Do We Need Intra-Daily Data to Forecast Daily Volatility?

Denisa Banulescu[*], Bertrand Candelon[†] and Christophe Hurlin [‡]

First version: March 2012

Abstract

Considering mixed data sampling (MIDAS) regressions, we analyze the influence of the sampling frequency of intra-daily predictors on the accuracy of the volatility forecasts. We propose various in-sample and out-of-sample comparisons of daily, weekly and bi-weekly volatility forecasts issued from MIDAS regressions based on intra-daily regressors sampled at different frequencies. First, we show that increasing the frequency of the regressors improves the forecasting abilities of the MIDAS model. In other words, using regressors sampled at 5 minutes gives more accurate forecasts than using regressors sampled at 10 minutes, etc. These results are robust to the choice of the loss function (MSE, Qlike, etc.) and to the choice of the forecasting horizon. Third, the MIDAS regressions with high-frequency regressors (sampled between 5 minutes and 30 minutes) provide more accurate in-sample forecasts than a GARCH model based on daily data. However, except the one-period-ahead forecasts of the calm period, the out-of-sample forecasts of MIDAS models are not significantly different from the GARCH forecasts, whatever the sampling frequency used, confirming that the direct use of high-frequency data does not necessarily improve volatility predictions.

JEL classification: C22; C53; G12

Keywords: Volatility Forecasting; MIDAS; High-Frequency Data

[*]Maastricht University and University of Orléans (LEO, UMRS CNRS 7332). Email: georgiana.banulescu@univ-orleans.fr

[†]Maastricht University, Department of Economics. Email: b.candelon@maastrichtuniversity.nl

[‡]University of Orléans (LEO, UMRS CNRS 7332). Email: christophe.hurlin@univ-orleans.fr

1 Introduction

Since the seminal ARCH/GARCH models of Engle (1982) and Bollerslev (1986), various approaches have been proposed to measure and forecast the volatility of financial asset returns. One approach consists in integrating some high frequency information relatively to the frequency of interest. In particular, various authors advocate the use of intra-daily information in order to produce daily volatility forecasts (Ding and Granger, 1996; Engle, 2000; Ghysels et al., 2006, among others).

This movement to higher frequency volatility models can be viewed as a natural consequence of the availability of higher frequency returns, but not only. The use of high frequency information is also related to the news impact curve introduced by Engle and Ng (1993). Ghysels and Chen (2010) show that intra-daily news has an impact on future daily volatility. They show that moderately "good (intra-daily) news" diminishes the next day volatility while both "very good news" (unusual high intra-daily positive returns) and "bad news" (negative returns) increase the volatility, with the latter having a more disturbing impact. Note that all their models are based on five-minute intra-daily information set. On the contrary, if we increase too much the frequency of the information set used to forecast the daily volatility, microstructure noises may lead to poor forecasts. The question is to know if there exists an "optimal" frequency for the information set used to forecast daily or lower frequency volatilities.

An obvious way to gauge the influence of the sampling frequency of intra-daily regressors on the daily, weekly or bi-weekly volatility forecasts is to use a mixed data sampling regression (MIDAS thereafter) proposed by Ghysels et al. (2006).[1] MIDAS regressions can be viewed as parsimoniously parameterized regressions that allow estimating the link function between data sampled at different frequencies.

Our aim is to propose a comparison of the forecasting abilities of various MIDAS regressions of the daily (or weekly and bi-weekly) volatility over intra-daily regressors sampled at different frequencies. In their seminal paper, Ghysels et al. (2006) consider various MIDAS regressions that differ in the specification of regressors (squared returns, absolute returns,

[1] Another way would consist in using the UHF-GARCH of Engle (2000). The salient feature of UHF-GARCH is the use of ultra-high-frequency data irregularly spaced to predict volatility. The random time of data arrival is then supposed to carry information. This model can be used to identify the timing of trades and then to quantify the impact of this timing on volatility.

realized volatility, realized power and return ranges) to determine which is the *best predictor* of volatility.[2] Here, the goal of our study is to determine the *best sampling frequency* of a *given predictor* .

Indeed, if many theoretical and empirical studies have been devoted to the influence of the sample frequency on the realized measures of volatility (realized volatility, power and bi-power variations processes etc.), to the best of our knowledge no study has tackled the choice of the predictor's sampling frequency for such measures. Let us assume that the daily realized volatility, defined as the sum of intra-daily squared returns sampled at a given frequency $m1$, is our measure of interest. When it comes to forecasting it with a MIDAS regression based on intra-daily regressors (the absolute returns in our case), does the sampling frequency of the regressors $m2$ necessarily correspond to that of the data used to compute the realized volatility? Given the weight function in the MIDAS regression, this issue is not trivial and may have strong consequences on the forecasting performance of the model. More generally it raises the question: do we need intra-daily information to improve the forecast of daily or lower frequency volatilities?

In this paper, we propose in-sample and out-of-sample comparisons of the volatility forecasts issued from various MIDAS based on intra-daily regressors sampled at different frequencies for the S&P500 over the period 29/10/2004 to 31/12/2008. To compare these forecasts, we use the robust loss functions proposed by Patton (2011) and consider the realized volatility as the proxy of the true volatility process. The use of a proxy may be a perturbing factor in assessing the forecasts precision. That is why we consider robust loss functions which deliver the same forecasts ranking as if we were using the true volatility. We complete the analysis with standard tests for predictive accuracy (Diebold and Mariano, 1995). Finally, we compute the loss function as function of the sampling frequency and compare it to the loss function of the benchmark GARCH model.

First, we show that increasing the frequency of the regressors improves the forecasting abilities of the MIDAS model. In other words, using regressors sampled at 5 minutes gives more accurate forecasts than using regressors sampled at 10 minutes, using regressors sampled at 10 minutes gives more accurate forecasts than using regressors sampled at 15 minutes, and so on. The differences are significant after a frequency threshold: when one compares the 5-minute MIDAS to the others, the differences are significant for sampling

[2]Here, following Ghysels and Forsberg (2004) we limit our analysis to the intra-daily absolute returns.

frequencies larger than 30 minutes. The minimum of the loss functions is always obtained for a frequency of the regressors equal to five minutes. Second, these results are robust to the choice of the loss function (MSE, Qlike, etc.) and to the choice of the forecasting horizon (daily, weekly or bi-weekly). Third, for a sampling frequency of the explanatory variables equal to or higher than 30 minutes, the in-sample MIDAS daily volatility forecasts are significantly more accurate than those obtained from a daily GARCH model. This difference decreases with the frequency of regressors: with sampling frequency lower than 30 minutes the loss function differences between MIDAS and GARCH are not significant. Besides, the differences are not significant for longer forecasting horizons (weekly and bi-weekly). Finally, the out-sample results are more ambiguous, especially during crisis periods. The differences in terms of predictive accuracy are rarely statistically significant, putting hence MIDAS and GARCH forecasting abilities on equal footing. These results confirm those obtained by Ghysels et al. (2006). Surprisingly, the direct use of high-frequency data does not improve the volatility predictions.

The paper is structured as follows. Section 2 describes the methodology used, the MIDAS models, the sampling frequency puzzle and the comparison criteria and tests. Section 3 analyses the influence of the sampling frequencies and multi-period volatility forecasts and reports the results of the in-sample and out-of sample analysis. Section 4 proposes an in-sample comparison based on a constant weighting scheme. Section 5 concludes.

2 Methodology

To fix notation, let daily sampling be denoted by t and daily returns (from time $t-1$ to time t) by $r_{t,t-1} = log(P_t) - log(P_{t-1})$. Let us assume that there are m equally spaced observations per trade day, which means that $1/m$ is the (intra-daily) sampling period for a given day, t. When the data are sampled at a higher frequency, say m-times in a day, we keep the same notations as Ghysels et al. (2006) and denote the return over this interval as $r_{t,t-1/m} = log(P_t) - log(P_{t-1/m})$. For instance, the case $m = 78$ corresponds to a five-minute sampling frequency, or more precisely to 78 five-minute intervals within a trading day, considered between the trading hours of 9:30 am and 4:00 pm. In this case $r_{t,t-1/78}$ corresponds to the last 5-min return of day $t-1$, $r_{t-1/78,t-2/78}$ corresponds to the return of the penultimate 5-min period of day $t-1$, etc.

Conditionally on a MIDAS model, our goal is to evaluate the influence of the intra-daily information captured by the regressors on the forecasting accuracy of the volatility over some future horizons H, namely at daily ($H = 1$), weekly ($H = 5$) and bi-weekly ($H = 10$) horizons. The choice of the measure of volatility that has to be predicted is obviously crucial. Many alternatives (realized volatility, power and bi-power variations, high/low range etc.) can be considered. We follow Andersen and Bollerslev (1998) and consider that the volatility can be approximated, with some discretization error, by the sum of intra-daily squared returns.[3] Formally, the realized volatility $RV_{t,t-H}^{(Hm)}$ is defined as follows:

$$RV_{t,t-H}^{(Hm)} = \sum_{j=1}^{Hm} [r_{t-(j-1)/m,t-j/m}]^2, \tag{1}$$

where the superscript between parentheses indicates the number of high-frequency data used to compute the variable. For $H = 1$, $RV_{t,t-1}^{(m)}$ corresponds to the daily realized volatility, for $H = 5$, $RV_{t,t-5}^{(5m)}$ corresponds to the weekly realized volatility and so on. Whatever the choice of horizon H, all the realized measures are based on intra-daily squared returns sampled at frequency m. Note that the case where no intra-daily data are used corresponds to $m = 1$ and then, RV simply corresponds to the sum of daily squared returns.

2.1 MIDAS Models and Sampling Frequency

The general specification of a MIDAS volatility model is:

$$RV_{t+H,t}^{(Hm_1)} = \mu_H + \phi_H \sum_{k=0}^{k^{\max}} b_H(k, \theta_H) X_{t-k/m_2,t-(k+1)/m_2}^{(m_2)} + \varepsilon_{Ht}, \tag{2}$$

where $X_{t-k/m_2,t-(k+1)/m_2}^{(m_2)}$ denotes an intra-daily regressor sampled at frequency m_2 and $b_H(k, \theta_H)$ is a lag coefficient parameterized by a set of parameters θ_H. Many intra-daily regressors can be considered here: intra-daily squared returns, intra-daily absolute returns, etc. Following Ghysels and Forsberg (2004) and Ghysels et al. (2006), we opt for the intra-daily absolute returns $|r_{t-k/m_2,t-(k+1)/m_2}|$.[4] In this specification, the volatility for the period t to $t + H$ (for instance daily volatility if $H = 1$) is explained by the right-hand side

[3] Andersen et al. (2001,2003) and Barndorff-Nielsen and Shephard (2002, 2004) show that realized volatility, like the daily squared return, is a conditionally unbiased estimator of the true daily conditional variance of the equity returns. Besides, it is a more efficient estimator than the daily squared return.

[4] Ghysels and Forsberg (2004) mention that there are some reasons that determine a high persistence of absolute returns, such as the immunity of jumps, the better sampling error behavior or the good predictability features of this variable.

forecasting variables that are sampled at intra-daily frequency (for instance 5 minutes if $m_2 = 78$). Note that the realized volatility is computed using m_1 high-frequency squared returns, and this frequency might be different from m_2.

Our goal is to study the influence of the sampling frequency m_2 of the regressors on the forecasting ability of the MIDAS model. In other words, the idea is to determine whether or not high frequency information improves the short-term forecasts of daily volatility (for $H = 1$), and if so, for what frequency. Let us measure the forecasting ability by a loss function (MSFE, Qlike etc.) of the form $L\left(\widehat{RV}_{t+H,t}, \sigma^2_{t+H}\right)$ where σ^2_{t+H} is a proxy of the true volatility and $\widehat{RV}_{t+H,t}$ is the forecasted realized volatility. Our analysis could be summarized by determining the sign of the derivative $\partial L\left(\widehat{RV}_{t+H,t}, \sigma^2_{t+H}\right)/\partial m_2$. This sign is not obvious. It is well known that high-frequency data suffer from microstructure artifacts, such as bid-ask bounce, jumps, autocorrelation and irregular or missing data, all of which can lead to biases in the volatility estimates. The impact of microstructure noises on realized variance has been largely studied (see McAleer and Medeiros, 2008). However, no study has been devoted to the impact of these microstructure noises in a regression model in which they affect the regressors.[5] The trade-off between increasing the information set and the potential biases could be largely different in this case.

In their seminal paper, Ghysels et al. (2006) used MIDAS regressions to examine whether future (daily, weekly or bi-weekly) volatility is well predicted by past daily squared returns, absolute daily returns, realized daily volatility, realized daily power, and daily range. Since all of the regressors are used within a framework with the same number of parameters and the same maximum number of lags, the results from MIDAS regressions are directly comparable. Hence, the MIDAS setup allows determining if one of the regressors dominates others. The logic is similar here, except that we consider the *same regressor* (*i.e.* the absolute intra-daily returns) for *various sampling frequencies*. MIDAS is clearly a suitable framework for such an analysis since the number of parameters is the same whatever the sampling period of the regressors. Therefore, as in Ghysels et al. (2006), our results from MIDAS regressions are directly comparable.

There are several other benefits of using MIDAS regressions in this perspective: (*i*) it

[5] In our MIDAS specification, these noises may affect both the dependent variable (realized variance) and the regressors. But, by keeping constant the sampling frequency $m1$ used to compute the realized variance while varying the frequency m_2, we can disentangle the both influences on volatility forecasting abilities.

is a very convenient and simple specification of the link function between some variables sampled at different frequencies, (*ii*) the weights of lagged predictors are parameterized by a flexible function that only depends on two parameters whatever the lag order chosen. These parameters are estimated from the data. Hence, the weight profile on the lagged predictors is captured by the estimated shape of the function with no additional pre-testing or lag-selection procedures.

Ghysels et al. (2004, 2007) consider many possibilities regarding the form of the weight function, and we focus on one particular specification based on Beta function:

$$b_H(k, \theta_H) = \frac{f(k/k^{\max}, \theta_1; \theta_2)}{\sum_{j=1}^{k^{\max}} f(j/k^{\max}, \theta_1; \theta_2)}, \tag{3}$$

where $\theta_H = (\theta_1, \theta_2)'$, $k^{\max}$ is the maximum lag order considered in specification (2), $f(z, a, b) = z^{a-1}(1-z)^{b-1}/\beta(a, b)$ and $\beta(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, where $\Gamma(.)$ denotes the Gamma function. This specification of the weight function has several interesting features: (*i*) the coefficients are positive and such a condition guarantees that the weights are non-negative, and consequently the volatility process itself is non-negative; (*ii*) the selection of $k^{\max}$ can be done by considering a large value and letting the weights vanish (the only cost of this action is the loss of initial data in the sample); (*iii*) the weights are normalized to add up to one in order to identify the scale parameter ϕ_H. The shape of the weight function depends on the value of θ. For instance, with $\theta_1 = 1$ and $\theta_2 > 1$ the vector of weights has a slowly decaying pattern. For further details, see Ghysels et al. (2007). As most of the previous studies, in the rest of the paper we will assume that $\theta_1 = 1$ and denote $\theta_H = \theta_2$, $f(k/k^{\max}; \theta_H) = \theta_H z(1-z)^{\theta_H - 1}$.

For each specification and sampling frequency m_2, the parameters (μ_H, ϕ_H, θ) are estimated by Nonlinear Least Squares (NLS) method (see Ghysels et al. 2004, 2006 for more details). The corresponding Matlab codes are available upon request.

2.2 Comparing the Volatility Forecasts

The influence of the sampling frequency (of the regressors) on the volatility forecasts is evaluated by the ability of the MIDAS model to produce good in-sample fit and out-of-sample forecasts of realized volatility. Most of the forecasting volatility literature has paid attention almost exclusively to the accuracy of one-period-ahead forecasts (Andersen and Bollerslev, 1998; Hansen and Lunde, 2005), whereas portfolio management, risk management (Value-at-Risk forecasts for instance), option pricing and hedging applications often emphasize the importance of weekly, bi-weekly or even monthly forecasts. Thus, we consider three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). Note that for $H > 1$, the MIDAS can be considered as a self-assessment between the direct and iterated forecasting methods (see Marcelino et al., 2006; Ghysels et al., 2009) since these forecasts use the entire historical information-set of daily or intra-daily data (as in our case), but provide a direct long-horizon volatility.

For a given frequency m_2, the MIDAS forecasts for a horizon H are simply defined by:

$$\widehat{RV}_{t+H,t}^{(Hm_1,m_2)} = \widehat{\mu}_H + \widehat{\phi}_H \sum_{k=0}^{k^{\max}} b_H(k, \widehat{\theta}_H) \; |r_{t-k/m_2,t-(k+1)/m_2}|, \tag{4}$$

where $\left(\widehat{\mu}_H, \widehat{\phi}_H, \widehat{\theta}_H\right)$ denote the NSL estimates of the parameters $(\mu_H, \phi_H, \theta_H)$ obtained conditionally on intra-daily regressors sampled at a frequency m_2.

By varying the sample frequency m_2 (for instance at 5 minutes, 10 minutes and so on) we obtain various in-sample and out-of-sample forecasts of the future volatility for a given horizon H. To assess the accuracy of these forecasts, we use a loss function that penalizes deviations from the ex-post realization of the volatility. Since the true volatility is not observable, even ex-post, the realized volatility is used as a proxy. We assume that this proxy is conditionally unbiased: such an assumption crucially depends on the sampling frequency m_1 of intra-daily squared returns used to compute the daily (for $H = 1$) realized volatility. As mentioned by Patton (2011), for liquid stocks and/or index tracking stocks, the realized volatility can been plausibly considered free from market microstructure effects even when it is computed using 5-minute returns. However, the use of a conditionally unbiased proxy does not necessarily guarantee that the loss function leads to the same outcome and the same ranking among models (or frequencies in our case) as if the true latent variable was used. Patton (2011) derives necessary and sufficient conditions on the functional form of "robust" loss function, in the sense that it preserves the ranking of various forecasts when using a noisy volatility proxy. He proposes a general form for the homogeneous robust loss functions, indexed by a single parameter b, that encompasses in particular the MSE and the Qlike loss functions.

$$L\left(\widehat{RV}_{t+H,t}^{(Hm_1,m_2)}, RV_{t+H,t}^{(Hm_1)}; b\right) \tag{5}$$
$$= \begin{cases} \frac{1}{(b+1)(b+2)}\left(\widehat{RV}_{t+H,t}^{b+2} - RV_{t+H,t}^{b+2}\right) - \frac{1}{b+1}\widehat{RV}_{t+H,t}^{b+1}\left(\widehat{RV}_{t+H,t} - RV_{t+H,t}\right), & \text{for } b \notin \{-1,-2\} \\ RV_{t+H,t} - \widehat{RV}_{t+H,t} + \widehat{RV}_{t+H,t}\ \log\left(\widehat{RV}_{t+H,t}/RV_{t+H,t}\right), & \text{for } b = -1 \\ \left(\widehat{RV}_{t+H,t}/RV_{t+H,t}\right) - \log\left(\widehat{RV}_{t+H,t}/RV_{t+H,t}\right) - 1, & \text{for } b = -2 \end{cases}$$

Note that the Mean Squared Error (MSE) loss function is obtained when $b = 0$ and the Qlike loss function is obtained when $b = -2$, up to additive and multiplicative constants. Here, we propose to compare the in-sample and out-of sample volatility forecasts obtained for different sampling frequencies m_2 on the basis of the average loss function $\overline{L}(m_1, m_2)$:

$$\overline{L}(m_1, m_2) = \frac{1}{T}\sum_{t=1}^{T} L\left(\widehat{RV}_{t+H,t}^{(Hm_1,m_2)}, RV_{t+H,t}^{(Hm_1)}; b\right), \tag{6}$$

where T denotes the sample size. In order to check the robustness of our conclusion we will consider three values for the parameter b: $b = 0$ (MSE), $b = -1$ and $b = -2$ (Qlike). Finally, in comparing these forecasts we test for the statistical significance of the loss function differences and implement the test of Diebold, Mariano (1995)(henceforth DM). Note that two alternative MIDAS models with two different sampling frequencies m_2 cannot be considered as nested models. The use of DM is then appropriate.

As a benchmark, we also consider the volatility forecasts issued from a simple GARCH model on daily data. For $H > 1$, we consider iterated volatility forecasts $\widehat{h}_{t+H,t}$ designed as follows:

$$\widehat{h}_{t+H,t} = \sum_{k=1}^{H} \widehat{h}_{t+k,t+k-1}, \tag{7}$$

where $\widehat{h}_{t+k,t+k-1}$ denotes the estimated conditional variance for the day $t+k$ conditional on the information available at time $t+k-1$ with $\widehat{h}_{t+k,t+k-1} = \widehat{\alpha}_0 + \widehat{\alpha}_1(r_{t+k-1,t+k-2} - \widehat{\mu})^2 + \widehat{\beta}\widehat{h}_{t+k-1,t+k-2}$, for $k = 1$, and $\widehat{h}_{t+k,t+k-1} = \widehat{\alpha}_0 + \widehat{\alpha}_1\widehat{h}_{t+k-1,t+k-2} + \widehat{\beta}\widehat{h}_{t+k-1,t+k-2}$, for $1 < k \leq H$.

3 Sampling Frequencies and Multi-Period Volatility Forecasts

We present in this section the empirical results obtained in-sample and out-of-sample, starting with a data description.

3.1 Data

The data set used in this paper consists of tick by tick price data for the S&P 500 index spanning the period 29/10/2004 to 31/12/2008. The data are reported from 9:30 am to 4:00 pm every trading day. From the high frequency price series we derive 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15 returns, respectively. In addition, we compute the daily measure of realized volatility and the daily return by using the five-minute frequency of the price data. In order to avoid the contribution of the volatility that comes from the overnight or holiday closures, the series of returns are computed by using only open-to-close data. We therefore focus only on trading day volatility. The data set contains hence 1603 trading days with 78/39/26/13/6/2 observations of 5-min/10-min/15-min/30-min/1h05/3h15 returns per day.

The period from 29/12/2004 to 27/12/2006 (a total of 500 observations) is exploited for the purpose of the in-sample analysis. For the out-of-sample analysis we present the results issued from two specific periods, one during a relatively calm volatility period (2007) and the other during the 2008 crisis. All forecasts are pseudo out-of-sample and a rolling sample estimates scheme is used. The parameter estimates are hence constantly updated and repeated forecasts out-of-sample are performed. We use approximately two years of observations (500 observations) to initialize the process, and the next year (250 observations) to compare the forecasts. Note that all MIDAS regressions are based on non-overlapping samples of intra-daily absolute returns and realized volatility, respectively, and run for different forecasting horizons ($H = 1$, 5, 10 days).

3.2 In-Sample Analysis

Table 1 reports in-sample regression diagnostics and the estimated parameters of the MIDAS models based on intra-daily regressors sampled at different frequencies between 5-min and 3h15. Results are reported for the three forecasting horizons ($H = 1$, 5, 10 days). In order to compare the various models, the lag order $k^{\max}$ is fixed such that the information used to estimate the weight function covers a period of 40 days, whatever the sampling frequency of regressors.[6] For instance, for a 5-min regressor we use a k_{max} equal to 78×40 lags, where 78 represents the number of five-minute intervals within a trading day.

First, all the estimated parameters $\widehat{\mu}_H$, $\widehat{\phi}_H$ and $\widehat{\theta}_H$ decrease with the sampling frequency. For instance, for $H = 1$ (top panel), the estimated parameter $\widehat{\theta}_H$ decreases from 34.555 for a 5-minutes frequency to 16.116 for a 3h15 frequency. This drop has an impact on the shape of the weight function: the smaller the value of parameter θ_H, the higher the slope of the weight function. Consequently, the use of high frequency intra-daily regressors leads to endogenously give more importance to the information available during the first past day. Column 5 of Table 1 reports the sum of the estimated weights $b_H(k, \widehat{\theta}_H)$ related to the first past day, expressed as a percentage of the sum of estimated weights for $k = 0$ to $k^{\max}$. Columns 6 and 7 report the same information for the weights placed on the second to the fifth day of lags and beyond the fifth day of lags, respectively. The weight pattern changes with the sampling frequency of the regressors. For a 5-min frequency we have 58.3 percent of the total weight placed on the first day of lags, whereas the weight related to the first day of lags decreases progressively to 53.2, 49.9, 46.1, 40.4 and 33.3 percent for the explanatory variables sampled at 10-min, 15-min, 30-min, 1h05 and 3h15, respectively. A reverse result is observed for the lags in the subsequent days. For instance, for the lags on days 2-5 the part of the associated weight is 40.7, 45, 47.5, 50.1, 53.1 and 54.9 percent for a sampling frequency of the regressors equal to 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15, respectively.

Table 1 also reports the p-values of the Portmanteau tests for serial correlation in the residuals and squared residuals, respectively. Ghysels et al. (2006) advocate that, generally, the errors of MIDAS models may feature autocorrelation or heteroskedasticity. In our case, the p-values of the Portmanteau test indicate that, with the exception of the models using regressors sampled at a frequency of 1h05 and 3h15, respectively, all residuals appear uncorrelated. Similarly, the second block of p-values shows that squared residuals are not serially correlated. Note that, in order to avoid autocorrelation in the residuals, all regressions were computed by employing non-overlapping prediction samples.

[6] Increasing the lag length beyond 40 days has no significant effect on the results.

For each MIDAS model, Figure 2 displays the profile of the estimated weight function $b_H(k, \hat{\theta})$ for a daily predicting horizon. The sum of weights is normalized to one and each subplot presents their relative share across all the lags over the last 15 days. Note that for a given frequency, each point corresponds to an estimated weight $b_H(k, \hat{\theta})$ related to a lagged regressor $|r_{t-k/m_2,t-(k+1)/m_2}|$. Thus, for a 5-minute frequency, $15 \times 78 = 1170$ weights are reported (since for each day, we consider 78 intra-daily lagged regressors sampled at 5 minutes), whereas for a 1h05 frequency, $6 \times 15 = 90$ weights are reported. Whatever the frequency used, we observe a smoothly declining weighting scheme that vanishes approximately around 12 days. Figure 2 also confirms the idea that the higher the sampling frequency of the regressors is, the more the model endogenously gives weight to the recent past information in the forecasts of the future daily volatility.

The influence of the sampling frequency on the MIDAS forecasting ability is summarized in Table 2. This table reports the t-statistics from DM test of equal predictive accuracy based on the estimates issued from six MIDAS models for $H = 1$, with regressors sampled at different frequencies and a GARCH(1,1) model. The corresponding p-values of the bilateral test are reported in brackets and the significant t-statistics for a 5% level are in bold. The last line of each panel corresponds to the average loss function of each model. The DM test is based on the loss functions defined for $b = 0$ (MSE), $b = -1$ and $b = -2$ (Qlike), respectively. For the sake of simplicity, we focus on the first block of results ($b = 0$). The sign of the t-statistics indicates which forecasting model performs better. Considering hence two models, model A (in line) and model B (in column), a positive t-statistic indicates that model B forecasts produce larger average losses than model A forecasts, while a negative sign indicates the opposite. For instance, the first value in the table (situated on the second line and first column) represents the t-statistic issued from the comparison of the predictive capacity of a MIDAS regression with the explanatory variable sampled at 5-min and a MIDAS regression with the explanatory variable sampled at 10-min. The value of this t-statistic (0.211) is positive, which means that the MIDAS predictions with the 5-min intra-daily regressor outperform the MIDAS predictions with the 10-min intra-daily regressor. However the difference here is not statistically significant for standard risk levels since the p-value is equal to 0.833.

First, the DM t-statistics issued from the comparison of each two MIDAS models are all positive and become statistically significant from a certain frequency of regressors. This indicates that the higher the frequency of explanatory variable, the better the predictive abilities of MIDAS models. These results confirm the intuition that in order to predict the daily realized volatility computed with 5-minute intra-daily returns (m_1 frequency), the best intra-daily MIDAS regressors are those sampled at 5 minutes (m_2 frequency), or at least at a frequency higher than 30 minutes (since the t-statistics are no significant in the first column). They are also robust to the choice of the loss function. For a scale parameter of the loss function equal to -1 or -2, it is worth noting that the results are very similar to those obtained for the MSE ($b = 0$). Second, for a sampling frequency of the explanatory variables higher or equal to 30 minutes, the MIDAS forecasts significantly outperform the GARCH forecasts for a 5% level. This result is also robust to the choice of the loss function.

To summarize, if the sampling frequency of the explanatory variables is higher or equal to 30 minutes, MIDAS models significantly dominate GARCH model and in the case of MIDAS models their predictive ability increases with the frequency of regressors. Figure 5 confirms these findings. The left side of the figure displays the average loss function issued from each of the six MIDAS models with different sampling frequency of regressors as well as the average loss function of the daily GARCH model. This representation is done for the three loss functions considered. On the right side of the figure we report the p-values of the DM test of equal predictive accuracy for each MIDAS and our GARCH benchmark. The shaded areas mark the cases for which the null hypothesis of the DM test is rejected. If we consider the case of the MSE loss function ($b = 0$), the loss function of MIDAS model is always inferior to the GARCH loss function, whatever the intra-daily frequency of regressors. However, this difference is statistically significant only for a sampling frequency higher or equal to 30 minutes. For a sampling frequency of 1h05 and 3h15, respectively, the average loss function of MIDAS increases and the difference between the MIDAS and GARCH predictions is no more statistically significant.

Overall, the in-sample analysis confirms that the sampling frequency matters for the prediction of daily volatility. The higher the intra-daily frequency of explanatory variables is, the more accurate a MIDAS model seems to be in terms of prediction of volatility. The three loss functions indicate that the best intra-daily forecasting variables used to predict the daily realized volatility computed with 5-minute intra-daily returns are those that are also sampled at 5 minutes, since the minimum of the loss functions is always obtained for this frequency.

3.3 Out-of-Sample Analysis

In the previous subsection, we show that increasing the frequency of the regressors improves the in-sample forecasting abilities of the MIDAS model. Besides, we observe that the in-sample MIDAS daily volatility forecasts (for a sampling frequency equal or smaller to 30 minutes) are significantly more accurate than those obtained from a daily GARCH model. The aim of this section is to check if these two results remain valid for the out-of-sample forecast.

We consider two out-of-sample forecast exercises: one in times of crisis (2008) and the other during a quieter volatility period (2007). For each period, pseudo out-of-sample forecasts are computed, by using rolling sample estimates and constantly updating the parameter estimates. Three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$), are successively considered. As for the in-sample exercise, the forecasting variables are sampled at different intra-daily frequencies ranging from five minutes to three hours. Note that MIDAS volatility models produce directly multi-period volatility forecasts, while for the GARCH model we use dynamic forecasts (cf. equation 7). The out-of-sample forecasts are compared through Patton's loss function and DM test for equal predictive ability. However, for the sake of simplicity, we present only the results for the MSE loss function ($b = 0$) computed for each forecasting horizon H. The other results are available upon request.

Table 3 reports the DM test t-statistics based on the forecasts issued from the six MIDAS models and the GARCH model for the calm period of January 2007 to December 2007. The same results, but for the crisis period of January 2008 to December 2008 are reported in Table 4. Corresponding p-values are in brackets and boldface numbers indicate that the difference of predictive ability between the model situated on the line and the model in column is significant for a 5% level. As explained in the previous subsection, the sign of the t-statistics indicates which forecasts perform better. A positive t-statistic indicates hence that the model in column outperforms the model in line in terms of forecasting ability, while a negative sign indicates the opposite. Note that the three blocks of results correspond to the three forecasting horizons (one day, one week and two weeks, respectively).

For the 2007 sample, we remark that the one-period-ahead results ($H = 1$) confirm those obtained in-sample. First, with some exceptions, the DM t-statistics issued from the comparisons of each two MIDAS models are positive and become statistically significant from a certain intra-daily frequency of regressors. Second, for a sampling frequency higher or equal to 1h05, the MIDAS models have a statistically significant advantage relative to the GARCH model in terms of predictive accuracy at a 5 % risk level. However, for longer forecasting horizons ($H = 5$ or 10), these differences are no more statistically significant. Using the same approach, we now focus on the forecasting results related to the crisis period (year 2008). Clearly, the crisis affects the forecasting performance of the models and most of t-statistics are not statistically significant which means that (ignoring the sign of the t-statistics) the MIDAS models do not significantly outperform the benchmark in terms of forecasting accuracy.

Figure 7 displays the average loss function computed from each of the six MIDAS models and the GARCH average loss function over the calm (the left panel) and the crisis (the right panel) periods. Figure 8 reports the p-values of the DM test of equal predictive accuracy issued from the comparison of each MIDAS with intra-daily regressors sampled at different frequencies of time and the GARCH model for the considered periods of calm (the left panel) and crisis (the right panel), respectively. For the calm forecasting period the average loss function of the MIDAS models is, with one exception, monotonically increasing with the sampling frequency of MIDAS regressors. Besides, it is almost all the time inferior to the average loss function of the GARCH benchmark. It confirms hence our first results that increasing the frequency of the regressors improves the in-sample forecasting abilities of the MIDAS model. However, whatever the frequency used, it seems that the out-of-sample MIDAS volatility forecasts are not all the time significantly different from those obtained by using a GARCH model, especially for the multi-period-ahead forecasts. This latter result is even more stressed in the crisis period. In this case, the average loss of the MIDAS modes is always larger than the average loss of the benchmark, but the difference is rarely significant. These results put MIDAS and GARCH forecasting abilities on equal footing.

Therefore, during crisis periods the forecasting ability of the MIDAS models is not so different (even worse) from the forecasting ability of a simple GARCH, whatever the sampling frequency of the regressors. On contrary, for relatively calm period when no important shock occurs, MIDAS models may significantly outperform the GARCH benchmark.

4 Sampling Frequency and Weight Specification

The previous analyses point out the influence of the sampling frequency of intra-daily regressors on the daily, weekly or bi-weekly volatility forecasts. However, in the MIDAS model, the weights $b_H(k, \widehat{\theta}_H)$ related to the lagged regressors, are endogenously determined by the estimated parameter $\widehat{\theta}_H$. Thus, for each frequency m_2, the estimated weight function changes. It is important to note the implication of these changes in the weight function. For a given lag order $k^{\max}$, the more important the sampling frequency of the regressors m_2, the more recent information is introduced in the model. For instance, let us consider a simple example with $k^{\max} = 10$, $\theta_1 = 1$ and $\theta_2 = 5$ for two sampling frequencies, respectively $m_2 = 78$ (5 minutes) and $m_2 = 26$ (15 minutes). Figure 1 displays the corresponding weight function. The first nine weights are positive and the tenth is negligible (<3.94e-05). For a 5-minute sampling frequency, the corresponding specification of the model (2) implies that the realized volatility is explained by the lagged intra-daily absolute returns (at five minutes) observed over the last $9 \times 5 = 45$ minutes. When the sampling frequency is fixed to fifteen minutes ($m_2 = 26$), the same model implies that the realized volatility depends on the lagged intra-daily absolute returns (at fifteen minutes) observed over the last $9 \times 15 = 135$ minutes. Therefore, the change in m_2 not only affects the frequency of the explanatory variables, but also the information set used to forecast the volatility.

In order to get the same information set, a solution would consist in adapting the lag order $k^{\max}$ for each sampling frequency. For instance, by fixing $k^{\max} = 3$ for fifteen minutes ($m_2 = 26$), we would get the same information set as for the five-minute frequency. However, even if we adjust the maximum lag order, the estimated weights for the same lagged period (say a day) may be different given the frequency used. Thus, in order to control for this effect, we now compare the MIDAS models while imposing the same weights as those obtained for the 5-minute frequency. First, we estimate the weights for the 5-minute frequency. Second, for all the other frequencies, the weights are not estimated but computed by summation of the weights obtained for the 5-minute frequency. For instance, for the 15-minute frequency, the weight related to the lag $k = 1$ (that corresponds to the absolute returns observed for the last 15 minutes) is defined as the sum of three first lags ($k = 1, 2, 3$) estimated for the 5-minute frequency. Consequently, in this case the differences of forecasting abilities between the various MIDAS may only come from the differences of information included in the intra-daily regressors. The pattern of the weight function for all considered frequencies of the regressors is displayed on Figure 3.

The left side of Figure 6 displays the average loss function issued from each of the six MIDAS models with different sampling frequency of regressors as well as the average loss function of the daily GARCH model. This representation is done for the three loss functions considered ($b = 0$, -1 and -2). The right side of the figure displays the p-values of the DM test of equal predictive accuracy for MIDAS and our GARCH benchmark. The shaded areas mark the cases for which the null hypothesis of the DM test is rejected. The results are quite close similar to those obtained for the previous in-sample analysis with estimated weight functions (cf. section 3.2). In particular, the loss function is increasing with the sampling frequency. So, it confirms that increasing the frequency of the regressors improves the forecasting abilities of the MIDAS model.

5 Conclusion

Is there any way to improve the volatility predictions by varying the frequency of the forecasting variables? Do we need intra-daily data to forecast daily volatility? And if so, which is the optimal intra-daily data sampling frequency of the forecasting variables?

In this paper we try to give some pertinent answers to these questions. Therefore, we study the impact of the past absolute intra-daily returns on the next period realized volatility computed with 5-min intra-daily information and measured at horizons ranging from one day to two weeks. We consider various MIDAS models with intra-daily forecasting variables sampled at different frequencies ranging from five minutes to three hours. The corresponding in-sample and out-of-sample forecasts are then compared through a robust loss function (Patton, 2011) and standard forecast comparison tests (Diebold and Mariano, 1995).

Our conclusions can be summarized in four points. First, we find that the higher the sampling frequency of the regressors, the better the MIDAS model in terms of predicting ability. These results are robust to the choice of the loss function and the forecasting horizon and hold both in- and out-of-sample. Second, for a sampling frequency of the explanatory variables equal or higher than 30 minutes, the MIDAS in-sample daily volatility forecasts are significantly more accurate than those obtained with a simple GARCH model based on daily returns. For the MIDAS, the minimum of the loss functions are always obtained for a frequency of the regressors equal to five minutes. Third, for multi-period-ahead forecasts, even if the pattern of the loss function is the same, the differences between each MIDAS model and its benchmark in terms of predictive accuracy are not statistically significant anymore. Fourth, the out-of-sample results are more ambiguous since the losses related to the MIDAS forecasts are always more important than those related to the GARCH forecasts, especially during crisis periods. However, the difference is rarely significant, putting hence MIDAS and GARCH forecasting abilities on equal footing.

Our findings confirm those of Engle and Ng (1993) and Ghysels and Chen (2010). During a destabilizing period, using high-frequency data agitate even more the state of volatility. As a consequence, the forecasting ability of our volatility models derail. On the contrary, when no important shock occurs, we find that MIDAS models could outperform the GARCH benchmark.

Table 1: Regression diagnostics and estimated parameters of MIDAS models with intra-daily regressors

Horizon	Frequency	μ_H	ϕ_H	θ_H	Day 1	Days 2-5	Days 6-15	> 15 Days	Q(10)	$Q^2(10)$
1 day	5-min	-1.64E-05	0.107	34.555	0.583	0.407	0.010	0.000	0.055	0.983
	10-min	-1.59E-05	0.076	30.012	0.532	0.450	0.018	0.000	0.066	0.987
	15-min	-1.64E-05	0.063	27.326	0.499	0.475	0.026	0.000	0.078	0.986
	30-min	-1.25E-05	0.041	24.445	0.461	0.501	0.038	0.000	0.056	0.956
	1h05	-8.64E-06	0.025	20.464	0.404	0.531	0.065	0.000	<0.0001	0.967
	3h15	6.12E-07	0.011	16.116	0.333	0.549	0.117	0.001	<0.0001	0.861
1 week	5-min	-2.65E-05	0.404	37.594	0.614	0.379	0.007	0.000	0.153	0.698
	10-min	-1.37E-05	0.270	36.549	0.603	0.389	0.008	0.000	0.146	0.630
	15-min	-9.56E-06	0.218	38.146	0.619	0.375	0.006	0.000	0.088	0.598
	30-min	-1.36E-05	0.158	27.691	0.504	0.472	0.025	0.000	0.016	0.193
	1h05	-1.96E-05	0.110	21.165	0.414	0.526	0.057	0.002	0.034	0.134
	3h15	3.05E-05	0.044	11.848	0.258	0.535	0.174	0.034	0.002	0.295
2 weeks	5-min	-2.38E-07	0.712	27.640	0.503	0.472	0.025	0.000	0.970	0.755
	10-min	3.40E-06	0.507	23.358	0.446	0.509	0.043	0.001	0.985	0.636
	15-min	1.06E-05	0.411	28.885	0.519	0.460	0.021	0.000	0.973	0.777
	30-min	1.23E-05	0.288	23.104	0.443	0.512	0.045	0.001	0.986	0.911
	1h05	-9.39E-05	0.257	6.836	0.159	0.439	0.262	0.141	0.890	0.459
	3h15	8.98E-05	0.081	5.174	0.122	0.375	0.275	0.228	0.286	0.219

This Table presents the estimated parameters of the MIDAS regressions with the explanatory variable (e.g. intra-daily absolute returns) sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15, respectively. A summary of the weight function is also displayed; thus column "Day 1" reports how much weight is placed on the first day of intra-daily lags of the predictor, column "Days 2-5" presents how much weight is placed on the second to the fifth day of intra-daily lags and so on. The columns "$Q(10)$" and "$Q^2(10)$" report the p-values of Portmanteau tests of serial correlation in the residuals and squared residuals, respectively. We use a sample of 500 observations to run the MIDAS regressions. The three panels correspond to the three prediction horizons considered ($H = 1$ day, $H = 5$ days, $H = 10$ days).

Table 2: Comparison of the predictive ability of the MIDAS models based on Patton loss functions (in sample analysis/ $H = 1$)

$b = 0$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
(MSE)	5-min							
	10-min	0.211						
		(0.833)						
	15-min	1.320	1.667					
		(0.187)	(0.096)					
	30-min	1.525	1.228	0.475				
		(0.127)	(0.219)	(0.634)				
	1h05	**2.595**	**2.552**	1.951	1.529			
		(0.009)	**(0.011)**	(0.051)	(0.126)			
	3h15	**2.789**	**2.673**	**2.218**	**2.567**	0.919		
		(0.005)	**(0.008)**	**(0.027)**	**(0.010)**	(0.358)		
	GARCH	**3.111**	**3.082**	**2.717**	**2.587**	1.430	0.952	
		(0.002)	**(0.002)**	**(0.007)**	**(0.010)**	(0.153)	(0.341)	
	Average loss function	2.05E-10	2.06E-10	2.10E-10	2.13E-10	2.27E-10	2.35E-10	2.47E-10
$b = -1$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	-0.140						
		(0.889)						
	15-min	1.204	2.121					
		(0.229)	(0.034)					
	30-min	1.517	1.377	0.380				
		(0.129)	(0.168)	(0.704)				
	1h05	**3.332**	**3.312**	**2.522**	**2.474**			
		(0.001)	**(0.001)**	**(0.012)**	**(0.013)**			
	3h15	**3.510**	**3.452**	**2.844**	**3.337**	1.159		
		(0.000)	**(0.001)**	**(0.004)**	**(0.001)**	(0.247)		
	GARCH	**3.290**	**3.473**	**3.013**	**2.778**	1.086	0.425	
		(0.001)	**(0.001)**	**(0.003)**	**(0.005)**	(0.277)	(0.671)	
	Average loss function	4.32E-06	4.30E-06	4.45E-06	4.50E-06	4.98E-06	5.21E-06	5.34E-06
$b = -2$	Frequency	5-min .	10-min	15-min	30-min	1h05	3h15	GARCH
(Qlike)	5-min							
	10-min	-0.830						
		(0.406)						
	15-min	0.548	2.112					
		(0.584)	(0.035)					
	30-min	0.576	1.282	0.011				
		(0.565)	(0.200)	(0.991)				
	1h05	**3.008**	**3.585**	**2.644**	**3.351**			
		(0.003)	**(0.000)**	**(0.008)**	**(0.001)**			
	3h15	**3.457**	**3.891**	**2.965**	**3.718**	1.026		
		(0.001)	**(0.000)**	**(0.003)**	**(0.000)**	(0.305)		
	GARCH	**2.512**	**3.282**	**2.480**	**2.357**	0.313	-0.346	
		(0.012)	**(0.001)**	**(0.013)**	**(0.018)**	(0.754)	(0.729)	
	Average loss function	0.124	0.121	0.126	0.126	0.144	0.150	0.147

This Table presents the t-statistics from DM test of equal predictive accuracy based on the estimates issued from a MIDAS with intra-daily regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1), for the SP500 index over the period January 2005 to December 2007. The p-values of the bilateral test are reported in brackets. The sign of the t-statistics indicates which forecasts perform better. For instance, the first value of the t-statistic (which is a positive one) indicates that the MIDAS predictions with 10-min intra-daily regressor (model in line) underperform the MIDAS predictions with 5-min intra-daily regressor (model in column); however, the result is not statistically significant. Note that the horizon of estimation is one day, and the three blocks of results correspond to three different scalar parameters of the loss function, namely $b = 0$ (MSE), -1, -2 (Qlike).

Table 3: Comparison of the forecasting ability of the MIDAS models based on Patton loss function (out of sample analysis/ calm period/ $b = 0$)

$H = 1$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	0.797						
		(0.426)						
	15-min	1.031	0.178					
		(0.302)	(0.859)					
	30-min	-0.929	-1.048	-1.289				
		(0.353)	(0.295)	(0.197)				
	1h05	1.223	0.769	0.736	**2.113**			
		(0.221)	(0.442)	(0.462)	**(0.035)**			
	3h15	**3.184**	**3.158**	**2.994**	**2.966**	**2.350**		
		(0.001)	**(0.002)**	**(0.003)**	**(0.003)**	**(0.019)**		
	GARCH	**2.758**	**2.685**	**2.636**	**2.844**	**2.288**	1.161	
		(0.006)	**(0.007)**	**(0.008)**	**(0.004)**	**(0.022)**	(0.246)	
	Average loss function	1.54E-09	1.56E-09	1.56E-09	1.48E-09	1.62E-09	1.97E-09	2.10E-09
$H = 5$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	1.369						
		(0.171)						
	15-min	1.674	0.515					
		(0.094)	(0.606)					
	30-min	-0.709	-1.495	**-2.242**				
		(0.478)	(0.135)	**(0.025)**				
	1h05	**1.965**	0.986	0.545	**2.328**			
		(0.049)	(0.324)	(0.585)	**(0.020)**			
	3h15	**2.150**	1.721	1.882	**2.344**	1.685		
		(0.032)	(0.085)	(0.060)	**(0.019)**	(0.092)		
	GARCH	1.744	1.057	1.036	1.919	0.755	-0.620	
		(0.081)	(0.291)	(0.300)	(0.055)	(0.450)	(0.535)	
	Average loss function	2.59E-08	2.85E-08	2.96E-08	2.51E-08	3.07E-08	3.61E-08	3.40E-08
$H = 10$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	1.655						
		(0.098)						
	15-min	**2.242**	-0.157					
		(0.025)	(0.875)					
	30-min	1.547	-0.474	-0.694				
		(0.122)	(0.636)	(0.488)				
	1h05	**2.373**	1.625	1.706	1.729			
		(0.018)	(0.104)	(0.088)	(0.084)			
	3h15	**2.231**	1.559	1.790	1.837	0.990		
		(0.026)	(0.119)	(0.073)	(0.066)	(0.322)		
	GARCH	1.529	0.720	0.894	1.090	-0.051	-1.127	
		(0.126)	(0.471)	(0.371)	(0.276)	(0.959)	(0.260)	
	Average loss function	1.00E-07	1.11E-07	1.10E-07	1.08E-07	1.26E-07	1.40E-07	1.25E-07

This Table presents the t-statistics from DM test of equal predictive accuracy based on the forecasts issued from a MIDAS with intra-daily regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1), for the SP500 index over the period January 2007 to December 2007 (calm period). The p-values of the bilateral test are reported in brackets. The sign of the t-statistics indicates which forecasts perform better. For instance, the first value of the t-statistic (which is a positive one) indicates that the MIDAS predictions with 10-min intra-daily regressor (model in line) underperform the MIDAS predictions with 5-min intra-daily regressor (model in column); however, the result is not statistically significant. Note that the scalar parameter of the loss function is fixed to 0 (MSE) and and the three blocks of results correspond to three forecasting horizons (one day, one week and two weeks, respectively).

Table 4: Comparison of the forecasting ability of the MIDAS models based on Patton loss function (out of sample analysis/ crisis period/ $b = 0$)

$H = 1$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	1.240						
		(0.215)						
	15-min	-0.571	-1.760					
		(0.568)	(0.078)					
	30-min	0.775	-0.054	0.994				
		(0.438)	(0.957)	(0.320)				
	1h05	0.832	0.384	1.006	0.554			
		(0.405)	(0.701)	(0.314)	(0.580)			
	3h15	0.732	0.187	0.954	0.289	-0.258		
		(0.464)	(0.852)	(0.340)	(0.772)	(0.796)		
	GARCH	-0.193	-0.458	-0.107	-0.515	-0.609	-0.642	
		(0.847)	(0.647)	(0.915)	(0.607)	(0.543)	(0.521)	
	Average loss function	1.20E-07	1.25E-07	1.18E-07	1.25E-07	1.28E-07	1.27E-07	1.16E-07
$H = 5$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	0.482						
		(0.630)						
	15-min	-1.571	**-2.193**					
		(0.116)	**(0.028)**					
	30-min	0.918	0.598	**2.455**				
		(0.359)	(0.550)	**(0.014)**				
	1h05	-0.114	-0.553	1.652	-1.058			
		(0.909)	(0.580)	(0.099)	(0.290)			
	3h15	0.116	-0.231	**1.992**	-0.784	0.379		
		(0.908)	(0.818)	**(0.046)**	(0.433)	(0.705)		
	GARCH	-0.549	-0.697	-0.056	-0.924	-0.611	-0.873	
		(0.583)	(0.486)	(0.956)	(0.356)	(0.541)	(0.383)	
	Average loss function	2.05E-06	2.12E-06	1.85E-06	2.20E-06	2.03E-06	2.07E-06	1.83E-06
$H = 10$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	1.240						
		(0.215)						
	15-min	-0.571	-1.760					
		(0.568)	(0.078)					
	30-min	0.775	-0.054	0.994				
		(0.438)	(0.957)	(0.320)				
	1h05	0.832	0.384	1.006	0.554			
		(0.405)	(0.701)	(0.314)	(0.580)			
	3h15	0.732	0.187	0.954	0.289	-0.258		
		(0.464)	(0.852)	(0.340)	(0.772)	(0.796)		
	GARCH	-0.193	-0.458	-0.107	-0.515	-0.609	-0.642	
		(0.847)	(0.647)	(0.915)	(0.607)	(0.543)	(0.521)	
	Average loss function	1.20E-07	1.25E-07	1.18E-07	1.25E-07	1.28E-07	1.27E-07	1.16E-07

This Table presents the t-statistics from DM test of equal predictive accuracy based on the forecasts issued from a MIDAS with intra-daily regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1), for the SP500 index over the period January 2008 to December 2008 (crisis period). The p-values of the bilateral test are reported in brackets. The sign of the t-statistics indicates which forecasts perform better. For instance, the first value of the t-statistic (which is a positive one) indicates that the MIDAS predictions with 10-min intra-daily regressor (model in line) underperform the MIDAS predictions with 5-min intra-daily regressor (model in column); however, the result is not statistically significant. Note that the scalar parameter of the loss function is fixed to 0 (MSE) and and the three blocks of results correspond to three forecasting horizons (one day, one week and two weeks).

Table 5: Comparison of the forecasting ability of the MIDAS models based on Patton loss function (in sample analysis/ same weights/ $H = 1$)

	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
$b = 0$ (MSE)	5-min							
	10-min	0.254						
		(0.800)						
	15-min	1.463	1.702					
		(0.144)	(0.089)					
	30-min	1.522	1.272	0.529				
		(0.128)	(0.204)	(0.597)				
	1h05	**2.264**	**2.293**	**1.786**	1.427			
		(0.024)	**(0.022)**	**(0.074)**	(0.154)			
	3h15	**2.951**	**2.949**	**2.528**	**2.871**	1.374		
		(0.003)	**(0.003)**	**(0.011)**	**(0.004)**	(0.169)		
	GARCH	**3.111**	**3.093**	**2.717**	**2.473**	1.257	0.524	
		(0.002)	**(0.002)**	**(0.007)**	**(0.013)**	(0.209)	(0.600)	
	Average loss function	2.05E-10	2.06E-10	2.11E-10	2.14E-10	2.28E-10	2.40E-10	2.47E-10
$b = -1$	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	-0.191						
		(0.849)						
	15-min	1.103	**2.016**					
		(0.270)	**(0.044)**					
	30-min	1.398	1.400	0.482				
		(0.162)	(0.161)	(0.630)				
	1h05	**3.020**	**3.205**	**2.565**	**2.500**			
		(0.003)	**(0.001)**	**(0.010)**	**(0.012)**			
	3h15	**3.468**	**3.674**	**3.171**	**3.540**	1.455		
		(0.001)	**(0.000)**	**(0.002)**	**(0.000)**	(0.146)		
	GARCH	**3.290**	**3.503**	**3.059**	**2.687**	0.872	0.004	
		(0.001)	**(0.000)**	**(0.002)**	**(0.007)**	(0.383)	(0.997)	
	Average loss function	4.32E-06	4.30E-06	4.44E-06	4.52E-06	5.03E-06	5.34E-06	5.34E-06
$b = -2$ (Qlike)	Frequency	5-min	10-min	15-min	30-min	1h05	3h15	GARCH
	5-min							
	10-min	-0.872						
		(0.383)						
	15-min	0.321	1.943					
		(0.748)	(0.052)					
	30-min	0.426	1.258	0.131				
		(0.670)	(0.209)	(0.896)				
	1h05	**2.836**	**3.706**	**2.862**	**3.519**			
		(0.005)	**(0.000)**	**(0.004)**	**(0.000)**			
	3h15	**3.086**	**3.956**	**3.124**	**3.618**	0.974		
		(0.002)	**(0.000)**	**(0.002)**	**(0.000)**	(0.330)		
	GARCH	**2.512**	**3.371**	**2.636**	**2.428**	0.160	-0.565	
		(0.012)	**(0.001)**	**(0.008)**	**(0.015)**	(0.873)	(0.572)	
	Average loss function	0.124	0.120	0.126	0.126	0.146	0.152	0.147

This Table presents the t-statistics from DM test of equal predictive accuracy based on the forecasts issued from a MIDAS with intra-daily regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15, respectively) and a GARCH(1,1), for the SP500 index over the period January 2005 to December 2006. The p-values of the bilateral test are reported in brackets. The sign of the t-statistics indicates which forecasts perform better. For instance, the first value of the t-statistic (which is a positive one) indicates that the MIDAS predictions with 10-min intra-daily regressor (model in line) underperform the MIDAS predictions with 5-min intra-daily regressor (model in column), and the result is not statistically significant. Note that the scalar parameter of the loss function varies from $b = 0$ (MSE) to $b = -2$ (Qlike) and for all MIDAS regressions we impose the same weights that those obtained for the 5-minute frequency. Therefore, for the regressors sampled at a frequency higher than 5-min, the weights are computed by cumulating the weight function issued from the 5-min MIDAS regression. For instance, in a 30-min MIDAS regression, the weight for the first 30-min absolute return is the sum of the first 6 observations of the weight function issued from a 5-min MIDAS regression, and so on.

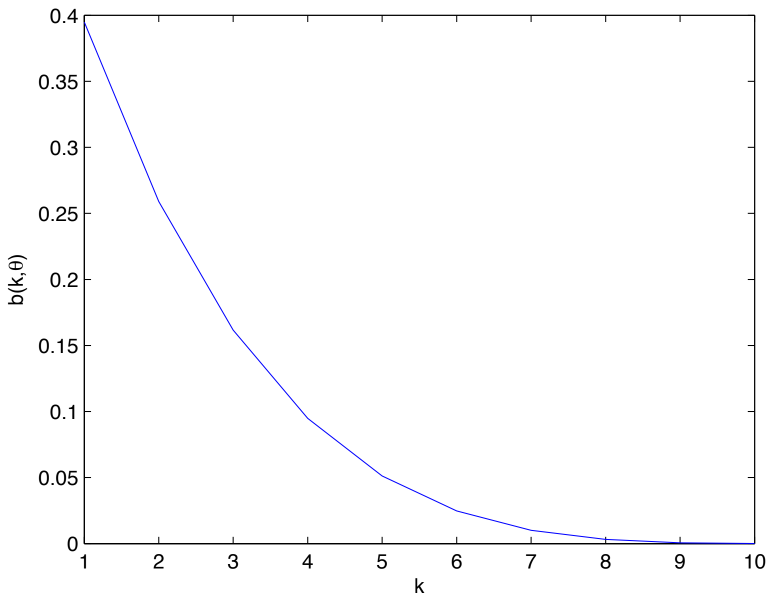

Figure 1: Simulated weight function

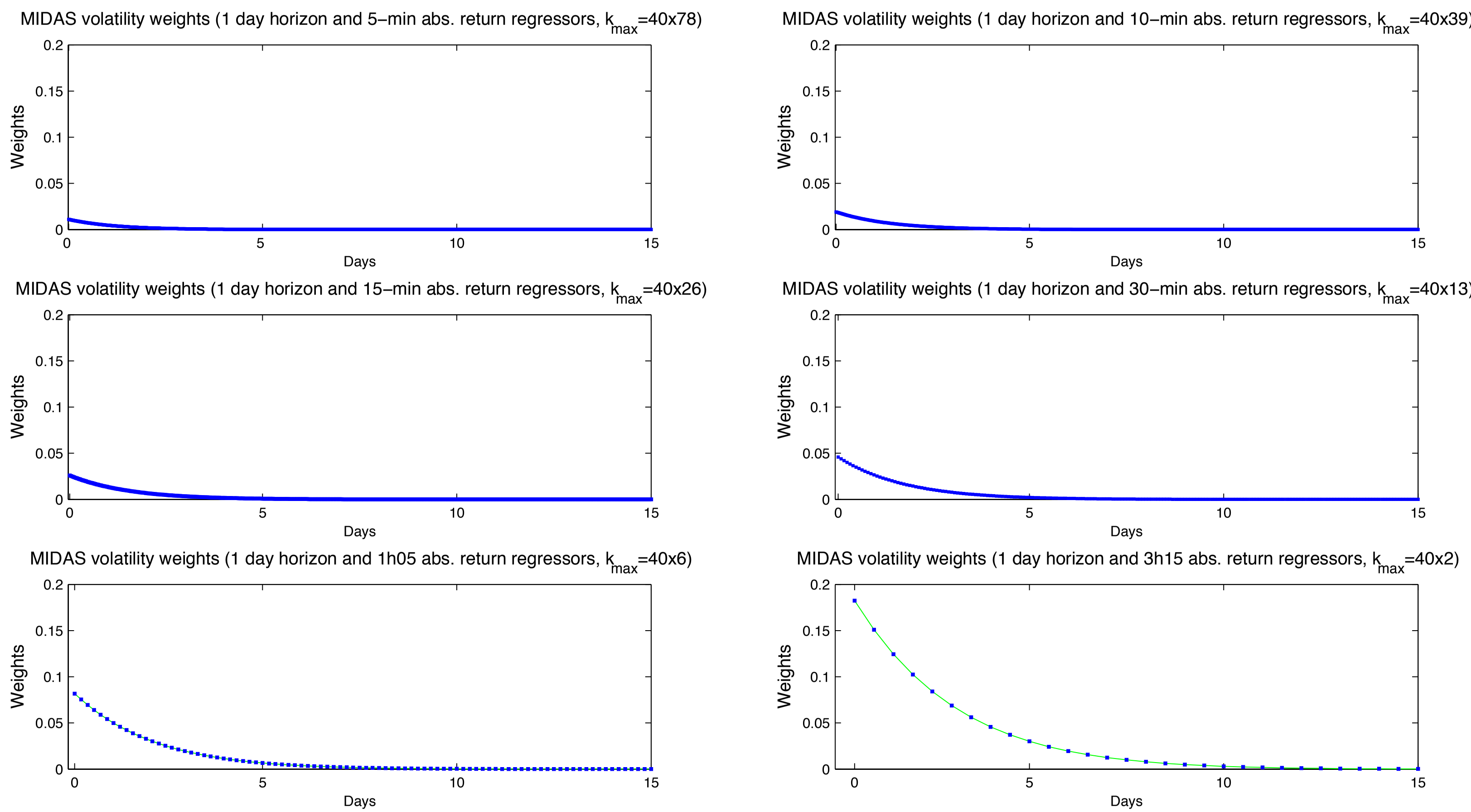

Figure 2: This Figure displays the pattern of the weight function for different intra-daily frequencies of the regressors in a MIDAS regression with 40 days of lags.We display the 15 first days of lags, since the weights are negligible for the next ones. Note that, for a 5-min absolute returns, 15 daily lags means effectively 40*78 of 5-min intra-daily lags observations. The same logic is used for the other intra-daily absolute return regressors sampled at 10-min, 15-min, 30-min, 1h05 and 3h15, respectively.

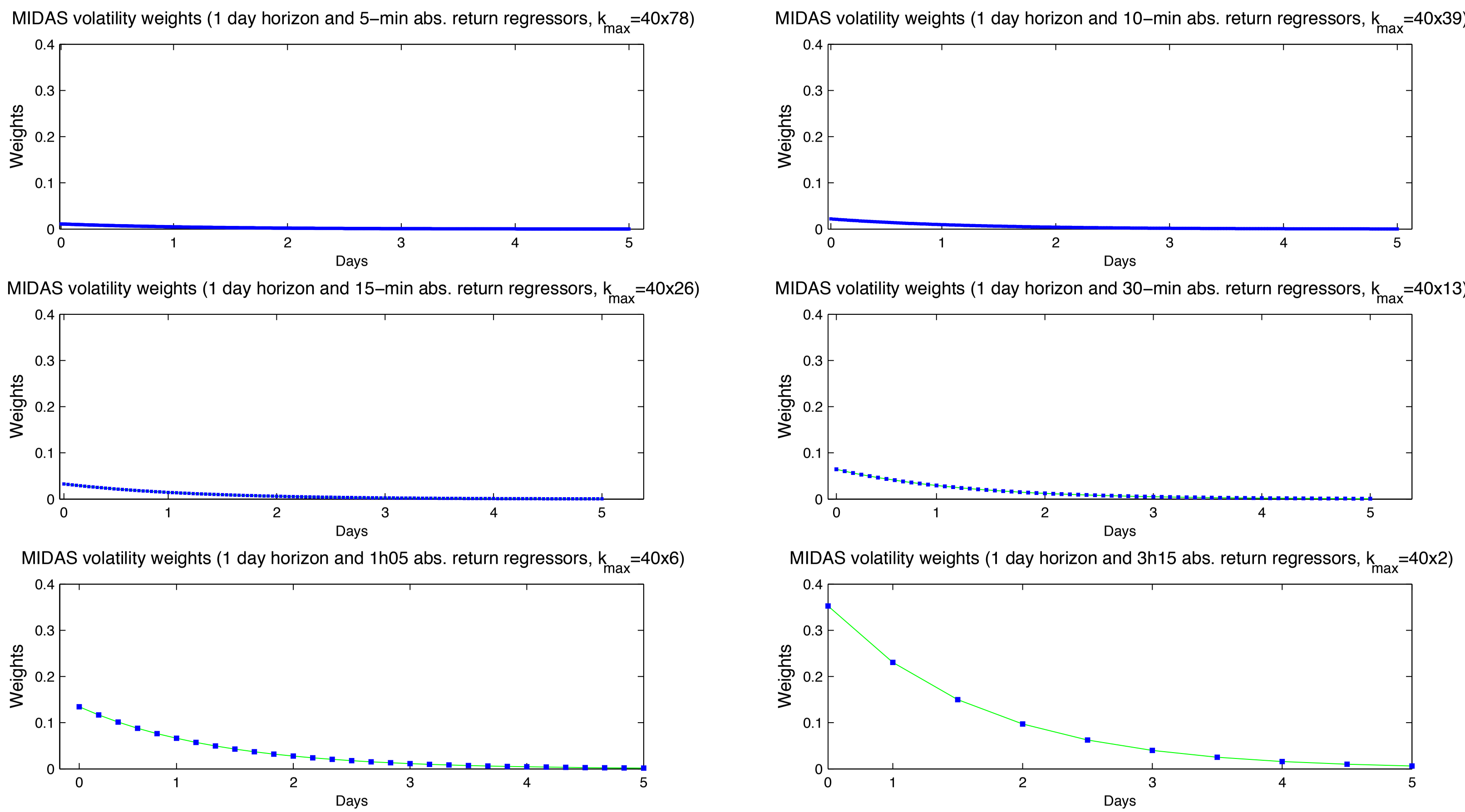

Figure 3: This Figure displays the pattern of the weight function for different intra-daily frequencies of the regressors in a MIDAS regression with 40 days of lags.We display the 5 first days of lags, since the weights are negligible for the next ones. Note that, for a 5-min absolute returns, 5 daily lags means effectively 5*78 of 5-min intra-daily lags. The same logic is used for the other intra-daily absolute return regressors sampled at 10-min, 15-min, 30-min, 1h05 and 3h15, respectively. Note that this Figure corresponds to the case when we impose the same weights that those obtained for the 5-minute frequency for all MIDAS regressions.

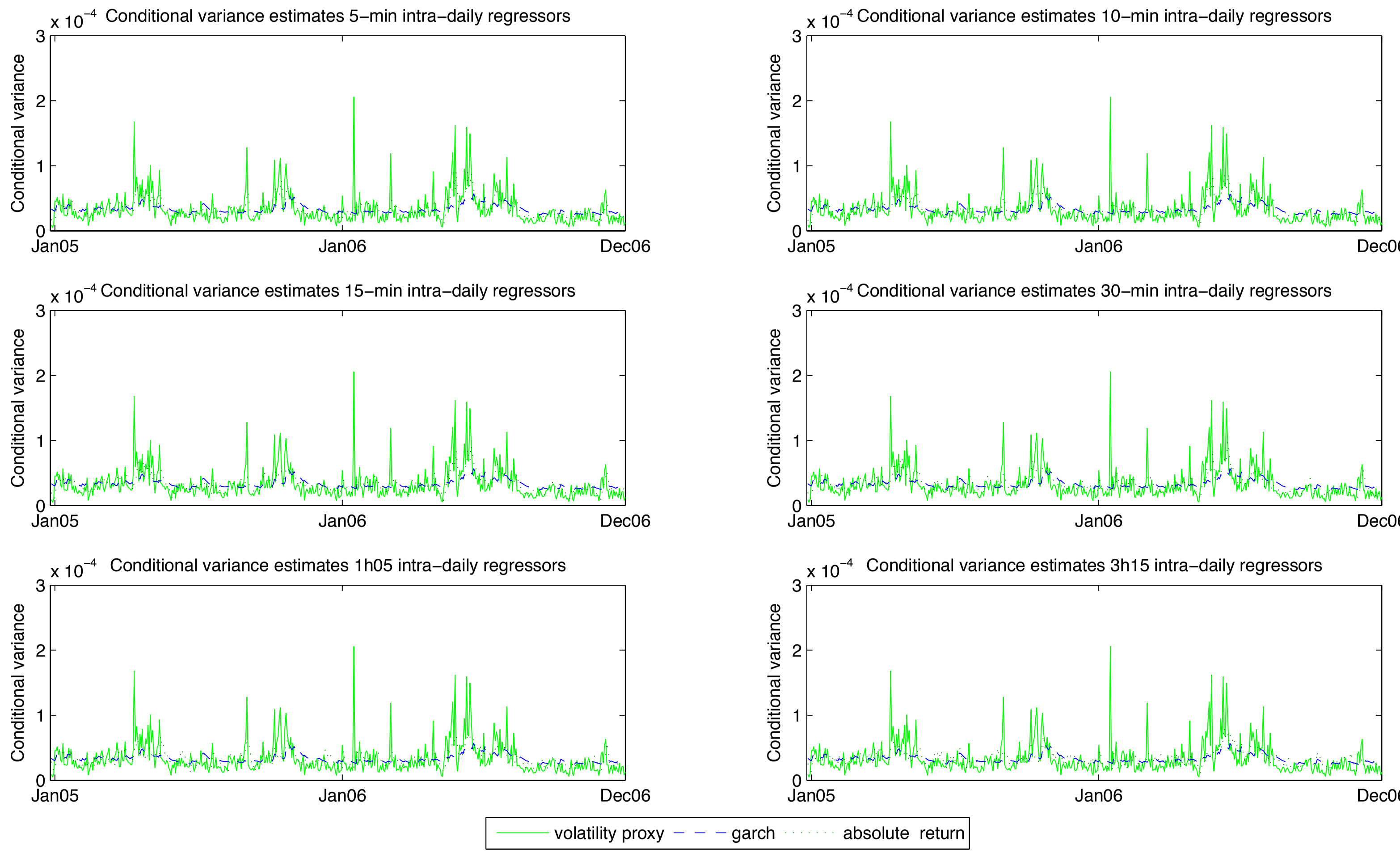

Figure 4: This Figure displays the daily estimated volatility issued from a MIDAS model using different intra-daily frequencies of the regressors (absolute returns) and the daily conditional variance estimates issued from GARCH. The volatility proxy used is the realized volatility.

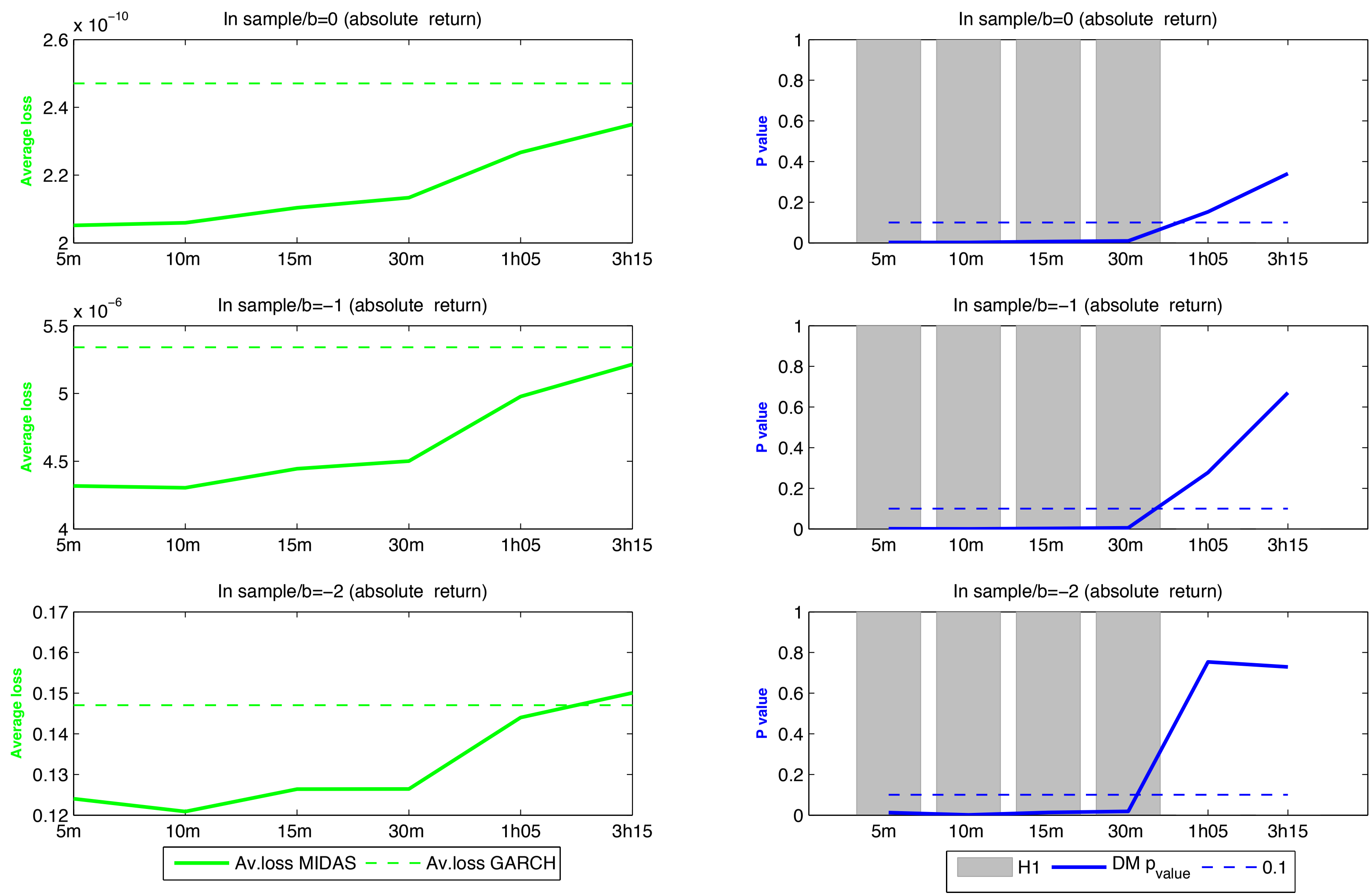

Figure 5: The left side of this Figure displays the average loss function of each MIDAS model and the GARCH average loss function (considered as benchmark). On the right side we present the associated p-values of the DM test of equal predictive accuracy issued from the statistical comparison of a MIDAS with intra-daily absolute returns as regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1). Different scale parameters of the loss function are used. The shaded areas mark the situations for which the null hypothesis of the DM test is rejected.

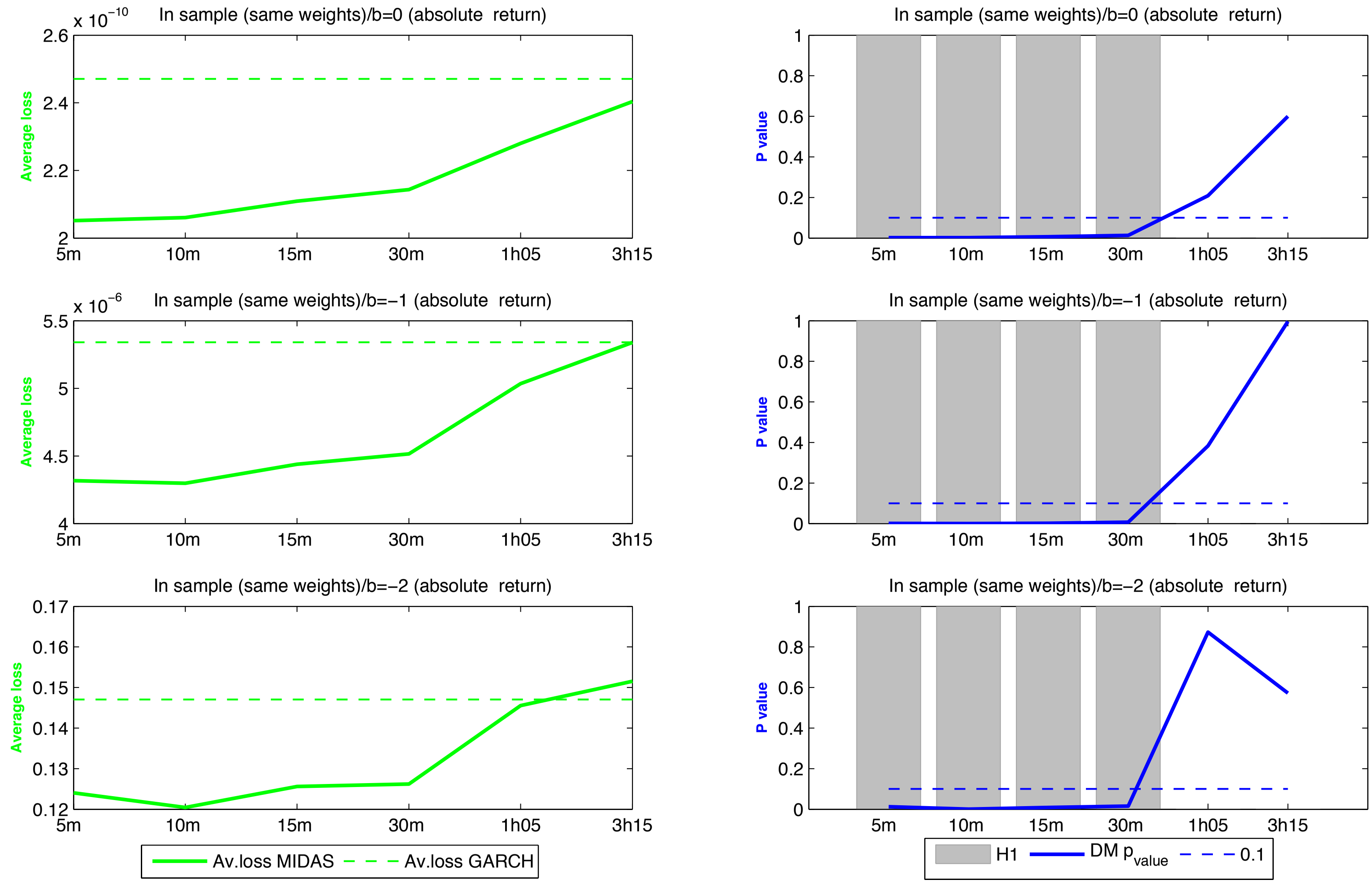

Figure 6: The left side of this Figure displays the average loss function of each MIDAS model and the GARCH average loss function (considered as benchmark). On the right side we present the associated p-values of the DM test of equal predictive accuracy issued from the comparison of a MIDAS with intra-daily absolute returns as regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1). Different scale parameters of the loss function are used. The shaded areas mark the situations for which the null hypothesis of the DM test is rejected. Note that this Figure corresponds to the case when the weight functions have the same profile over a given period of time, by imposing the same weights that those obtained for the 5-minute frequency for all MIDAS regressions.

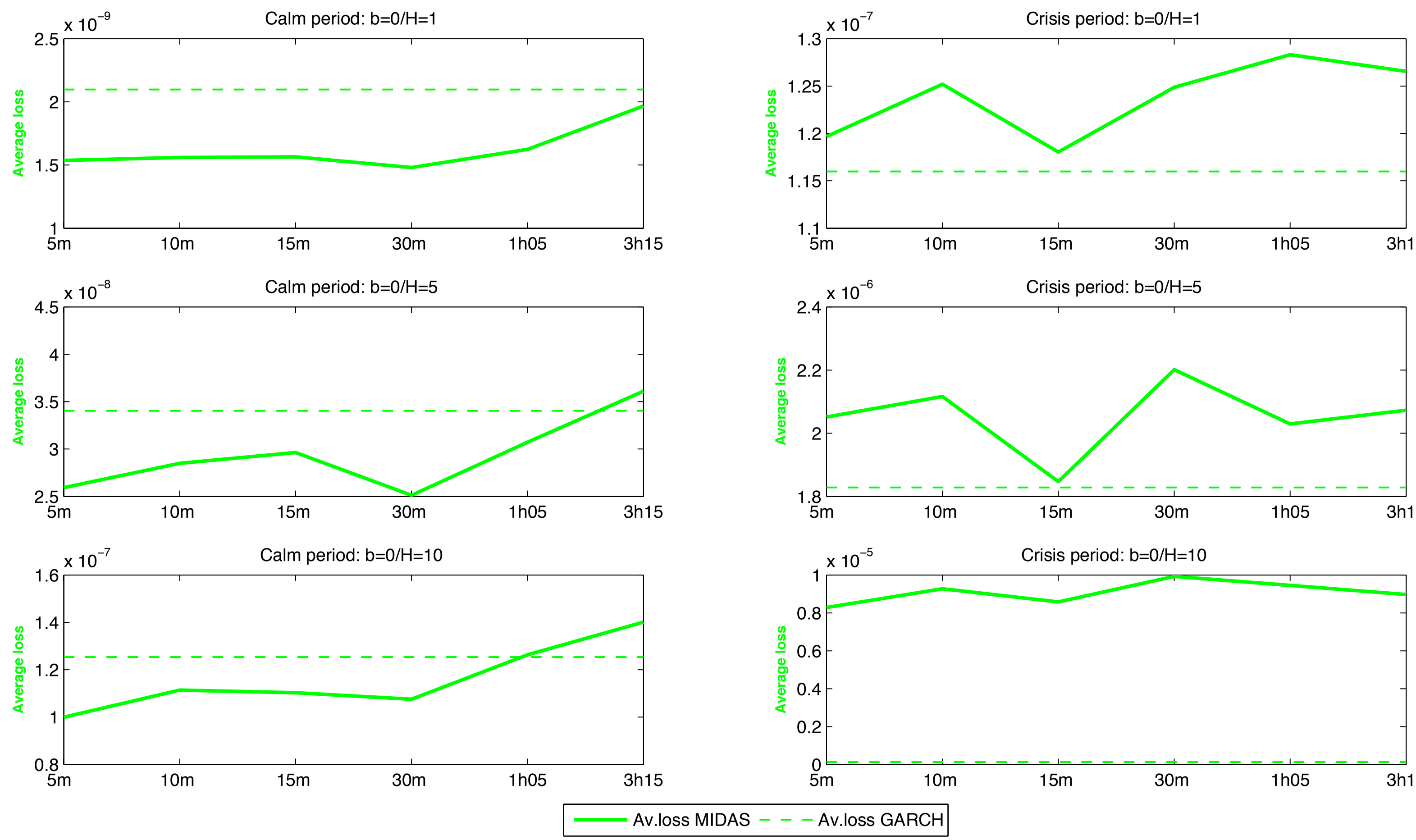

Figure 7: This Figure displays the average loss function of each MIDAS model and the GARCH average loss function (considered as benchmark) over the calm (2007) and the crisis (2008) forecasting periods. Three forecasting horizons are considered ($H = 1$, $H = 5$, $H = 10$) and the scale parameter of the loss function are fixed at 0 (MSE).

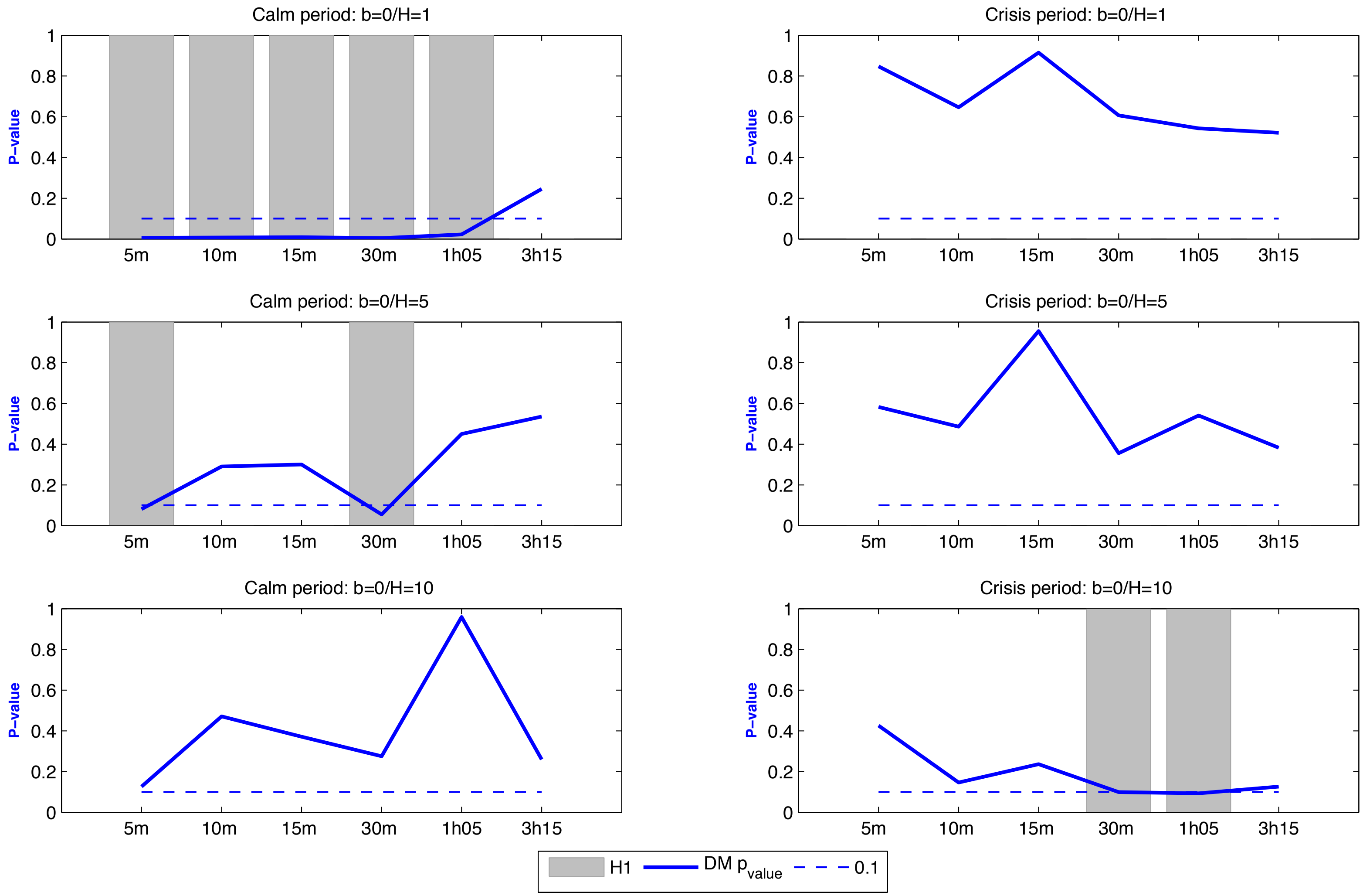

Figure 8: This Figure displays the p-values of the DM test of equal predictive accuracy for the forecasts issued from the comparison of each MIDAS with intra-daily regressors (sampled at 5-min, 10-min, 15-min, 30-min, 1h05 and 3h15) and a GARCH(1,1) over the calm (2007) and the crisis (2008) periods. Three forecasting horizons are considered ($H = 1$, $H = 5$, $H = 10$) and the scale parameter of the loss function are fixed at 0 (MSE). The shaded areas mars the situations for which the null hypothesis of the DM test is rejected.

References

[1] Andersen, T., Bollerslev, 1998. "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts", *International Economic Review*, 39, 885-905.

[2] Andersen, T., Bollerslev, T., Diebold, F.X., Labys, P., 2001." The distribution of exchange rate volatility", *Journal of American Statistical Association*, 96, 42-55.

[3] Andersen, T., Bollerslev, T., Diebold, F.X., Labys, P., 2003. "Modeling and forecasting realized volatility", *Econometrica*, 71, 529-626.

[4] Andersen, T., Bollerslev, T., Diebold, F.X., Labys, P., 2002. *"Parametric and nonparametric volatility measurement"*. In: Hansen, L.P., Ait-Sahalia, Y. (Eds.), Handbook of Financial Econometrics. North-Holland, Amsterdam, forthcoming.

[5] Barndorff-Nielsen, O., Shephard, N., 2002." Econometric analysis of realized volatility and its use in estimating stochastic volatility models", *Journal of the Royal Statistical Society*, Series B 64, 253-280.

[6] Barndorff-Nielsen, O., Shephard, N., 2004. "Power and bipower variation with stochastic volatility and jumps (with discussion)", *Journal of Financial Econometrics*, 2, 1-48.

[7] Bollerslev, T., 1986." Generalized autoregressive conditional heteroskedasticity", *Journal of Econometric*, 31, 307-327.

[8] Diebold, F.X., Mariano, R.S., 1995. "Comparing predicting accuracy", *Journal of Business and Economic Statistics*, 13(3), 253-263.

[9] Ding, Z., Granger, C., 1996. "Modeling volatility persistence of speculative returns: A new approach", *Journal of Econometrics*, 73, 185-215.

[10] Engle, R.F., 1982." Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation", *Econometrica*, 50, 987-1008.

[11] Engle, R., 2000. "The econometrics of ultra-high-frequency data", *The Econometric Society*, 1, 1-22.

[12] Engle, R., Ng, V., 1993. "Measuring and testing the impact of news on volatility", *Journal of Finance*, 71, 1749-1778.

[13] Ghysels, E., Chen, X., 2010. "News - good or bad - and its impact on volatility predictions over multiple horizons", Working paper.

[14] Ghysels, E., Forsberg, L., 2004. "Why do absolute returns predict volatility so well?", Working paper.

[15] Ghysels, E., Rubia, A., Valkanov, R., 2009." Multi-Period forecasts of volatility: Direct, Iterated, and Mixed-data approaches", Working paper.

[16] Ghysels, E., Santa-Clara, P., Volkanov, R., 2004. "The MIDAS touch: mixed data sampling regression", Working paper.

[17] Ghysels, E., Santa-Clara, P., Volkanov, R., 2006. "Predicting volatility: getting the most out of return data sampled at different frequencies", *Journal of Econometrics*, 131, 59-95.

[18] Ghysels, E., Sinko, A., Valkanov, R., 2007. "MIDAS regression: Further results and new directions", *Econometric reviews*, 26, 53-90.

[19] Hansen, P., Lunde, A., 2005. "A forecast comparison of volatility models: Does anything beat a GARCH (1,1)?", *Journal of Applied Econometrics*, 20, 873-889.

[20] Marcelino, M., Stock, J., Watson, M., 2006. "A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series", *Journal of Econometrics*, 135, 499-526.

[21] McAleer, M., Medeiros, M.C., 2008. "Realized volatility: a review", *Econometric Reviews*, 27 (1-3), 10-45.

[22] Patton, A. J., 2011. "Volatility forecast comparison using imperfect volatility proxies", *Journal of Econometrics*, 260, 246-256.