Do We Need Ultra-High Frequency Data to Forecast Variances?*

Denisa Banulescu[†], Bertrand Candelon[‡], Christophe Hurlin[§], Sébastien Laurent[¶]

September 29, 2015

Abstract

In this paper we study various MIDAS models in which the future daily variance is directly related to past observations of intraday predictors. Our goal is to determine if there exists an optimal sampling frequency in terms of volatility prediction. Via Monte Carlo simulations we show that in a world without microstructure noise, the best model is the one using the highest available frequency for the predictors. However, in the presence of microstructure noise, the use of ultra high-frequency predictors may be problematic, leading to poor volatility forecasts. In the application, we consider two highly liquid assets (*i.e.*, Microsoft and S&P 500). We show that, when using raw intraday squared log-returns for the explanatory variable, there is a "high-frequency wall" or frequency limit above which MIDAS-RV forecasts deteriorate. We also show that an improvement can be obtained when using intraday squared log-returns sampled at a higher frequency, provided they are pre-filtered to account for the presence of jumps, intraday periodicity and/or microstructure noise. Finally, we compare the MIDAS model to other competing variance models including GARCH, GAS, HAR-RV and HAR-RV-J models. We find that the MIDAS model provides equivalent or even better variance forecasts than these models, when it is applied on filtered data.

JEL classification: C22; C53; G12

Keywords: Variance Forecasting; MIDAS; High-Frequency Data

*The authors thank Eric Ghysels and the participants at the Ph.D. course on MIDAS models jointly organized by CESAM, the National Bank of Belgium (NBB) and the Center for Operations Research and Econometrics (CORE) in June 2013. We also thank Gilbert Colletaz and Christophe Boucher for helpful comments on the paper, as well as the participants at the 61^{st} Congress of the French Economic Association in Paris, 2012, at the 6^{th} International Workshop on Methods in International Finance Network in Sydney, 2012, at the 6^{th} International Conference on Computational and Financial Econometrics in Oviedo, 2012, at the 21^{st} Symposium of the Society for Nonlinear Dynamics and Econometrics in Milan, 2013, at the 3^{rd} Spring International Conference of the French Finance Association in Lyon, 2013. The usual disclaimers apply.

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

[‡]Insti7/IPAG chaire on Financial Stability and Systemic Risks. Email: candelonb@gmail.com

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

[¶]IAE Aix-en-provence, GREQAM. Email: sebastien.laurent@iae-aix.com

1 Introduction

The mixed data sampling (henceforth MIDAS) regression model, introduced in Ghysels et al. (2004), allows to forecast a measure of the daily variance (*e.g.*, realized variance) by considering past intraday log-returns. In their seminal paper, Ghysels et al. (2006) consider various MIDAS regressions with different daily (squared returns, absolute returns, realized variance, realized power and return range) and intradaily regressors (squared returns, absolute returns), to examine whether one specification dominates the others. The goal of our study is different and consists in determining, for a *given* intradaily predictor, whether a *sampling frequency* (or a range of frequencies) dominates the others.[1] The objective is then to identify the best sampling frequency, using out-of-sample forecast evaluation criteria.

This issue is not straightforward. On the one hand, not using the readily available high-frequency observations to perform variance forecasts implies a loss of information through the temporal aggregation. On the other hand, if the sampling frequency of the predictors is increased too much, the market microstructure noise (bid-ask bounce, screen fighting, jumps, and irregular or missing data) may lead to less accurate variance forecasts.

This question has to be distinguished from the well-documented discussion about the optimal sampling frequency of the returns used to compute realized estimators of daily variance (see Hansen and Lunde, 2004; Aït-Sahalia and Mancini, 2008; Garcia and Meddahi, 2006; Ghysels et al., 2006, among others). Our goal consists in focusing on the optimal sampling frequency for the purpose of variance *prediction*, and not for variance *measurement*.

Consider a MIDAS variance model whose aim is to predict a measure of variance over some future horizon. This variance measure is typically a realized measure (realized variance, realized kernel etc.), based on intradaily returns sampled at a frequency m_2. In order to forecast variance, we adopt exactly the same approach as Ghysels et al. (2006) and consider intradaily predictors (squared returns, intradaily bipower variation, etc.) sampled at a frequency m_1, where m_1 may be different from m_2. The discussion concerns only the sampling frequency of the predictors, m_1.

In a related paper, Ghysels and Sinko (2011) study a regression prediction problem with variance measures that are contaminated by market microstructure noise and examine optimal sampling for the purpose of variance prediction. They observe that, in general, discussions about the impact of microstructure have mostly focused on measurement. Ghysels and Sinko (2011) focus instead on prediction in a regression framework, and therefore they can consider estimators that are suboptimal in the mean squared error (MSE) sense, since their covariation with the predictor is the object of interest. The authors consider univariate MIDAS regressions for the evaluation

[1] In this study, we limit our analysis to the MIDAS specifications in which the future variance is directly related to past observations of intraday predictors, as in Ghysels et al. (2006). An alternative consists in using high-frequency data to compute daily realized measures (realized variance, two-scale estimator, realized kernel, etc.) which are introduced, in a second step, into a MIDAS regression model, as in Ghysels et al. (2006) and Ghysels and Sinko (2011). This choice will be discussed in Section 5.

of the prediction performance and derive the optimal frequency in terms of prediction MSE. Their dependent variable is defined as the two scales estimator of the weekly variance (Aït-Sahalia et al., 2005), and computed from the 5-minute, 1-minute or 2-second returns. One of the main differences with our study is that the authors consider various MIDAS specifications for which the predictors also correspond to realized estimators (plain vanilla, two scales estimator, Zhou, 1996, etc.), constructed using different sampling frequencies (from two seconds to ten minutes). Thus, high-frequency data are aggregated into daily realized measures, which are then used as predictors of future variance. In contrast, our goal is to analyze the direct impact of the intradaily predictors on the variance forecasts, and ultimately to evaluate the usefulness of the mixing of frequencies in this context. To this aim, we consider MIDAS models in which we directly project future realized variance onto high-frequency regressors, as in Ghysels et al. (2006).[2]

To address these issues, we propose a Monte Carlo simulation study. Considering a noise-free diffusion process, we generate returns series at different sampling frequencies m_1 and daily realized variance measures, using the same set of continuous-time structural parameters. Then, we apply simple MIDAS specifications in which daily realized variance is predicted by past intradaily squared log-returns sampled at a frequency m_1, ranging from one minute to one day. The variance forecasts are compared based on the robust loss function proposed by Patton (2011) and the model confidence set (MCS) test introduced by Hansen et al. (2011). This test aims at identifying among the set of competing models (*i.e.*, sampling frequencies), the subset of models that are equivalent in terms of forecasting ability and their outperformance of all the other models for a given confidence level. Several results stand out. First, we observe that a higher sampling frequency for the regressors implies giving more weight to the most recent observations of the regressors. Second, we show that increasing the frequency of the regressors always improves the forecasting abilities of the MIDAS model. The average loss increases when the regressors are sampled less frequently, regardless of the choice of the loss function. These differences are statistically significant and the MCS test always concludes that the MIDAS model with the highest available sampling frequency significantly stands out in terms of forecasting performances. Nevertheless, opting for ultra-high frequency regressors is not optimal in the presence of microstructure noise. In this specific case, we need to use pre-averaged data to improve the MIDAS performances.

The sensitivity of MIDAS variance models to the choice of the sampling frequency m_1 is also investigated on real data, *i.e.*, log-returns of the S&P 500 index and Microsoft over the period of October 29, 2004 to December 31, 2008. The empirical results obtained for these two assets allow us to draw some interesting conclusions. First, when using raw intraday returns, variance forecasts are not statistically different for sampling frequencies of the predictors (m_1) ranging from five minutes to one hour. Besides, it turns out that ultra-high-frequency

[2]Surprisingly, Ghysels et al. (2006) find that the forecasts directly using high-frequency data do not outperform those based on daily regressors (although the daily regressors are themselves obtained through the aggregation of high-frequency data). One related question is to understand whether this result depends on the sampling frequency of the high-frequency data.

regressors (*i.e.*, $m_1 < 5$min) do not necessarily provide useful information to improve the variance forecasts because the loss function increases. The shape of the loss function indicates the presence of a "high-frequency wall", *i.e.*, a limit frequency beyond which the quality of the forecasts deteriorates. This result is due to the presence of microstructure noise, jumps and intraday periodicity in the regressors. When the MIDAS regression model is applied to filtered data (Lee and Mykland, 2008; Boudt et al., 2011; Lahaye et al., 2011), the conclusion in favor of the use of the highest available frequency remains valid. This point is crucial and indicates that the mixing frequency may require the use of filtered series. Indeed, the weighting scheme in MIDAS models does not allow, by itself, to underweight the observations affected by jumps or other market microstructure noise. These results are robust to the choice of variance measure (realized variance, realized kernel), forecasting horizon and sample period (calm/crisis).

Finally, we compare the performance of the MIDAS model (applied to filtered or unfiltered regressors) to other competing variance models, namely the GARCH(1,1), the Student Generalized Autoregressive Score (GAS) model (Creal et al., 2013), the Heterogeneous Autoregressive Realized Volatility-based (HAR-RV) model (Corsi, 2009) and the HAR-RV adjusted for jumps (Andersen et al., 2007). We show that MIDAS models are providing comparable or even better variance forecasts when filtered high-frequency data are used.

The chapter is structured as follows. Section 2 introduces the notations, the MIDAS model and the sampling frequency puzzle. Section 3 proposes a Monte Carlo simulation. In Section 4, we perform an empirical analysis and study the influence of the jumps and the intraday periodicity on the MIDAS performances. We also compare MIDAS to other competing variance models. Section 5 concludes.

2 Modeling Strategies

2.1 Notation

To set the notation, let p_t denote the price for a financial asset sampled at daily frequency, and the corresponding daily return be defined by $r_{t,t-1} \equiv log(p_t) - log(p_{t-1})$. The equally spaced series of continuously compounded returns is assumed to be observed m times per day (or to have an horizon of $1/m$), and be computed as $r^{(m)}_{t,t-1/m} \equiv log(p_t) - log(p_{t-1/m})$, where $t = 1/m, 2/m, ...$ Throughout the analysis, we consider that the trading day spans the time period from 9:30 am to 16:00 pm, covering for instance $m = 390$ 1-minute equally spaced intervals and $m = 78$ 5-minute equally spaced intervals. $r_{t,t-1/78}$ corresponds to the last 5-minute return of the day $t-1$, $r_{t-1/78,t-2/78}$ corresponds to the return of the penultimate 5-minute period of day $t-1$, and so on.

2.2 MIDAS variance Models and Sampling Frequency

MIDAS models for variance predictions have been introduced in a number of recent studies, including Ghysels et al. (2005, 2006), Ghysels and Sinko (2011), Ghysels and Valkanov (2012), Chen and Ghysels (2011), among others.

The general specification of the MIDAS variance model is given by:

$$\sigma^2_{t+H,t} = \mu_{H,m_1} + \phi_{H,m_1}\Omega_{H,m_1}(L^{1/m_1})X^{(m_1)}_{t,t-1/m_1} + \varepsilon_t, \tag{1}$$

where $\sigma^2_{t+H,t}$ is a measure of variance evaluated over some future horizon H, and $X^{(m_1)}_{t,t-1/m_1}$ denotes an intradaily regressor sampled at frequency m_1. The distributed lag polynomial is defined as:

$$\Omega_{H,m_1}(L^{1/m_1}) = \sum_{k=0}^{k_{max}} L^{k/m_1}\omega_{H,m_1}\left(k, \theta_{H,m_1}\right), \tag{2}$$

where $\omega_{H,m_1}\left(k, \theta_{H,m_1}\right)$ corresponds to the lag coefficient associated with $X^{(m_1)}_{t,t-1/m_1}$, θ_{H,m_1} is a finite set of parameters, L is the lag operator such that $L^{1/m_1}X^{(m_1)}_{t,t-1/m_1} = X^{(m_1)}_{t-1/m_1,t-2/m_1}$, and k_{max} denotes the maximum number of lagged coefficients. In this specification, the low-frequency variance (for instance, daily variance if $H = 1$, weekly variance if $H = 5$, etc.) is predicted by the right-side intradaily forecasting factors which are sampled at a high-frequency m_1 (for instance, five minutes if $m_1 = 78$). Several intradaily regressors can be considered with this aim (*e.g.*, intradaily squared returns, intradaily bipower variation). Following Ghysels et al. (2006) we consider the intradaily squared returns $r^{(m_1)2}_{t,t-1/m_1}$, while other alternatives will be used to appraise the robustness of our results.[3]

Since $\sigma^2_{t+H,t}$ is unobservable, we rely on a proxy. For simplicity, we adopt the realized variance (Andersen and Bollerslev, 1998a), defined for the period t to $t+H$ as follows:[4]

$$RV^{(m_2)}_{t+H,t} = I_{H,m_2}(L^{1/m_2})r^{(m_2)2}_{t+H,t+H-1/m_2}, \tag{3}$$

where the distributed lag polynomial in L^{1/m_2} is defined such that $I_{H,m_2}(L^{1/m_2}) = \sum_{j=0}^{Hm_2} L^{j/m_2}$, and m_2 accounts for the sampling frequency of the squared returns used to compute the realized variance. Notice that the frequencies m_1 and m_2 may be different. In fact, the choice of m_2 is related to the variance *measurement* issue (see Hansen and Lunde, 2004; Aït-Sahalia and Mancini, 2008; Garcia and Meddahi, 2006; Ghysels et al.,

[3] Another alternative would be to accommodate large volumes of data in a parsimonious way by using intraday aggregate measures of volatility (*e.g.*, 5-minute realized volatility, etc.) as regressors. However, the aim of this paper is rather to observe the behavior of ultra-high-frequency data when estimating/forecasting volatility and not to propose the best volatility model and/or predictor of volatility.

[4] A large number of alternative estimators (*e.g.*, realized bipower variation, realized kernel, etc.), that deal with issues such as jumps and other market microstructure noise, have been proposed, especially by Barndorff-Nielsen and Shephard (2004a), Barndorff-Nielsen et al. (2008), Zhang (2006), Hansen and Horel (2009), *inter alios*. Some of them will be considered in the section devoted to the robustness analysis of our findings.

2006, among others), *i.e.*, the consistency of the estimator defined by the realized measure.[5] On the contrary, the choice of the sampling frequency of the predictors, m_1, is related to the variance *prediction* issue. Indeed, this choice determines the regressors in the MIDAS variance model, and as a consequence, its forecasting abilities.

Under these assumptions, the MIDAS-RV regression becomes:

$$RV_{t+H,t}^{(m_2)} = \mu_{H,m_1} + \phi_{H,m_1} \Omega_{H,m_1}(L^{1/m_1}) r_{t,t-1/m_1}^{(m_1)2} + \varepsilon_t. \quad (4)$$

One advantage of this specification is that it preserves the information contained in high-frequency data (Ghysels and Valkanov, 2012) without computing daily aggregates such as realized variance for the regressors. In this context, we aim to determine the influence of the sampling frequency m_1 on the forecasting performances of the MIDAS model. In a related study, Ghysels et al. (2006) compare several MIDAS specifications based on *different* intradaily or daily variance regressors (*e.g.*, squared returns, absolute returns, realized volatility, realized power, and range). The logic is similar here, except that we consider the *same intradaily regressor, i.e.,* $X_{t,t-1/m_1}^{(m_1)}$, for *various sampling frequencies.* For instance, we compare various MIDAS models where the same predictor is sampled at one minute ($m_1 = 390$), two minutes ($m_1 = 195$), five minutes ($m_1 = 78$), and so on. The question is whether increasing the sampling frequency m_1 systematically improves the quality of the variance forecasts, and ultimately if we need ultra-high frequency data in order to forecast daily variances.

3 Monte Carlo Simulation Study

We first propose a Monte Carlo simulation study in order to determine the influence of the sampling frequency of the regressors on the predictive abilities of MIDAS-RV models. We begin by describing the simulation setup and then we follow this by discussing the results.

3.1 Monte Carlo Design

Let us assume that instantaneous log-returns, $\mathrm{d}p_t$, are generated by the continuous-time martingale

$$\mathrm{d}p(t) = \sigma(t)\mathrm{d}W_p(t), \quad (5)$$

where $W_p(t)$ denotes a standard Wiener process, and $\sigma(t)$ is given by a separate continuous-time diffusion process. For $\sigma(t)$, we use the diffusion limit of the GARCH(1,1) process introduced by Nelson (1990), *i.e.*,

$$\mathrm{d}\sigma^2(t) = \theta(\omega - \sigma^2(t))\mathrm{d}t + (2\lambda\theta)^{1/2}\sigma^2(t)\mathrm{d}W_\sigma(t), \quad (6)$$

[5] Since this study is not meant to determine the optimal sampling frequency m_2, in the rest of the paper, the daily RV will always be computed by summing up *5-minute* squared returns (*i.e.*, $m_2 = 78$), as recommended by Andersen and Bollerslev (1998a).

where $\omega > 0$, $\theta > 0$, $0 < \lambda < 1$, and the Wiener processes, $W_p(t)$ and $W_\sigma(t)$, are independent. Drost and Werker (1996) and Drost and Nijman (1993) prove that the exact discretization for stochastic variance processes is in line with the weak GARCH(1,1) representation meaning that a weak GARCH process can be identified at any discrete-time sampling frequency from the parameters of a continuous GARCH and vice versa.[6] $(\omega, \theta, \lambda)$ is set to $(0.636, 0.035, 0.296)$ as in Andersen and Bollerslev (1998a) (these parameters have been calibrated by the authors to fit the daily GARCH estimates for the Deutschemark-U.S. Dollar (DM-$) spot exchange rates).

Given this data generating process (DGP), we draw large series of continuous-time log-returns and compute log-returns series sampled at different frequencies by applying the temporal aggregation proprieties of flow variables, *i.e.*, $r^{(m)}_{t,t-1/m} = \int_{t-1/m}^{t} \mathrm{d}p(\tau)\mathrm{d}\tau$. We consider various frequencies m_1 ranging from five seconds to one day. Next, we compute discrete realized variance series by summing up simulated 5-minute squared return series (see Eq. (3)). In so doing, we obtain all the necessary elements to run MIDAS-RV regressions as defined in Eq. (4).

One notable advantage of this procedure is that it allows to generate intradaily log-returns series at different sampling frequencies, m_1, and daily RVs, using the same data generating process and the same set of continuous time structural parameters. In addition, this process is calibrated to reproduce the main features of typical real financial series. The Monte Carlo simulation exercise is based on 10,000 replications and the daily/intradaily series are simulated for a period of 1,000 days. The parameters of all the competing MIDAS-RV models are estimated by Nonlinear Least Squares (NLS).[7] In order to allow for a fair comparison between models, the maximum lag order k_{max} is fixed, such that the past information used to predict volatility covers a period of 30 days, whatever the sampling frequency of the regressor. A daily forecasting horizon ($H = 1$) is used in all simulations.

3.2 Weight Function and Sampling Frequency

One of the key features of MIDAS models is that it provides a parsimonious specification. This property is particularly important in our context, as the inclusion of high-frequency data might imply a significant increase in the number of lagged forecasting variables and hence the number of unrestricted parameters to be estimated (Ghysels and Valkanov, 2012). For instance, running unrestricted regressions based on the intraday information over the last 30 days implies estimating 30×390 parameters for a 1-minute regressor, 30×78 parameters for a 5-minute regressor, and so on. Nevertheless, the MIDAS model projects directly future variance onto an important number of high-frequency lagged regressors while considering a small number of parameters. The trick consists in using a suitable parametrization for the weights $\omega_{H,m_1}(k, \theta_{H,m_1})$ to circumvent the problem

[6] Meddahi and Renault (1998) show that the strong GARCH setting does not have a closed form with respect to temporal and contemporaneous aggregations.

[7] For more details about the estimation procedure, see Ghysels et al. (2004).

of parameter proliferation. Therefore, as noted by Ghysels et al. (2006), the parametrization $\omega_{H,m_1}(k, \theta_{H,m_1})$ becomes one of the most important ingredients in a MIDAS regression.

Two specifications of the weight function are generally considered, namely the exponential Almon lag and the Beta lag (Ghysels et al., 2007). These specifications have several interesting features: *i*) the distributed lag polynomial is tightly parameterized and prevents the proliferation of parameters as well as additional pre-testing or lag-selection procedures;[8] *ii*) the coefficients are positive, which guarantees non-negative weights and consequently non-negative variance forecasts; *iii*) the data-driven weights are normalized to add up to one in order to identify the scale parameter ϕ_{H,m_1}. There is no clear theoretical a priori for assuming that one specification is better than the other. However, Chen and Tsay (2011) and Frale and Monteforte (2011) find that the Beta function is more suitable for an important number of time lags, as Almon could be very computationally demanding in such a context. For this reason, we adopt the Beta lag polynomial

$$\omega_{H,m_1}(k, \theta_{H,m_1}) = \frac{f(k/k^{\max}, \theta_1; \theta_2)}{\sum_{j=0}^{k^{\max}} f(j/k^{\max}, \theta_1; \theta_2)}, \tag{7}$$

where $\theta_{H,m_1} = (\theta_1, \theta_2)'$ is a vector of positive parameters, $f(z, a, b) = z^{a-1}(1-z)^{b-1}/\mathrm{B}(a, b)$, with $\mathrm{B}(.)$ the Beta function defined as $\mathrm{B}(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$, and $\Gamma(.)$ representing the Gamma function. Depending on the value of the parameter θ_1, this weight function can take many shapes, including flat weights, gradually declining weights, as well as hump-shaped patterns. The second parameter, θ_2, determines the decreasing speed of the weighting shape. The smaller the parameter θ_2, the smoother the weighting scheme. In other words, θ_2 determines the proportion of the total weight associated with the more recent past observations.

Table 1 presents the outline of the regression diagnostics for the 10,000 replications considered in the Monte Carlo experiment. The first four columns represent the average values of the parameter estimates for various MIDAS-RV models, each associated with a particular frequency m_1. Results first suggest that the estimates for the constant term of the model, μ_{H,m_1}, and the first parameter of the weight function, θ_1, decrease with m_1. On the contrary, the estimates for the scale parameter, ϕ_{H,m_1}, and the second parameter of the weight function, θ_2, increase with the sampling frequency. These changes imply a deformation of the weight function that gives more weight to the most recent observations. This is confirmed by the next four columns of Table 1. Column "Day 1" reports $\sum_{k=0}^{K} \omega_{H,m_1}(k, \bar{\bar{\theta}}_{H,m_1})$ for a value of K corresponding to one day and $\bar{\bar{\theta}}_{H,m_1} = \left(\bar{\bar{\theta}}_1, \bar{\bar{\theta}}_2\right)'$ is the vector containing the average estimates of the parameters θ_1 and θ_2 over the 10,000 replications. Similarly, Column "Days 2-5" presents how much weight is given to the information on the second to the fifth lagged day, and so on. Results confirm that more weight is given to the most recent observations of the variance predictor when the sampling frequency m_1 increases. The proportion of the weights allocated to the observations of the first

[8] The selection of $k^{\max}$ can be done by considering a large value and letting the weights vanish.

lagged day represents 88.19% when the regressors are sampled at one minute, and this proportion decreases progressively to 15.67% when the regressors are sampled at 3h15.

Table 1: Regression diagnostics and estimated weights of MIDAS models with intradaily regressors

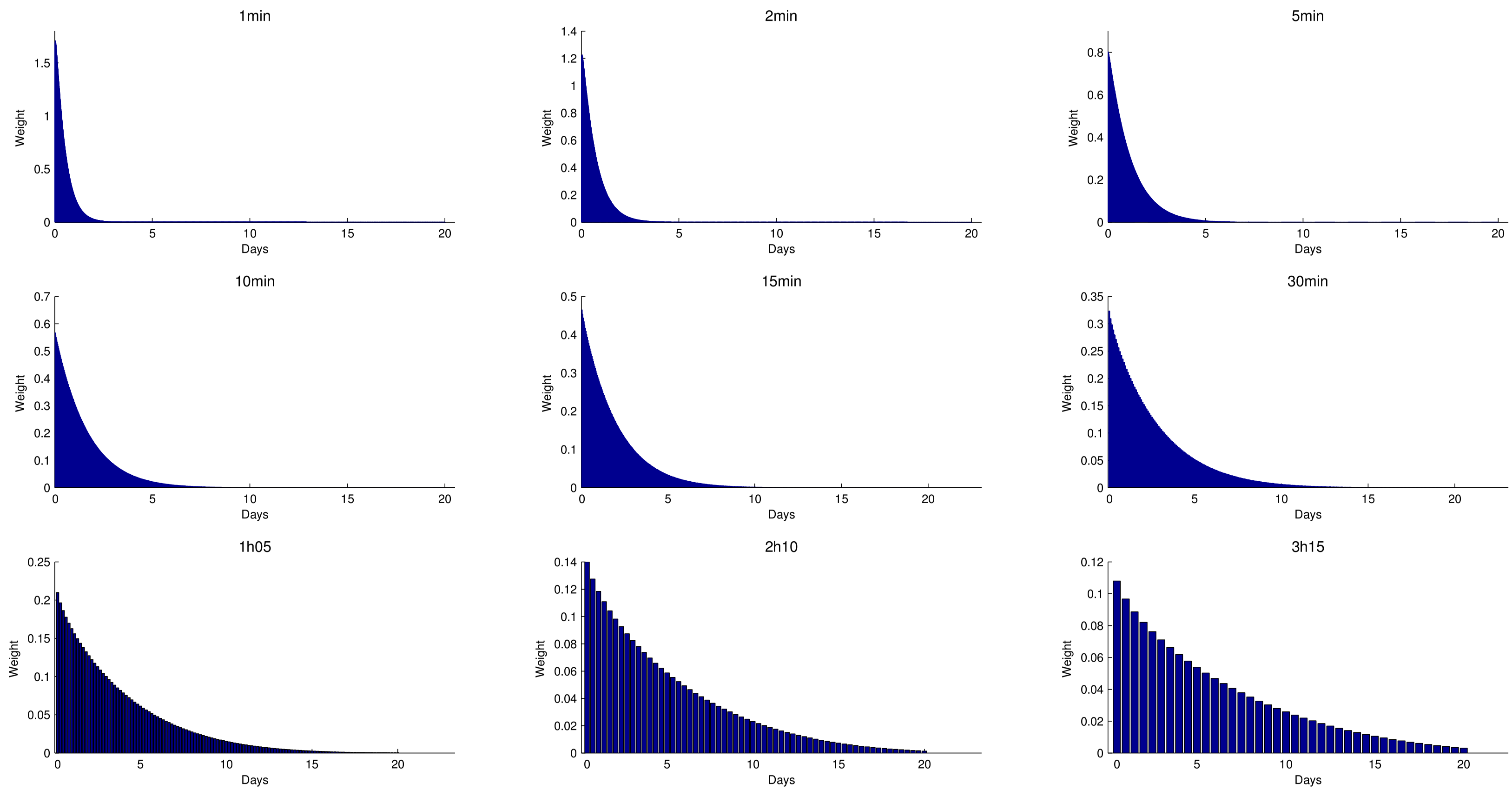
Frequency	μ	ϕ	θ_1	θ_2	Day 1	Days 2-5	Days 6-15	> 15 Days	%Q(12)	$\%Q^2(12)$	MSE
5s	0.0178	4545.0808	0.8603	154.1350	0.9962	0.0038	0.0000	0.0000	62	100	0.0201
15s	0.0212	1505.7378	1.0088	119.8657	0.9825	0.0175	0.0000	0.0000	65	100	0.0213
30s	0.0242	748.8773	1.0299	90.5192	0.9508	0.0492	0.0000	0.0000	61	100	0.0224
1min	0.0287	371.5729	1.0715	66.4510	0.8819	0.1181	0.0000	0.0000	56	100	0.0241
2min	0.0342	184.1507	1.0500	45.6410	0.7720	0.2278	0.0003	0.0000	52	100	0.0263
3min	0.0392	121.6448	1.0639	37.8144	0.6995	0.2993	0.0012	0.0000	55	100	0.0280
5min	0.0465	72.1224	1.0585	29.1585	0.6034	0.3909	0.0056	0.0000	47	100	0.0307
10min	0.0597	35.2234	1.0158	19.4822	0.4760	0.4944	0.0292	0.0004	100	100	0.0350
15min	0.0702	23.0368	1.0052	15.7625	0.4114	0.5315	0.0554	0.0017	100	100	0.0392
30min	0.0914	11.0603	0.9686	10.7978	0.3195	0.5468	0.1218	0.0119	100	100	0.0458
1h05	0.1278	4.7577	0.9549	7.2350	0.2331	0.5126	0.2045	0.0498	100	100	0.0566
2h10	0.1704	2.1779	0.9422	5.2898	0.1801	0.4577	0.2527	0.1095	100	100	0.0703
3h15	0.1992	1.3618	0.8898	4.1149	0.1567	0.4105	0.2655	0.1673	100	100	0.0771
1day	0.2630	0.5803	0.8417	3.0120	0.1252	0.3505	0.2667	0.2576	100	100	0.0913

Note: This table reports average values (over 10,000 replications) of the parameter estimates for the daily MIDAS-RV model ($H = 1$) with regressors sampled at five seconds, 15 seconds, 30 seconds, one minute, two minutes, three minutes, five minutes, 10 minutes, 15 minutes, 30 minutes, 1h05, 2h10, 3h15 and one day (Eq. 4). Column "Day 1" reports the sum of the weights associated with the first lagged day of the predictors, column "Days 2-5" present how much weight is given to the information of the second to the fifth lagged day of the predictors, and so on. The next two columns, namely $\%Q(12)$ and $\%Q^2(12)$, correspond to the frequencies of rejection of the null hypothesis of no serial correlation at the 5% significance level for the Ljung-Box test applied on respectively residuals and squared residuals. The last column reports the average Mean Squared Error (MSE).

To illustrate the deformation of the Beta polynomial shape, Figure 1 displays $\widehat{\phi}_{m_1}\omega_{m_1}(k,\widehat{\theta}_{m_1})$, *i.e.*, the product of the weight (determined by the Beta function) and the scale parameter estimate (see Eq. 1), as a function of the sampling frequency m_1. For ease of comparison, the weights are displayed for a only 20 day window. First, in all of the cases, the weight function is gradually declining and there is no hump-shaped pattern. Second, the slope of the weight function becomes smaller when the predictors are sampled at a lower frequency. If the weights vanish after approximately three days in the case of a regressor sampled at one minute, the weights of the observations associated with the 20^{th} lagged day are still positive for a regressor sampled at three hours.

The sampling frequency of the predictors also has an impact on the quality of the fit of daily realized variances (*i.e.*, $H = 1$). Columns $\%Q$ and $\%Q^2$ in Table 1 report rejection frequencies of the null hypotheses of no serial correlation in the residuals and squared residuals (respectively) at the 5% significance level using a Ljung-Box test with 12 lags. Serial correlation and heteroskedasticity are detected in most cases.

Figure 1: Scaled weight function



Note: This figure displays the scaled weights pattern based on average parameter estimates obtained in a Monte Carlo simulation study, for the nine MIDAS models with regressors sampled at a frequency ranging from one minute to 3h15. The scaled weights are obtained by multiplying the Beta weight function $\widehat{\omega}_{m_1}(k, \widehat{\theta}_{m_1})$ by the scale parameter, $\widehat{\phi}_{m_1}$ (see Eq. 1). For ease of comparison, the weights are represented over the first 20 lagged days whatever the sampling frequency of the variance regressor.

This is in line with Ghysels et al. (2006), who also find significant autocorrelation and heteroskedasticity in the residuals of daily MIDAS-RV regressions.[9] However, results suggest that the problem of serial correlation in the residuals is less pronounced when using ultra-high frequency returns.

For instance, when the regressors are sampled at five minutes, the residuals do not feature autocorrelation in 53% of the simulated samples, and this percentage decreases to 35% when the regressors are sampled at 15 seconds. Finally, increasing the sampling frequency m_1 always tends to improve the in-sample goodness of fit, as indicated in the last column of Table 1 where the average MSE is reported.

Notice that all the MIDAS-RV models are directly comparable in terms of MSE since they all have the same number of estimated parameters whatever the frequency m_1. When the sampling frequency increases from five minutes to one minute, the gain in terms of average MSE reaches 27.38%. These gains are statistically significant. Since we have many competing models (*i.e.*, frequencies m_1), we focus on multiple comparison-based tests and use the Model Confidence Set (MCS) approach introduced by Hansen et al. (2011). This test allows identifying, among an universe of competing forecasting models, the subset of models that are equivalent in terms of forecasting ability, and which outperform all the other models at a confidence level α. Interestingly, we find that the MCS test systematically selects the ultra-high frequency MIDAS-RV specification, regardless of the loss-function used (KLIC, AIC or BIC).[10]

3.3 Out-of-Sample Analysis

To check whether the previous results remain valid out-of-sample, we subsequently focus on the influence of the sampling frequency of the variance predictors, m_1, on the predictive abilities of the MIDAS-RV model. At each replication, we compute a sequence of $T = 500$ daily realized variance forecasts, $\{\widehat{RV}^{(m_2)}_{t+1,t}\}_{t=1}^{T}$, for each MIDAS-RV specification. The forecasts sequences are obtained with a rolling window approach and the parameter estimates are updated every 50 days.

In order to compare these forecasts, we must use a loss function, defined as a general function of the variance forecasts and the true variance. In our simulation framework, the variance can be measured by the daily integrated variance, $IV_{t,t-1} = \int_{t-1}^{t} \sigma^2(\tau)\mathrm{d}\tau$. However, in practice, the integrated variance is not observable and we have to use a proxy. To reproduce the real conditions of application of the MIDAS-RV models, we also use a variance proxy to define the loss function, *i.e.*, the realized variance $RV^{(m_2)}_{t+H,t}$.[11] However, it is well known that the use of a proxy may distort the ranking of models based on loss functions. Andersen and Bollerslev (1998a)

[9] However, Ghysels et al. (2006) find no significant autocorrelation for longer forecasting horizons (from one week to four weeks). We obtain similar results in our simulations (not reported).

[10] We set the significance level to $\alpha = 25\%$ and use 10,000 bootstrap resamples (with block length of five observations) to obtain the distribution under the null of equal empirical fit (Hansen et al., 2011). The MCS p-values are qualitatively similar when we choose different block lengths. These results are available under request.

[11] Notice that the results obtained with the integrated variance (not reported) are qualitatively the same than those obtained with the realized variance.

and Andersen et al. (2005) show that the comparison of losses, even based on a conditionally unbiased proxy, may lead to a different outcome than the one obtained if the true latent variable had been used. More recently, Hansen and Lunde (2006a), Patton and Sheppard (2009), Patton (2011), Laurent et al. (2013) have also insisted on the possible distortions observed in the ranking of volatility forecasts induced by the use of a noisy proxy.[12] For these reasons, we adopt the family of robust and homogeneous loss functions proposed by Patton (2011), *i.e.*,

$$L(\widehat{\sigma}^2, \sigma^2; b) = \begin{cases} \frac{1}{(b+1)(b+2)}(\widehat{\sigma}^{2(b+2)} - \sigma^{2(b+2)} - \frac{1}{b+1}\widehat{\sigma}^{2(b+1)}(\widehat{\sigma}^2 - \sigma^2), & \text{for } b \notin \{-1,-2\} \\ \sigma^2 - \widehat{\sigma}^2 + \widehat{\sigma}^2 \ \log \frac{\widehat{\sigma}^2}{\sigma^2}, & \text{for } b = -1 \\ \frac{\widehat{\sigma}^2}{\sigma^2} - \log \frac{\widehat{\sigma}^2}{\sigma^2} - 1, & \text{for } b = -2 \end{cases} \tag{8}$$

with b a scalar parameter, σ^2 a measure of the true variance (*i.e.*, the realized variance in our case) and $\widehat{\sigma}^2$ the predicted variance measure. This loss function encompasses in particular the MSE and the QLIKE loss functions when $b = 0$ and $b = -2$, respectively.

Evaluating the influence of the sampling frequency of the predictors on the predictive abilities of the MIDAS-RV model reduces to determining the sign of the derivative of the average loss function given by:

$$L_{m_1} = T^{-1} \frac{\partial \sum_{t=1}^{T} L(\widehat{RV}_{t+1,t}^{(m_2)}, RV_{t+1,t}^{(m_2)}; b)}{\partial m_1}. \tag{9}$$

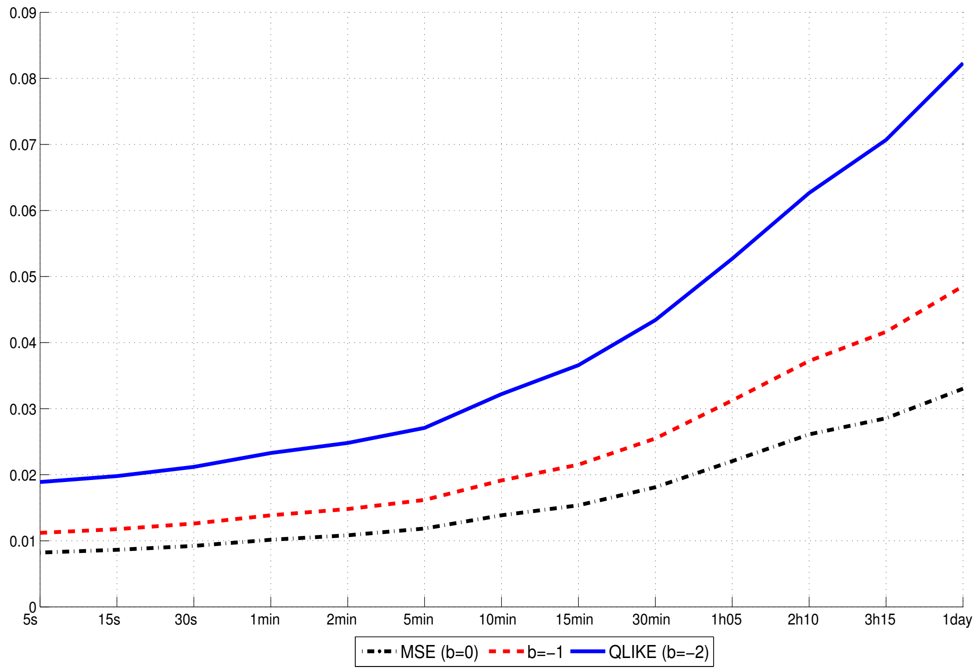
Since the sign of this derivative cannot be determined analytically we proceed by numerical analysis. Figure 2 displays the average (over the 10,000 replications) of the loss function $T^{-1} \sum_{t=1}^{T} L(\widehat{RV}_{t+1,t}^{(m_2)}, RV_{t+1,t}^{(m_2)}; b)$, as a function of the frequency m_1.

In order to assess the robustness of our results, we consider three values for the parameter b, namely 0 (MSE), -1 and -2 (QLIKE). The main conclusion is that the average loss decreases with the sampling frequency of the predictors, regardless of the loss function specification. For instance, the MSE increases progressively from 0.0085 (for a 1-minute regressor) to 0.0286 (for a regressor sampled twice a day). The use of the highest available frequency for the predictors is hence favored not only in-sample but also out-of-sample.

Besides, these gains are found to be statistically significant using the MCS test of Hansen et al. (2011) (as discussed in the previous section). Table 2 reports the MCS results for one replication of the Monte Carlo experiment. For each sampling frequency m_1, we display the average loss function along with the corresponding MCS p-value. The entries in bold correspond to the cluster of the best MIDAS-RV models as identified by the MCS test.

[12] The robustness of the forecasts ranking has also an impact on the statistical inference used to asses the predictive accuracy. If the loss function ensures consistency of the ranking, the variability of the variance proxy is only likely to reduce the power of the test, but not its asymptotic size, which means that for a robust loss function it is always possible to recover asymptotically the true ranking. For more details, see Laurent et al. (2013).

Figure 2: MIDAS average loss function



Note: This figure displays the average loss function (y-axis) associated with the MIDAS-RV forecasts based on various sampling frequencies (m_1) of the predictors (x-axis). Three different specifications of the robust loss function are considered, *i.e.*, Eq. (8) for $b = \{0, -1, -2\}$.

Table 2: Model Confidence Set test

	MSE ($b = 0$)		$b = -1$		QLIKE ($b = -2$)	
Frequency	Av. loss	MCS p-value	Av. loss	MCS p-value	Av. loss	MCS p-value
15s	**0.0091**	**0.7796**	**0.0104**	**1**	**0.0191**	**1**
30s	**0.0090**	**0.7796**	**0.0108**	**0.7396**	**0.0195**	**0.8013**
1min	**0.0083**	**1**	**0.0107**	**0.8894**	**0.0201**	**0.6808**
2 min	**0.0112**	**0.4386**	0.0123	0.1566	**0.0210**	**0.3583**
3min	**0.0087**	**0.7796**	**0.0108**	**0.8894**	**0.0196**	**0.8013**
5min	0.0120	0.1855	0.0136	0.0243	0.0237	0.0044
10min	0.0153	0.0613	0.0174	0.0026	0.0290	0.0001
15min	0.0200	0.1094	0.0208	0.0062	0.0330	0.0001
30min	0.0228	0.0613	0.0247	0.0026	0.0405	< 0.0001
1h05	0.0249	0.0613	0.0280	0.0026	0.0450	< 0.0001
2h10	0.0259	0.0613	0.0268	0.0021	0.0438	< 0.0001
3h15	0.0270	0.0613	0.0356	0.0016	0.0597	< 0.0001
1day	0.0385	0.0415	0.0500	0.0011	0.0818	< 0.0001

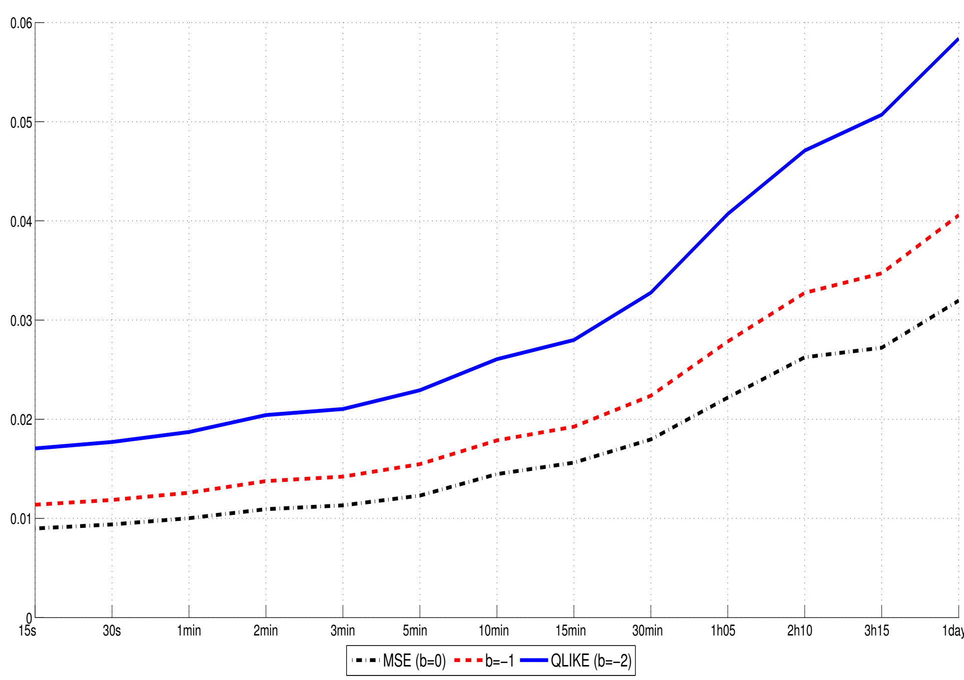
Note: This table presents the Model Confidence Set (MCS) results obtained for three different loss functions, *i.e.*, Eq. (8) for $b = \{0, -1, -2\}$. For each MIDAS specification the average value of the loss function is reported (first column) along with the corresponding p-value (second column) resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples (with block length of five daily observations) are used to obtain the distribution under the null of equal predictive accuracy. The entries in bold refer to the best MIDAS-RV forecasts according to the MCS test.

For each of the three loss functions, the MCS test confirms that the use of the ultra-high-frequency regressors leads to a significant improvement in the forecasting performances. This result is not specific for the particular replication reported in Table 2. The average value of the MCS p-values obtained over all the replications is not informative. Alternatively, it is possible to count the number of replications for which the MIDAS specification with the highest sampling frequency outperforms the other models. We find that in 98% of replications, the MCS approach selects the 15-second MIDAS model to be the best. This proportion reaches 56% when we consider the clusters of outperforming models including also 30-second to 3-minute MIDAS regressors.

3.4 DGP Sampling Frequency

In the previous experiment we considered a continuous data generating process and concluded in favor of the use of the highest frequency available for the predictors. Following Visser (2011) and Hecq et al. (2012), we now consider a DGP for 1-second log-returns where the conditional variance varies every consecutive five minutes according to a discrete-time GARCH(1,1) with parameters $(\alpha_0, \alpha_1, \beta) = (2.4693e-07, 0.0057, 0.9941)$ but is constant during every 5-minute intervals.

Figure 3: 5-minute DGP: MIDAS-RV average loss function



Note: See Figure 2. Notice that the conditional variance of the simulated 1-second log-returns (before aggregation) varies every five minutes according to a discrete-time GARCH(1,1) but is constant during every consecutive 5-minute intervals.

As in the previous simulation, out-of-sample forecasts as previously and analyze the forecasting accuracy of MIDAS variance models. Figure 3 displays the average loss functions (over 10,000 replications) as a function of the sampling frequency m_1.

Results suggest that using a sampling frequency m_1 greater than five minutes (*i.e.*, either 15 seconds, 30 seconds, one, two or three minutes) does not improve much the quality of the fit. However, using data sampled at a much lower frequency than five minutes leads to a huge loss of information and therefore important increases of the average losses, irrespective of the choice of the loss function.

3.5 Microstructure Noise

The main conclusion of the previous Monte Carlo simulation is that ultra-high-frequency log-returns are not always useful in the context of MIDAS-RV. Another disadvantage of using ultra-high-frequency data is that at these frequencies, the true price process is likely to be contaminated by microstructure effects arising from market frictions, such as the bid-ask bounce or the discreteness of prices. This phenomenon produces spurious variations in asset prices and induces autocorrelation in high-frequency log-returns (see, Hansen and Lunde, 2006b; Zhou, 1996; Aït-Sahalia et al., 2005).

The consequence of this noise on the realized variance is known (*i.e.*, it is upward/downward biased) but its impact on MIDAS-RV has not been investigated so far. We argue that this noise has an impact on the optimal frequency of the variance predictors when relying on raw data and therefore leads to a loss of information.

To study the impact of microstructure noise on MIDAS-RV models, we first simulate 1-second log-returns using the same approach that the one described in the previous simulation, except that the dynamics of the discrete-time GARCH(1,1) model is at the 1-minute frequency and not 5-minute. To contaminate the log-returns by noise, a normal random variable with mean 0 and variance $10^{-3} \times IntegratedQuarticity$ is added to every 1-second log-return, knowing that the daily integrated quarticity is defined as follows: $IQ_{t,t-1} = \int_{t-1}^{t} \sigma^4(\tau)\mathrm{d}\tau$.[13]

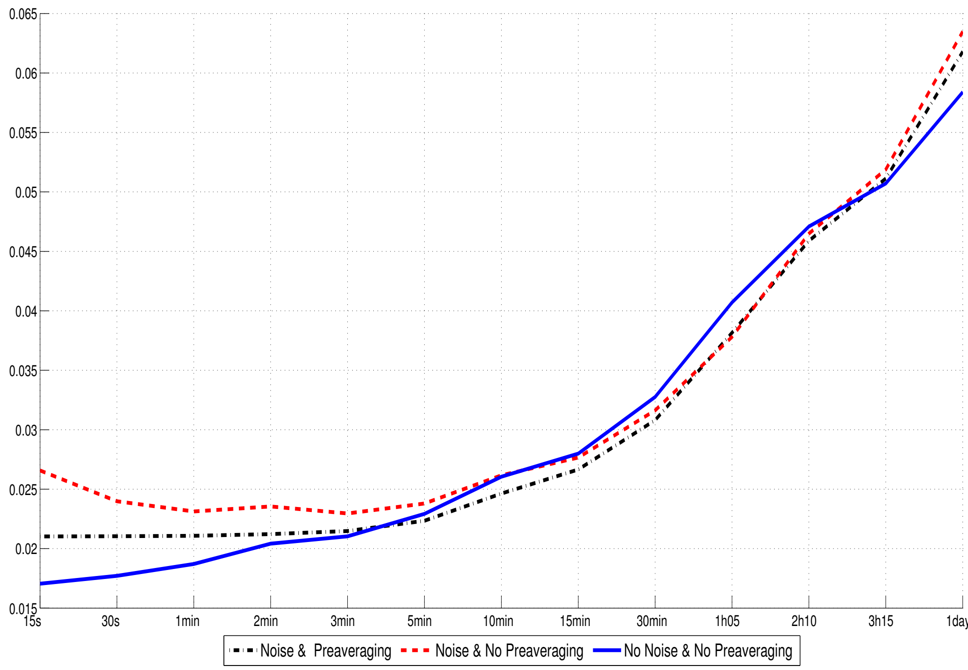
This Monte Carlo experiment is based on 10,000 replications with the regressors and the realized variance simulated for 1,000 days (500 days are used for the purpose of the in-sample estimation and 500 days for the out-of-sample analysis). Figure 4 displays the average QLIKE for 500 out-of-sample one-step-ahead forecasts (over 10,000 replications) as a function of the sampling frequency m_1. Three models are considered. The (blue) solid line corresponds to the case where the MIDAS-RV is estimated on non-contaminated log-returns. The (red) dashed line corresponds to the case where the MIDAS-RV is estimated on contaminated log-returns. Recall that realized variance (the endogenous variable) is calculated on 5-minute returns, a frequency at which the simulated noise is negligible on realized variance.

Results clearly suggest that microstructure noise deteriorates the fit of MIDAS-RV models when using ultra-high-frequency returns. The optimal frequency is between 30 seconds and five minutes but the average QLIKE is about 40% greater than in the case without noise.

[13] Including different noise structures (*i.e.*, i.i.d noise vs. autocorrelated noise) remains an interesting direction for future research.

To account for the presence of microstructure noise in the context of non-parametric volatility estimators, it is standard practice to pre-filter ultra-high-frequency log-returns using the pre-averaging technique introduced by Podolskij et al. (2009) and Jacod et al. (2009). To the best of our knowledge, pre-averaging has never been used in the context of MIDAS models. The (black) dashed line corresponds to the case where the MIDAS-RV is estimated on contaminated but pre-averaged log-returns. Pre-averaging proves to be useful in the context of MIDAS models especially when relying on data sampled at frequencies higher than 15 minutes. Interestingly, for frequencies between 15 seconds and five minutes, the QLIKE of this model is stable and does not blow up, as in the case of the MIDAS-RV model estimated on contaminated log-returns (red dashed line).

Figure 4: MIDAS-RV average QLIKE on 1-minute log-returns



Note: This figure displays the average QLIKE (y-axis) associated with the MIDAS-RV forecasts based on various sampling frequencies (m_1) of the predictors (x-axis). Three different MIDAS specifications are considered: (i) the regressors are not contaminated by noise, (ii) the regressors are contaminated by noise, (iii) the regressors are contaminated by noise but based on pre-averaged returns. In this experiment the variance is assumed to be constant within each 1-minute interval.

4 Application

The main conclusion of Section 3 is that the choice of the optimal sampling frequency m_1 for the predictors in MIDAS-RV models is not obvious. We have seen two cases where the use of the highest available frequency does not necessarily improve the quality of the fit or the predictions. Therefore, a "high-frequency wall" might exist (*i.e.*, a frequency limit above which MIDAS-RV forecasts deteriorate or do not improve). In the Monte-Carlo simulation, only two features of the DGP have been considered to justify the presence of this "high-frequency wall", *i.e.*,

- that the process is not a pure continuous-time model but rather a model where the conditional variance is constant by pieces of for instance one or five minutes;
- and/or the presence of microstructure noise.

It has also been largely documented in the literature that high-frequency log-returns are characterized by the presence of strong intraday periodicity in volatility and jumps. Intraday periodicity (Wood et al., 1985; Harris, 1986; Andersen and Bollerslev, 1997, 1998b; Hecq et al., 2012) can be defined as the cyclical pattern of variance within the trading day, *i.e.*, the fact that variance is typically more important at the opening and closing of the trading day and lower in the middle of the day while jumps correspond to large discontinuities in prices. Unlike intraday periodicity, jumps are not regular (most of the time they appear as the result of an unexpected news arrival) and are known to affect largely variance estimates and forecasts. For more details about the properties and the detection of jumps, see Bates (1996), Barndorff-Nielsen and Shephard (2004b, 2006), Lee and Mykland (2008), Boudt et al. (2011), Lahaye et al. (2011), among many others.

In the application, we propose to investigate the impact of these two additional features of the data on MIDAS-RV models in an application on two highly liquid assets, one exchange-traded fund (ETF) and one quoted share. The use of an ETF is justified by the increasing importance of these assets in the fund management industry.[14]

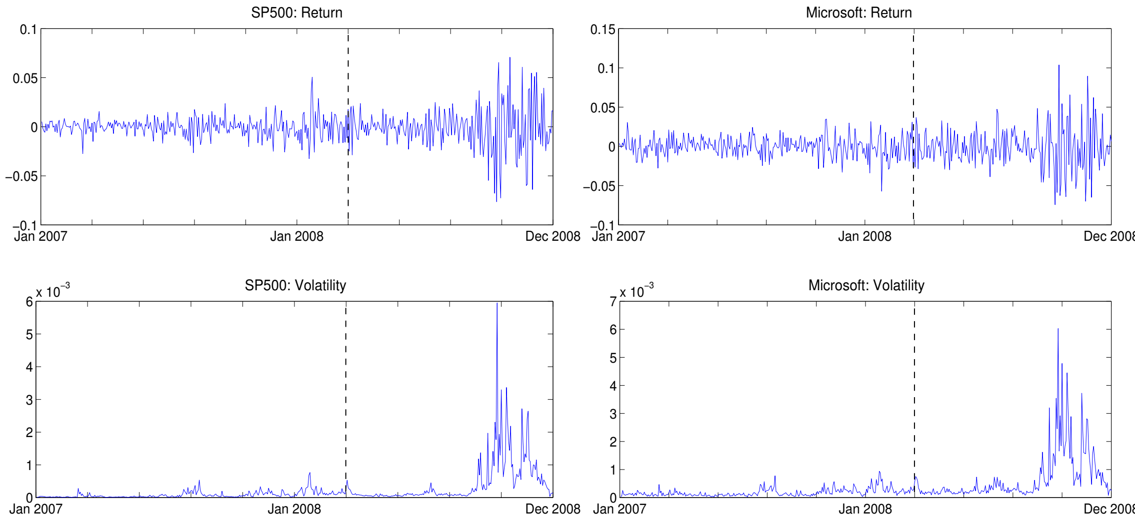
4.1 Data

The dataset consists of tick-by-tick prices and quotations from NYSE Trade and Quote (TAQ) database for Microsoft (MSFT) and one ETF (provided by SPDR ETFs) that tracks the S&P 500 index, spanning the period from September 2, 2004 to December 31, 2008. The price and quote series are reported every trading day from 9:30 am to 4:00 pm and rigorously cleaned using a set of baseline rules proposed by Barndorff-Nielsen et al. (2009). In order to avoid the effect of variance that comes from the overnight or holiday closures, all the variables are computed by using open-to-close data and focusing hence only on the effective trading day variance. The equally spaced intraday returns are subsequently derived from the high-frequency price series. The dataset contains hence 1,101 trading days with 390/78/39/26/13/6/2/1 observations per day of respectively 1-minute/5-minute/10-minute/15-minute/30-minute/1h05/3h15/1-day log-returns.

To compute the variance forecasts, we consider a rolling sample estimation scheme. The parameter estimates are updated every 50 days. For a fair comparison of the MIDAS models, the lag order k^{max} is fixed such that the information used to estimate the parameters covers a period of 70 days, regardless of the sampling frequency

[14]At the end of August 2011, 2 982 ETFs worldwide were managing USD 1 348 bn, which represents 5.6% of the assets in the fund management industry. Additionally, the total ETF turnover on-exchange via the electronic order book was 8.5% of the equity turnover (Fuhr, 2011).

Figure 5: Daily returns and realized variance



Note: The figure reports the daily realized variance and return series for S&P 500 and Microsoft, respectively. The vertical line splits the sample into the relatively calm period of 2007 and the crisis period of 2008.

of the regressors. For instance, for a 5-minute regressor we use a k^{max} equal to 78×70 lags, where 78 represents the number of 5-minute intervals within a trading day.

Finally, the out-of-sample sample covers two years, *i.e.*, 2007 and 2008. To test the robustness of the results upon the state of financial markets, the sample is split into two periods. The first one corresponds to the relatively calm variance period of 2007, and the second one to the financial crisis of 2008 (the end of this period corresponding to the peak of the crisis).

4.2 Optimal sampling frequency for MIDAS-RV on raw data

We first consider one-step-ahead forecasts of MIDAS-RV models estimated on raw data, sampled at different frequencies m_1 ranging between one minute and one day. Three horizons (H) are considered for the endogenous variable $RV_{t+H,t}^{(m_2)}$, *i.e.*, one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$).

Table 3 reports the MCS test for both the calm and crisis periods. For each horizon H, the average QLIKE is reported along with the p-value of the MCS test. The entries in bold correspond to the best models selected by the MCS procedure. The striking result is that the loss function does not smoothly decrease with the sampling frequency and seems to indicate the presence of a "high-frequency wall". In particular, the use of ultra-high-frequency regressors leads to a deterioration in the quality of variance forecasts.

Consider the example of S&P 500 during the calm period (Panel A). The loss function has a convex shape and its minimum is reached for a predictor sampled at five minutes, whatever the forecasting horizon considered. Using 1-minute log-returns leads to a deterioration in the quality of the variance forecasts. This deterioration is statistically significant because MIDAS-RV estimated on 1-minute log-returns does not belong to the MCS set of optimal models. For the crisis period, the MCS test selects the 5-minute frequency as optimal for $H = 1$ and $H = 5$, and 10- and 15-minute frequencies for the two-week horizon. For Microsoft all the models but the one estimated on 1-minute log-returns are found to be statistically equivalent and superior during the calm period (panel A) for $H = 1$.

All in all, these results question the usefulness of ultra-high-frequency data in the context of MIDAS-RV models.

Table 3: MIDAS sampling frequency puzzle

Panel A: Calm period (2007)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value
1min	0.2423	0.0945	0.2161	0.0157	0.2203	0.0663	0.0971	0.0932	0.2357	0.2327	0.1083	0.0095
5min	0.2152	**1.0000**	0.1369	**1.0000**	0.1904	**1.0000**	0.0823	**0.9635**	0.2061	**0.7692**	0.0861	0.1837
10min	0.2203	**0.3562**	0.1412	**0.6192**	0.1968	**0.3828**	0.0878	**0.4825**	0.2013	**1.0000**	0.0950	0.1770
15min	0.2265	0.0945	0.1447	**0.5429**	0.2060	0.0663	0.0881	**0.4825**	0.2245	0.2327	0.0948	0.0985
30min	0.2152	**0.9989**	0.1407	**0.6192**	0.1920	**0.8240**	0.0833	**0.9635**	0.2140	**0.6019**	0.0874	0.1837
1h05	0.2254	**0.3562**	0.1447	**0.6192**	0.2017	**0.3828**	0.0933	**0.4825**	0.2187	**0.6019**	0.0930	0.1837
3h15	0.2713	0.0232	0.1449	**0.6192**	0.2471	0.0663	0.0815	**1.0000**	0.2518	0.2327	0.0732	**1.0000**
1day	0.2672	0.0945	0.1627	**0.5429**	0.2359	0.0663	0.0971	**0.4825**	0.2529	0.2327	0.0943	0.1837

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value
1min	0.2185	0.1704	0.8640	0.1235	0.2337	0.1648	0.7399	0.1036	0.5308	0.0651	0.2231	0.1622
5min	0.2033	**1.0000**	0.2891	0.1998	0.2182	**1.0000**	0.7152	0.1079	0.4144	0.1915	0.1914	**1.0000**
10min	0.2166	0.1704	0.1848	**1.0000**	0.2340	0.1648	0.6791	0.1079	0.3245	**0.4715**	0.2839	0.1160
15min	0.2172	0.1704	0.1960	**0.5976**	0.2306	0.1648	0.1820	**1.0000**	0.3146	**1.0000**	0.9670	0.0218
30min	0.2317	0.0138	0.7694	0.1998	0.2662	0.0406	0.2397	0.1746	0.3408	0.1915	0.2670	0.1160
1h05	0.2428	0.0138	0.2566	0.1998	0.3081	0.0406	0.2958	0.1079	0.3850	0.1915	0.5070	0.1160
3h15	0.2468	0.1704	0.2612	0.1998	0.3437	0.0075	0.2485	0.1746	0.4352	0.0651	0.2213	**0.3685**
1day	0.3305	0.0138	0.2539	0.1998	0.3190	0.0283	0.2719	0.1746	0.3911	0.1915	0.2419	0.1622

Note: This table presents the MCS test results for the S&P 500 and Microsoft. The results are reported for three forecasting horizons, *i.e.*, one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). The QLIKE is reported along with the p-value of the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing models includes eight MIDAS specifications with regressors sampled at a frequency ranging from one minute to one day.

4.3 Breaking the Wall

We have suggested several explanations for the existence of this "high-frequency wall", *i.e.*, an underlying DGP whose conditional variance is constant by pieces of *e.g.*, one or five minutes or the presence of intraday periodicity, jumps and microstructure noise.

While no solution for breaking this wall might exist for the first one, filtering the raw lag-returns might help to improve the performance of MIDAS-RV in presence of intraday periodicity, jumps and microstructure noise. This is precisely the purpose of the this section.

4.3.1 Intraday periodicity

Figure 6 illustrates the intraday periodicity in the variance for the S&P 500 and Microsoft series, by plotting the average squared log-returns for each 1-minute, 5-minute, 30-minute and 1h05 interval, respectively. A clear U-shaped pattern is identifiable, as first noted by Wood et al. (1985), suggesting that the variance is systematically high at the opening, declines to a low point at midday and then increases at the end of the trading day.

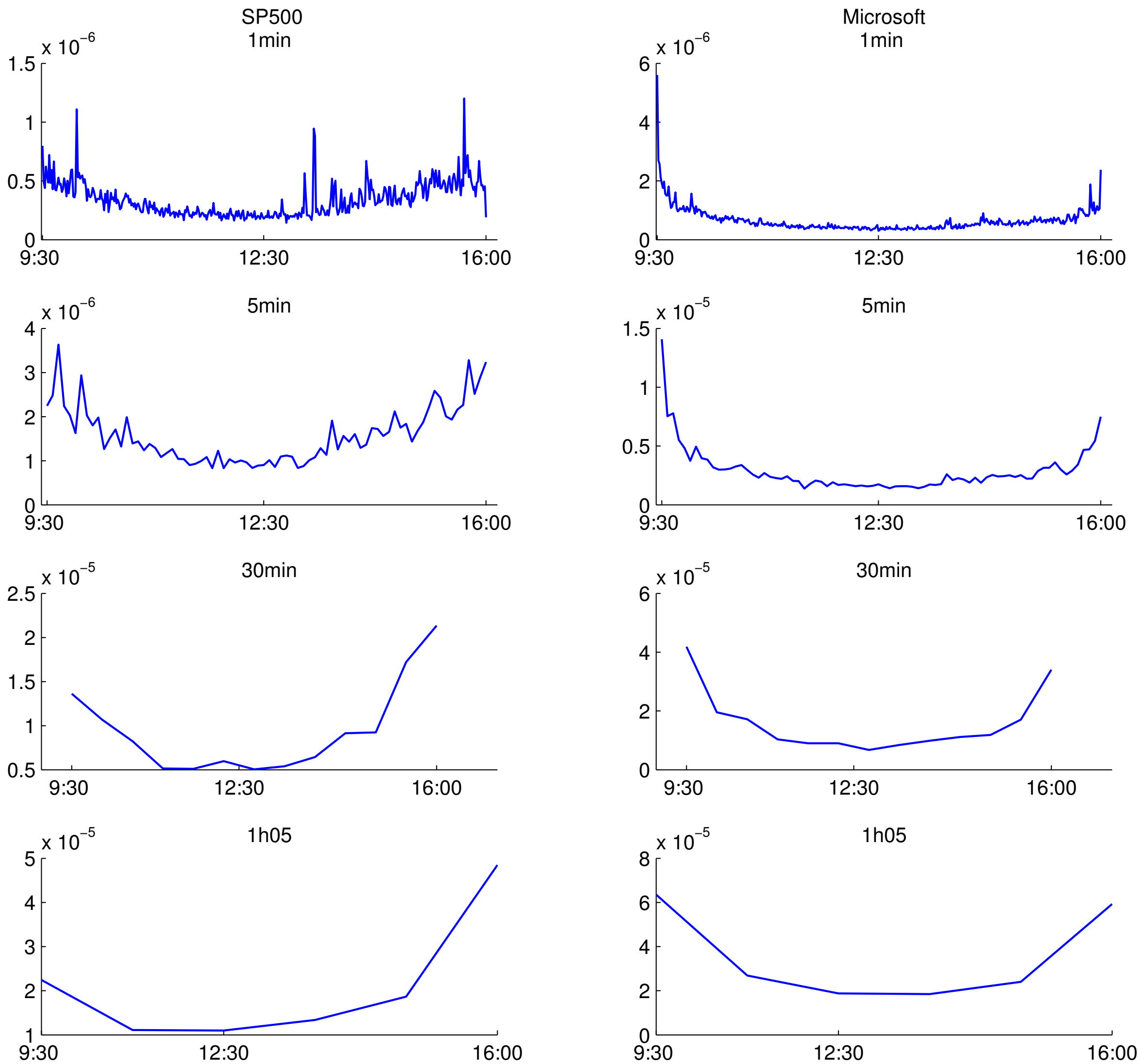
To estimate the intraday periodicity in volatility we rely on the non-parametric weighted standard deviation (WSD) of Boudt et al. (2011), a non-parametric estimator that is robust to additive jumps. If r_i denotes a raw return (sampled at a certain frequency), the corresponding periodicity adjusted return is obtained by dividing r_i by $\hat{f}_i^{\text{WSD}}$, *i.e.*, $r_i/\hat{f}_i^{\text{WSD}}$, where $\hat{f}_i^{\text{WSD}}$ is the estimated WSD of Boudt et al. (2011) for the i^{th} return.

4.3.2 Jumps

To filter out the jumps in the regressors of the MIDAS-RV model, we first apply a modified version of the jump test of Lee and Mykland (2008) proposed by Boudt et al. (2011). More specifically, we assume that the log-price process $\log p(s)$ follows a *Brownian SemiMartingale with Finite Activity Jumps* (BSMFAJ) diffusion $\mathrm{d}p(s) = \mu(s)\mathrm{d}s + \sigma(s)\mathrm{d}w(s) + \kappa(s)\mathrm{d}q(s)$, where $\mu(s)$ is the drift, $\sigma(s)$ is the spot volatility, $w(s)$ is a standard Brownian motion, the occurrence of jumps is governed by a finite activity counting process $q(s)$ and the size of the jumps is given by $\kappa(s)$.

The idea behind the jump test of Lee and Mykland (2008) is that in the absence of jumps, instantaneous returns are increments of Brownian motion and, therefore, standardized returns that are too large to plausibly come from a standard Brownian motion must reflect jumps. In their original paper, Lee and Mykland (2008) standardize every intraday returns r_i by a robust estimate of the spot volatility, denoted $\hat{s}_i$, that assumes that the volatility is constant on a local window spanning between several hours to one or two days before or around the tested return. Their original statistic for jumps is $\mathrm{J}_i = \frac{|r_i|}{\hat{s}_i}$, where $\hat{s}_i$ is the averaged bi-power variation belonging to the local window. To control for the size of the multiple jump tests Lee and Mykland (2008) use the extreme value theory result that the maximum of n i.i.d. realizations of the absolute value of a standard normal random variable is asymptotically (for $n \to \infty$) Gumbel distributed. More specifically, in the absence of jumps, the probability that the maximum of any set of n J-statistics exceeds $g_{n,\alpha} = -\log(-\log(1-\alpha))b_n + c_n$, with $b_n = 1/\sqrt{2\log n}$ and $c_n = (2\log n)^{1/2} - [\log \pi + \log(\log n)]/[2(2\log n)^{1/2}]$, is about α. Lee and Mykland

Figure 6: Intraday periodicity



Note: This figure displays the average squared log-returns for each 1-minute, 5-minute, 30-minute and 1h05 interval, for S&P 500 and Microsoft.

(2008)'s proposal is that all returns for which the J test statistic exceeds this threshold $g_{n,\alpha}$ should be declared to be affected by jumps. In the application, we set $\alpha = 1\%$ and n to the total number of observations in the sample.

However, for such long windows, the assumption of constant volatility is at odds with the overwhelming empirical evidence that the intraday variation in market activity causes intraday volatility to be strongly time-varying and even displays discontinuities (see Figure 6). For this reason, we implemented the modified version proposed by Boudt et al. (2011) that accounts for the presence of intraday periodicity, *i.e.*, $\mathrm{FJ}_i^{\mathrm{WSD}} = \frac{|r_i|}{\hat{f}_i^{\mathrm{WSD}} \hat{s}_i^{\mathrm{WSD}}}$, where $\hat{f}_i^{\mathrm{WSD}}$ is the estimated WSD of Boudt et al. (2011) for the i^{th} return (which is standardized such that its square has mean one in the local window).

Periodicity and jumps adjusted returns are computed as $(r_i/\hat{f}_i^{\text{WSD}}) \times I(\text{FJ}_i^{\text{WSD}} < g_{n,1\%}) + I(\text{FJ}_i^{\text{WSD}} > g_{n,1\%})$, where $I(.)$ is an indicator function.

4.3.3 Microstructure noise

As explained above, pre-averaging (Podolskij et al., 2009; Jacod et al., 2009) is a powerful technique to robustify volatility estimators to the presence of microstructure noise. Instead of noisy intraday returns (r_t), the authors suggest using pre-averaged returns $(\tilde{r}_t)$ which, by the law of large numbers, asymptotically lose the noise component. More precisely, $\tilde{r}_t$ is approximated by an average of staggered returns r_t in a neighborhood of t, the noise being hence averaged away. The pre-averaging approach depends on a bandwidth parameter, or window length, that increases with the sample and indicates the weighting scheme to be put into effect. The order of the window size is chosen to lead to optimal convergence rates $(n^{-1/4})$.

To the best of our knowledge, pre-averaging has never been used in the context of MIDAS models. The balanced pre-averaging has been applied on 1-minute and 5-minute returns previously filtered for intraday periodicity and jumps, since it delivers according to Christensen et al. (2010) the best rate of convergence.

4.3.4 Results

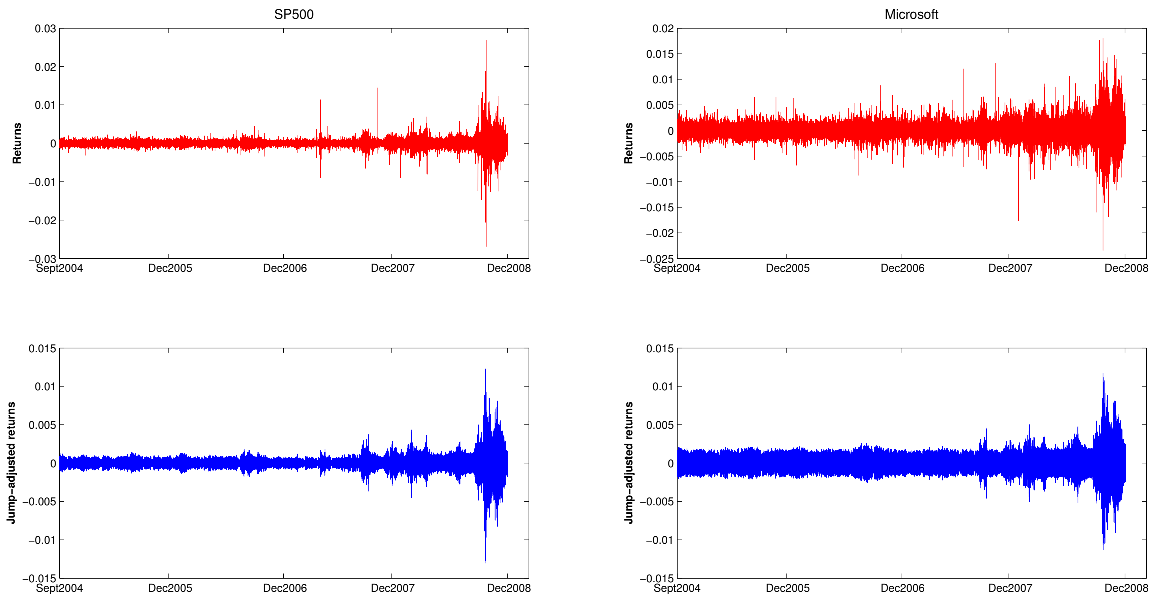
Figure 7 displays the filtered 1-minute return series for S&P 500 and Microsoft, adjusted for intraday periodicity and jumps. The correction procedure purges the intraday periodicity, identifies and smoothes the jumps, but preserves the variance dynamics. The procedure is applied to the 5-minute return series as well.

The MCS test is subsequently applied on the MIDAS out-of-sample obtained with both filtered and unfiltered data. The results are reported in Table 4. We notice a significant improvement in the MIDAS variance forecasts when using high-frequency predictors filtered for intraday periodicity and jumps. For the calm period (panel A), the S&P 500 forecasts obtained with filtered (both for jumps and periodicity) 1-minute predictors always belong to the set of superior forecasting models as identified by MCS. During the crisis period, similar results are obtained with filtered 5-minute regressors for short horizons ($H = 1$ or $H = 5$). These results prove the importance of using the filtered data, especially for short forecasting horizons.

Figures 8 and 9 display the average QLIKE for the calm and crisis periods, and the three forecasting horizons, for both S&P 500 and Microsoft. First, we remark that the gains related to the use of filtered data for intraday periodicity are generally lower than the gains related to data filtered both for periodicity and jumps. Second, considering filtered data during a relatively calm period, we get a loss function which smoothly decreases with the sampling frequency m_1 as in the Monte Carlo experiment.

To complete our analysis, we also perform MIDAS variance forecasts based on regressors filtered for pe-

Figure 7: Intraday returns and intraday jump-adjusted returns



Note: This figure displays the 1-minute intraday return series (in red) as well as the 1-minute return series filtered for intraday periodicity and jumps (in blue). (Lee and Mykland, 2008; Boudt et al., 2011).

riodicity, jumps and microstructure noise (through the pre-averaging technique). The results are available in Appendix 6.1 and Appendix 6.2. For instance, we observe that 1-minute pre-averaged regressors improve Microsoft variance forecasts during both the calm and crisis periods. These results apply also for the S&P 500 variance forecasts at short forecasting horizons. They become more puzzled for long forecasting horizons (*e.g.*, two weeks), as well as for the crisis period.

Table 4: MIDAS sampling frequency puzzle: intraday periodicity and jumps adjustments

Panel A: Calm period (2007)

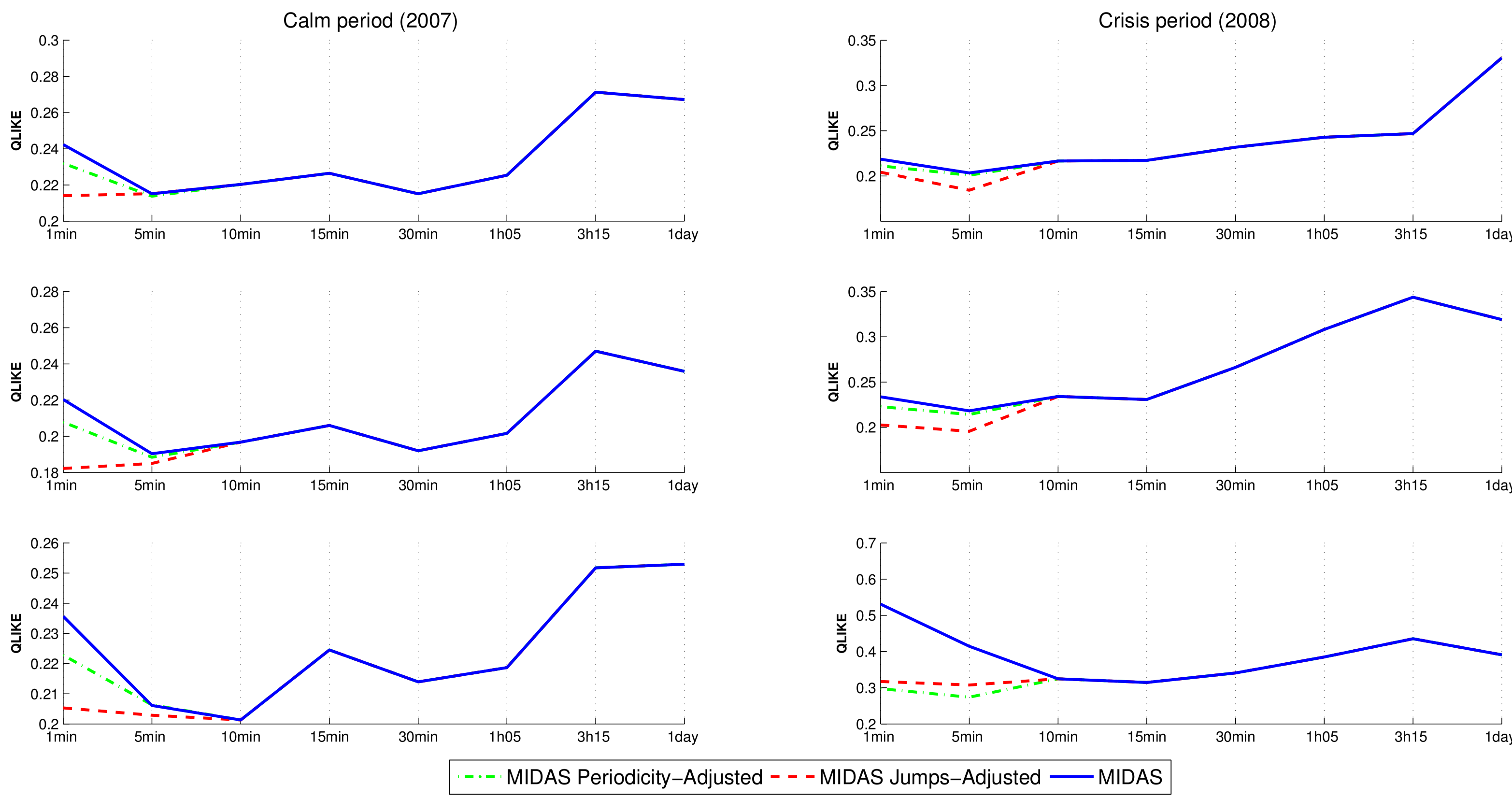
	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2423	0.1403	0.2161	0.0084	0.2203	0.1032	0.0971	0.1282	0.2357	0.1433	0.1083	0.0119
1min Per.Adj	0.2320	0.1403	0.1521	0.0584	0.2078	**0.5087**	0.0993	0.0944	0.2228	**0.6283**	0.1104	0.0105
1min Jumps.Adj	0.2141	**0.9962**	0.1631	0.0084	0.1823	**1.0000**	0.1105	0.0083	0.2053	**0.9284**	0.1196	0.0018
5min RAW	0.2152	**0.9482**	0.1369	**1.0000**	0.1904	**0.7536**	0.0823	**0.9896**	0.2061	**0.9284**	0.0861	**0.2639**
5min Per.Adj	0.2138	**1.0000**	0.1399	**0.7404**	0.1884	**0.8198**	0.0831	**0.9896**	0.2063	**0.9284**	0.0873	**0.2639**
5min Jumps.Adj	0.2152	**0.9962**	0.1380	**0.7762**	0.1850	**0.8198**	0.0830	**0.9896**	0.2029	**0.9324**	0.0888	**0.2639**
10min	0.2203	**0.3497**	0.1412	**0.7404**	0.1968	**0.5087**	0.0878	**0.5991**	0.2013	**1.0000**	0.0950	**0.2542**
15min	0.2265	0.1403	0.1447	**0.6187**	0.2060	0.1032	0.0881	**0.5603**	0.2245	**0.2596**	0.0948	0.1450
30min	0.2152	**0.9962**	0.1407	**0.7413**	0.1920	**0.8198**	0.0833	**0.9896**	0.2140	**0.8609**	0.0874	**0.2639**
1h05	0.2254	**0.3497**	0.1447	**0.7404**	0.2017	**0.5589**	0.0933	**0.5603**	0.2187	**0.6557**	0.0930	**0.2639**
3h15	0.2713	0.0414	0.1449	**0.7404**	0.2471	0.1032	0.0815	**1.0000**	0.2518	0.1433	0.0732	**1.0000**
1day	0.2672	0.1403	0.1627	**0.6187**	0.2359	0.1032	0.0971	**0.5991**	0.2529	**0.2596**	0.0943	**0.2639**

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2185	0.0205	0.8640	0.0189	0.2337	0.0354	0.7399	0.0113	0.5308	0.0944	0.2231	0.0616
1min Per.Adj	0.2112	0.0390	0.1635	0.1167	0.2227	**0.3369**	0.1465	0.1439	0.2975	0.0999	0.1679	**1.0000**
1min Jumps.Adj	0.2042	0.0390	0.1587	**1.0000**	0.2026	**0.4552**	0.1362	**1.0000**	0.3172	0.0999	0.1758	**0.6321**
5min RAW	0.2033	0.0390	0.2891	0.0189	0.2182	**0.3369**	0.7152	0.0113	0.4144	0.0999	0.1914	0.2172
5min Per.Adj	0.2008	0.0390	0.1673	0.1041	0.2142	**0.4552**	0.1534	0.1439	0.2737	**1.0000**	0.4426	0.0145
5min Jumps.Adj	0.1842	**1.0000**	0.1598	**0.8721**	0.1956	**1.0000**	0.1528	0.1439	0.3073	0.0999	0.1735	**0.6321**
10min	0.2166	0.0205	0.1848	0.0189	0.2340	0.0354	0.6791	0.0113	0.3245	0.0999	0.2839	0.0240
15min	0.2172	0.0390	0.1960	0.1041	0.2306	0.0354	0.1820	0.0309	0.3146	0.0999	0.9670	0.0145
30min	0.2317	0.0204	0.7694	0.0189	0.2662	0.0212	0.2397	0.0113	0.3408	0.0999	0.2670	0.0145
1h05	0.2428	0.0204	0.2566	0.0189	0.3081	0.0204	0.2958	0.0113	0.3850	0.0999	0.5070	0.0145
3h15	0.2468	0.0205	0.2612	0.0189	0.3437	0.0037	0.2485	0.0177	0.4352	0.0403	0.2213	0.2172
1day	0.3305	0.0204	0.2539	0.0189	0.3190	0.0120	0.2719	0.0113	0.3911	0.0999	0.2419	0.0616

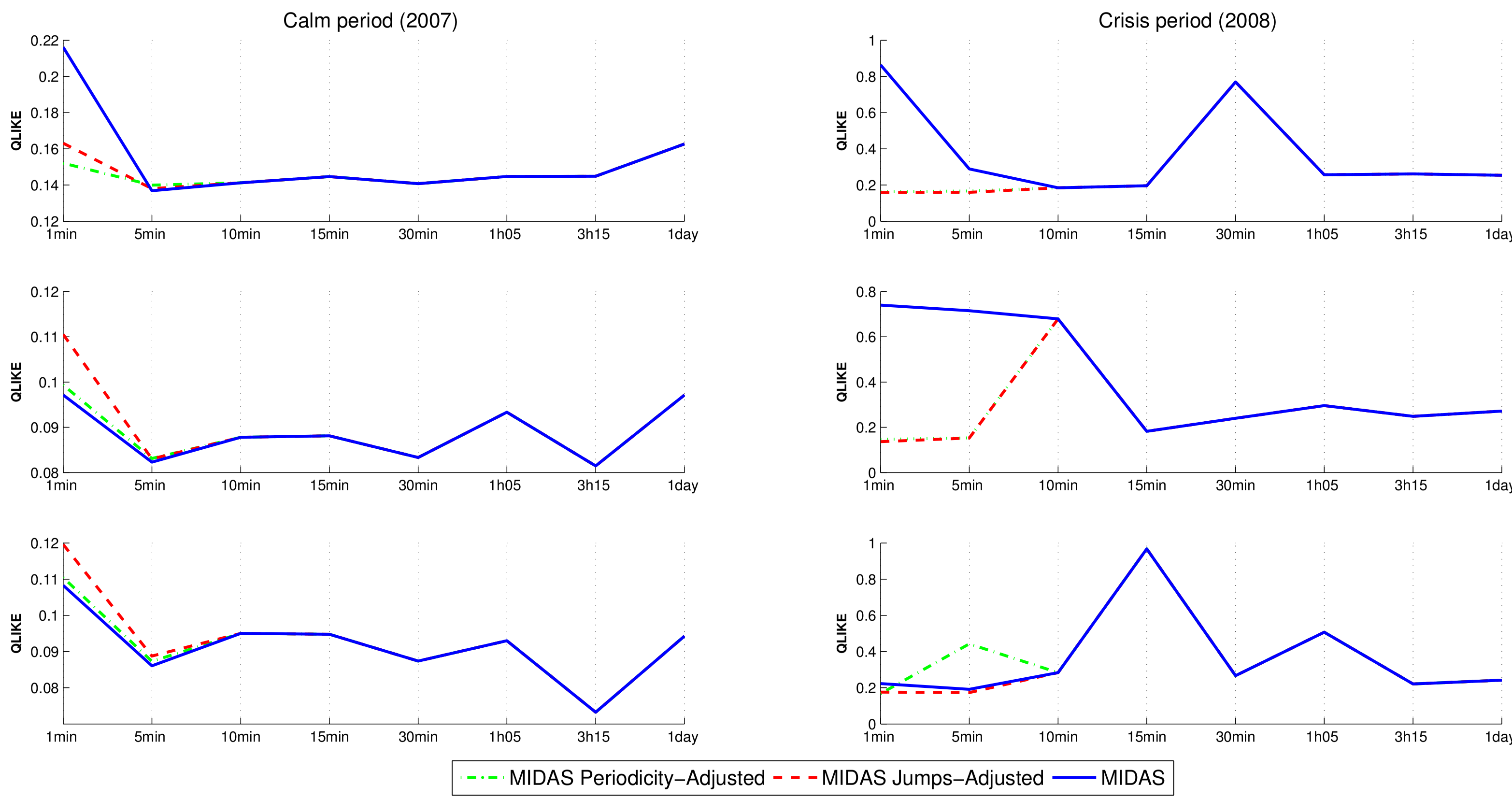
Note: This table presents the MCS test results obtained for two assets (S&P 500 and Microsoft) during both calm and crisis periods. The results are reported for three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). For each of them, we present the average value of the QLIKE loss function along with the corresponding p-value resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing variance models includes eight MIDAS specifications with regressors sampled at a frequency ranging from one minute to one day, as well as four MIDAS models with 1- and 5-minute regressors adjusted for intraday periodicity and jumps.

Figure 8: S&P 500 average QLIKE



Note: S&P 500 - This figure displays the average QLIKE for the MIDAS-RV forecasts for various sampling frequencies (m_1) of the predictors for three forecasting horizons. The left panel corresponds to the calm period (2007) and the right panel to the crisis period (2008). The solid blue line corresponds to the MIDAS-RV model with raw data sampled at frequencies 1-min to 1-day. The black and red dotted lines correspond respectively to the MIDAS-RV models on periodicity and jumps and periodicity filtered log-returns sampled at frequencies 1-min and 5-min.

Figure 9: Microsoft average QLIKE



Note: See Figure 8.

4.4 MIDAS and Other Competing Variance Models

In this section, we compare the predictive accuracy of the MIDAS-RV forecasts with those obtained for four widely used variance models based on daily and/or intradaily data, *i.e.*, the GARCH model, the Generalized Autoregressive Score (GAS), the Heterogeneous Autoregressive Realized Volatility-based model (HAR-RV) and the HAR-RV adjusted for jumps (HAR-RV-J).

i) The first competing model is the popular GARCH(1,1) model, pioneered by Engle (1982) and Bollerslev (1986), *i.e.*:

$$r_{t+1,t} = c + z_{t+1,t}\sqrt{h_{t+1,t}}, \tag{10}$$

$$h_{t+1,t} = \alpha_0 + \alpha_1(r_{t,t-1} - c)^2 + \beta_1 h_{t,t-1}. \tag{11}$$

ii) The second model is the GAS model, recently introduced by Harvey (2013) and Creal et al. (2013). This model is designed to better treat large outliers. We consider the Student GAS specification where the one step-ahead conditional variance is defined as follows:

$$h_{t+1,t} = w_0 + a_1 u_{t,t-1} h_{t,t-1} + \phi_1 h_{t,t-1}, \tag{12}$$

with $u_{t,t-1} = ((v+1)z^2_{t,t-1})/(v-2+z^2_{t,t-1}) - 1$, and $z_t \sim t(0,1,v)$.

Notice that for the GARCH and GAS models, the variance forecasts for $H > 1$ are obtained as $\sum_{i=1}^{H} h_{t+i,t}$ and not directly from $r_{t+H,t}$ as opposed to the MIDAS-RV model.

iii) The third competing model is the HAR-RV model, proposed by Corsi (2009):

$$RV_{t+1,t} = \alpha_0 + \alpha_1 RV_{t,t-1} + \alpha_2 RV^w_{t,t-1} + \alpha_3 RV^m_{t,t-1} + \varepsilon_{t+1}, \tag{13}$$

where $RV_{t+1,t}$ is the daily realized variance (Eq. 3) and by convention, $RV^w_{t,t-1} = \frac{1}{5}\sum_{i=0}^{4} RV_{t-i,t-i-1}$ and $RV^m_{t,t-1} = \frac{1}{22}\sum_{i=0}^{21} RV_{t-i,t-i-1}$. This model is conceived as an additive cascade of different variance components defined over different time horizons of one day, one week (w), and one month (m), respectively. The HAR-RV is therefore a constrained version of the MIDAS-RV model with intradaily squared return regressors and a particular weight structure. Indeed, given the definition of the realized volatility, Eq. (13) can be rewritten as a weighted sum of past observations of the intraday squared returns. For more details, see Appendix 6.3.

iv) Andersen et al. (2007) extended the classical HAR-RV framework by taking into account the lagged effect of jumps. The HAR-RV-J model (where J stands for jumps) is formally defined as following:

$$RV_{t+1,t} = \alpha_0 + \alpha_1 RV_{t,t-1} + \alpha_2 RV^w_{t,t-1} + \alpha_3 RV^m_{t,t-1} + \gamma_1 J_{t,t-1} + \gamma_2 J^w_{t,t-1} + \gamma_3 J^m_{t,t-1} + \varepsilon_{t+1}, \tag{14}$$

where $J_{t,t-1} = I_t \times (RV_{t,t-1} - BV_{t,t-1})$ is a random variable that is nonzero for the intervals in which jumps do occur and zero otherwise, BV is the daily realized bipower variation (Barndorff-Nielsen and Shephard, 2004b) which is defined as:

$$BV_t = \mu_1^{-2} \sum_{l=2}^{m} |r_{t,l}||r_{t,l-1}|, \tag{15}$$

with $\mu_1 = \sqrt{(2/\pi)} \approx 0.79788$, and $I_t \equiv I\left(Z_t > \Phi_{0.999}\right)$, where Z_t is defined as:

$$Z_t = \frac{m^2(RV_{t,t-1} - BV_{t,t-1})RV_{t,t-1}^{-1}}{[(\mu_1^{-4} + 2\mu_1^{-2} - 5)max\{1, TQ_{t,t-1}(m)BV_{t,t-1}^{-2}\}]^{1/2}} \tag{16}$$

with $TQ_{t,t-1}$ the tri-power quarticity, a robust estimator of the integrated variance, and $\Phi_{0.999}$ the 99.9% quantile of the standard normal distribution.[15]

The MCS procedure is now applied on 18 models, namely the eight MIDAS models with regressors sampled between 1-min and one day, the six MIDAS specifications with 1- and 5-minute regressors adjusted for intraday periodicity, jumps and/or microstructure noise, and the four competing variance models (HAR-RV, HAR-RV-J, Student GAS, GARCH). The results are summarized in Table 5. The global conclusion is that the MIDAS models provide (at least for these two assets) comparable, or even better, variance forecasts than the other competing models. In terms of the loss function, the models are dominated during the calm period by the MIDAS models, except for the daily Microsoft forecasts. For the calm period, we find that the forecasts provided by different MIDAS specifications are statistically comparable to those issued from the HAR-RV and HAR-RV-J models. The GAS model provides comparable forecasts only in the case of S&P 500 for a forecasting horizon of two weeks. During the crisis, the best forecasts are generally provided by the MIDAS models with 1- or 5-minute filtered predictors, and the cluster of superior forecasting models no longer includes the HAR-RV and HAR-RV-J models. For longer horizons, the GAS and the GARCH provide similar results to those obtained with the MIDAS model. These findings confirm the intuition that high-frequency data can be used to successfully forecast volatility, provided that these data are filtered for periodicity and jumps. For these two assets, MIDAS models outperform in many cases standard variance models such as the GARCH model, or even the HAR-RV, HAR-RV-J or GAS models.

[15] $TQ_t \equiv m\mu_{4/3}^{-3} \sum_{l=3}^{m} |r_{t,l}|^{4/3}|r_{t,l-1}|^{4/3}|r_{t,l-2}|^{4/3}$, where $\mu_{4/3} \equiv 2^{2/3}\Gamma(7/6)\Gamma(1/2)^{-1}$.

Table 5: Comparing competing variance models

Panel A: Calm period (2007)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2423	0.1617	0.2161	0.0103	0.2203	0.1444	0.0971	0.1726	0.2357	0.1830	0.1083	0.0046
1min Per.Adj	0.2320	0.1617	0.1521	0.0704	0.2078	**0.6375**	0.0993	0.1285	0.2228	**0.6896**	0.1104	0.0039
1min Jumps.Adj	0.2141	**0.9961**	0.1631	0.0103	0.1823	**1.0000**	0.1105	0.0059	0.2053	**0.9262**	0.1196	0.0031
1min Jumps.Adj_Preav	0.2155	0.1617	0.1560	0.0103	0.1900	**0.7983**	0.1000	0.0059	0.4243	0.1830	0.1054	0.0039
5min	0.2152	**0.9773**	0.1369	**0.7738**	0.1904	**0.7983**	0.0823	**0.9888**	0.2061	**0.9262**	0.0861	**0.2702**
5min Per.Adj	0.2138	**1.0000**	0.1399	**0.7738**	0.1884	**0.8186**	0.0831	**0.9888**	0.2063	**0.9262**	0.0873	**0.2702**
5min Jumps.Adj	0.2152	**0.9773**	0.1380	**0.7738**	0.1850	**0.8186**	0.0830	**0.9888**	0.2029	**0.9320**	0.0888	**0.2702**
5min Jumps.Adj_Preav	0.2149	**0.9961**	0.1560	0.0213	0.1899	**0.7983**	0.1001	0.0059	0.4286	0.1830	0.1052	0.0039
10min	0.2203	0.1617	0.1412	**0.7738**	0.1968	**0.6375**	0.0878	**0.5281**	0.2013	**1.0000**	0.0950	0.2100
15min	0.2265	0.1617	0.1447	**0.3101**	0.2060	0.1444	0.0881	**0.5148**	0.2245	**0.3046**	0.0948	0.1429
30min	0.2152	**0.9961**	0.1407	**0.7738**	0.1920	**0.8186**	0.0833	**0.9888**	0.2140	**0.8471**	0.0874	**0.2702**
1h05	0.2254	0.1617	0.1447	**0.6354**	0.2017	**0.7171**	0.0933	**0.5148**	0.2187	**0.6896**	0.0930	**0.2702**
3h15	0.2713	0.0630	0.1449	**0.7738**	0.2471	0.1444	0.0815	**1.0000**	0.2518	0.1830	0.0732	**1.0000**
1day	0.2672	0.1617	0.1627	**0.3101**	0.2359	0.1444	0.0971	**0.5148**	0.2529	0.1830	0.0943	**0.2702**
HAR-RV	0.2176	0.1617	0.1345	1.0000	0.1941	**0.7726**	0.0868	**0.5148**	0.2172	**0.6896**	0.0943	0.2100
HAR-RVjumps	0.2187	0.1617	0.1359	0.7738	0.1973	**0.7171**	0.0883	**0.5148**	0.2226	**0.4406**	0.0960	0.1429
GARCH	0.3240	0.0630	0.2208	0.0027	0.2849	0.1444	0.1677	0.0002	0.2935	0.1830	0.1747	>0.0001
GAS	0.3161	0.0630	0.1884	0.0103	0.2669	0.1444	0.1292	0.0059	0.2688	**0.3046**	0.1296	0.0046

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2185	0.0304	0.8640	0.0265	0.2337	0.0539	0.7399	0.0121	0.5308	0.0315	0.2231	0.0883
1min Per.Adj	0.2112	0.0698	0.1635	0.1823	0.2227	**0.4921**	0.1465	0.1455	0.2975	0.1162	0.1679	**1.0000**
1min Jumps.Adj	0.2042	0.0698	0.1587	**1.0000**	0.2026	**0.6140**	0.1362	**1.0000**	0.3172	0.1162	0.1758	**0.6477**
1min Jumps.Adj_Preav	0.2430	0.0304	0.1698	0.1823	0.2596	0.0258	0.1725	0.0322	0.3517	0.1162	0.1918	**0.4436**
5min	0.2033	0.0698	0.2891	0.0265	0.2182	**0.4921**	0.7152	0.0121	0.4144	0.1162	0.1914	**0.3105**
5min Per.Adj	0.2008	0.0698	0.1673	0.1339	0.2142	**0.6140**	0.1534	0.1455	0.2737	**0.3766**	0.4426	0.0207
5min Jumps.Adj	0.1842	**1.0000**	0.1598	**0.8786**	0.1956	**1.0000**	0.1528	0.1455	0.3073	0.1162	0.1735	**0.6477**
5min Jumps.Adj_Preav	0.2418	0.0304	0.1780	0.0271	0.2627	0.0258	0.1757	0.0239	0.3502	0.1162	0.1943	**0.3105**
10min	0.2166	0.0304	0.1848	0.0265	0.2340	0.0539	0.6791	0.0121	0.3245	0.1162	0.2839	0.0207
15min	0.2172	0.0698	0.1960	0.0271	0.2306	0.0539	0.1820	0.0322	0.3146	0.1162	0.9670	0.0207
30min	0.2317	0.0304	0.7694	0.0265	0.2662	0.0258	0.2397	0.0121	0.3408	0.1162	0.2670	0.0207
1h05	0.2428	0.0304	0.2566	0.0265	0.3081	0.0258	0.2958	0.0121	0.3850	0.0315	0.5070	0.0207
3h15	0.2468	0.0304	0.2612	0.0265	0.3437	0.0058	0.2485	0.0153	0.4352	0.0315	0.2213	**0.3105**
1day	0.3305	0.0304	0.2539	0.0265	0.3190	0.0162	0.2719	0.0121	0.3911	0.0318	0.2419	0.0359
HAR-RV	0.2102	0.0698	0.1781	0.0271	0.2449	0.0539	0.1754	0.0153	0.3563	0.0318	0.2346	0.0207
HAR-RVjumps	0.2148	0.0698	0.1719	0.1339	0.2472	0.0539	0.1669	0.0717	0.3569	0.0315	0.2205	0.0359
GARCH	0.2291	0.0698	0.2205	0.0265	0.2282	**0.4921**	0.2041	0.0153	0.2723	**0.2778**	0.2208	**0.3105**
GAS	0.2363	0.0304	0.2447	0.0265	0.2173	**0.6140**	0.2142	0.0153	0.2348	**1.0000**	0.2226	**0.3105**

Note: This table presents the MCS test results obtained for two assets (S&P 500 and Microsoft) during both calm and crisis periods. The results are reported for three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). For each of them, we present the average value of the QLIKE loss function along with the corresponding p-value resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing variance models includes eight MIDAS specifications with regressors sampled at a frequency ranging from one minute to one day, six MIDAS models with 1- and 5-minute regressors adjusted for intraday periodicity, jumps and/or microstructure noise, the HAR-RV, HAR-RV-J, GARCH and GAS models.

4.5 Robustness Check

In this section we examine the robustness of our results. The most frequent criticism on both MIDAS and HAR-RV models concerns the absence of some constraints ensuring the positivity of the variance process. A straightforward solution consists in predicting the logarithm of the variance proxy, *i.e.*,

Log-MIDAS:

$$\log(RV^{(m_2)}_{t+H,t}) = \mu_{H,m_1} + \phi_{H,m_1}\Omega_{H,m_1}(L^{1/m_1})\log(X^{(m_1)}_{t,t-1/m_1}) + \varepsilon_t. \tag{17}$$

Log-HAR-RV:

$$\log(RV_{t+H,t}) = \alpha_0 + \alpha_1\log(RV_{t,t-1}) + \alpha_2\log(RV^w_{t,t-1}) + \alpha_3\log(RV^m_{t,t-1}) + \varepsilon_{t+1}. \tag{18}$$

Log-HAR-RV-J:

$$\begin{aligned}\log(RV_{t+H,t}) &= \alpha_0 + \alpha_1\log(RV_{t,t-1}) + \alpha_2\log(RV^w_{t,t-1}) + \alpha_3\log(RV^m_{t,t-1}) \\ &\quad +\gamma_1\log(J_{t,t-1}+1) + \gamma_2\log(J^w_{t,t-1}+1) + \gamma_3\log(J^m_{t,t-1}+1) + \varepsilon_{t+1}.\end{aligned} \tag{19}$$

To forecast the logarithm of the realized measure of volatility, we follow exactly the same procedure as for the level of variance. Next, in order to compare the log-variance forecasts with the level of variance proxy, the following transformation is required (Andersen et al., 2003):

$$\widehat{RV'}_{t+H,t} = \exp(\log(\widehat{RV}_{t+H,t}) + \frac{1}{2}Var(e_{t+H,t})), \tag{20}$$

where $\log(\widehat{RV}_{t+H,t})$ is the forecast of the log of realized volatility and $Var(e_{t+H,t})$ is the variance of the forecasting errors.

The results of the MCS-based comparison procedure are reported in Table 6. Once again, during the calm period, the standard models are generally dominated by the log-MIDAS models. For shorter forecasting horizons (one day and one week), the cluster also includes the log-HAR-RV-J model. Only in the case of S&P 500, the GAS model provides statistically comparable forecasts for an horizon of one and two weeks. During the crisis, the daily log-MIDAS-RV with 5-minute regressors pre-filtered for intraday periodicity, jumps and microstructure noise has the smallest QLIKE. For the one-week forecasting horizon the better forecast fit is given by the log-HAR-RV-J model for both Microsoft and S&P 500. The subset of superior forecasting models (as identified by the MCS) encompasses a smaller number of log-MIDAS specifications than in the calm period, the log-HAR-RV-J model (for one day and one week forecasting horizons) and the Student GAS model (for one and two week-ahead S&P 500 forecasts).

Table 6: Log version - MCS Test

Panel A: Calm period (2007)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2406	**0.3927**	0.1553	0.0793	0.2299	**0.4631**	0.0998	0.0862	0.2454	0.0472	0.1136	0.0186
1min Per.Adj	0.2350	**0.8643**	0.1546	0.0793	0.2294	**0.4631**	0.0997	**0.3683**	0.2057	**1.0000**	0.1253	0.0069
1min Jumps.Adj	0.2395	**0.3927**	0.1660	0.0793	0.2118	**0.7967**	0.1262	0.0034	0.2152	**0.9008**	0.1263	0.0011
1min Jumps.Adj_Preav	0.2295	**0.6360**	0.1656	0.0793	0.2174	**0.4805**	0.1126	0.0618	1.0253	0.0014	0.1145	0.0186
5min	0.2251	**0.8643**	0.1403	**0.9546**	0.1987	**0.7967**	0.0862	**0.8161**	0.2149	**0.9219**	0.0915	**0.6301**
5min Per.Adj	0.2252	**0.8643**	0.1419	**0.9142**	0.2051	**0.7967**	0.0913	**0.6182**	0.2178	**0.9008**	0.0924	**0.6301**
5min Jumps.Adj	0.2246	**0.8643**	0.1409	**0.9546**	0.2097	**0.7571**	0.0902	**0.6182**	0.2079	**0.9820**	0.0948	**0.5343**
5min Jumps.Adj_Preav	0.2105	**1.0000**	0.1517	0.0793	0.1931	**1.0000**	0.1022	0.0862	0.2305	**0.9008**	0.1056	0.0246
10min	0.2252	**0.8643**	0.1432	**0.9142**	0.2116	**0.4805**	0.0902	**0.6182**	0.2294	0.0472	0.0905	**0.6301**
15min	0.2324	**0.4502**	0.1451	**0.6697**	0.2180	**0.4631**	0.0893	**0.5495**	0.2289	**0.6145**	0.0906	**0.5343**
30min	0.2198	**0.8643**	0.1369	**1.0000**	0.2158	**0.6218**	0.0811	**0.8161**	0.2118	**0.9820**	0.0834	**0.8619**
1h05	0.2348	**0.5145**	0.1420	**0.9546**	0.2193	**0.4805**	0.0837	**0.8161**	0.2104	**0.9820**	0.0800	**0.8792**
3h15	0.2644	0.0592	0.1508	**0.6697**	0.2498	0.2145	0.0773	**1.0000**	0.2523	0.0472	0.0785	**1.0000**
1day	0.2973	0.0592	0.1805	0.0793	0.2385	**0.3504**	0.1069	**0.4387**	0.2287	**0.9008**	0.0880	**0.8619**
HAR-RV	0.2444	0.1056	0.1528	0.0793	0.2622	0.0124	0.1179	0.0014	0.3308	0.0014	0.1425	0.0007
HAR-RVjumps	0.2205	**0.8643**	0.1386	**0.9546**	0.2161	**0.4805**	0.0922	**0.5495**	0.2615	0.0375	0.1056	0.0186
GARCH	0.3240	0.0592	0.2208	0.0697	0.2849	**0.3504**	0.1677	0.0014	0.2935	0.0472	0.1747	0.0011
GAS	0.3161	0.0592	0.1884	0.0793	0.2669	**0.3504**	0.1292	0.0862	0.2688	**0.0472**	0.1296	0.0186

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.3628	0.0001	0.1881	0.0953	0.5074	0.0024	0.2244	0.0182	0.6579	0.0011	0.3851	0.0047
1min Per.Adj	0.3553	0.0001	0.2575	0.0139	0.5223	0.0024	0.2093	0.0182	0.4240	0.1263	0.4045	0.0039
1min Jumps.Adj	0.3369	0.0001	0.2407	0.0276	0.4799	0.0115	0.3244	0.0182	0.6406	0.0024	0.4123	0.0047
1min Jumps.Adj_Preav	0.2926	0.0048	0.2369	0.0139	0.4875	0.0115	0.3434	0.0182	0.6936	0.0024	0.4136	0.0047
5min	0.2530	0.0048	0.2085	0.0832	0.3928	0.0115	0.2029	0.0182	0.5136	0.0024	0.1682	**1.0000**
5min Per.Adj	0.2544	0.0048	0.1712	0.0953	0.4154	0.0115	0.2683	0.0182	0.5072	0.0024	0.1737	**0.3670**
5min Jumps.Adj	0.2083	0.0295	0.2011	0.0953	0.3872	0.0115	0.1986	0.0182	0.5550	0.0024	0.3325	0.0047
5min Jumps.Adj_Preav	0.1514	**1.0000**	0.1471	**1.0000**	0.2404	**0.5737**	0.2175	0.0182	0.3625	0.1263	0.2851	0.1848
10min	0.2323	0.0295	0.1699	0.0953	0.2952	0.0115	0.1624	0.0634	0.4903	0.0024	0.3197	0.0047
15min	0.1930	0.0295	0.1591	**0.3878**	0.2065	**0.9076**	0.1946	0.0182	0.4679	0.0580	0.1820	**0.2582**
30min	0.1926	0.0295	0.1725	0.0953	0.2992	0.0115	0.2407	0.0182	0.2814	**0.3369**	0.2797	0.1362
1h05	0.1857	0.0295	0.2122	0.0276	0.2854	0.1576	0.2594	0.0182	0.3575	0.1263	0.2628	0.1029
3h15	0.1993	0.0295	0.2173	0.0495	0.2552	0.0115	0.2300	0.0182	0.3809	0.1060	0.2717	0.1810
1day	0.2765	0.0001	0.2273	0.0139	0.2544	0.0645	0.2042	0.0182	0.3243	0.1263	0.2302	0.1848
HAR-RV	0.1748	0.1520	0.1662	0.0953	0.2649	0.0115	0.1821	0.0182	0.4382	0.0024	0.2644	0.0047
HAR-RVjumps	0.1667	0.2427	0.1497	**0.8194**	0.2039	**1.0000**	0.1426	**1.0000**	0.3294	0.1263	0.1965	0.1848
GARCH	0.2291	0.0048	0.2205	0.0953	0.2282	**0.5737**	0.2041	0.0182	0.2723	0.2280	0.2208	0.1848
GAS	0.2363	0.0295	0.2447	0.0276	0.2173	**0.9076**	0.2142	0.0182	0.2348	**1.0000**	0.2226	0.1848

Note: This table presents the MCS test results obtained for the two assets under analysis (S&P 500 and Microsoft) during both calm and crisis periods. The results are reported for three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). For each of them, we present the average value of the QLIKE loss function along with the corresponding p-value resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing variance models includes eight log-MIDAS specifications with regressors sampled at a frequency ranging from one minute to one day, six log-MIDAS models with 1- and 5-minute regressors adjusted for intraday periodicity, jumps and/or microstructure noise, the log-HAR-RV, log-HAR-RV-J, GARCH and GAS models.

Another robustness check exercise consists in changing the measure of variance to be predicted. It is well-documented that the realized variance estimator may become biased and inconsistent in the presence of market microstructure noise. A large number of alternative proxies of variance (*e.g.*, realized bipower variation, realized kernel, etc.) that deal with issues such as jumps and other market microstructure noise, have consequently been introduced by Barndorff-Nielsen and Shephard (2004a), Zhang (2006), Barndorff-Nielsen et al. (2008), Hansen and Horel (2009), *inter alios*. To assess the robustness of our results, we also consider the realized kernel (Barndorff-Nielsen et al., 2008) as dependent variable in our specifications and obtain similar results (Appendix 6.4). Our results are also robust to the choice of the predictors of variance (*e.g.*, intradaily bipower variation instead of squared intradaily returns). For a synthesis of all these extra findings see Appendix 6.5.

The last exercise imagined is meant to highlight the pure effect of sampling frequency. For this, we first estimated a MIDAS-RV model with regressors sampled at one minute and imposed then the weights obtained previously on lower frequency regressors of volatility. These weights would not be optimal, but it is a good way to isolate the differences in predictability attributed to the difference in sampling frequency versus change in optimal weights with respect to the different data sample.[16]

Table 7: Same weights

	ϕ_{H,m_1}		QLIKE	
	S&P 500	MSFT	S&P 500	MSFT
1min	334.4542	332.3547	0.1547	0.1064
5min	73.1739	76.4844	0.1513	0.1124
10min	36.6549	40.2871	0.1600	0.1230
15min	25.5069	27.8444	0.1596	0.1285
30min	12.9734	14.0814	0.1581	0.1457
1h05	5.7102	6.3409	0.1891	0.1908
3h15	1.6805	1.9722	0.2866	0.3689
1day	0.8570	0.8085	0.3401	1.1116

Note: This table presents the estimates for the scale parameter, ϕ_{H,m_1}, as well as the QLIKE loss function for the daily MIDAS-RV model with regressors sampled at frequency ranging from one minute to one day. The weights from the model using 1-minute returns are imposed on lower frequency returns.

Thus, Table 7 presents the estimates for the scale parameter, ϕ_{H,m_1}, as well as the QLIKE loss function for the daily MIDAS-RV model with regressors sampled at frequency ranging from one minute to one day. As previously indicated, the weights from the model using 1-minute regressors are imposed on lower frequency regressors. We observe the decreasing pattern of the parameter estimates for the two assets as in Section 3.2. The same declining shape is also remarked for the QLIKE loss function. The more frequent information we use, the more we gain in precision.

[16]We thank the anonymous referee for this suggestion.

5 Conclusion

This paper analyses the forecasting performance of MIDAS-RV models in which future variances are directly related to past intraday log-returns. These predictors are usually constructed from tick-by-tick data and, consequently, the econometrician needs to choose a sampling frequency. The question we raise is whether ultra-high frequency data is needed to forecast variances.

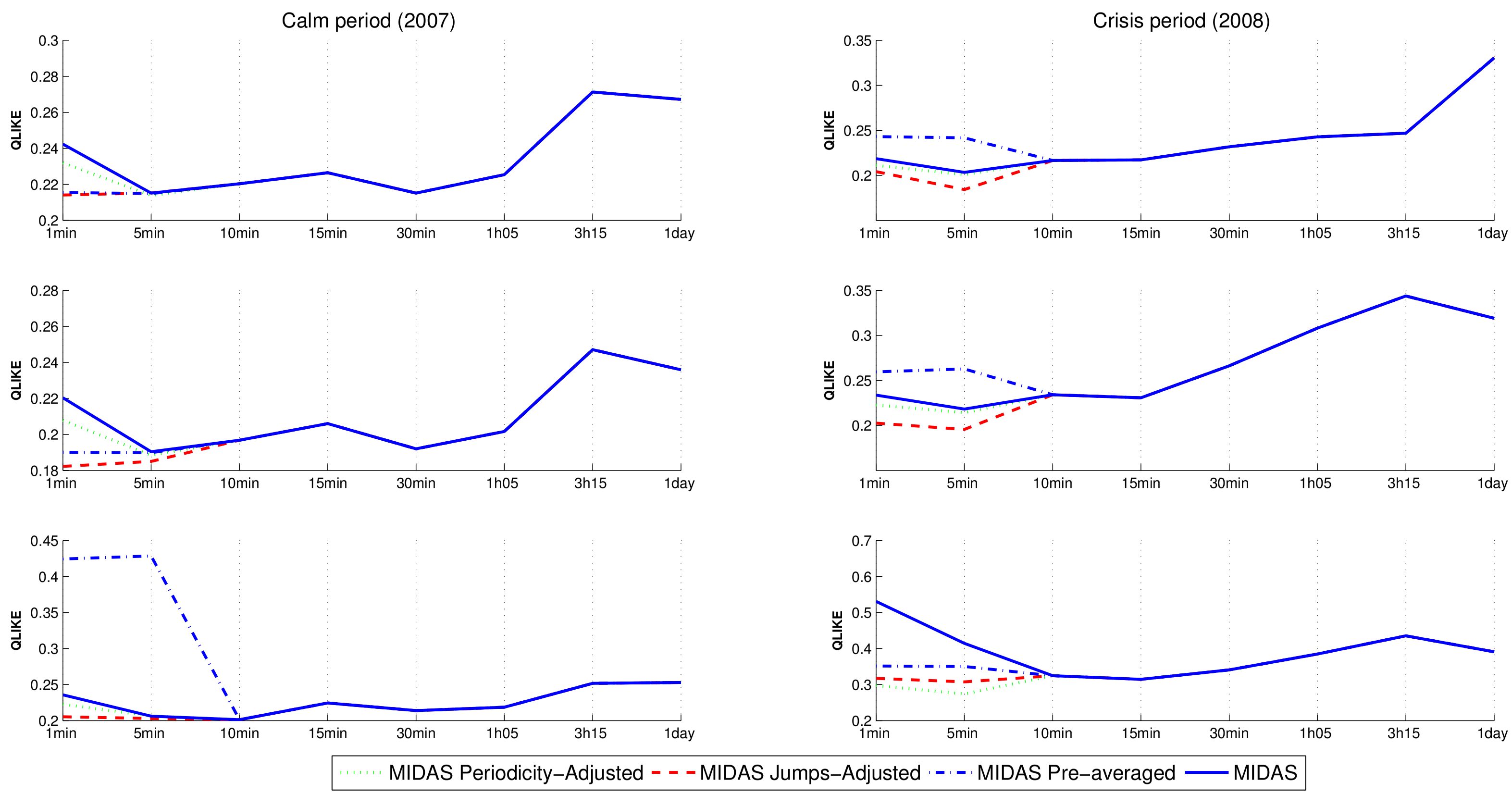
The main findings of our study are the following. First, we show in a Monte Carlo simulation study that, in a world without jumps, periodicity in volatility and microstructure noise, there is an advantage in using the highest available frequency for the predictors. The information content of very high-frequency data improves significantly the quality of the MIDAS forecasts. Second, when considering two highly liquid assets (namely Microsoft and S&P 500) contaminated with typical market microstructure noise and intraday periodicity, we find that the use of very high-frequency predictors may become problematic. In particular, we show that there may exist a "high-frequency wall", *i.e.*, a limit frequency above which the MIDAS forecasts may be less accurate. This result clearly illustrates the influence of the jumps and the intraday periodicity on the *prediction* of volatility, and not only on its *measurement*. Third, we discuss the potential solutions to combine the gains issued from high-frequency predictors and the negative impact of microstructure noise. A first solution consists in augmenting the MIDAS model by modifying the weighting scheme in order to limit the influence of the contaminated observations. A second solution consists in applying the MIDAS regression model on filtered data. Here, we adopt the latter solution and show that estimating MIDAS-RV models on filtered log-returns leads to significantly better out-of-sample forecasts. Finally, we compare the MIDAS model to other competing variance models including GARCH, GAS, HAR-RV and HAR-RV-J models. Results suggest that, for both assets, MIDAS models yield better forecasts in most cases and importantly never yield inferior forecasts, provided they are applied on filtered data.

A future research direction will be to compare the approach taken in this paper, where realized variance is directly related to past intraday data, as in Ghysels et al. (2006), with that of Ghysels et al. (2006) or Ghysels and Sinko (2011), where daily realized measures (that are potentially robust to microstructure noise and jumps) are introduced in a MIDAS-RV model.

6 Appendices

6.1 Appendix A: Pre-averaged MIDAS regressors – S&P 500

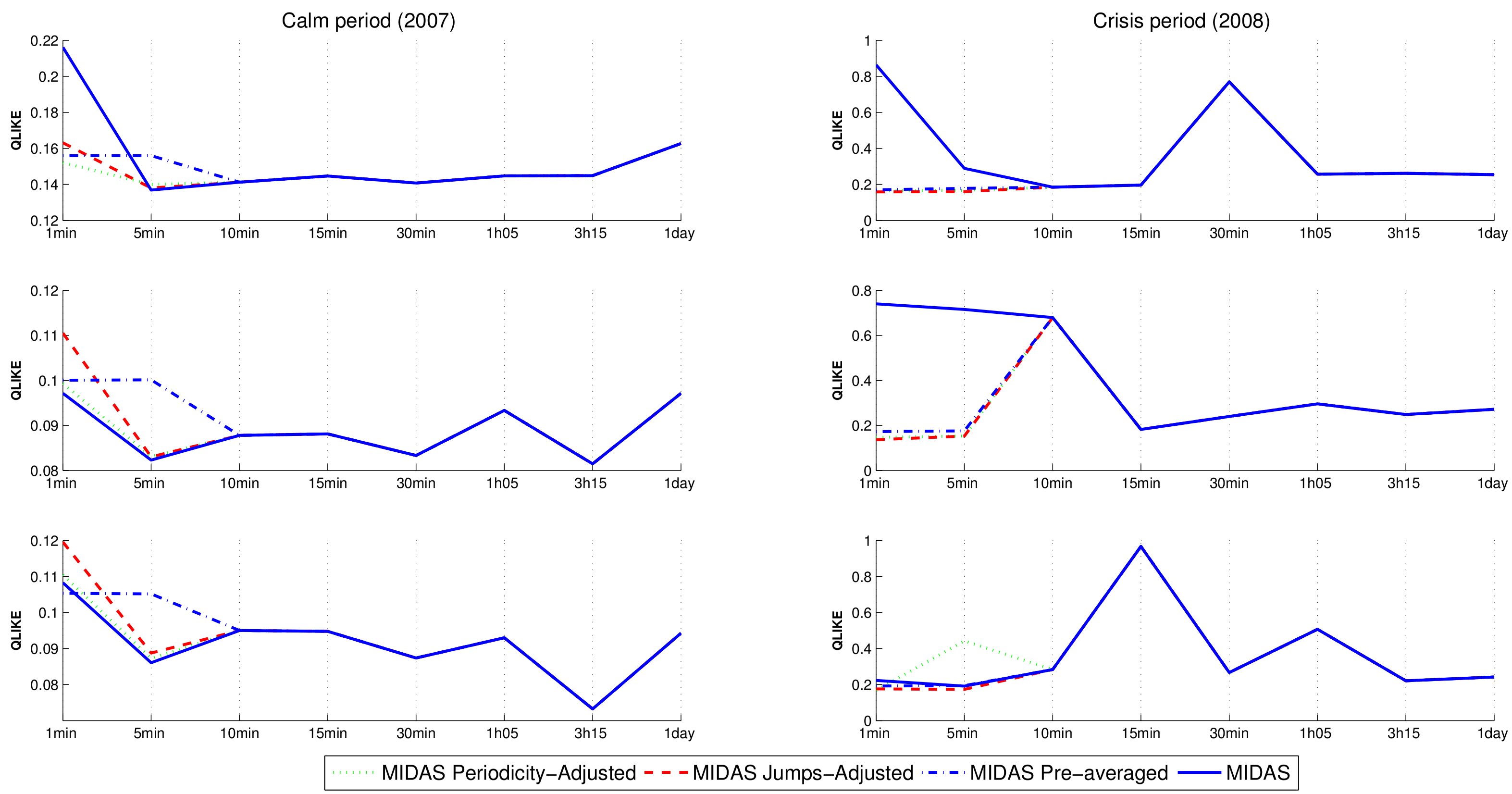
Figure 10: S&P 500 average QLIKE



Note: S&P 500 - This figure displays the average QLIKE for the MIDAS-RV forecasts for various sampling frequencies (m_1) of the predictors for three forecasting horizons.The left panel corresponds to the calm period (2007) and the right panel to the crisis period (2008). The solid blue line corresponds to the MIDAS-RV model with raw data sampled at frequencies 1-min to 3h15. The black, red and blue dotted lines correspond respectively to the MIDAS-RV models on periodicity, jumps and periodicity and noise, jumps and periodicity filtered log-returns sampled at frequencies of 1-min and 5-min.

6.2 Appendix B: Pre-averaged MIDAS regressors – Microsoft

Figure 11: Microsoft average QLIKE



Note: See Figure 10.

6.3 Appendix C: HAR-RV versus MIDAS

In this appendix, we show that the HAR-RV model proposed by Corsi (2009) can be written as a weight-constrained form of the MIDAS model with regressors sampled at a frequency m_2. The HAR-RV model is defined as:

$$RV_{t+1,t}^{(m_2)} = \alpha_0 + \alpha_1 RV_{t,t-1}^{(m_2)} + \alpha_2 RV_{t,t-1}^{(m_2)w} + \alpha_3 RV_{t,t-1}^{(m_2)m} + \varepsilon_{t+1}, \tag{21}$$

where $RV_{t+1,t}$ is the daily realized variance given by:

$$RV_{t+1,t}^{(m_2)} = I_{m_2}(L^{1/m_2}) r_{t+1,t+1-1/m_2}^{(m_2)2}, \tag{22}$$

with $I_{m_2}(L^{1/m_2}) = \sum_{j=0}^{m_2-1} L^{j/m_2}$, and m_2 the sampling frequency of the squared returns used to compute the realized variance. By convention

$$RV_{t,t-1}^{(m_2)w} = \frac{1}{5}\sum_{i=0}^{4} RV_{t-i,t-i-1}^{(m_2)}, \tag{23}$$

and

$$RV_{t,t-1}^{(m_2)m} = \frac{1}{22}\sum_{i=0}^{21} RV_{t-i,t-i-1}^{(m_2)}. \tag{24}$$

To complete the explanation, we include Eq. (22), Eq. (23) and Eq. (24) into the definition of the model and obtain:

$$\begin{aligned} RV_{t+1,t}^{(m_2)} &= \alpha_0 + (\alpha_1 + \tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3) RV_{t,t-1}^{(m_2)} + (\tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3)\textstyle\sum_{i=1}^{4} RV_{t-i,t-i-1}^{(m_2)} + \\ &\quad \tfrac{1}{22}\alpha_3 \textstyle\sum_{i=5}^{21} RV_{t-i,t-i-1}^{(m_2)} + \varepsilon_{t+1} \\ &= \alpha_0 + (\alpha_1 + \tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3) I_{m_2}(L^{1/m_2}) r_{t,t-1/m_2}^{(m_2)2} + (\tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3)\textstyle\sum_{i=1}^{4} I_{m_2}(L^{1/m_2}) r_{t-i,t-i-1/m_2}^{(m_2)2} \\ &\quad + \tfrac{1}{22}\alpha_3 \textstyle\sum_{i=5}^{21} I_{m_2}(L^{1/m_2}) r_{t-i,t-i-1/m_2}^{(m_2)2} + \varepsilon_{t+1}. \end{aligned} \tag{25}$$

Finally, the HAR-RV model takes the form of a daily MIDAS-RV model with squared return regressors sampled at a frequency m_2:

$$\begin{aligned} RV_{t+1,t}^{(m_2)} &= \alpha_0 + (\alpha_1 + \tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3)\textstyle\sum_{j=0}^{m_2-1} L^{j/m_2} r_{t,t-1/m_2}^{(m_2)2} + (\tfrac{1}{5}\alpha_2 + \tfrac{1}{22}\alpha_3)\textstyle\sum_{i=1}^{4}\sum_{j=0}^{m_2-1} L^{j/m_2} r_{t-i,t-i-1/m_2}^{(m_2)2} \\ &\quad + \tfrac{1}{22}\alpha_3 \textstyle\sum_{i=5}^{21}\sum_{j=0}^{m_2-1} L^{j/m_2} r_{t-i,t-i-1/m_2}^{(m_2)2} + \varepsilon_{t+1}. \end{aligned} \tag{26}$$

6.4 Appendix D: MIDAS-RK Specification

Table 8: MIDAS-RK specification

Panel A: Calm period (2007)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value
1min	0.2488	0.1361	0.2008	0.1436	0.2165	0.2038	0.1096	0.0519	0.2306	0.1975	0.1132	0.1338
1min Per.Adj	0.2403	0.1361	0.2029	0.0985	0.2039	**0.5292**	0.1115	0.0519	0.2199	**0.2987**	0.1142	0.1944
1min Jumps.Adj	0.2211	**0.9069**	0.2164	0.0169	0.1738	**1.0000**	0.1244	0.0034	0.1960	**0.7304**	0.1302	**0.2673**
1min Jumps.Adj_Preav	0.2209	0.1361	0.1993	0.1436	0.1792	**0.6843**	0.1087	0.0519	0.4065	0.1975	0.1086	0.2243
5min	0.2204	**0.9069**	0.1841	**0.5802**	0.1836	**0.6843**	0.0914	**0.9359**	0.2044	**0.4371**	0.0903	**0.3099**
5min Per.Adj	0.2192	**1.0000**	0.1886	**0.3606**	0.1810	**0.6843**	0.0929	**0.9359**	0.1998	**0.6074**	0.0900	**0.3534**
5min Jumps.Adj	0.2231	0.1361	0.1870	**0.3606**	0.1771	**0.7382**	0.0929	**0.9359**	0.1944	**1.0000**	0.0921	**0.5883**
5min Jumps.Adj_Preav	0.2200	**0.9375**	0.1993	0.1436	0.1787	**0.7382**	0.1090	0.0455	0.4063	0.1975	0.1084	0.1260
10min	0.2254	0.1361	0.1888	**0.3606**	0.1905	**0.5292**	0.0993	**0.4303**	0.2127	0.1975	0.0980	0.2278
15min	0.2313	0.1361	0.1899	**0.3606**	0.1988	0.2045	0.0967	**0.6397**	0.2194	**0.2987**	0.0953	**0.2709**
30min	0.2228	0.1361	0.1852	**0.3606**	0.1866	**0.6843**	0.0928	**0.9359**	0.2079	**0.6074**	0.0912	**0.4657**
1h05	0.2339	0.1361	0.1900	**0.3606**	0.1946	**0.5724**	0.1027	**0.4303**	0.2115	**0.4371**	0.0963	**0.3038**
3h15	0.2766	0.0837	0.1857	**0.3606**	0.2366	0.2038	0.0883	**1.0000**	0.2417	0.1975	0.0762	**0.3093**
1day	0.2683	0.1361	0.2042	**0.3606**	0.2278	0.2045	0.1051	**0.4303**	0.2439	0.1975	0.0991	0.1999
HAR-RV	0.2242	0.1361	0.1809	**1.0000**	0.1916	**0.5597**	0.0979	**0.5051**	0.2127	**0.2987**	0.0982	**0.2659**
HAR-RVjumps	0.2292	0.1361	0.1837	**0.3606**	0.1953	**0.5292**	0.1007	0.1194	0.2178	0.1975	0.1030	0.0312
GARCH	0.3195	0.0837	0.2203	0.0985	0.2663	0.2038	0.1265	0.0519	0.2708	0.1975	0.1236	0.0090
GAS	0.3102	0.0837	0.2026	**0.3606**	0.2466	0.2038	0.1043	**0.6397**	0.2451	0.1975	0.0954	**0.2823**

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value	QLIKE	p-value
1min	0.2342	0.0130	0.3377	0.0497	0.2458	0.0906	1.1943	0.0035	0.3280	0.0787	0.3076	0.0036
1min Per.Adj	0.2255	0.0269	0.4581	0.0497	0.2353	**0.5226**	0.5386	0.0227	0.3120	0.0897	0.4533	0.0036
1min Jumps.Adj	0.2119	0.0269	0.1886	**0.8972**	0.2155	**0.5555**	0.1502	**1.0000**	0.3344	0.0897	0.1797	**1.0000**
1min Jumps.Adj_Preav	0.3830	0.0094	0.1976	**0.2843**	0.2973	0.0021	0.1858	0.0656	0.4030	0.0787	0.1991	**0.4496**
5min	0.2193	0.0269	0.1976	0.0973	0.2301	**0.5226**	0.1963	0.0656	0.2995	0.0897	0.9667	0.0034
5min Per.Adj	0.2150	0.0269	0.1932	**0.2843**	0.2280	**0.5555**	0.1646	0.1629	0.2873	**0.3441**	0.4717	0.0036
5min Jumps.Adj	0.1805	**1.0000**	0.1878	**1.0000**	0.2065	**1.0000**	0.1649	0.1433	0.3243	0.0897	0.1805	**0.9300**
5min Jumps.Adj_Preav	0.2587	0.0094	0.2083	0.0497	0.3011	0.0019	0.1880	0.0518	0.4033	0.0787	0.2012	**0.3384**
10min	0.2306	0.0130	0.2083	0.0497	0.2423	0.0906	0.3447	0.0035	0.3428	0.0787	0.3525	0.0034
15min	0.2319	0.0130	0.3678	0.0497	0.2448	0.0906	0.7383	0.0035	0.3320	0.0787	0.5074	0.0034
30min	0.2491	0.0094	0.2585	0.0497	0.2887	0.0197	0.2169	0.0227	0.3600	0.0787	0.2417	0.0259
1h05	0.2606	0.0094	0.2300	0.0497	0.3137	0.0368	0.6801	0.0035	0.4043	0.0787	0.3155	0.0034
3h15	0.2635	0.0130	0.2847	0.0497	0.3670	0.0019	0.2551	0.0227	0.4520	0.0351	0.2275	**0.3384**
1day	0.3943	0.0094	0.2848	0.0497	0.3481	0.0021	0.2806	0.0227	0.4085	0.0787	0.2490	0.1867
HAR-RV	0.2355	0.0269	0.2022	0.0973	0.2735	0.0906	0.1890	0.0518	0.3950	0.0383	0.2519	0.0036
HAR-RVjumps	0.2418	0.0130	0.1968	**0.2843**	0.2742	0.0884	0.1818	0.0656	0.3943	0.0351	0.2397	0.0036
GARCH	0.2445	0.0269	0.2285	0.0497	0.2385	**0.5226**	0.1907	0.1433	0.2772	**0.3441**	0.1976	**0.7651**
GAS	0.2533	0.0130	0.2589	0.0497	0.2293	**0.5555**	0.2112	0.0518	0.2426	**1.0000**	0.2126	**0.3384**

Note: This table presents the MCS test results obtained for S&P 500 and Microsoft during both calm and crisis periods. The results are reported for three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). For each of them, we present the average value of the QLIKE loss function along with the corresponding p-value resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing variance models includes eight MIDAS-RK specifications with regressors (squared return) sampled at a frequency ranging from one minute to one day, six MIDAS-RK models with 1- and 5-minute regressors adjusted for intraday periodicity, jumps an/or microstructure noise, the HAR-RV, HAR-RV-J, GARCH and GAS models.

6.5 Appendix F: MIDAS with Bipower Variation Return Regressors

Table 9: MIDAS with bipower variation return regressors

Panel A: Calm period (2007)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2506	0.1766	0.2468	0.0173	0.2298	0.1358	0.0972	0.0728	0.2417	**0.2714**	0.1089	0.0073
1min Per.Adj	0.2391	**0.2366**	0.1534	0.0507	0.2158	**0.6589**	0.0994	0.0728	0.2307	**0.4804**	0.1112	0.0073
1min Jumps.Adj	0.2127	**0.9454**	0.1589	0.0173	0.1816	**1.0000**	0.1047	0.0180	0.2062	**1.0000**	0.1149	0.0016
1min Jumps.Adj_Preav	0.2155	0.2366	0.1560	0.0173	0.1900	**0.8142**	0.1000	0.0180	0.4248	**0.2714**	0.1053	0.0073
5min	0.2089	**1.0000**	0.1363	**0.7362**	0.1890	**0.8142**	0.0803	**0.7513**	0.2095	**0.9505**	0.0860	**0.5805**
5min Per.Adj	0.2097	**0.9454**	0.1379	**0.7362**	0.1887	**0.8142**	0.0796	**1.0000**	0.2086	**0.9505**	0.0849	**0.5805**
5min Jumps.Adj	0.2145	**0.8360**	0.1386	**0.7362**	0.1885	**0.8142**	0.0815	**0.7513**	0.2085	**0.9505**	0.0881	**0.3249**
5min Jumps.Adj_Preav	0.2137	**0.9454**	0.1566	0.0173	0.1891	**0.8142**	0.1003	0.0507	0.4424	**0.2575**	0.1046	0.0191
10min	0.2171	0.2366	0.1397	**0.7362**	0.1987	**0.7734**	0.0888	0.1794	0.2161	**0.6993**	0.0934	**0.3249**
15min	0.2212	0.2366	0.1507	0.0764	0.2016	**0.7734**	0.0928	0.0728	0.2230	**0.6387**	0.0969	0.0191
30min	0.2200	0.2366	0.1499	0.2474	0.1970	**0.8142**	0.0916	0.0728	0.2302	**0.4804**	0.0957	0.1700
1h05	0.2625	0.2366	0.1456	**0.5887**	0.2101	**0.6589**	0.0964	0.0728	0.2229	**0.6993**	0.1037	0.0191
3h15	0.3410	0.0160	0.1502	**0.3013**	0.3065	0.1358	0.0833	**0.7513**	0.2993	**0.2714**	0.0766	**1.0000**
1day	0.3642	0.1766	0.1737	0.0507	0.2769	**0.6447**	0.1063	0.0728	0.4990	0.0720	0.1032	**0.3249**
HAR-RV	0.2176	0.2366	0.1345	**1.0000**	0.1941	**0.8142**	0.0868	0.0740	0.2172	**0.7722**	0.0943	0.1349
HAR-RVjumps	0.2187	0.2366	0.1359	**0.7362**	0.1973	**0.7734**	0.0883	0.0728	0.2226	**0.4804**	0.0960	0.0191
GARCH	0.3240	0.1527	0.2208	0.0173	0.2849	0.1358	0.1677	0.0011	0.2935	**0.2714**	0.1747	0.0002
GAS	0.3161	0.1385	0.1884	0.0173	0.2669	0.1358	0.1292	0.0180	0.2688	**0.2714**	0.1296	0.0160

Panel B: Crisis period (2008)

	H=1				H=5				H=10			
	S&P 500		MSFT		S&P 500		MSFT		S&P 500		MSFT	
	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value	QLIKE	*p*-value
1min	0.2261	0.0318	0.8523	0.0034	0.2428	0.0132	0.7263	0.0029	0.3122	0.1570	1.2599	0.0066
1min Per.Adj	0.2200	0.0318	0.1990	0.1032	0.2330	0.0472	0.5763	0.0205	0.3000	**0.4979**	0.1664	**1.0000**
1min Jumps.Adj	0.2095	0.0318	0.1605	**1.0000**	0.2031	**0.5605**	0.1343	**1.0000**	0.3207	**0.4979**	0.1748	**0.5802**
1min Jumps.Adj_Preav	0.2431	0.0318	0.1717	0.2419	0.2609	0.0132	0.1741	0.0205	0.3524	0.1570	0.1927	**0.3513**
5min	0.4821	0.0318	0.1722	0.1032	0.2314	**0.4355**	0.1586	0.0577	0.8549	0.0389	0.2625	0.0204
5min Per.Adj	0.1858	0.0318	0.1663	0.2419	0.2025	**0.5605**	0.1508	0.1618	0.2866	**0.4979**	0.1702	**0.6597**
5min Jumps.Adj	0.1747	**1.0000**	0.1620	**0.8316**	0.1937	**1.0000**	0.1519	0.1170	0.3097	**0.4979**	0.1721	**0.6200**
5min Jumps.Adj_Preav	0.2465	0.0318	0.1927	0.1032	0.2712	0.0132	0.1848	0.0205	0.3625	0.1570	0.2002	0.1029
10min	0.2233	0.0318	0.2059	0.0034	0.2293	0.1082	0.3507	0.0029	0.3277	0.1570	0.6099	0.0139
15min	0.2139	0.0318	0.1743	0.1032	0.2249	0.1082	0.2972	0.0205	0.5412	0.0389	0.3097	0.0139
30min	0.1937	0.0318	0.2547	0.0034	0.2454	0.0132	0.2603	0.0163	0.8991	0.0389	0.2655	0.0142
1h05	0.2339	0.0318	0.2638	0.0034	0.2985	0.0132	0.3464	0.0029	0.3637	0.1570	1.0577	0.0126
3h15	0.2937	0.0318	0.2880	0.0034	0.2843	0.0132	0.2497	0.0088	0.9281	0.0389	0.2327	0.0204
1day	0.3798	0.0215	0.3876	0.0034	0.4460	0.0132	0.3400	0.0088	0.4760	0.0389	0.3069	0.0142
HAR-RV	0.2102	0.0318	0.1781	0.1032	0.2449	0.0472	0.1754	0.0205	0.3563	0.0389	0.2346	0.0142
HAR-RVjumps	0.2148	0.0318	0.1719	0.2419	0.2472	0.0132	0.1669	0.0525	0.3569	0.0389	0.2205	0.0204
GARCH	0.2291	0.0318	0.2205	0.1032	0.2282	**0.4355**	0.2041	0.0205	0.2723	**0.4979**	0.2208	0.1029
GAS	0.2363	0.0318	0.2447	0.1032	0.2173	**0.5605**	0.2142	0.0205	0.2348	**1.0000**	0.2226	0.1029

Note: This table presents the MCS test results obtained for S&P 500 and Microsoft during both calm and crisis periods. The results are reported for three forecasting horizons, namely one day ($H = 1$), one week ($H = 5$) and two weeks ($H = 10$). For each of them, we present the average value of the QLIKE loss function along with the corresponding p-value resulting from the MCS test. The confidence level for the MCS test is set to $\alpha = 25\%$ and 10,000 bootstrap resamples are used, with block length of five observations, to obtain the distribution under the null of equal predictive accuracy. The set of the competing variance models includes eight MIDAS specifications with regressors (bipower variation) sampled at a frequency ranging from one minute to one day, six MIDAS models with 1- and 5-minute regressors adjusted for intraday periodicity, jumps and/or microstructure noise, the HAR-RV, HAR-RV-J, GARCH and GAS models.

References

Aït-Sahalia, Y., Mancini, L., 2008. Out of sample forecasts of quadratic variation. Journal of Econometrics 147, 17–33.

Aït-Sahalia, Y., Mykland, P. A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial studies 18 (2), 351–416.

Andersen, T. G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T. G., Bollerslev, T., 1998a. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39 (4), 885–905.

Andersen, T. G., Bollerslev, T., 1998b. Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. Journal of Finance 53 (1), 219–265.

Andersen, T. G., Bollerslev, T., Diebold, F. X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. The Review of Economics and Statistics 89 (4), 701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.

Andersen, T. G., Bollerslev, T., Meddahi, N., 2005. Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. Econometrica 73, 279–296.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2008. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76, 1481–536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: Trades and quotes. The Econometrics Journal 12 (3), C1–C32.

Barndorff-Nielsen, O. E., Shephard, N., 2004a. Econometric analysis of realised covariation: High frequency based covariance, regression and correlation in financial economics. Econometrica 72, 885–925.

Barndorff-Nielsen, O. E., Shephard, N., 2004b. Power and bipower variation with stochastic volatility and jumps (with discussion). Journal of Financial Econometrics 2, 1–48.

Barndorff-Nielsen, O. E., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of financial Econometrics 4 (1), 1–30.

Bates, D. S., 1996. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. Review of financial studies 9 (1), 69–107.

Bollerslev, T., 1986. Generalized autoregressive heteroskedasticity. Journal of Econometrics 31, 307–327.

Boudt, K., Croux, C., Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. Journal of Empirical Finance 18 (2), 353–367.

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

Chen, Y.-c., Tsay, W.-J., 2011. Forecasting commodity prices with mixed-frequency data: An OLS-based generalized ADL approach. Working paper.

Christensen, K., Kinnebrock, S., Podolskij, M., 2010. Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. Journal of Econometrics 159 (1), 116–133.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics 7 (2), 174–196.

Creal, D. D., Koopman, S. J., Lucas, A., 2013. Generalized autoregressive score models with applications. Journal of Applied Econometrics 28 (5), 777–795.

Drost, F. C., Nijman, T. E., 1993. Temporal aggregation of GARCH processes. Econometrica: Journal of the Econometric Society, 909–927.

Drost, F. C., Werker, B. J. M., 1996. Closing the garch gap: Continuous time GARCH modeling. Journal of Econometrics 74 (1), 31–57.

Engle, R. F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 45, 987–1007.

Frale, C., Monteforte, L., 2011. FaMIDAS: A mixed frequency factor model with MIDAS structure. Banca d'Italia.

Fuhr, D., 2011. ETF landscape. Industry highlights. BlackRock Report.

Garcia, R., Meddahi, N., 2006. Comment on realized variance and market microstructure noise. Journal of Business & Economic Statistics 24, 184–191.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2004. The MIDAS touch: Mixed data sampling regression models. Working paper.

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

Ghysels, E., Santa-Clara, P., Valkanov, R., 2006. Predicting volatility: getting the most out of return data sampled at different frequencies. Journal of Econometrics 131 (1), 59–95.

Ghysels, E., Sinko, A., 2011. Volatility forecasting and microstructure noise. Journal of Econometrics 160 (1), 257–271.

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

Ghysels, E., Valkanov, R., 2012. Forecasting volatility with MIDAS. In L. Bauwens, C. Hafner, and S.Laurent (Eds.), Handbook of Volatility Models and Their Applications, 383–401.

Hansen, P. R., Horel, G., 2009. Quadratic variation by markov chains. Working paper.

Hansen, P. R., Lunde, A., 2004. Technical appendix: An unbiased measure of realized variance. Brown University, 2004.

Hansen, P. R., Lunde, A., 2006a. Consistent ranking of volatility models. Journal of Econometrics 131, 97–121.

Hansen, P. R., Lunde, A., 2006b. Realized variance and market microstructure noise. Journal of Business & Economic Statistics 24 (2), 127–161.

Hansen, P. R., Lunde, A., Nason, J. M., 2011. The model confidence set. Econometrica 79, 456–497.

Harris, L., 1986. A transaction data study of weekly and intradaily patterns in stock returns. Journal of Financial Economics 16, 99–117.

Harvey, A. C., 2013. Dynamic Models for Volatility and Heavy Tails. Cambridge University Press.

Hecq, A., Laurent, S., Palm, F. C., 2012. Common intraday periodicity. Journal of Financial Econometrics 10 (2), 325–353.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: the pre-averaging approach. Stochastic Processes and their Applications 119 (7), 2249–2276.

Lahaye, J., Laurent, S., Neely, C. J., 2011. Jumps, cojumps and macro announcements. Journal of Applied Econometrics 26 (6), 893–921.

Laurent, S., Rombouts, J. V., Violante, F., 2013. On loss functions and ranking forecasting performances of multivariate volatility models. Journal of Econometrics 173 (1), 1–10.

Lee, S. S., Mykland, P. A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial studies 21 (6), 2535–2563.

Meddahi, N., Renault, E., 1998. Aggregations and marginalization of GARCH and stochastic volatility models. Université de Montréal. Working paper.

Nelson, D. B., 1990. ARCH models as diffusion approximations. Journal of econometrics 45 (1), 7–38.

Patton, A. J., 2011. Volatility forecast comparison using imperfect volatility proxies. Journal of Econometrics 160 (1), 246–256.

Patton, A. J., Sheppard, K., 2009. Evaluating volatility and correlation forecasts. In: Handbook of financial time series. Springer, pp. 801–838.

Podolskij, M., Vetter, M., et al., 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Bernoulli 15 (3), 634–658.

Visser, M. P., 2011. GARCH parameter estimation using high-frequency data. Journal of Financial Econometrics 9 (1), 162–197.

Wood, R. A., McInish, T. H., Ord, J. K., 1985. An investigation of transactions data for NYSE stocks. The Journal of Finance 40 (3), 723–739.

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. Bernoulli 12 (6), 1019–1043.

Zhou, B., 1996. High-frequency data and volatility in foreign-exchange rates. Journal of Business & Economic Statistics 14 (1), 45–52.