Julie Bennett provided research assistance. The views expressed here are the authors' alone do not reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

1 Introduction

Aware of the forward-looking behavior of financial markets, economists often use financial variables as predictors for events such as downside growth risks (Adrian, Boyarchenko and Giannone, 2019) or recessions (Estrella and Mishkin, 1998; Chauvet and Potter, 2005; Liu and Moench, 2016; Bauer and Mertens, 2018). Because financial variables are sampled at a higher frequency (daily) than most economic indicators (monthly, quarterly), financial predictors are typically time aggregated before they are used to estimate the forecast model.[1] In this paper, we propose a mixed-frequency strategy to exploit the predictive content of high frequency financial variables for low frequency economic events. By applying a Bayesian estimation strategy, we deliver accurate event-probability forecasting with a measure the uncertainty around the probability forecasts.

Mixed-data sampling regressions are typically estimated using (nonlinear) least squares (Ghysels, Sinko and Valkanov, 2007; Ghysels, Santa-Clara and Valkanov, 2006; Clements and Galvão, 2008) and, for binary dependent variables, likelihood-based estimators are applied (Audrino, Kostrov and Ortega, 2019). Bayesian methods have been developed to accommodate mixed data sampling of both the conditional mean and the conditional variance (Pettenuzzo et al., 2016) and to deal with a large number of predictors (Mogliani and Simoni, 2020). Our Bayesian estimation strategy combines the Gibbs sampler developed for probit models (Albert and Chib, 1993) with a Metropolis step to draw the parameters of the beta function that parsimoniously describes the aggregating weights. Our choice of a beta function to aggregate high-frequency data accommodates applications with large numbers of leads and lags of the high-frequency variable and is compatible with cases where the binary dependent variable is available at quarterly frequency and the regressor at the daily frequency as in Galvao (2013) and Ghysels, Plazzi, Valkanov, Rubia and Dossani (2019). Our approach differs from the use of Almon lags in Pettenuzzo et al. (2016) and the use of unrestricted weighting schemes in Carriero, Clark and Marcellino (2020).

[1] If the aim is to predict economic activity, instead of binary macroeconomic events, Andreou, Ghysels and Kourtellos (2013), Galvao (2013) and Pettenuzzo, Timmermann and Valkanov (2016) are examples where time aggregation is not applied to financial variables.

The mixed-data sampling probit model is applied to answer three empirical research questions regarding the use of high-frequency variables to predict low-frequency events. First, we revisit the predictive content of daily/weekly term spreads for monthly/quarterly U.S. recession phases (Chauvet and Potter, 2005; Liu and Moench, 2016; Bauer and Mertens, 2018). Second, we evaluate whether weekly financial variables have additional predictive ability over monthly/quarterly economic activity variables for contractionary GDP periods with elevated downside risks as in Adrian et al. (2019) and Plagborg-Moller, Reichlin, Ricco and Hasenzagl (2020). Third, we assess whether we could have anticipated the June 2020 NBER announcement of a peak in February 2020 using the weekly economy index, proposed by Lewis, Mertens and Stock (2020).

An important characteristic of mixed-frequency models is that forecasts for low-frequency events can be updated each time new observations of the high-frequency data become available. We evaluate whether these updates improve forecasting performance and find improved accuracy when using the weekly spread to predict quarterly GDP contractions.

The evaluation of binary event probabilities is directly linked to the relative losses incurred by false negatives versus false positives. We consider three different loss functions. The first two loss functions—the area under the receiver operating characteristic curve (AUROC; as applied in Berge and Jorda (2011) and Liu and Moench (2016)), and the diagonal of the elementary score (Bouallegue, Haiden and Richardson, 2018)—are designed to measure the ability of the forecasting model to accurately classify the event. We also consider a logarithm score, more in line with the quantile forecasting evaluation and the predictive scores in Adrian et al. (2019) and Plagborg-Moller et al. (2020). We show that the ranking across models changes with the loss function employed. Comparison with simple forecasting rules is also useful to assess the marginal contribution of alternative forecasting models. In this paper, relative performance is measured with respect to the unconditional probability forecast (equal to the frequency of the event) using measures of skill employed in the climatology literature (Bouallegue, Magnusson, Haiden and Richardson, 2019).

We describe our mixed-frequency approach to model binary dependent variables in Section

2. Details of our econometric implementation, including the Bayesian estimation strategy and probability forecasting loss functions are described in Section 3. Section 4 provides empirical results and discussion of our three empirical applications. Section 5 concludes with a summary assessment of the proposed forecasting model and implications of the empirical results for macroeconomic forecasting.

2 The MIDAS-Probit

2.1 Setup

Define a monthly binary variable, $S_t = \{0, 1\}$, where $S_t = 1$ indicates that the economic event of interest occurs. In our main example below, we set $S_t = 1$ for months within a recession phase. Define a latent variable, y_t^*, such that $y_t^* \geq 0$ if $S_t = 1$ and $y_t^* < 0$ if $S_t = 0$. Then, a single-regressor probit model to predict event S_t h-months-ahead is:

$$\Pr[y_{t+h}^* \geq 0 | \Omega_t] = \Phi(\beta_{0,h} + \beta_{1,h} z_t), \tag{1}$$

where $\Phi(.)$ is the CDF of a standard normal density and the information set, Ω_t, consists only of a single monthly variable, z_t.[2] Direct multi-step-ahead monthly forecasts can be obtained by estimating (1) for $h = 1, ..., H$ by changing the indicator variable as $S_{t+1}, ..., S_{t+H}$, implying that parameters $\left(\beta_{0,h}, \beta_{1,h}\right)$ change with the horizon.

While events of interest, such as recessions, are observed either monthly or quarterly, many predictors—because they are financial variables—are available at a daily or higher frequency. For example, Estrella and Mishkin (1998) provide evidence of the predictive content of the slope of the yield curve for U.S. recessions using probit models. The ability of the spread between the 10-year Treasury bond yield and the 3-month Treasury bill rate to predict NBER recessions in probit

[2] The model can be easily generalized to include lags of the predictor (e.g., Liu and Moench (2016) suggested to add both z_t and z_{t-6} when predicting recessions using the term-structure spread) and additional regressors as we show below.

models has been further evaluated by Chauvet and Potter (2005), Kauppi and Saikkonen (2008) and Liu and Moench (2016).

The application of the standard probit shown in eq. (1) requires the aggregation of any high-frequency data to match the sampling of S_t. For example, daily data can either be averaged over the month or represented by the last day of the low-frequency period. However, this temporal aggregation assumes that the high-frequency timing of innovations to the predictors are unimportant and can result in the loss of this information. In the next few sections, we propose an alternative specification of a mixed-frequency probit that use high-frequency data directly, preserving information about the high-frequency data.

2.2 The Single-Predictor MIDAS-Probit

For expositional simplicity, we will continue to characterize the one-predictor case; the generalization to multiple predictors at possibly multiple frequencies appears below. Assume that the high-frequency variable is sampled m times more frequently than S_t (and, consequently, y_t^*); for example, if the low-frequency variable is monthly and the high-frequency predictor is daily, $m = 21$ trading days (on average over the year). Thus, for each realization of S_t, the information set Ω_t includes one low-frequency period of high frequency observations: $z_t^{(m)}, z_{t-\frac{1}{m}}^{(m)}, z_{t-\frac{2}{m}}^{(m)}, ...z_{t-\frac{m-1}{m}}^{(m)}$.[3]

We can preserve the probit-type formulation for the $h-$period-ahead direct multi-step forecast by writing:

$$\Pr[y_{t+h}^* \geq 0|\Omega_t] = \Phi\left(\beta_{0,h} + \beta_{1,h}\sum_{k=1}^{K} \varpi\left(k;\boldsymbol{\theta}_h\right) z_{t-((k-1)/m)}^{(m)}\right), \tag{2}$$

where K is number of lags at the sampling frequency of $z_t^{(m)}$. The functions, $w\left(k;\boldsymbol{\theta}_h\right)$, attributing weights to each of the (high frequency) lags of $z_t^{(m)}$ up to K, are often referred to as **MI**xed **DA**ta **S**ampling (MIDAS) functions (Ghysels et al., 2007; Ghysels et al., 2006; Clements and Galvão, 2008; Andreou, Ghysels and Kourtellos, 2010; Kuzin, Marcellino and Schumacher, 2013). If the number

[3]Notice that the integer timing t is still in the low-frequency variable. The high-frequency variable is observed in fractions of the low-frequency periods. Thus, $t - 1/m$ indexes one high-frequency period before the t-th observation of the low-frequency variable.

of high frequency lags K is equal to m, then $\sum_{k=1}^{K} \varpi\left(k;\boldsymbol{\theta}_h\right) z_{t-(k-1)/m}^{(m)}$ is the predictor aggregated to the frequency of the binary dependent variable S_t. If $\varpi\left(k;\boldsymbol{\theta}_h\right) = 1/m$, the aggregation scheme would average the high-frequency data (say daily) within the low-frequency period (say a month).

To identify the slope parameter, $\beta_{1,h}$, the weights are constrained to sum to 1:

$$\varpi\left(j;\boldsymbol{\theta}_h\right) = \frac{f(k,\boldsymbol{\theta}_h)}{\sum_{k=1}^{K} f(k,\boldsymbol{\theta}_h)}.$$

While the weights can, in principle, take on a number of functional forms, we employ a beta function:

$$f(k,\boldsymbol{\theta}_h) = \frac{\kappa^{\theta_1-1}\left(1-\kappa\right)^{\theta_2-1}\Gamma\left(\theta_{1,h}+\theta_{2,h}\right)}{\Gamma\left(\theta_{1,h}\right)\Gamma\left(\theta_{2,h}\right)}; \kappa = k/(K+1), \tag{3}$$

where $\Gamma\left(.\right)$ is a gamma function and $\boldsymbol{\theta}_h$ is a vector of the two parameters that govern the shape of the weighting function (Ghysels et al., 2007; Andreou et al., 2010). The beta function collapses to the equal weighting aggregation scheme if $K = m$ and $\theta_{1,h} = \theta_{2,h} = 1$. In typical empirical MIDAS applications, however, the number of high-frequency lags of the predictor is set such that $K \geq m$ (Galvao, 2013; Pettenuzzo et al., 2016; Carriero et al., 2020). If $\theta_1 < \theta_2$, weights are decreasing across lags, so less weight is assigned to values further in the past.

The MIDAS approach is a parsimonious method of employing many lags of the high-frequency variable, as we only need to estimate three parameters $(\beta_{1,h}, \theta_{1,h}, \theta_{2,h})$ for each predictor. In contrast, Foroni, Marcellino and Schumacher (2015) argue to use the UMIDAS specification that estimates the coefficient for each lag separately. Most applications of UMIDAS are for small differences in frequency ($m = 3$) as in that case parameter proliferation is limited (Foroni et al., 2015) or for cases where Bayesian methods are employed to deal with the large number of parameters (as, for example, Carriero et al. (2020)).

2.3 An Example

The model presented in the previous section mixes two variables at two frequencies. Before we present the general version of the MIDAS probit, we provide an example that may help illuminate the frequency mixing problem. Suppose that $z_t^{(m)}$ is the *weekly-sampled* term spread observed on Saturday and that the indicator variable, S_t, is observed monthly and equal to 1 if the US economy is in a recession as set by the NBER turning points. Thus, the ratio of the sampling frequencies, m, is approximately 13/3, based on 13 weeks in a quarter consisting of 3 months. Based on the Liu and Moench (2016) suggestion that longer lags improve the recession forecasting performance of the spread, we consider lags of $z_t^{(m)}$ covering approximately the last seven months, setting $K = 32$. Following Chauvet and Potter (2005), we compute one-year-ahead recession probability forecasts (i.e., S_{t+12}) using the spread, as this is the horizon that the spread is more informative. As $K > m$, the MIDAS term is simply the weighted sum of the 32 weeks of high-frequency data observed before the forecast origin.[4]

We estimated a MIDAS probit model for this example ($h = 12$, $m = 13/3$, $K = 32$) by applying the estimation approach described in Section 3.1 and using data from 1962M1 to 2020M4. Figure 1 shows the posterior mean (and 68% bands) of the beta weighting function estimated for the MIDAS probit model in (2). The red line in Figure 1 indicates the case that $\theta_{1,h} = \theta_{2,h} = 1$, that is, equal weights. As we are able to fully characterize the uncertainty on the estimation of the weights due to our Bayesian estimation strategy, Figure 1 provides clear evidence that the optimal weighting function applied to the weekly spread differs from equal weighting. Weights are higher for shorter lags (up to one month), although weights for longer lags are non-zero (including after 7 months).

2.4 The General MIDAS-Probit Model

The model introduced above extends to a mix of any number of variables of multiple frequencies under the assumption that the binary dependent variable is sampled at the lowest common fre-

[4] As we will show below, it is also not necessary that the forecast origin coincides with an observation of the low-frequency data.

quency. The general form of the MIDAS-probit can be estimated using an $(N+1) \times 1$ vector of predictors that includes a constant and N same- or higher-frequency regressors.

Define $Z_{nt}(K_n, \boldsymbol{\theta}_n)$ as the weighted sum of K_n lags of the high-frequency variable $z_{nt}^{(m)}$ using a beta weighting function with parameters $\boldsymbol{\theta}_n$—that is, $Z_{nt}(K_n, \boldsymbol{\theta}_n) = \sum_{k=1}^{K_n} \varpi(k; \boldsymbol{\theta}_i) z_{n,t-(k-1)/m}^{(m)}$. Then, define:

$$\mathbf{Z}_t(\Theta) = [1, Z_{1t}(K_1, \boldsymbol{\theta}_1), ..., Z_{Nt}(K_N, \boldsymbol{\theta}_N)]', \tag{4}$$

where $\Theta = (\boldsymbol{\theta}_1', ..., \boldsymbol{\theta}_N')$. The general form of the MIDAS-probit is:

$$P[y_{t+h}^* \geq 0 | \Omega_t] = \Phi[\mathbf{Z}_t(\Theta_h)' \boldsymbol{\beta}_h], \tag{5}$$

where $y_{t+h}^* \geq 0$ if $S_{t+h} = 1$, $y_{t+h}^* < 0$ if $S_{t+h} = 0$ for $t = 1, ..., T-h$, $\boldsymbol{\beta}_h$ is a $(N+1) \times 1$ vector of slopes, and the parameters are indexed by horizon to produce a direct multi-step forecast.

Note that the MIDAS-probit can be written as a regression model:

$$y_{t+h}^* = \mathbf{Z}_t(\Theta_h)' \boldsymbol{\beta}_h + u_{t+h}, \tag{6}$$

where $u_{t+h} \sim N(0,1)$. Changing the distributional assumption for u_{t+h} can change the model from probit to logit, etc.

3 Econometric Implementation

In this section, we briefly describe the steps required to estimate the model, form forecasts, and evaluate the forecasts. While we highlight some steps in the econometric process, we leave most of the relatively standard parts of the algorithm to Appendix A.

3.1 Estimation

MIDAS models are usually estimated using nonlinear least squares as in Ghysels et al. (2006) and Clements and Galvão (2008) or, in the case of binary dependent variables, by maximum likelihood (Audrino et al., 2019). As these techniques require numerical optimization, algorithms have been proposed to improve convergence (Ghysels and Qian, 2019). When MIDAS regressions are estimated using Bayesian weights, authors typically choose weighting schemes that are linear in the parameters such as the Almon weighting function (Pettenuzzo et al., 2016) and the unrestricted function (Carriero et al., 2020).

We propose an alternative estimation algorithm: a Gibbs sampler (Gelfand and Smith, 1990; Casella and George, 1992) with a Metropolis-in-Gibbs step (Chib and Greenberg, 1995) to sample the MIDAS hyperparameters that govern the weights. As described in Greenberg (2013, ch. 8), Gibbs sampling is the standard method to estimate probit models using Bayesian methods and should be equivalent to maximum likelihood under some conditions. As it is clear from our representation of estimates in Figure 1, an advantage of the Bayesian estimation is that we are able to fully characterize uncertainty in the MIDAS weights and slopes, facilitating inference regarding the relevance of the use of mixed frequencies.

To start, we require a prior for the slopes, $\boldsymbol{\beta}_h$, and the MIDAS hyperparameters, Θ_h. Conditional on known MIDAS weights, the latent variable in the probit can be written as a linear regression. Thus, we assume a standard conjugate zero-mean Normal prior on the slope coefficients. We further assume that the MIDAS hyperparameters have (restricted) Gamma priors and are centered around the belief that the high-frequency data is equally weighted. For most of our specifications, we impose a joint restriction across each predictor's MIDAS hyperparameters, $\theta_{2,n,h} \geq \theta_{1,n,h}$, so that older data are given (weakly) less weight. Because we assumed a standard Probit, the variance of the latent variable is assumed to be fixed at 1.

The sampler is decomposed into three blocks: (i) the slope parameters, $\boldsymbol{\beta}_h|\Theta_h, \left\{y^*_{t+h}\right\}_{t=1}^{T-h}$; (ii) the MIDAS weight hyperparameters, $\Theta_h|\boldsymbol{\beta}_h, \left\{y^*_{t+h}\right\}_{t=1}^{T-h}$; and (iii) the latent variable, $y^*_{t+h}|\boldsymbol{\beta}_h, \Theta_h$

for $t = 1, ..., T - h$. As mentioned above, the slope parameters have conjugate normal posteriors. The MIDAS hyperparameters are drawn via the MH-in-Gibbs step, assuming a Gamma proposal density truncated so that the aforementioned restriction, $\theta_{2,n,h} \geq \theta_{1,n,h}$, holds. We draw the latent variable sequentially from independent truncated normal densities, where the direction of the truncation depends on the value of the observed binary indicator. Details for the algorithm and the derivation of the posterior hyperparameters are included in Appendix A.

We compute intervals for the event probabilities $P[y^*_{t+h} \geq 0|\Omega_t]$ by constructing a direct forecast for the event probability at each draw of the sampler and each integer forecast origin ($t = 1, ..., T - h$). Then, for each forecast origin, we compute the mean as the predicted probability and use quantiles (the 16^{th} and 84^{th} quantiles) to compute intervals. To compute out-of-sample forecasts, for $P[y^*_{T+h} \geq 0|\Omega_T]$, we use draws of $\boldsymbol{\beta}_h$ and Θ_h obtained by estimating (5) to compute event probability forecasts conditional on $\mathbf{Z}_T(\Theta)$.

3.2 Forecasting Low-Frequency Variables Using High-Frequency Predictors

In the model descriptions in the previous sections, the parameters are defined and estimated assuming that the in-sample forecast origin coincides with the observation of the low-frequency variable, that is, for each $t = 1, 2, ..., T - h$ (at low-frequency), we observe the event indicator S_{t+h} and the conditioning information set Ω up to t, which is in the case of a single predictor:

$$\Omega_t = \left\{ z_t^{(m)}, z_{t-1/m}^{(m)}, ..., z_{t-(m-1)/m}^{(m)}, z_{t-1}^{(m)}, z_{t-1-1/m}^{(m)}, ..., z_{t-1-K/m}^{(m)} \right\}.$$

However, one of the advantages of MIDAS models is that forecasts can be updated between observations of the dependent variable, at the highest frequency available. For example, an intra-period forecast origin contains information

$$\Omega_{t-j/m} = \left\{ z_{t-j/m}^{(m)}, ..., z_{t-(j-K/m)}^{(m)} \right\},$$

where $j = 1, ..., m-1$ represents high frequency information that is available after the observation of S_{t+h-1} but before the observations of S_{t+h}. Clements and Galvão (2008) show how MIDAS regressions can employ intra-quarter monthly data to nowcast quarterly GDP growth. They estimate separate MIDAS regression models for each monthly horizon.

We apply a different strategy based on the approach in Ghysels et al. (2019): we generate a forecast at each intra-period observation of the high-frequency variable using the model parameters estimated at the last observation of the low-frequency variable. Thus, the model is estimated just once per each low-frequency horizon but updated within the quarter using new high-frequency information.

Given a draw from the sampler, we have estimates of $(\boldsymbol{\beta}_h, \Theta_h)$ constructed using S_{t+h} and Ω_t where $t = 1, ..., T-h$. As described earlier, we can use these coefficients to compute the h-step-ahead event probability forecast $P[y^*_{T+h} \geq 0|\Omega_T]$. However, these forecasts assume we are able to observe all intra-quarter high-frequency data between $T-1$ and T. We could instead anticipate the computation of this h-step ahead event probability forecasts by considering conditional information sets at the higher frequency: $\Omega_{T-1+(1/m)}, \Omega_{T-1+(2/m)}, ..., \Omega_T$. This means that the last observation available to compute the probability forecasts do not coincide with the end of the low-frequency period. For example, if a weekly predictor is applied to forecast quarterly events, then, the conditioning information set does not coincide with the end of the quarter as the estimation sample employed to compute $(\boldsymbol{\beta}_h, \Theta_h)$. For the case of a single regressor, these muti-step direct forecasts are computed as:

$$\begin{aligned} P[y^*_{T+h} &\geq 0|\Omega_{T-(j/m)}] = \Phi\left(\beta^{(i)}_{0,h} + \beta^{(i)}_{1,h}\sum_{k=1}^{K}\varpi\left(k;\boldsymbol{\theta}^{(i)}_h\right)z^{(m)}_{T-(j/m)-(k-1)/m}\right), \\ \text{for } j &= 0, ..., m-1, \end{aligned} \tag{7}$$

where the supercripts (i) denote a sampler draw for $(\boldsymbol{\beta}_h, \Theta_h)$, indicating that as before we are able to compute intervals that assess sampling uncertainty for each high-frequency forecasted event

probability.

The sequence of probabilistic forecasts $P[y^*_{T+h} \geq 0|\Omega_{T-(j/m)}]$ for $j = 0, ..., m-1$ all target the outcome S_{T+h}. An evaluation of the accuracy for each j is a way to see the relevance of intra-quarter data news on h-step-ahead event probability forecasting. The evaluation is similar to the one in Banbura, Giannone, Modugno and Reichlin (2013), who use a filter with fixed parameters to update low-frequency nowcasts as new high-frequency data is released. We provide an evaluation of our benchmark approach described in (7) with an approach that would estimate the forecasting model for high-frequency horizons in Appendix B.

3.3 Evaluation of Event Probability Forecasts

The main output of the models described in Section 2 are forecasts for event probabilities. In the literature, the performance of recession probability forecasts from the standard probit has been evaluated using different methods. The pseudo-R^2 (Estrella and Mishkin, 1998) compares the log-likelihood function of the model with predictors to a model that includes only an intercept. Thus, the pseudo-R^2 measures in-sample fit and is less applicable for evaluating out-of-sample predictions. Chauvet and Potter (2005) uses the out-of-sample Briers score but Benedetti (2010) argues that it is not suitable for rare events. Because NBER recessions occur in only 10.8% percent of the out-of-sample months since 1977M1, we forgo analysis with the Briers score.

An alternative is the AUROC [see Berge and Jorda (2011) and Liu and Moench (2016)], which has a number of useful features. It does not require forecasted probabilities to be converted into binary events, does not rely on a specific loss function, and has become the measure of choice for classification problems (Berge and Jorda, 2011). The AUROC, however, is not a proper score: deviations from true probabilities can improve the score in some circumstances. Bouallegue et al. (2019) argue that the AUROC is a measure of potential skill in classification of the binary events, but they suggest alternative proper scores for rare events: the logarithm (or ignorance) score (LS) and the diagonal elementary score (DES).

The LS assumes a loss function more appropriate for rare events than the Brier's score. Assume we observe predicted probabilities, $P_r = P[y^*_{T+h+r} \geq 0|\Omega_{T+r}]$, computed over the out-of-sample period for $r = 1, ..., R$ observations. If the observed outcomes are $S_r = S_{T+h+r}$, then the out-of-sample LS score is:

$$LS(h) = \frac{1}{R} \sum_{r=1}^{R} - \ln |1 - S_r - P_r| ,$$

where the weight assigned to mistaken classification depends on $\ln(P_r)$.

The DES assumes the objective is to maximize the classification between events and non-events (Bouallegue et al., 2018).[5] Bouallegue et al. (2019) recommends the DES over the LS when binary events are rare but have high-impact consequences and false positives do not cause large losses. The out-of-sample DES is:

$$DES(h) = \frac{1}{R} \sum_{r=1}^{R} \pi \mathbf{I}[P_r > \pi](1 - S_r) + (1 - \pi)\mathbf{I}[P_r \leq \pi]S_r,$$

where $\mathbf{I}[.]$ is an indicator function.

Following Bouallegue et al. (2019), we convert three of the metrics above to skill scores that compare the performance between the probability forecasts P_r and a reference forecast equal to the constant probability forecast using π (or the unconditional forecast). Skill scores measure the reliability and resolution of the probabilistic forecasts when using these proper score functions (Bouallegue et al., 2018) used to evaluate the performance of economic forecasters (Galbraith and van Norden, 2012). The skill score measures for the LS and DES are:

$$\begin{aligned} LSS(h) &= 1 - \frac{LS(h)}{LS^{unc}(h)}, \\ DESS(h) &= 1 - \frac{DES(h)}{DES^{unc}(h)}. \end{aligned}$$

The LSS can be negative, where the $DESS$ is always positive. In both cases, these skill scores

[5] The score is equivalent to the Kuipers score when the threshold is equal to the unconditional probability of the event ($\pi = \frac{1}{R} \sum_{r=1}^{R} S_r$).

measure gains with respect to the unconditional probability forecast. In the case of the AUROC, one can convert values to a skill score as:

$$ROCS = 2AUROC - 1.$$

We apply these measures of the accuracy of event probabilistic forecasts in the next section.

4 Empirical Applications

In this section, we apply the MIDAS-Probit model to three empirical macroeconomics research questions of interest. The empirical exercises aim to exploit whether MIDAS-Probit models are a new relevant addition to event probability forecasting in macroeconomics.

4.1 NBER Recession Probabilities using the Spread

As in Estrella and Mishkin (1998), Chauvet and Potter (2005) and Kauppi and Saikkonen (2008), we use the NBER chronology of business cycles, where the peak defines the last month of an expansion phase and the trough defines the last month of a recession phase. Thus, we define $S_t = 1$ for the month after the peak through the month at the trough. We use the spread between the 10-year Treasury bond and the 3-month Treasury Bill to construct one-year-ahead forecasts. Our main innovation in this exercise is the use of weekly-sampled spread data using eq. (2) compared with the monthly-sampled spread data using eq. (1) as a means of improving the forecast.[6] As in Liu and Moench (2016), we use the AUROC to evaluate both in-sample and out-of-sample predictive performance.

We set the number of weekly lags for the spread in (2) to 32 and consider just one monthly lag

[6] As our aim is to evaluate possible gains from disaggregation (as in Andreou et al. (2010)), we do not consider intra-month information sets in this exercise. Our main results are for the weekly-sampled spread. If we use the daily spread, as suggested by Andreou et al. (2013), to forecast output growth, the AUROC is virtually the same as with weekly data. Thus, as in Galvao (2013), we prefer to use weekly instead of daily data.

in (1), as in Chauvet and Potter (2005), and also both z_t and z_{t-6} as in Liu and Moench (2016).[7] The full-sample estimates use data from 1962M1 to 2020M4. For the (recursive) out-of-sample experiments, the forecasting origins begin in 1976M1 and go through 2019M4 ($R = 520$). At each new forecasting origin, we re-estimate all parameters of both (1) and (2) using all of the data available at that time and the algorithms described in detail in the appendix, and we then compute forecasts using direct forecasting as explained in Section 2.1.

Table 1 shows the AUROC results for both the full-sample and the out-of-sample experiments. The full-sample AUROC values are computed by using the time series of predicted probabilities conditional on each draw of the parameters. The values in the table are the means across all 5000 draws and the values in brackets are the 68% intervals using the corresponding quantiles of the AUROC empirical density.

For out-of-sample forecasts, both the parameters and conditioning information set change for each computed probability forecast. Thus, we first compute the posterior mean of the probability forecasts at each forecast origin (integrating out parameter uncertainty), then compute the AUROC for the implied time series of the predicted probability. To account for parameter uncertainty, Table 1 also reports the 68% coverage of the AUROC, computed similarly to the full-sample predictions, but with the caveat that the AUROC headline value is not necessarily centered in the indicated interval. Results in Table 1 clearly indicate that using weekly data improves forecasting performance: The headline AUROC of the MIDAS-Probit specification increases in comparison with both monthly specifications for the out-of-sample exercise. Moreover, the 68% intervals for the AUROC values do not overlap.

In summary, this first exercise suggests that the MIDAS-Probit is a useful addition to modelling approaches to extract high-frequency information from financial variables to predict recession

[7] Chauvet and Potter (2005) and Kauppi and Saikkonen (2008) suggest that a dynamic version of the probit—i.e., if lags of the latent variable y_t^* are included as predictors—produces more accurate recession probabilities at short-horizons. The dynamic model accounts for serial correlation in the business cycle phases but makes multi-step forecasting more complicated. We can accommodate serial correlation by adding economic activity as shown in our second empirical exercise. We leave study of the mixed frequency dynamic probit for future research.

events.

4.2 Evaluation of Forecasts for Quarterly Contraction Events

4.2.1 The High-Impact Rare Event

Recessions are high-impact economic events, but historical chronologies for recession periods are not always available. Even when they are available, additional assumptions are required to define recession phases at quarterly frequency. As Adrian et al. (2019) has suggested, a key input for stabilization policy is the identification of periods when GDP growth is "vulnerable," as future outcomes may lead to an economic contraction. We consider an alternative high-impact event that is based on the currently-observed, quarterly-sampled year-on-year U.S. GDP growth ($g_t = 100[(GDP_t/GDP_{t-4}) - 1]$). If $g_t < 0.5$, we set $S_t = 1$ and interpret it as a contraction quarter.[8] Because the contraction event is constructed from a moving average of the past year's data, the timing of this event is not simultaneous to the NBER turning points. While there is a contraction event for all of the NBER recession phases, the timing of the events are, in general, delayed.

During the out-of-sample period (from 1985Q2 up to 2020Q1), the unconditional probability of a contraction quarter is 7.8%, so this is an even rarer event than the monthly NBER recession, and not far from the 5% growth at risk quantile in Adrian et al. (2019). As a consequence, to compare probabilistic forecasts for the contraction event, we consider the LSS and DES score measures described in Section 3.3 in addition to the AUROC. For interpretability, we consider skill measures.

4.2.2 The Set of Predictors and Economic Implications

The literature suggests that both the yield curve spread and the Chicago Fed Financial Condition Index (NFCI) are potential predictors for contraction events. The NCFI anticipates periods when

[8] Contractions are normally periods of negative growth rates. But to define a simple rule using observed GDP data, we have a trade-off between a rule that finds contractions due to weather-related outliers (by defining negative growth periods in quarterly GDP growth) and one that finds no contraction periods even if the NBER has identified a recession (if the contraction rule is applied to year-on-year growth rate). We settle on a rule based on year-on-year growth rates (to avoid outliers) but with a threshold that is slightly larger than zero (0.5%), so it does not completely miss a NBER recession period.

growth is at risk (Adrian et al., 2019), although not all of these periods materialize as a contraction phase. The yield curve spread is often negative before recessions (Chauvet and Potter, 2005), so it may also lead contraction events. Both the spread and the NFCI are available weekly; thus, we use high-frequency data on both these financial variables when computing contraction probabilities using the MIDAS Probit specification described below.

Berge and Jorda (2011) provide evidence that the monthly Chicago Fed National Activity Index (CFNAI) is informative about U.S. recession phases. Thus, we consider the CFNAI as an alternative low-frequency predictor. Including an economic activity variable addresses whether high-frequency financial variables improve forecasting performance when the target is a low-frequency event such as a contraction (as also suggested by Plagborg-Moller et al. (2020)). This analysis extends the empirical results for GDP growth in Andreou et al. (2013) and Galvao (2013).

4.2.3 The MIDAS-Probit Specification

We forecast one-quarter-ahead and one-year-ahead using two benchmark specifications—a probit with the quarterly average CFNAI and a MIDAS probit with monthly CFNAI with 12 lags. We then construct forecasts for the same horizons that include various weekly financial variables and the CFNAI from:

$$P[y^*_{t+h} \geq 0|\Omega_t] = \Phi\left(\beta_{0,h} + \beta_{1,h}\sum_{k=1}^{K_1} \varpi\left(k;\boldsymbol{\theta}_{1,h}\right) z^{(m=3)}_{1,t-\frac{k-1}{3}} + \beta_{2,h}\sum_{k=1}^{K_2} \varpi\left(k;\boldsymbol{\theta}_{2,h}\right) z^{(m=13)}_{2,t-\frac{k-1}{13}}\right), \qquad (8)$$

where z_1 is the monthly CFNAI and z_2 is the weekly financial variable. We include one financial variable at a time. We set the number of lags for the monthly CFNAI to 12 months ($K_1 = 12$), the number of lags of spread to 32 weeks (approx. 7 months) and the one for the NFCI to 16 weeks.[9]

Financial variables typically predict macro events at long horizons. Thus, for $h = 1$, we apply the high-frequency lag operator from $z^{(m=13)}_{2,t-3}$ instead of $z^{(m=13)}_{2,t}$, implying that the most up-to-date

[9] We experimented with alternative numbers of lags for all predictors. Preliminary evidence on forecasting performance support our described choices.

information on the financial variable is one-year before the event. Preliminary results suggest that this choice improves forecasting performance for both financial variables, supporting the evidence in Chauvet and Potter (2005) and Liu and Moench (2016).

4.2.4 Out-of-Sample Exercise and Predicted Probabilities

We compute the out-of-sample probability forecasts for contraction events from 1985Q2 through 2020Q1 ($R = 140$).[10] Because there is no excessive variability in the coefficients $\boldsymbol{\beta}_h, \boldsymbol{\Theta}_h$, we re-estimate the models every 10 quarters with increasing samples.[11] Using the posterior mean draws for $\boldsymbol{\beta}_h, \boldsymbol{\Theta}_h$, we vary the information set each week within a quarter—that is, for each quarterly forecast origin, we compute the intra-quarter probability forecasts for $j = 1, ..., 12$ described in (7). Figure 2 describes how the intra-quarter information for both the weekly, $z_{2t}^{(m=13)}$, and the monthly, $z_{1t}^{(m=3)}$, regressors are employed for data up to T to compute $P[y_{T+h}^* \geq 0 | \Omega_{T-(j/13)}]$ for $j = 0, ..., 12$. The Figure describes the last observations available to estimate $(\boldsymbol{\beta}_h, \boldsymbol{\Theta}_h)$ in blue, the last observations in the conditioning information set if at the end of the quarter in green ($j = 0$), and the intra-quarter information sets in orange. The available information on $z_{1t}^{(m=3)}$ does not change every week as $z_{2t}^{(m=13)}$ and Figure 2 shows how the monthly variable is updated over the quarter.

Figure 3 shows the means of the posterior probabilities for each weekly forecast origin (and shaded areas represent the contraction events that will be observed h-quarters later). Each of the h-step-ahead forecasts within a quarter predict the same outcome at $T + h + r$, where S_r indicates the target outcome for $r = 1, .., R$ as in Section 3.3. The predictive probabilities are computed for four models: (i) a probit model with quarterly CFNAI, (ii) a mixed-frequency version with only the monthly CFNAI, (iii) the monthly CFNAI and weekly spread, and (iv) the monthly CFNAI and the weekly NFCI. The top panel shows results for one-quarter-ahead forecasts ($h = 1$) and the

[10] The in-sample period employed to estimate each forecasting model varies with the availability of the predictors. Monthly CFNAI data is available from 1968, the weekly NFCI from 1971, and weekly spread from 1962.

[11] This strategy has no negative impact on forecasting performance compared with estimating at each new quarterly origin.

bottom panel for one-year-ahead forecasts ($h = 4$).

For $h = 1$, all models attribute at least 50% probability to the contraction events in some periods. Either model using only the CFNAI produces a large number of false positives at a threshold of 20%. Including a financial variable usually sharpens the probability predictions. The average predictive probability $\frac{1}{R}\sum_{i=1}^{R} P_i$ is about 17% using only the CFNAI, 9% if the spread is included, and 12% if the NFCI is included. It is not clear, however, that financial variables improve the predictive classification of contraction periods one quarter ahead.

For $h = 4$, only the MIDAS-Probit that includes the weekly spread produces event probabilities above 50% for any forecasting origin. One-year-ahead forecasted probabilities from models using only the CFNAI do not vary much over time (standard deviation of about 0.04 using the monthly CFNAI) in comparison with predictions obtained using financial variables (standard deviation of 0.06 with NFCI and 0.12 using the spread). For one-year-ahead, the inclusion of financial variables clearly reduces the frequency of false positives.

4.2.5 Comparing Predicted Probabilities using Skill Scores

Figure 4 presents three different skill scores (ROCS, top panel; LSS, center panel; and DESS, bottom panel) computed for the out-of-sample period for each intra-quarter information set ($j = 0, ..., 12$ with R observations for each j). The values in the left panel are for $h = 1$ and the ones in the right panel for $h = 4$. In all cases more accurate forecasts lead to higher skill values.

The ranking across forecasting models changes with the loss function employed. Using the ROCS, models with financial variables usually do not outperform the quarterly probit. If we use instead the LSS, we find that financial variables improve forecasts, especially at the beginning of the quarter when only information on the past quarter is available. Finally, we see a larger gap between the models with financial variables in comparison with models with only the CFNAI using the DESS. The main reason is that this loss function defines occurrence of an event using the unconditional probability (7.8%); as a consequence, a larger number of false positives are classified

using the CFNAI in contrast with models that also include one of the financial variables.[12]

Monthly CFNAI only has predictive ability (as measured by the AUROC and the DES) over the quarterly CFNAI for one-quarter-ahead forecasts. The impact of varying the conditioning information set on the relative performance also changes with the quarterly horizon. Because our outcome measure is constructed from a moving average of the past year's data, models with the spread tend to perform worse as more data is accumulated over the quarter. Intra-quarter information has a large impact on ROCS and LSS than using the DES. Finally, using the latest available information may improve forecasts using the NFCI but not using the spread, where older data may improve forecasting performance.

4.2.6 Contraction in 2020Q1?

The 68% interval for the probability of a contraction in 2020Q1 using information up to 2019Q1 (one-year-ahead forecasts) is [3%,7%] if the the NFCI is the weekly indicator included in the MIDAS Probit specification in (8), and [22%,33%] if the spread is included instead.

As these intervals provide us with dissimilar information on the probability of a contraction in 2020Q1, the evaluation described earlier may help us to choose which probabilistic forecast to trust. If we remove 2020Q1 from the evaluation sample, the ROCS averaged over all j is 0.42 using the spread and 0.24 using the NFCI. The DES skill also favors the spread as the value is 0.35 for the spread and 0.20 for the NFCI. The LS inverts the ranking in favour of the NFCI as the value is -0.03 (so worse than the unconditional) using the spread but 0.03 using the NFCI. If the classification of the contraction event is important (either the potential or actual), then we should favour the prediction obtained with the spread. The probability of a contraction is then clearly above the unconditional probability but values are not as high as the indication obtained during previous events.

[12] The model with NFCI performs better under LS than the model with the spread. This corroborates the evidence using quantile regressions to favor the NFCI since the LS is linked to quantile-weighted scores. If we are mainly interested in maximizing the frequency of correct classification of contraction events, then the DES—for which the weekly spread is the chosen financial variable—is more appropriate.

As the financial market could not have anticipated the economic effects of the COVID-19 pandemic in 2019, which is clearly the main cause of the contraction observed in 2020Q1, we see these probabilistic predictions as a weak indication of a downturn in 2020Q1 even if there were no pandemic.

4.3 Classifying Turning Points using the Weekly Economic Index

Lewis et al. (2020) propose the weekly economic index (WEI) to monitor the U.S. economy in real time. Indeed, since March 21, 2020 the weekly measure of economic activity has been updated and released every week. The index is computed using 10 different weekly time series, including some traditional business cycle indicator variables, such as initial claims of unemployment insurance. As the NBER turning points chronologies are available at monthly frequency, the MIDAS Probit is a candidate approach to extract recession probabilities information from the WEI that may improve real-time classification of turning points.

To obtain accurate real-time recession probabilities from the weekly WEI, we modify the MIDAS-Probit specification in (2) to accommodate two relevant features of the data. First, the NBER Business Cycle Dating Committee calls peaks and troughs with a delay. On June 9th, 2020 the committee called a peak for February 2020, implying a recession starting in March 2020. As a consequence, March, April, and May are unclassified when computing recession probabilities forecasts during these months. Thus, we estimate the model using the information on the binary variable only up through 2018, implying that the model outputs real-time classifications, varying only with changes in the availability of the information set as described in Section 2.4.

Second, the WEI may lag business cycle phases because it includes variables related to unemployment claims, which is classified as lagging the reference cycle in Stock and Watson (1999). Thus, we use both past and future information on weekly WEI to compute the probability of being in recession in month t. As our aim is to identify turning points that are published with a delay of a few months with an economic activity variable available with a delay of days, using available

"future" information from leads (as proposed in Andreou et al. (2013)) is feasible and may improve accuracy.

4.3.1 The MIDAS-Probit Specification

The MIDAS-Probit model to extract information from leads and lags of weekly WEI, $z_t^{(m)}$, to compute pseudo-real-time recession probabilities is:

$$P[y_t^* \geq 0 | \Omega_{t+(K_f/m)}] = \Phi\left(\beta_0 + \beta_f \sum_{k=1}^{K_f} \varpi\left(k; \boldsymbol{\theta}_f\right) z_{t+(k/m)}^{(m)} + \beta_b \sum_{j=0}^{K_b-1} \varpi\left(k; \boldsymbol{\theta}_b\right) z_{t-(k/m)}^{(m)}\right), \quad (9)$$

where $\varpi\left(k; \boldsymbol{\theta}_b\right)$ weights contemporaneous and lag weekly values of $z_t^{(m)}$, and $\varpi\left(k; \boldsymbol{\theta}_f\right)$ weights K_f lead values. In the algorithm described in Section 3, we impose restrictions on the parameters of the beta functions such that larger weights are given for $x_t^{(m)}$ at weekly periods that are near t. The restrictions are such that the $\varpi\left(k; \boldsymbol{\theta}_f\right)$ weighting function decreases with the lead/lag horizon. We assume that $K_f = 7$ and $K_b = 8$, resulting in a two-sided window of 15 weeks.

Figure 5 shows posterior mean estimates and 68% coverage bands for both weighting functions (lead and lag) with the weights normalized to sum up to 1. The first lead $(t+1)$ is assign proportionately more weight than the first lag $(t-1)$, consistent with unemployment claims being a lagging indicator and requiring future values to identify turning points.

4.3.2 Weekly Recession Probabilities

As we discussed in the earlier exercises, there is a trade-off between correctly predicting rare events and inducing false positives. Figure 6 shows the posterior mean of the predicted probabilities for the model in (9) estimated with data up to 2018M12 computed for all intra-period (weekly) information sets as described in Section 3.3 including 68% bands for predicted probabilities. One can clearly see that when $S_t = 0$, the predicted probabilities are usually below the unconditional value (10%), with the exception of the last three months of 2015. This supports our claim that the combination

of the MIDAS-Probit and the WEI provides an effective way to monitor recessions in real time.

4.3.3 Turning Points Predictive Performance

Figure 7 shows the predicted probabilities of recession of the weekly WEI lead/lag MIDAS probit around two events: the 2009M6 trough (left) and the 2020M2 peak (right). The model is estimated using the NBER recession dates as the binary variable and the weekly WEI indicator using data through 2018M12. Thus, values in the left panel are in-sample estimates; values in the right panel are out-of-sample. Dates on the x-axis are forecast origins in weeks; intra-period information sets are as described in Section 3.3. We also present predicted probabilities using an alternative specification that includes only WEI lags – that is, the MIDAS probit is re-estimated with $\beta_f = 0$ in eq. (9). Dashed lines represent the 68% confidence bands for both specifications. The blue line in Figure 7 represents the turning point date.

The left plot in Figure 7 clearly indicates that the MIDAS-Probit specification with both leads and lags outperforms the one with only lags: the model with both leads and lags reduces the number of weeks that the predictions indicate false positives from July onwards. The predicted probabilities from the MIDAS-Probit with both leads and lags are below 90% for the first time in 2009M7, a few weeks after the trough. If, instead, only past WEI values are employed, the predicted probabilities are below 90% only at the end of August, despite the fact that they are in-sample estimates. This evidence supports the use of the MIDAS-Probit specification in (9) to predict recessions using the WEI in real-time. When the model is applied out-of-sample for the 2020M2 peak, the probability of recession goes above 90% on February 15^{th}. If we consider only past WEI values instead, then the model indicates a turning point only on April 4^{th}, lagging the NBER business cycle judgment on the turning point date.

But does this mean that the MIDAS-Probit with the WEI is able to anticipate the NBER 2020 peak announcement? As the WEI data were published every week in April and May with limited delay, we can say we could have observed the recession probabilities above 90% as the value

for April 4^{th} was announced—two months after the NBER announcement in early June. Jeremy Piger publishes smoothed monthly recession probabilities using a model with monthly economic indicators based on Chauvet and Piger (2008). Evidence from the release dates of their probabilities in ALFRED suggests that the early May release included a 100% probability of recession in March, suggesting a peak in February. Thus, using a weekly economic index published with limited delay can indeed identify turning points faster in real-time. Our contribution here is to show how the MIDAS-Probit can be helpful to filter information from WEI to predict turning points.

5 Conclusions

In this paper, we propose a new tool for macroeconomic forecasting of key low-frequency events. The MIDAS-Probit model is effective in delivering high-frequency probability forecasts for macroeconomic events by exploiting the predictive content of high-frequency financial and economic indicators. We provide empirical evidence that weekly-sampled financial variables—in particular, the yield curve spread—help predict contraction events at a one-year horizon. We also show how to filter the information of the WEI to obtain weekly-updated NBER recession probabilities that are able to accurately anticipate turning points.

We consider three different loss functions when evaluating the additional predictive content of financial variables to predict contractions. One of them, the diagonal score, has been recently proposed in the climatology literature by Bouallegue et al. (2018) as a proper score to evaluate the classification ability of alternative forecasts of rare events. The diagonal score heavily penalizes false positives and suggests that the yield curve spread predictive content is superior to the financial condition index to predict contractionary GDP quarters one year in advance.

References

Adrian, T., Boyarchenko, N. and Giannone, D. (2019). Vulnerable growth, *American Economic Review* **109**(4): 1263–1289.

Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data, *Journal of the American Statistical Association* **88**(422): 669–679.

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

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

Audrino, F., Kostrov, A. and Ortega, J.-P. (2019). Predicting U.S. bank failures with MIDAS logit models, *Journal of Financial and Quantitative Analysis* **54**(6): 2575–2603.

Banbura, M., Giannone, D., Modugno, M. and Reichlin, L. (2013). Now-casting and the real-time data flow, *Handbook of Economic Forecasting, volume 2A*, Elsevier, chapter 4, pp. 195–237.

Bauer, M. D. and Mertens, T. M. (2018). Economic forecasts with the yield curve, *FRBSF Economic Letter* **07**.

Benedetti, R. (2010). Scoring rules for forecast verification, *Monthly Weather Review* **138**(1): 203–211.

Berge, Travis, J. and Jorda, O. (2011). Evaluating the classification of economic activity into recessions and expansions, *American Economic Journal: Macroeconomics* **3**(2): 246–277.

Bouallegue, Z. B., Haiden, T. and Richardson, D. S. (2018). The diagonal score: definitions, properties, and interpretations, *Quarterly Journal of the Royal Meteorological Society* **144**(714): 1463–1473.

Bouallegue, Z. B., Magnusson, L., Haiden, T. and Richardson, D. S. (2019). Monitoring trends in ensemble forecast performance focusing on surface variables and high-impact events, *Quarterly Journal of the Royal Meteorological Society* **145**(721): 1741–1755.

Carriero, A., Clark, T. and Marcellino, M. (2020). Nowcasting tail risks to economic activity with many indicators, *FRB of Cleveland Working Paper n. 20-13R* .

Casella, G. and George, E. I. (1992). Explaining the Gibbs sampler, *American Statistician* **46**(3): 167–174.

Chauvet, M. and Piger, J. (2008). A comparison of real-time performance of business cycle dating methods, *Journal of Business and Economic Statistics* **26**(1): 42–49.

Chauvet, M. and Potter, S. (2005). Forecasting recessions using the yield curve, *Journal of Forecasting* **24**(2): 77–103.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *American Statistician* **49**(4): 327–335.

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

Estrella, A. and Mishkin, F. S. (1998). Predicting U.S. recessions: Financial variables as leading indicators, *The Review of Economics and Statistics* **80**(1): 45–61.

Foroni, C., Marcellino, M. and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials, *Journal of the Royal Statistical Society: series A* **178**(1): 57–82.

Galbraith, J. W. and van Norden, S. (2012). Assessing gross domestic product and inflation probability forecasts derived from bank of england fan charts, *Journal of the Royal Statistical Society, series A* **175**(3): 713–727.

Galvao, A. B. (2013). Changes in predictive ability with mixed frequency data, *International Journal of Forecasting* **29**(3): 395–410.

Gelfand, A. R. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**(410): 398–409.

Ghysels, E., Plazzi, A., Valkanov, R., Rubia, A. and Dossani, A. (2019). Direct versus iterated multi-period volatility forecasts: Why MIDAS is king, *Annual Review of Financial Economics* **11**: 173–195.

Ghysels, E. and Qian, H. (2019). Estimating MIDAS regressions via OLS with polynomial parameter profiling, *Econometrics and Statistics* **9**: 1–16.

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

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

Greenberg, E. (2013). *Introduction to Bayesian Econometrics*, 2nd edn, Cambridge University Press.

Kauppi, H. and Saikkonen, P. (2008). Predicting U.S. recessions with dynamic binary response models, *Review of Economic and Statistics* **90**(4): 777–791.

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

Lewis, D., Mertens, K. and Stock, J. (2020). U.S. economic activity during the early weeks of the SARS-Cov-2 outbreak, *Federal Reserve Bank of New York Staff Reports* **April**.

Liu, W. and Moench, E. (2016). What predicts US recessions?, *International Journal of Forecasting* **32**(4): 1138–1150.

Mogliani, M. and Simoni, A. (2020). Bayesian MIDAS penalized regressions: estimation, selection and prediction, *Journal of Econometrics* **forthcoming**.

Pettenuzzo, D., Timmermann, A. and Valkanov, R. (2016). A MIDAS approach to modelling first and second moment dynamics, *Journal of Econometrics* **193**(2): 315–334.

Plagborg-Moller, M., Reichlin, L., Ricco, G. and Hasenzagl, T. (2020). When is growth at risk?, *Brooking Papers on Economic Activity, BPEA Conference Drafts* **March**.

Stock, J. H. and Watson, M. W. (1999). Business cycle fluctuations in US macroeconomic time series, *Handbook of Macroeconomics, vol 1*, Elsevier, pp. 3–64.

A Bayesian Estimation of MIDAS Probit models

A.1 The Metropolis-in-Gibbs Algorithm

The MCMC sampler for the mixed-frequency probit in (5) can be broken down into blocks: the block for the slope coefficients, $\boldsymbol{\beta}_h$; the beta weighting function parameters, Θ_h; and the latent data, $\left\{\mathbf{y}^* = \mathbf{y}_T^{*h} = y_{1+h}^*, y_{2+h}^*, ..., y_{T-h+h}^*\right\}$.

Table 1: Priors for Estimation		
Parameter	Prior Distribution	Hyperparameters
$\boldsymbol{\beta}_h$	$N\left(\mathbf{m}_0, \mathbf{M}_0\right)$	$\mathbf{m}_0 = \mathbf{0}_{n+1}$; $\mathbf{M}_0 = \mathbf{I}_{n+1}$
$\boldsymbol{\theta}_i$	$\Gamma\left(\mathbf{d}_0, \mathbf{D}_0\right)$	$\mathbf{d}_0 = \mathbf{1}_2$; $\mathbf{D}_0 = \mathbf{1}_2; \Delta_i = 1$

As described in Table 1, we adopt the standard normal prior for the slope coefficients and the constant, and we make use of the identification restriction that $var(u_{t+h}) = 1$ as indicated in (6). The priors for the parameters of the weighting function are gamma distributed (as these parameters should be positive) and constructed to center around the belief that the high-frequency data is equally weighted.[13] Table 1 also includes the hyperparameter Δ_i that is designed to control the acceptance of the metropolis step.

A.1.1 Drawing $\boldsymbol{\beta}_h$ conditional on $\Theta_h, \mathbf{y}^*$

Conditional on Θ_h and $\mathbf{y}^*$, (6) is a linear regression. Let $\mathbf{Z}$ represent the $(T-h) \times (1+n)$ matrix of stacked $\mathbf{z}_t(\Theta_h)'$ vectors. Then, given the prior $N(\mathbf{m}_0, \mathbf{M}_0)$, a draw of $\boldsymbol{\beta}_h$ can be made from $\boldsymbol{\beta}_h|\Theta_h, \mathbf{y}^* \sim N(\mathbf{m}, \mathbf{M})$, where

$$\mathbf{M} = \left(\mathbf{M}_0^{-1} + \mathbf{Z}'\mathbf{Z}\right)^{-1}$$

and

$$\mathbf{m} = \mathbf{M}\left(\mathbf{M}_0^{-1}\mathbf{m}_0 + \mathbf{Z}'\mathbf{y}^*\right).$$

A.1.2 Drawing Θ_h conditional on $\boldsymbol{\beta}, \mathbf{y}$

Obtaining a draw of $\boldsymbol{\theta}_i$ ($i = 1, ..., n$, where the h subscript is removed for simplicity here) can be accomplished using a Metropolis-in-Gibbs step (Chib and Greenberg, 1995) to sample from the nontractable posterior distribution. The Metropolis step requires a candidate draw from a proposal density which is accepted with a probability that depends on both the likelihood and parameters' prior distribution.

We utilize a Gamma proposal density, whose hyperparameters depend on the previous accepted

[13] We could adopt a diffuse prior over the $\boldsymbol{\theta}$ hyperparameters. This would be an improper prior and would be invalid for computation of marginal likelihoods.

draw. In other words, for the j iteration, we draw a candidate $\boldsymbol{\theta}_i^{[j]} = \left(\theta_{i,1}^{[j]}, \theta_{i,2}^{[j]}\right)'$ from

$$\theta_{i,1}^{[j]} \sim \Gamma\left(\sqrt{\Delta_i}\theta_{i,1}^{[j-1]}, \left(\Delta_i \theta_{i,1}^{[j-1]}\right)^2\right) \tag{10}$$

$$\theta_{i,2}^{[j]} \sim \Gamma\left(\sqrt{\Delta_i}\theta_{i,2}^{[j-1]}, \left(\Delta_i \theta_{i,2}^{[j-1]}\right)^2\right) \tag{11}$$

where the superscript $j-1$ represents the draw from the previous iteration. The hyperparameter Δ_i is a scaling factor that can be tuned to achieve a reasonable acceptance rate. As we have prior information on the most likely shape of the beta weighting function (increasing, decreasing, hump-shaped) for a given empirical application, then we use this information to accept only candidate draws compatible with our prior view on the shape of the beta function. For example, if we think the weighting function should be decreasing as in Figure 1, then we repeat (10) and (11) until we find a candidate draw that satisfy $\theta_{i,1}^{[j]} \leq \theta_{i,2}^{[j]}$. These restrictions help with identification of the weighting function and slope parameters and are a clear advantage of our estimation strategy.

The candidate draw is then accepted with probability $A = \min\{\alpha, 1\}$, where

$$\alpha = \frac{f\left(\mathbf{y}^* | \boldsymbol{\theta}_i^{[j]}\right)}{f\left(\mathbf{y}^* | \boldsymbol{\theta}_i^{[j-1]}\right)} \frac{dG\left(\boldsymbol{\theta}_i^{[j]} | \mathbf{d}_0, \mathbf{D}_0\right)}{dG\left(\boldsymbol{\theta}_i^{[j-1]} | \mathbf{d}_0, \mathbf{D}_0\right)} \frac{dG\left(\boldsymbol{\theta}_i^{[j-1]} | \sqrt{\Delta_i}\boldsymbol{\theta}_i^{[j]}, \left(\Delta_i \boldsymbol{\theta}_i^{[j]}\right)^2\right)}{dG\left(\boldsymbol{\theta}_i^{[j]} | \sqrt{\Delta_i}\boldsymbol{\theta}^{[j-1]}, \left(\Delta_i \boldsymbol{\theta}^{[j-1]}\right)^2\right)}, \tag{12}$$

where $f(.|.)$ reflects the conditional likelihood whose log is

$$\ln f\left(\mathbf{y}^* | \boldsymbol{\theta}_i^{[j]}\right) = \sum_t \ln \phi[y_{t+h}^* - (\mathbf{Z}_t(\Theta_h^{[j]})' \boldsymbol{\beta}_h)]$$

and $dG(.|.,.)$ is the gamma pdf and $\phi()$ is the normal pdf. Note that these steps are designed to draw the vector $\boldsymbol{\theta}_i$, that is, the parameters of one weighting function at the time. This means that we draw $\boldsymbol{\theta}_i | \boldsymbol{\theta}_{\neq i}$. In the empirical applications covered in this paper, we either have $N = 1$ or $N = 2$.

A.1.3 Drawing $\mathbf{y}^*$ conditional on $\boldsymbol{\beta}_h, \Theta_h$

Given the parameters $\boldsymbol{\beta}_h$ and Θ_h and the observables, we draw the vector $\mathbf{y}^*$ element by element, that is, $y^*_{1+h}, ..., y^*_T$ (as regression errors are assumed to be iid). Each value is draw at each MCMC iteration from a truncated normal density, that is,

$$\begin{aligned} y^*_{t+h} &\sim \begin{cases} TN_{(-\infty,0]}\left(\mathbf{Z}_t(\Theta_h)'\boldsymbol{\beta}_h, 1\right) & \text{if} \quad S_{t+h} = 1 \\ TN_{(0,\infty)}\left(\mathbf{Z}_t(\Theta_h)'\boldsymbol{\beta}_h, 1\right) & \text{if} \quad S_{t+h} = 0 \end{cases} \\ \text{for } t &= 1, ..., T-h. \end{aligned}$$

A.1.4 The Predicted Probabilities

For each posterior draw of $\boldsymbol{\beta}_h, \Theta_h$ and $\mathbf{y}^*$, we compute predictive probabilities using:

$$P(S_{t+h} = 1) = P(y^*_{t+h} \geq 0) = \Phi[\mathbf{Z}_t(\Theta_h)'\boldsymbol{\beta}_h].$$

B Computation of High Frequency Probability Forecasts

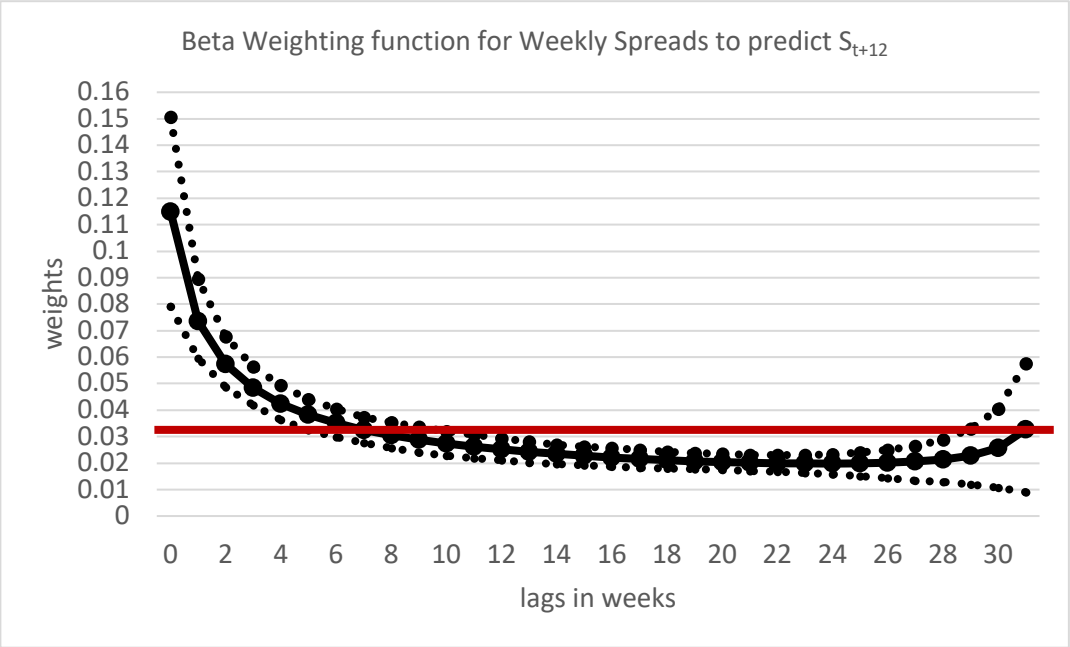
We describe here the results of a comparison between our benchmark approach described in (7) to compute high-frequency probability forecasts using the MIDAS-Probit with an approach that would estimate the forecasting model for each high-frequency horizon $\left(\boldsymbol{\beta}_{h-(j-m)}, \boldsymbol{\Theta}_{h-(j-m)} \text{ for } j = 0, ..., m-1\right)$.

Figure A1 shows 68% intervals for the predicted probabilities computed every week for a one-quarter ahead event with a mixed-frequency model that combines three different frequencies (described in detail in Section 4.2, see eq. (8)). We show the results using the approach in (7) in the top panel and the case where parameters are re-estimated for each $j = 0, ..., m-1$ in the bottom panel.

One can clearly see that the re-estimation leads to predicted probability intervals with higher variability (as one fits the weighting function and slope for each horizon), but it does not seem to improve the accuracy of the event probabilities, as the effect on the ability of the posterior mean

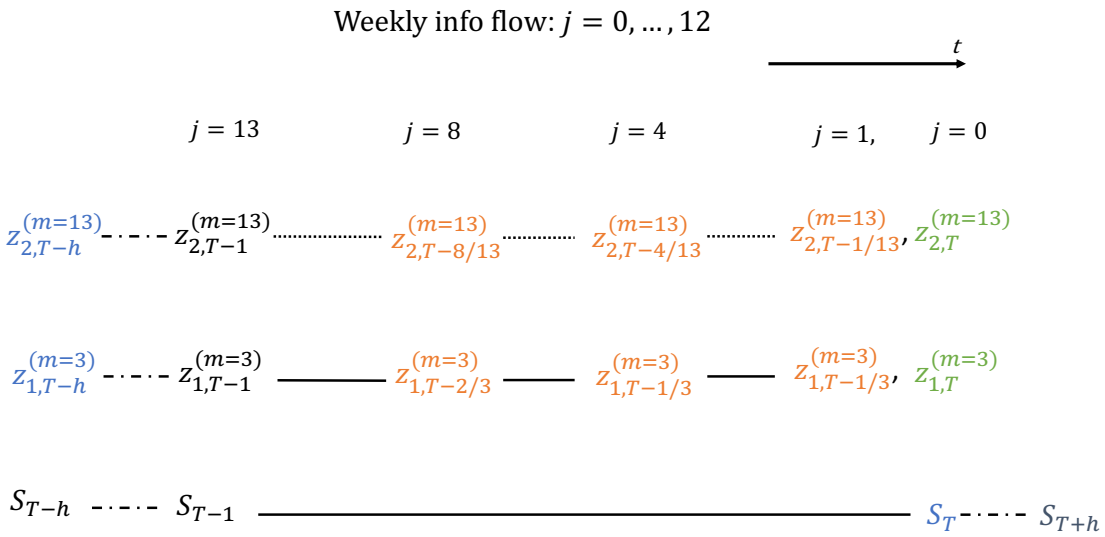
probability to classify the event is small (AUROC changes from 0.937 to 0.943). The re-estimation each week has the disadvantage of significantly increasing the computation time to obtain the weekly-updated forecasts.

Figure 1 Posterior Estimates of the Weighting Function for MIDAS-Probit with Weekly Spreads and NBER Recession Indicators.



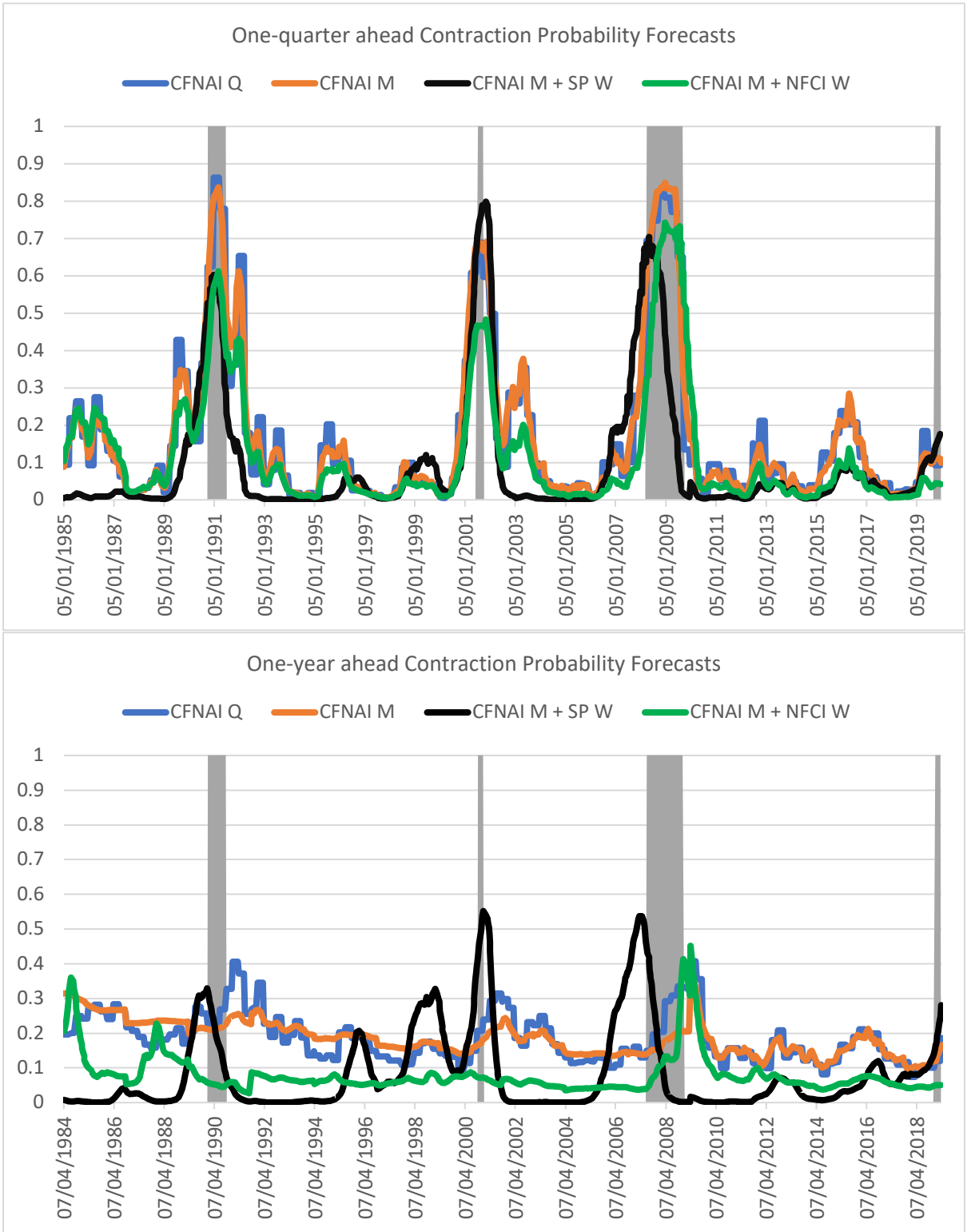
Notes: Dotted lines are 68% bands. Red line describes equal weights: $1/K$ where K=32. Sample period employed in the estimation: 1962M1-2020M4.

Figure 2: Flow of information up to end of quarter T.



Note: Flow for a quarterly binary dependent variable S_t, and two candidate predictors: z_1 is monthly and z_2 is weekly. Forecasting target is $Prob(S_{T+h} = 1)$. When estimating the parameters at T, values in blue are the last values available. Then values in green are the conditioning information set at T if only information up to the end of the quarter is considered. Values in orange are alternative conditioning sets using intra-quarter information for monthly (m=3) and weekly series (m=13). For comparison with forecasts using weekly information sets, conditioning monthly info indicated for *j=13* is repeated up to *j=9*, then the value for *j=8* is repeated up to *j=5*. Note that j=13 indicates also the latest information set if data is available up to quarter T-1. If h=1, the first two columns will be the same.

Figure 3: Posterior Mean Estimates of Out-of-Sample Quarterly Contraction Probabilities using the CFNAI, the Spread and the NFCI for weekly-updated information sets.



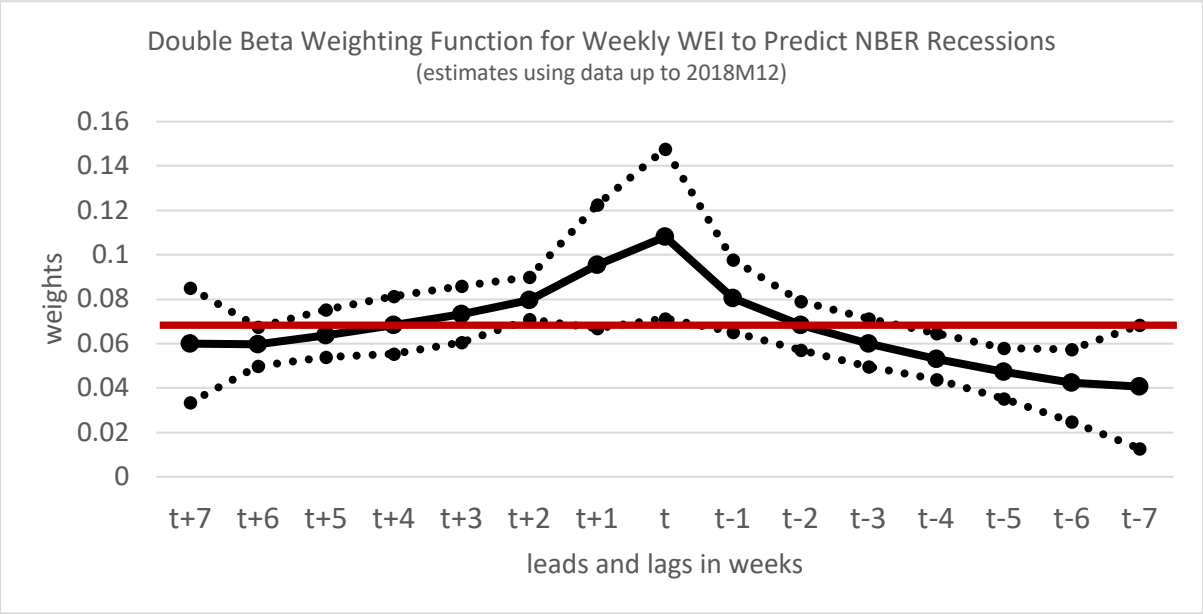
Note: Dates refer to *last weekly information set employed to compute the probability forecast*. Shaded dates indicate periods of contraction for the outcome variable observed either one-quarter (top plot) or one-year (bottom plot) later (from 1985Q2 to 2020Q1). These are computed with 4,000 draws (after initial 1,000 are removed). The model for the quarterly CFNAI (CFNAI) is a probit, all other predicted probabilities are computed using MIDAS probit specifications. CFNAI M includes 12 monthly lags of CFNAI. The other two specifications follow eq. (8) using either the weekly spread (SP W) or the NFCI (NFCI W).

Figure 4: Skill Scores for Out-of-Sample Quarterly Contraction Probabilities using the CFNAI, the Spread and the NFCI for weekly-updated information sets.

.

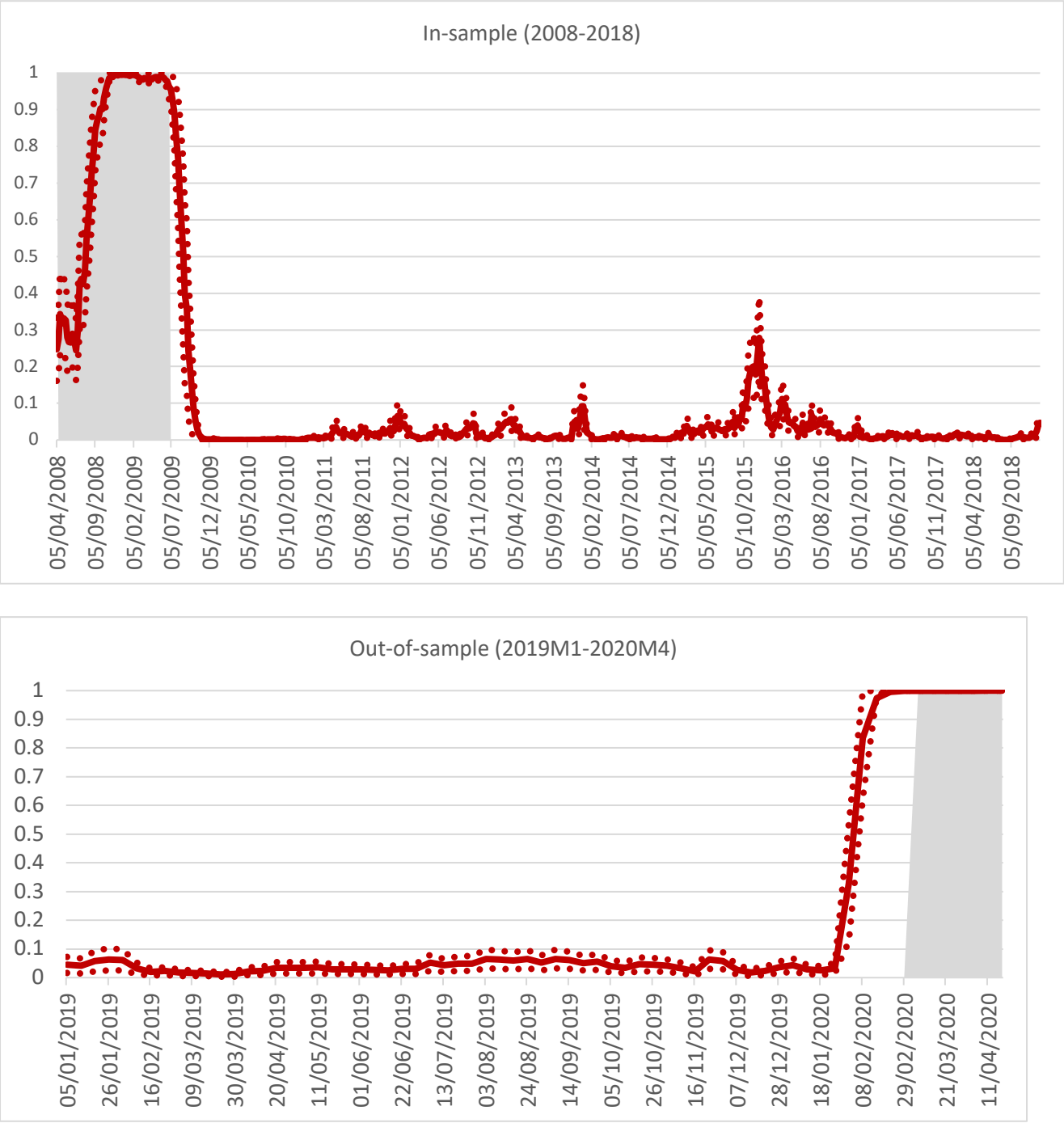
Notes: The out-of-sample period is from 1985Q2 to 2020Q1 (R=140). The horizonal axis describes *j* as in Figure 2. Check Figure 2 for how the information set on monthly and quarterly data is updated within the quarter. Details of the forecasting models are in the note to Figure 3 as these are skill scores computed for the predictions presented in Figure 3.

Figure 5: Posterior Estimates of the Weighting Functions for MIDAS-Probit with Weekly WEI



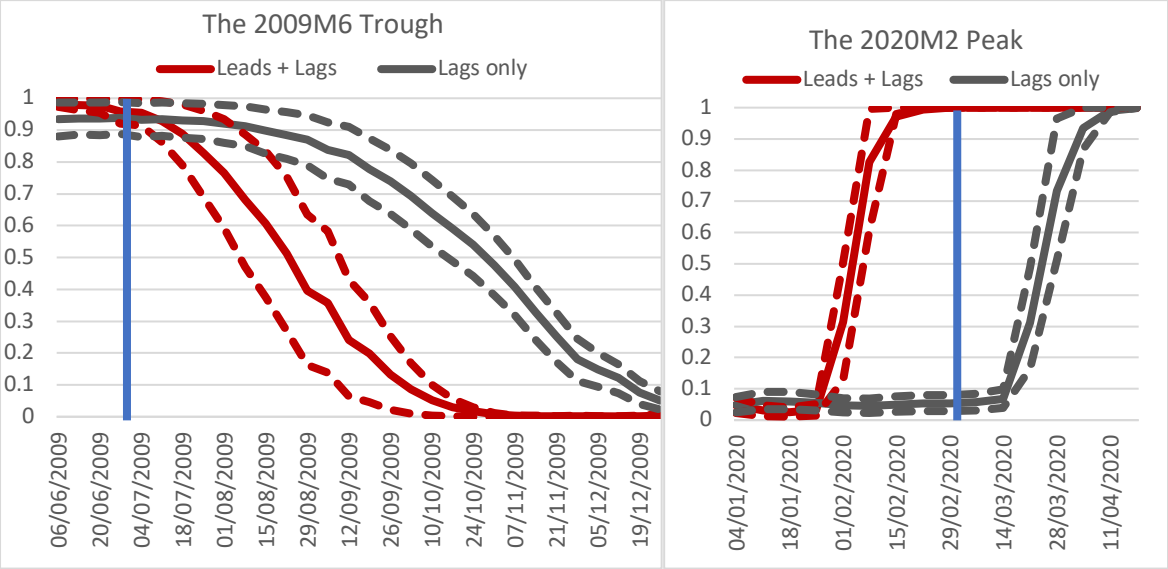
Notes: The above link two beta-weighting functions: one for the leads (+7 up to +1) and the other for contemporaneous and lags (0 up to -7). Weights are normalised to sum up to 1. Dotted lines are 68% bands. Sample period: 2008M1-2018M12. The red line indicates equal-weighting values.

Figure 6: Posterior Mean Estimates (and 68% bands) for Recession Probabilities using MIDAS-Probit model with WEI.



Notes: Dotted lines are 68% bands. Shaded areas indicate recession phases. These are predictions for the probability of a recession phase in the current month using the MIDAS-Probit specification with leads and lags. Parameters were estimated with data up to 2018M12. We use the WEI time series published in June 18th, 2020. These are computed with 7,000 draws (after initial 1,000 are removed).

Figure 7: MIDAS-Probit Probability of Recession around Turning Points



Note: These are predictions for the probability of a recession phase in the current month. Parameters were estimated with data up to 2018M12, so the first plot shows in-sample results while the second one is out-of-sample (NBER recession indicator for a peak not available). These are computed with 7,000 draws (after initial 1,000 are removed). The blue line indicates either the NBER trough (2009) or the NBER peak (2020).

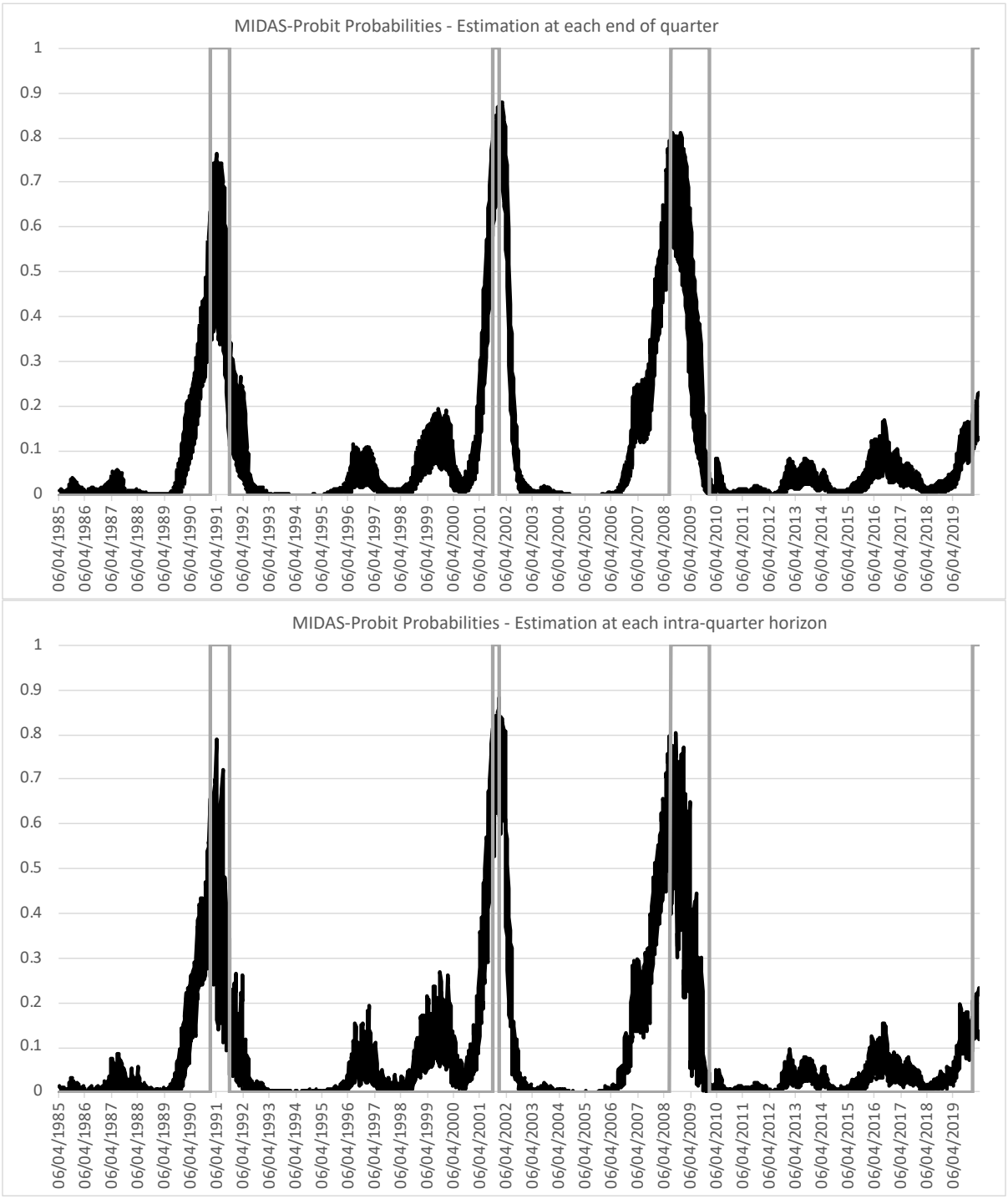
The predicted probability goes below 90% in the left plot in 18/07/09 for the leads + lags specification and in 22/08/09 for the lags only specification. The predicted probability goes above 90% in the right plot in 15/02/20 for the leads + lags specification and in 04/04/20 for the lags only specification. The leads-lag specification uses information up to 04/04/20 to identify a peak in 15/02/20.

Table 1: Full Sample and Out-of-Sample AUROC for one-year-ahead probabilities of recessions using the Spread: a comparison between Probit and MIDAS-Probit Specifications

	Full Sample AUROC	Out-of-Sample Period AUROC
Probit (z_t)	0.857	0.834
	[0.856 0.858]	[0.824 0.837]
Probit (z_t, z_{t-6})	0.881	0.859
	[0.880 0.883]	[0.841 0.861]
MIDAS-Probit	0.882	0.889
	[0.880 0.884]	[0.872 0.886]

Notes: The full sample is 1962M1-2020M4. The forecasting origins in the out-of-sample period is 1976M1-2019M4. Values in brackets are 68% bands computed using the posterior distribution of predicted probabilities. For the out-of-sample period, we first compute the posterior mean estimate of the predicted probability at each forecast origin before computing the AUC, which explains why the 68% bands are not centered. These are based on 4,000 draws (after initial 1,000 are removed).

Figure A1: Comparing the Effects of Intra-Quarter Re-estimation for MIDAS-Probit in Eq. (8) with the Spread as the monthly variable and h=1: 68% Intervals for the Probability of a Contraction



Note: Dates refer to t+1-(j/13). Re-estimation every week is employed to compute results in the bottom panel, while the parameters employed in the top panel are estimated only when full quarter information is available. These are based on 4,000 draws (after 500 are discarded).