

Direct versus Iterated Multi-Period Volatility Forecasts: Why MIDAS is King

ERIC GHYSELS, ALBERTO PLAZZI, ROSSEN VALKANOV,
ANTONIO RUBIA, AND ASAD DOSSANI*

This Draft: January 11, 2019

*Ghysels is with the Department of Economics and the Department of Finance, Kenan-Flagler Business School and Department of Economics, University of North Carolina, Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, CEPR and LFin UCLouvain, Email: eghysels@unc.edu. Plazzi is with the Institute of Finance, Università della Svizzera italiana and SFI, Via Buffi 13, 6900, Lugano, Switzerland, E-mail: alberto.plazzi@usi.ch. Valkanov is with the Rady School of Management, UCSD, Otterson Hall, 9500 Gilman Drive, MC 0093 La Jolla, CA 92093-0093, Email: rvalkanov@ucsd.edu. Rubia is with the University of Alicante, Campus de San Vicente, CP 03080, Spain, E-mail: antonio.rubia@ua.es. Dossani is with the Department of Finance and Real Estate, Colorado State University, Fort Collins, CO 80523, Email: asad.dossani@colostate.edu. We thank Yacine Aït-Sahalia, Andrew Patton, Allan Timmermann, and Mark Watson for helpful discussions of an earlier draft. This paper significantly expands upon an older paper, titled “Multi-Period Forecasts of Volatility: Direct, Iterated, and Mixed-Data Approaches” by Ghysels, Rubia, and Valkanov (2009). All remaining errors are our own.

Abstract

Multi-period-ahead forecasts of returns' variance are used in most areas of applied finance where long horizon measures of risk are necessary. Yet, the major focus in the variance forecasting literature has been on one-period-ahead forecasts. In this paper, we compare several approaches of producing multi-period-ahead forecasts within the GARCH and RV families – iterated, direct, and scaled short-horizon forecasts. We also consider the newer class of mixed data sampling (MIDAS) methods. We carry the comparison on 30 assets, comprising of equity, Treasury, currency, and commodity indices. While the underlying data is available at high-frequency (5-minutes), we are interested at forecasting variances 5, 10, 22, 44, and 66 days ahead. The empirical analysis, which is carried in-sample and out-of-sample with data from 2005 to 2018, yields the following results. For GARCH, iterated GARCH dominates the direct GARCH approach. In the case of RV, the direct RV is preferred to the iterated RV. This dichotomy of results emphasizes the need for an approach that uses the richness of high-frequency data and, at the same time produces a direct forecast of the variance at the desired horizon, without iterating. The MIDAS is such an approach and, unsurprisingly, it yields the most precise forecasts of the variance, in and out-of-sample. More broadly, our study dispels the notion that volatility is not forecastable at long horizons and offers an approach that delivers accurate out-of-sample predictions.

Keywords: Volatility forecasting, multi-period forecasts, mixed-data sampling

JEL: G17, C53, C52, C22

1 Introduction

In the extensive volatility literature, models are most often selected on the basis of their one-period-ahead forecast accuracy. This is true for the traditional ARCH/GARCH models of [Engle \(1982\)](#) and [Bollerslev \(1986\)](#), and the newer realized volatility (RV) approach of [Andersen and Bollerslev \(1998\)](#).¹ Financial decision-making, however, often calls for multi-period forecasts of return volatility, commonly extending from one week to several months into the future, depending on the application at hand (e.g., portfolio allocation, option pricing, risk management, or regulatory supervision) and the asset class of interest (e.g., equities, currencies, commodities, or Treasuries). As the underlying returns data are available at daily or intra-daily frequency, the question arises what is the best approach to construct multi-period volatility forecasts.

In this paper, we undertake a comprehensive empirical examination of multi-period forecasts of asset return volatility. Implicitly, we ask whether there is a “preferred” way to obtain multi-period forecasts of volatility – across asset classes and forecasting horizons? We carry this comparison within the framework of GARCH and RV forecasting models that are extensively used in academic and applied work. These approaches are known to produce precise one-period-ahead volatility forecasts ([Hansen and Lunde \(2005\)](#) and [Andersen et al. \(2003\)](#)). Yet, much less is known about their multi-period-ahead performance. In fact, the common belief is that volatility is difficult to forecast at horizons longer than ten days or so ([West and Cho \(1995\)](#), [Christoffersen and Diebold \(2000\)](#), [Corsi \(2009\)](#)). As we will see below, the perspective that we offer is markedly more optimistic: long-horizon volatility is more forecastable than previously suggested at horizons as long as three months.

There are various ways to approach a multi-period forecasting problem. Two widely used methods are the so-called “direct” and “iterated” approaches ([Marcellino et al. \(2006\)](#)). In the context of volatility forecasting, the direct approach involves estimating volatility at a specific horizon, say, monthly or quarterly frequency, and then using it to form direct predictions of volatility over the next period. The iterative approach is to first estimate a daily autoregressive volatility model and then iterate over the daily forecasts to obtain monthly or quarterly predictions. The direct and iterated approaches can be implemented when volatility is measured within either the GARCH or RV framework.

¹There are exceptions, notably [Diebold et al. \(1998\)](#), [Christoffersen and Diebold \(2000\)](#), and [Andersen et al. \(2001\)](#)

A more recent method to obtain multi-period volatility forecasts is the mixed-data sampling (MIDAS) approach introduced by Ghysels et al. (2005, 2006a). A MIDAS method uses daily volatility estimates to produce directly multi-period volatility forecasts. It can be viewed as a middle ground between the direct and the iterated approaches as it combines the forecasting model and the long-horizon method into one step and offers a convenient way of obtaining multi-period-ahead forecasts of volatility. An advantage of the MIDAS is that one can use a flexible parametric approach to parsimoniously specify the weights on the lagged regressors. Two weight specifications that we will use are the exponential Almon and Beta weights, whose parameters are estimated from the data, as in Ghysels et al. (2005, 2006a). One can also adopt a more restricted version of MIDAS, where the weighting scheme is preset to fit a linear regression framework, as in the HAR model of Corsi (2009). While other parametric specifications have been used, in this paper we focus on the exponential Almon, Beta, and HAR MIDAS specifications.

Yet another, perhaps less satisfactory, method to come up with multi-period forecasts is to scale up the estimated or forecasted daily variance by the horizon of interest. Despite the fact that implicit in the “scaling-up” approach is the strong and empirically untenable assumption that returns are i.i.d., its simplicity is one of the reasons it is popular among practitioners and extensively used in risk management platforms.²

We compare three GARCH methods – direct, iterated, and scaled-up – three RV methods – direct, iterated, and scaled-up – and three MIDAS – exponential Almon, Beta, and HAR – multi-period volatility forecasts. We do so using data of 30 futures contracts of assets from the following classes: US and Japan equity indices, interest rates, currencies, energies, metals, food, grains, and livestock. We have five-minute returns of these contracts and compute realized volatilities, i.e. the sum of intra-daily squared returns. The horizons of interest are weekly (5 days ahead), bi-weekly (10 days ahead), monthly (22 days ahead), bi-monthly (44 days ahead), and quarterly (66 days ahead). To our knowledge, such a study has not been carried out in the volatility literature.³

Comparing various multi-period variance forecasts necessitates a loss function that penalizes deviations of predictions from realizations and a test for predictive accuracy. As a loss function, we choose the maximum likelihood-based QLIKE, following Patton (2011)

²See J.P.Morgan/Reuters (1996) Technical Report (pp. 84) and its implementation in Riskmetrics.

³A few notable exceptions are Diebold et al. (1998), Christoffersen and Diebold (2000), and Andersen et al. (1999) but these studies are more limited in scope. Moreover, they do not consider MIDAS methods which, to preview the results, are particularly suitable for long-horizon volatility forecasting.

who shows that it is a consistent in the context of comparing volatility forecasts.⁴ We use the [Diebold and Mariano \(1995\)](#) and [West \(1996\)](#) test for predictive accuracy to formally gauge the superiority of one forecast over another.⁵

Our study yields the following in-sample and out-of-sample results. First, the iterated GARCH yields more precise in-sample forecasts than the direct GARCH method. For instance, at the 22-day horizon, the iterated GARCH produces the lowest QLIKE in about 83% of the assets. The direct GARCH is superior in only 7% of the assets, while in the remaining 10% the scaled-up method dominates. The same ranking holds true across horizons which leads us to conclude that, when volatility is modeled with a GARCH model, the iterated approach yields the best multi-period forecasts. Our results are consistent with those of [Sheppard \(2018\)](#) who considers various ways of filtering innovations to squared daily returns in the context of iterated and direct GARCH forecasts.

Second, within the class of autoregressive RV models, the direct RV yields the most precise in-sample forecasts. In 77% of the assets, it produces the lowest QLIKE at the 22-day horizon. The superiority of direct RV forecasts holds true all the way up to 66 days. In contrast to the GARCH results, RV iterated performs particularly poorly and has a higher QLIKE than even the scaled-up RV approach. Hence, we find that when autoregressive RV models are used to forecast the variance, the direct approach produces the best multi-period forecasts.

Third, the MIDAS models yield the most precise multi-period volatility forecasts in-sample across all approaches, including GARCH and RV. At the 5-day horizon, a MIDAS approach is the best forecast in 80% of the assets. Even at the 66-days horizon, a MIDAS has the lowest QLIKE in 63% of the cases. Furthermore, MIDAS specifications with flexible weights produce forecasts superior to those with preset weights. The difference

⁴[Patton \(2011\)](#) shows that the QLIKE is comparable and sometimes dominates the mean square forecasting error (MSE), while the mean absolute forecasting error (MAE) and other measures are not consistent.

⁵While the Diebold-Mariano-West test can be generalized and modified to account for parameter uncertainty and uncertainty stemming from a number of implicit choices made by the researcher when formulating a forecast, it remains the most widely used metric for predictive accuracy. The approaches by [Ghysels and Hall \(1990\)](#), [Hoffman and Pagan \(1989\)](#) and [West \(1996\)](#) explicitly address parameter uncertainty. [Giacomini and White \(2006\)](#) propose a test that can be viewed as a generalization, or a conditional version of [West's 1996](#) test. Rather than comparing the difference in average performance, [Giacomini and White \(2006\)](#) consider the conditional expectation of the difference across forecasting models. This conditioning approach allows not only for parameter uncertainty (as in [West \(1996\)](#)) but also uncertainty in a number of implicit choices made by the researcher when formulating a forecast, such as what data to use, the windows of in-sample estimation period, the length of the out-of-sample forecast, among others.

in performance is more noticeable at short horizons, which is surprising as [Corsi \(2009\)](#) proposed his linear model specification in the context of one-day-ahead forecasts. These results emphasize the benefit of using MIDAS with flexible weights, estimated from the data, as in [Ghysels et al. \(2006a\)](#).

Finally, the above results are also confirmed out-of-sample. Namely, iterated GARCH dominates direct GARCH, direct RV dominates iterated RV, and MIDAS with flexible weights produces the best overall out-of-sample forecast. These findings hold across assets at all horizons.

The paper is organized as follows. In section 2, we introduce the direct, iterated, and MIDAS multi-period forecasts. Section 3 discusses the loss function that the [Diebold and Mariano \(1995\)](#) and [West \(1996\)](#) test uses to evaluate the forecasts at various horizons. Section 4 presents the empirical results. In section 5, we conclude by offering directions for further research.

2 Multi-Period Variance Forecasts

The five minute log return for each asset is $r_{i,t-1/m_i,t} = \log(P_{i,t}) - \log(P_{i,t-1/m_i})$ where m_i is the number of times we observe a five-minute return of asset i in day t . For instance, for the future on the S&P500 index we have on average about 268 five-minute returns within a trading day.⁶ We drop the asset-specific subscript i to economize on notation when it is not necessary and write $r_{t-1/m,t}$.

We define daily log returns as $r_t = \log(P_t/P_{t-1})$ for the 30 indices that we consider. From T daily returns, we obtain $T_k = [T/k]$ non-overlapping k -period returns as $R_{\tau_k} = \log(P_t/P_{t-k}) = \sum_{j=0}^{k-1} r_{t-j}$, where $\tau_k = 1, \dots, T_k$. The long-horizon returns are computed without overlap in order to avoid mechanical serial correlation. We focus on multi-day returns sampled at the 5-day (R_{τ_5} , i.e. weekly), 10-day ($R_{\tau_{10}}$, i.e. bi-weekly), 22-day ($R_{\tau_{22}}$, i.e. monthly), 44-day ($R_{\tau_{44}}$, i.e. bi-monthly), and 66-day ($R_{\tau_{66}}$, i.e. quarterly) horizon.

We use increments of quadratic variation, Q_t , as a measure of variance for day t , consistent with the recent volatility forecasting literature ([Andersen and Bollerslev \(1998\)](#)). Since Q_t is not observable, we estimate it with the “realized variance”, defined as the sum

⁶A trading day is considered to begin at 6pm, due to CME operating hours. The average number of five-minute intervals within a day for each index return is specified in Table [A.1](#).

of five-minute squared returns, $\hat{Q}_t = \sum_{j=1}^m [r_{t-(j-1)/m}]^2$. Following the above notation, we compute multi-day realized variances \hat{Q}_{τ_5} , $\hat{Q}_{\tau_{10}}$, $\hat{Q}_{\tau_{22}}$, $\hat{Q}_{\tau_{44}}$, and $\hat{Q}_{\tau_{66}}$ as non-overlapping sums of the daily realized variance \hat{Q}_t .

We consider the following information sets in forming the variance forecasts. The history of daily returns at time t is $I_t^r = \{r_t, r_{t-1}, r_{t-2}, \dots, r_0\}$ and the history of k -period, non-overlapping returns is $I_{\tau_k}^R = \{R_{\tau_k}^k, R_{\tau_k-1}^k, R_{\tau_k-2}^k, \dots, R_0^k\}$. Similarly, we have information sets based on realized variances. The history of daily realized variances is collected in $I_t^{\hat{Q}} = \{\hat{Q}_t, \hat{Q}_{t-1}, \hat{Q}_{t-2}, \dots, \hat{Q}_0\}$ and the history of k -day realized variance is $I_{\tau_k}^{\hat{Q}} = \{\hat{Q}_{\tau_k}, \hat{Q}_{\tau_k-1}, \hat{Q}_{\tau_k-2}, \dots, \hat{Q}_0\}$. Note that we can use the information in I_t^r ($I_t^{\hat{Q}}$) to construct the information in $I_{\tau_k}^R$ ($I_{\tau_k}^{\hat{Q}}$) by summing the returns (realized variances) at the appropriate horizon. As we make it clear below, the data in $I_{\tau_k}^R$ and $I_{\tau_k}^{\hat{Q}}$ is used to form the direct forecasts (section 2.1) whereas I_t^r and $I_t^{\hat{Q}}$ is the basis for the iterated forecasts (section 2.2).

2.1 Direct GARCH and RV Forecasts

The direct GARCH and direct RV approaches that we consider are both autoregressive in nature and produce a one-period-ahead forecast of the k -period return variance. They use the data in the coarser information sets, $I_{\tau_k}^R$ and $I_{\tau_k}^{\hat{Q}}$. The main difference between them is that GARCH is constructed with k -period returns, while RV necessitates k -period realized variances.

Direct GARCH (GARCH-D): The direct GARCH is perhaps the simplest to implement. It involves estimating a GARCH(p, q) model of k -period returns in $I_{\tau_k}^R$, and then constructing a one-step-ahead forecast of k -period return variance. We denote this forecast by

$$V_{T_k+1|T_k}^{GARCH-D} = F(I_{\tau_k}^R; \theta^{GARCH-D},)$$

where $\theta^{GARCH-D}$ are the parameters of the model. GARCH models are simple, parsimonious and are known to produce accurate short-horizon forecasts (Hansen and Lunde (2005)). In a multi-period-ahead setting, the direct approach is likely to produce robust estimates which do not display a bias. However, given that the sample of non-overlapping k -period returns is small, they are subject to considerable estimation error.

In our empirical exercise, we fit a GARCH(1,1) model on the demeaned k -period returns. We also have results from more general GARCH(p,q) models, where p and q are chosen by either the Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC). However, the AIC- and BIC-chosen models fail to beat the GARCH(1,1) out-of-sample. This finding confirms that the [Hansen and Lunde \(2005\)](#) results hold at horizons longer than one-period ahead. Hence, we use the GARCH(1,1) in our analysis.

Direct RV (RV-D): We use the time series of k -day realized variance, $I_{T_k}^{\hat{Q}}$, to estimate a simple AR(p) model. We then construct a one-step ahead forecast of the k -period return variance, which is denoted

$$V_{T_{k+1}|T_k}^{RV-D} = F(I_{T_k}^{\hat{Q}}; \theta^{RV-D}),$$

where θ^{RV-D} are the parameters of the autoregressive model. To keep with the spirit of simplicity and robustness, we use an AR(1) model in the empirical implementation.⁷

2.2 Iterated GARCH and RV Forecasts

In the iterated forecasts, we use daily returns or realized variance in the finer information sets, I_t^r and $I_t^{\hat{Q}}$, to formulate iterated forecasts of the daily variance k -periods forward which are then aggregated to the horizon of interest. Since we are using daily data to estimate the underlying forecasting model, parameter uncertainty will be less of an issue than in the direct method, which would in turn increase forecast precision. However, since we are iterating the forecasts and summing them, model mis-specification errors might increase with the horizon. Hence, in general this method is thought of being bias-prone.

Iterated GARCH (GARCH-I): We start off with daily returns, which are then demeaned, and used to estimate a GARCH(1,1) model. The estimates of the daily GARCH are then iterated to produce daily k -periods forward forecasts, $V_{T+j|T}^{GARCH-I}$, for $j = 1, \dots, k$. We then construct the variance forecast over k periods as

$$V_{T+k|T}^{GARCH-I} = \sum_{j=1}^k V_{T+j|T}^{GARCH-I}.$$

⁷Results from a more general ARMA(p,q) model yield comparable in-sample and somewhat worse out-of-sample results.

Note that this forecast uses information at time T and the forecasts for days $T + 1$, $T + 2$, ..., $T + k$ would have to be iterated from the one-period daily forecasting model.⁸

Iterated RV (RV-I): We fit an AR(1) model on the daily quadratic variance estimates in $I_t^{\hat{Q}}$.⁹ The estimated autoregressive model is iterated forward to produce daily k -periods forward forecasts of the quadratic variance, $V_{T+j|T}^{RV-I}$, for $j = 1, \dots, k$. Similarly to the GARCH approach, we then sum the variance over k periods as

$$V_{T+k|T}^{RV-I} = \sum_{j=1}^k V_{T+j|T}^{RV-I}.$$

2.3 Scaled GARCH and RV Forecasts

We also consider a scaled version of the GARCH and RV approaches. Scaling involves estimating a daily volatility forecasting model up to time T , using it to form a one-period forecast and scaling it by the horizon k . This forecasting method assumes that log returns are i.i.d. Christoffersen and Diebold (2000) and Diebold et al. (1998), among others, document that this method is not appropriate in forecasting long-horizon volatility. However, its prominence in applied work is undisputed, largely due to J.P.Morgan/Reuters's 1996 widely adopted Riskmetrics approach and the suggestions in the Basel agreements.

Scaled GARCH (GARCH-S): We start from the GARCH(1,1) estimates of the GARCH-I model. However, instead of iterating forward the forecasts, we simply construct the variance forecast over k periods as

$$V_{T+k|T}^{GARCH-S} = k \times V_{T+1|T}^{GARCH-I}.$$

Scaled RV (RV-I): In a similar fashion, starting from the AR(1) estimates of the RV-I model, we construct the variance forecast over k periods as the scaled one-period forecast, or:

$$V_{T+k|T}^{RV-S} = k \times V_{T+1|T}^{RV-I}.$$

⁸As before, we find that a more general daily GARCH(p,q) model produces similar or even worse in- and out-of-sample results.

⁹We again experimented with more flexible ARMA(p,q) models but the AR(1) model provides a good balance between in-sample fit and out-of-sample performance.

2.4 MIDAS Volatility Forecasts

The idea behind the mixed-data sampling (MIDAS) approach is to use daily data and directly produce a multi-step ahead forecast. In the volatility forecasting context, the MIDAS predictors are daily quadratic variances in $I_t^{\widehat{Q}}$ while the quantity to predict is the k -period variance, \widehat{Q}_{τ_k} . Since the MIDAS approach is relatively new, we describe it more in detail. We start by formulating a MIDAS forecasting regression:

$$\widehat{Q}_{\tau_k} = \mu_k + \phi_k \sum_{j=0}^{j^{max}} b_k(j, \theta_k) \widehat{Q}_{t-j} + \varepsilon_{k,t} \quad (1)$$

where $b_k(j, \theta_k)$ is a parsimonious weighting function parameterized by a low-dimensional parameter vector θ_k , and j^{max} is a truncation value. The regression involves data sampled at different frequencies, since the realized volatility in equation (1) is measured at horizons ranging from five to 66 days, whereas the regressors are lags of the daily realized variances, \widehat{Q}_{t-j} . To make things more concrete, let's assume we are interested in forecasts at the monthly horizon. Then equation (1) relates the realized variance over, say, the month of December, 2010 with past daily realized variances up to the last day of November, 2010. The weights placed on the predictive lagged realized variances are estimated in-sample, and can be used to form a pseudo out-of-sample forecast.

The intercept μ_k , slope ϕ_k , and weighting scheme parameters θ_k are estimated with non-linear least squares. As noted before, the lag coefficients $b_k(j, \theta_k)$ are parameterized to be a low-dimensional function of underlying parameters θ_k . Without this parametric restriction, the number of parameters associated with the forecasters would proliferate significantly, leading to in-sample over-fit and poor out-of-sample forecasts. A suitable parameterization of $b_k(j, \theta_k)$ circumvents the problem of parameter proliferation, and is one of the most important ingredients in a MIDAS regression.

The weights $b_k(j, \theta_k)$ are scaled to add up to unity. This restriction allows us to estimate the parameter ϕ_k , which captures the overall predictability. In a volatility forecasting context, it is useful to ensure that $b_k(j, \theta_k)$ are non-negative. This second restriction guarantees that the volatility forecasts themselves are non-negative. We consider three

parameterizations of $b_k(j, \theta_k)$ that meet these restrictions. They can be written as

$$b_k(j, \theta_k) = \frac{f(j, \theta_k)}{\sum_{p=1}^{j^{max}} f(p, \theta_k)}, \quad (2)$$

where $f(j, \theta_k)$ is a positive function. Note that the parameters of the weights are horizon-specific, which is crucial as the volatility at different periods might have different dynamics.

We consider three specifications of $f(j, \theta_k)$ – the “Beta lag”, the “exponential Almon lag”, and the HAR weights – which have been suggested in prior literature (Ghysels et al. (2005), Ghysels et al. (2006a), and Corsi (2009)). Appendix A discusses the three specifications in detail. In the exponential and Beta lag, there are two parameters that can be estimated from the data, or $\theta_k = [\theta_{k,1}, \theta_{k,2}]$.¹⁰ The estimation of the MIDAS parameters is straightforward, either with non-linear least squares or quasi-maximum likelihood.

The HAR specification features three parameters and imposes a step-wise function on the daily lagged realized volatility (Corsi (2009)). There are two important differences between the former approaches and HAR weights. First, HAR weights are imposed a priori, rather than being estimated from the data. Corsi (2009) suggests a “one day, one week, one month” step function and argues that it works well empirically to forecast the volatility of stock returns. Another crucial difference is that the HAR weights are not horizon specific. The same lag structure is imposed on volatility forecasts at 5-day and 44-day horizons, which is clearly a limitation. By contrast, the parameters θ_k of the Beta lag and exponential Almon lag are estimated for each horizon and there is no a priori reason for them to be equal across horizons. Hence, the weights on the lags in (2) will also differ across horizons. However, the HAR is simple to estimate as it is a linear regression model.

To sum up, we consider three *MIDAS* weighting functions when constructing our variance forecasts which make use of the information set in $I_t^{\hat{Q}}$:

MIDAS Beta (MIDAS-B): $V_{T_k+1|T_k}^{MIDAS-B} = MIDAS(I_t^{\hat{Q}}; \theta^{MIDAS-B})$

MIDAS Exponential Almond (MIDAS-E): $V_{T_k+1|T_k}^{MIDAS-E} = MIDAS(I_t^{\hat{Q}}; \theta^{MIDAS-E})$

MIDAS HAR (MIDAS-H): $V_{T_k+1|T_k}^{MIDAS-H} = MIDAS(I_t^{\hat{Q}}; \theta^{MIDAS-H})$

We can think of the mixed-data regression (1) as combining the attractive features of

¹⁰The parameters are estimated without restrictions. For the Beta-lag specification, Ghysels et al. (2005) note that the constraint $\theta_{k,1} = 1$ and $\theta_{k,2} > 1$ insures a slowly decaying pattern of the weights and further reduces the dimension of the parameter space.

the iterated and direct forecasts. Notice that we can vary the forecast horizon by changing k , whereas the predictive variables remain the same and allow us to explore the richer information set $I_t^{\widehat{Q}}$. This is not true for the direct approach, where the predictive variables change with the horizon and forecasts are formed using information set $I_{\tau_k}^R$ or $I_{\tau_k}^{\widehat{Q}}$. In the MIDAS forecasts, it is not the regressors that change but the estimated shape of the lag function b_k , thus changing the weights placed on the predictive variables. Moreover, we form direct forecasts of future volatility at horizon k without iterating over forecasts. This is in contrast with the iterated approach. Therefore, the MIDAS approach allows us to side-step the iteration and aggregation issues associated with iterated forecasts as well as the inefficient use of lagged information that is characteristic of the direct approach.

While the MIDAS approach to formulate forecasts is quite general, we focused on regression (1) for several reasons. First, we could have extended the number of regression to include not only daily realized variances but other volatility forecasts such as daily squared returns, daily absolute returns, daily range measures (high-low), and others, as done in [Forsberg and Ghysels \(2006\)](#) and [Ghysels et al. \(2006a\)](#). We could have also used 5-minute squared or absolute returns from the raw data, as in [Ghysels et al. \(2006a\)](#). We settled on the daily realized variances as MIDAS predictors in order to make this forecast as directly comparable as possible with the RV forecasts. Moreover, the comparison of the squared daily return predictors to other predictors at shorter horizons have already been investigated extensively in [Forsberg and Ghysels \(2006\)](#) and [Ghysels et al. \(2006a\)](#). MIDAS regressions typically do *not* exploit an autoregressive scheme, so that \widehat{Q}_t is not necessarily related to lags of the left hand side variable. Instead, MIDAS regressions are first and foremost regressions and therefore the selection of the right hand side variables amounts to choosing the best predictor of future volatility from the set of several possible measures of past fluctuations in returns. In other words, MIDAS is a reduced-form forecasting device rather than a model of conditional volatility.

It should also be noted that there are other successful MIDAS regression specifications. One recent example is [Bollerslev et al. \(2018\)](#) who propose a MIDAS type approach which exploits similarities in volatilities among more than fifty commodities, currencies, equity indices, and fixed-income instruments spanning more than two decades. What sets apart their approach is that their MIDAS type regressions exploit these similarities through panel-based joint estimation. In this paper we do not use panel data regression methods that take advantage of cross-asset similarities.

3 Comparing the Forecasts

For each horizon k , we consider nine variance forecasts: $V_{T_k+1|T_k}^{GARCH-D}$, $V_{T_k+1|T_k}^{RV-D}$, $V_{T+k|T}^{GARCH-I}$, $V_{T+k|T}^{RV-I}$, $V_{T+k|T}^{GARCH-S}$, $V_{T+k|T}^{RV-S}$, $V_{T_k+1|T_k}^{MIDAS-B}$, $V_{T_k+1|T_k}^{MIDAS-E}$, and $V_{T_k+1|T_k}^{MIDAS-H}$. Hence, we need a method of ranking them and choosing the “best” one. To do that, we address three related issues. First, since the true k -period volatility that we are forecasting is unobservable, we will need to proxy for it. Second, in evaluating the forecasts, we have to agree on an appropriate loss function. Given the first issue, we require a loss function that produces robust rankings of the forecasts even in the absence of true volatility. Finally, we need to test whether two forecasts are statistically different from each other. We address these issues below.

3.1 Proxy for Unobservable Long-Horizon Volatility

As the true volatility is unobservable, we use the daily realized variance \widehat{Q}_t from 5-minute returns as a proxy for the population variance.¹¹ The multi-day realized variances \widehat{Q}_{τ_k} are obtained as non-overlapping sums of \widehat{Q}_t . Despite the use of high-frequency data, the estimated \widehat{Q}_{τ_k} will be a potentially noisy proxy of the true underlying volatility at that horizon. We have to keep that in mind when ranking the forecasts which makes choosing the appropriate loss function that much more important.

3.2 Ranking the Forecasts

The complication that arises in the context of ranking volatility forecasts is that volatility itself is not directly observable, but is rather estimated with noise. Ranking forecasts in the presence of imperfect volatility proxies has been studied by [Patton \(2011\)](#) who concludes that some loss functions preserve the ranking of the volatility forecasts.

We focus on the QLIKE criterion, which [Patton \(2011\)](#) shows being one of the few consistent loss functions, meaning that the ranking is robust to the use of an imperfect volatility proxy rather than the true volatility. The other robust loss function is the MSE. However, as noted by [Patton \(2011\)](#) and several authors before him, the MSE is more affected

¹¹[Andersen and Bollerslev \(1998\)](#) and subsequent work show that the realized variance is a much better proxy than squared returns.

by extreme observations than is the QLIKE.¹² The consistency and robustness to outliers of the QLIKE ensures that the error introduced from using a proxy rather than the true volatility does not change the ranking of our forecasting methods.

Using \widehat{Q}_s , we compute the feasible *QLIKE* at horizon k for a given model F which generates forecasts $V_{s,k}^F$ as

$$QLIKE(V_{s,k}^F, \widehat{Q}_s) = \log(V_{s,k}^F) + \frac{\widehat{Q}_s}{V_{s,k}^F}, \quad (3)$$

where s denotes the time of the forecast in the appropriate units.¹³ We rank the forecasts using the average QLIKE as an estimate of $E \left[QLIKE(V_{s,k}^F, \widehat{Q}_s) \right]$. The lower is the average QLIKE, the more precise is the forecast.

3.3 Testing the Forecasts

Consider two competing volatility forecasts, $F1$ and $F2$, for a given asset and horizon. We use Diebold and Mariano (1995) and West (1996) (DMW) to test the null hypothesis that the two forecasts deliver the same expected QLIKE, and hence are statistically indistinguishable. Let $v_{s,k}^{F1} \equiv QLIKE(V_{s,k}^{F1}, \widehat{Q}_s)$ and $v_{s,k}^{F2} \equiv QLIKE(V_{s,k}^{F2}, \widehat{Q}_s)$, and $d_{s,k} = v_{s,k}^{F1} - v_{s,k}^{F2}$. We test the null hypothesis that $E(d_{s,k}) = 0$, or

$$H_0 : E \left[QLIKE(V_{s,k}^{F1}, \widehat{Q}_s) - QLIKE(V_{s,k}^{F2}, \widehat{Q}_s) \right] = 0. \quad (4)$$

If model $F1$ is associated with a lower *QLIKE*, we test H_0 against the one-sided alternative hypothesis

$$H_1 : E \left[QLIKE(V_{s,k}^{F1}, \widehat{Q}_s) \right] < E \left[QLIKE(V_{s,k}^{F2}, \widehat{Q}_s) \right]. \quad (5)$$

The DWM test of the null (4) can be conducted with a Wald test that $E(d_{s,k}) = 0$ versus the alternative $E(d_{s,k}) < 0$.

¹²An older version of this paper reported results with the MSE. The results from that loss function are available upon request.

¹³That is, for the scaled and iterated approach that use daily information sets s equals $T + k$, whereas for the direct and MIDAS approaches s equals $T_k + 1$. The forecasts are sampled at the same times τ_k .

The DMW test can be carried out “in-sample”, by estimating the model parameters over the entire sample, or “out-of-sample”, by expanding an initial subsample and recursively estimating the parameters to produce forecasts, as in [West \(1996\)](#). We implement the DMW both in-sample and out-of-sample, as the two approaches have their own advantages, see [Hansen and Timmermann \(2015\)](#).

4 Data and Results

This section contains the empirical analysis. After a brief description of the data in [Section 4.1](#), we present the results of the in-sample exercise in [Section 4.2](#) and those of the out-of-sample analysis in [Section 4.3](#).

4.1 Data

We obtain data on 30 future contracts across eight asset classes: currencies, interest rates, stock indices, energy, metals, food and fiber, grains and oil seeds, life stock and meat. These assets span the four broad investable types of securities, namely, currencies, bonds, equities, and commodities. The data source is DTN IQFeed, an online provider of live and historical financial market data.¹⁴ The series start between September 2005 and February 2008, and all end on October 2018. The exact starting dates and further details are collected in [Table A.1](#). The futures prices of each contract are sampled at five minute intervals, and are then used to construct daily log returns and realized variance series as described above.

4.2 In-sample Results

We use the entire sample to produce forecasts with the nine approaches – three GARCH, three RV, and three MIDAS – and compute average QLIKE loss functions for each asset at a given horizon. We then rank the average QLIKE loss functions and use the DMW test to answer the following questions. First, we compare the performance of direct versus iterated

¹⁴The raw data take the form of continuous futures contracts. This is necessary in order to ensure that returns are correctly calculated. Prices are back-adjusted to create a continuous contract. This works by removing price gaps caused by a contract roll. The process starts at the end of the price series, and works its way back. This leaves current prices intact. Prices prior to the last roll date are adjusted.

forecasts within a category of forecasting approaches. In other words, we rank the three GARCH forecasts and test the first-best versus the second-best. Second, we discuss the performance of the scaled-up forecast in a category as it is often used in risk management and the finance industry. Within the GARCH category, we report the performance of the scaled-up daily GARCH forecast related to the Direct GARCH and Iterated GARCH. We do the same comparisons for the RV models.

Table 1 contains the results of the comparison within the group of the three GARCH models (Panel A) and three RV models (Panel B). All results are reported separately for the 30 assets and across the 5 horizons we consider. For a given asset-horizon combination, the number in the table refers to the best model, or the model with the lowest average in-sample QLIKE loss function. Entries in bold indicate that the best model’s average QLIKE is significantly lower (at the 5% c.l.) than the second best model’s average QLIKE, using the one-sided test outlined in Section 3.3. The bottom rows report the list of models and the fraction a given model is the best across assets for a given horizon, and across all horizons.

In Panel A of Table 1, we note that at the 5-day horizon, model 2 (namely, Iterated GARCH) is the best performing model in 57% percent of the assets and model 1 (Direct GARCH) is the best performing model in 23% of the cases. The Iterated GARCH is most frequently the best performing model at longer horizons as well; at 22-days, it dominates in 87% of the assets. Across horizons and assets, the Iterated GARCH dominates in 73% of the cases. Its performance is statistically significant particularly for currencies and interest rates. Hence, the approach that uses the wealth of daily information to generate daily GARCH forecasts, which are then iterated, delivers the bests forecasts. In about 50% of these cases, or 35% of the overall combinations, the iterated GARCH’s average QLIKE is statistically lower than that of the second best model. While there is a significant cross-sectional dependence across assets, we take the fractions to be a useful way of summarizing the overall performance of a model. The Direct and Scaled-up GARCH deliver rather comparable performances, ranking first in 23% and 20% at a 5-day horizon, respectively, and 13% and 14% across horizons. The takeaway from Panel A of Table 1 is that Iterated GARCH forecasts dominate Direct and Scaled-up GARCH forecasts at all horizons that we consider. The latter two are comparable in forecasting accuracy.

Turning now to the RV models in Panel B of Table 1, the evidence is completely different. At 5-day horizon, model 4 (Direct RV) is the best performing model in 90% of the assets and in 83% of the assets, its average QLIKE is statistically significantly lower (at

the 5% level) than that of the second best RV model, according to the DMW test. Looking across assets, the average QLIKE is statistically significantly lower than the second best RV model for all currencies, interest rates, energy, metals, grains, livestock, and most of food. The superior forecasting performance of the Direct RV model persists at longer horizons as well, although there is a systematic downward trend: at 10-days, it is the best model in 83% of the assets, while at 66-days, it dominates other RV models in 70% of the cases. The iterated RV performs particularly poorly at all horizons. It is clearly dominated by the Direct RV and by the Scaled-up RV model, which does better at longer horizons.

The picture that emerges from this table is quite surprising. Iterated GARCH produces better variance forecasts than Direct GARCH and, at the same time, Direct RV is the superior forecasting method to Iterated RV. These results seem to suggest two things. First, an approach that uses the wealth of daily information, as in the Iterated GARCH, delivers a superior performance. But then, why isn't the Iterated RV model also superior to the Direct RV? We conjecture that the autoregressive nature of the Iterated RV model is not flexible enough to capture the variance dynamics. In the forecasting context, the misspecification of the one-period model, while iterated forward, results in biased forecasts.

To test this conjecture, we extend the comparison to MIDAS models, which use daily forecasting information but also offer a flexible weighing scheme to directly forecast the realized variance at the horizon of interest. Hence, we expect the MIDAS approach to be less prone to aggregation of forecasting errors due to misspecification. We jointly compare the forecasting performance of the MIDAS, GARCH, and RV approaches. To be precise, we rank the average QLIKE of all nine models and carry a DMW test for the best model (the one with the lowest average QLIKE) against the model from the other two approaches with the lowest QLIKE. In other words, we ask if there is a best approach to forecast volatility.

In Table 2, we report the results of the rankings and tests for all 30 assets across the 5 horizons. We estimate the MIDAS models with non-linear least squares and a truncation j^{max} of 126 days. The model with the lowest average QLIKE is displayed. The format is the same as that of Table 1, including the rows that report the fraction a given model is the best across assets. At the 5-day horizon, MIDAS models produce the lowest average QLIKE in 87% of the assets. In 80% of the assets, the difference between the lowest average QLIKE of a MIDAS model and the average QLIKE of the best model among the other two groups (GARCH and RV) is statistically significant at the 5% level. MIDAS models generate the best forecasts across all assets and in particular for currencies, interest rates, energy, metals,

and food. Even at the 66-day horizon, MIDAS models delivers the best forecast in 73% of the cases. Across all horizons, a MIDAS model has the lowest average QLIKE in 80% of the assets.

It is worth noting that the Iterated GARCH model also yields precise forecasts for some assets. Specifically, it has the lowest average QLIKE in 11% of the cases, across horizons. In most of those cases, however, its average QLIKE was not statistically different from the second best model (which most often was a MIDAS). Surprisingly, the autoregressive RV approaches to forecast variances, both direct and integrated, did not fair well in this comparison. Hence, it is clear that while daily lagged variances can forecast future realized variance at long horizons, they have to be used within the MIDAS framework.

In Table 2, we notice a difference in performance within MIDAS models. The Exponential Almon lag specification does particularly well at short horizons. At the 5- and 10-day horizons, it is the best model for 53% and 57% of the assets, respectively. At a 66-day horizon, the Beta lag specification does best in 37% of the assets. The Corsi (2009) HAR specification does best at 22-day horizons. This is surprising as the HAR weights were conceived to forecast variances one-day-ahead and, hence, we expected this model to do well at short horizons. Overall, the Exponential Almon and Beta weights do better than the HAR, which is to be expected as the parameterized weights of these models are estimated from the data, whereas the HAR is specified a priori.

To illustrate the benefit of estimating the MIDAS weights from the data as opposed to setting them a priori, we plot in Figure 1 the estimated exponential Almon and Beta MIDAS weights against the predefined HAR MIDAS weights that underly the MIDAS volatility forecasts for one asset, the Euro/USD futures. In the plots on the left, we display the estimated MIDAS exponential Almon (top) and Beta (bottom) weights as well as the HAR weights for 5-day horizon forecasts. The exponential Almon and Beta weights are remarkably similar, which implies that both functions do equally well at capturing the forecasting relation at 5-day horizons. The HAR weights are somewhat different. Importantly, by construction, the HAR weights do not change with the forecasting horizon, while the estimated exponential Almon and Beta weights are horizon-specific. The two plots on the right in Figure 1 display the estimated Almon and Beta weights for 66-day horizon forecasts. They are markedly different than the 5-day weights. As we are forecasting the variance at a different horizon, there is no reason for the weight placed on past observations to be the same across horizons. The optimizer in the 66-day horizon case has chosen parameters that place most of the

weight on observations with lags of less than 10 days. The HAR weights are unchanged and are significantly different from the optimal MIDAS weights. Not surprisingly, as we saw in Table 2 for the Euro/USD futures, the optimal MIDAS weights produce lower QLIKE than the HAR MIDAS weights.

4.3 Out-of-sample Results

We construct the out-of-sample (OOS) exercise as follows. We estimate a volatility forecasting model for a given asset and horizon with an initial estimation subsample ending in December 2013. We use the estimates to form (pseudo) OOS volatility forecasts for the first 5, 10, and 22 days of 2014. Then, we expand the estimation subsample by the number of days corresponding to the forecasting horizon, re-estimate the model, and produce a new OOS forecast at all horizons. We limit the longest OOS horizon to 22 days, which gives us 57 non-overlapping, out-of-sample observations over which we compute the average QLIKE and conduct the DMW test.¹⁵ This is essentially the OOS procedure used by West (1996) which is by now standard in the forecasting literature.

Table 3 displays the OOS results in a format identical to Table 1. In Panel A, we sort the average QLIKE of the three GARCH approaches and report the best one along with significance at the 5% level. The OOS results are more nuanced. The Direct and Iterated GARCH perform equally well and are the best model in about a quarter of the assets across horizons. The Scaled-up GARCH performs surprisingly well and has the lowest OOS average QLIKE in the remaining 50% of the assets. We can clearly see that, within the GARCH approach, the scaled up daily forecasts perform as well as or better than the direct and iterated approaches.

The picture is markedly different when we turn to the results from the RV models, in Panel B of Table 3. The best performing RV approach is the direct one across horizons and nearly all assets. The forecasts from the iterated and scaled up RV approaches yield significantly higher QLIKES. These OOS results confirm and are even more clear than the in-sample results for RV in Panel B of Table 1. The conclusion from this panel is that, when we use autoregressive RV models to forecast future variance, the direct approach is the one to use.

¹⁵For 44 and 66-day horizons, we would have had 29 and 19 observations, clearly not enough for estimation and statistical inference that rely on asymptotics.

We compare the OOS forecasting performance of all nine models, including the three MIDAS ones, in Table 4. At the 5-day horizon, a MIDAS volatility model produces the best forecast in 77% of the assets. This is an impressive OOS performance, despite the fact that the number is somewhat less than the 87% percent observed in-sample (Table 2). And similarly to the in-sample results, the exponential Almon lag performs best across all MIDAS models: it yields the lowest QLIKE in 43% of the assets, which is slightly less than the 53% observed in-sample. While the Beta lag model is overall dominated by the exponential Almon lag, it does particularly well for equities at 22-day horizons.

It is important to point out that, compared to MIDAS and GARCH, the autoregressive RV models do not produce accurate OOS forecasts. Among the GARCH models, the scaled-up version continues to dominate particularly at long horizons. However, it is not clear to what extent these results are driven by small sample artifacts. After all, we only have 57 22-day OOS forecasts to compute the average QLIKE. From that perspective, the in-sample results might be equally, if not more, information about the relative rankings of the forecasts, as argued by [Hansen and Timmermann \(2015\)](#).

The overall message that emerges, in-sample and out-of-sample, is that the exponential Almon lag MIDAS model produces the lowest average QLIKE and thus the best forecasts, across assets and at most horizons.

5 Conclusion

To the question whether there is a preferred way to obtain multi-period volatility forecasts across assets, our answer is, yes, MIDAS type models. Furthermore, MIDAS specifications with flexible weights produce forecasts superior to those with preset weights.

The long-horizon forecasting aspect of our research is related to recent work in macroeconomics comparing various approaches of multi-period forecasts. [Marcellino et al. \(2006\)](#) compute direct and iterated forecasts of a multitude of macroeconomic variables and conclude that the iterated forecasts dominate the direct ones. [McCracken and McGillicuddy \(2018\)](#) revisit the [Marcellino et al. \(2006\)](#) analysis by including more recent data and find that direct forecasts dominate iterated ones especially during the Great Moderation period of 1984 to 2008. The lack of a consistent empirical finding is not surprising as the relative performance of one forecasting method over another hinges on the trade-off between bias and

estimation variance.¹⁶ Specifically, if the one-period model is known (no model uncertainty) and we are strictly concerned with parameter uncertainty, then iterating over the one-period forecast produces more efficient parameter estimates and more precise multi-period forecasts. However, in the more realistic case of misspecification in the one-period model (model uncertainty), directly forecasting the variance at the horizon of interest is more robust to biases arising from the misspecification. Because model uncertainty might lead to severe forecasting biases, the theoretical papers on the subject seem to favor direct over iterated multi-period forecasts (Bhansali (1999), Ing (2003), and Chevillon and Hendry (2005)).

We contribute in three aspects to this literature. First, the above studies are almost exclusively based on autoregressive models. Here, we introduce MIDAS as an alternative approach that is not autoregressive, is flexible and parametrically parsimonious, and yields good in and out-of-sample results. Second, the focus of these papers is on macroeconomic long-horizon forecasts, while in the current work, we are interested in long-horizon variance forecasts. Finally, we demonstrate that MIDAS variance forecasts are a useful middle-ground between the direct and iterated approaches. The ability of the MIDAS framework to produce a direct prediction of the long-horizon variance using higher frequency predictors can also be seen in the context of Merton (1980) who argued that high-frequency sampling is a dominant factor in volatility measurement and forecasting. This is also a feature of the iterated GARCH and RV approaches which use as forecasters daily squared returns or realized variances. However, unlike the iterated approaches, in a MIDAS the forecasted variable is the long-horizon variance, which allows us to side-step the need of aggregating the forecasts and introducing bias.

The insights from our data-driven exercise are of value especially because of the lack of theoretical guidance on long-horizon forecasting. Generally speaking, empirical results are conditional on the sample at hand and the design of the out-of-sample experiment. In our case, however, the findings are remarkably consistent across asset classes, horizons, and in and out-of-sample. At the very least, they speak to the method that should be used in multi-period volatility forecasts. More generally, our results might provide guidance for future theoretical work on long-horizon volatility forecasting.

¹⁶See Findley (1983), Findley (1985), Lin and Granger (1994), Clements and Hendry (1996), Bhansali (1999), Chevillon and Hendry (2005), and Marcellino et al. (2006) for further details.

References

- Andersen, T. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905.
- Andersen, T., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of exchange rate volatility. *Journal of American Statistical Association*, 96:42–55.
- Andersen, T., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.
- Andersen, T., Bollerslev, T., and Lange, S. (1999). Forecasting Financial Market Volatility: Sample Frequency vis-a-vis Forecast Horizon. *Journal of Empirical Finance*, 6:457–477.
- Bhansali, R. (1999). Parameter estimation and model selection for multistep prediction of a time series: A review. In Ghosh, S., editor, *Asymptotics, Nonparametrics, and Time Series*. Marcel Dekker.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327.
- Bollerslev, T., Hood, B., Huss, J., and Pedersen, L. H. (2018). Risk everywhere: Modeling and managing volatility. *The Review of Financial Studies*, 31(7):2729–2773.
- Chevillon, C. and Hendry, D. (2005). Non-parametric direct multi-step estimation for forecasting economic processes. *International Journal of Forecasting*, 21(1):201–218.
- Christoffersen, P. and Diebold, F. (2000). How Relevant is Volatility Forecasting for Financial Risk Management? *Review of Economics and Statistics*, 82(1):12–22.
- Clements, M. and Hendry, D. (1996). Mutli-Step Estimation for Forecasting. *Oxford Bulletin of Economics and Statistics*, 58(1):657–684.
- Corsi, F. (2009). A Simple Long Memory Model of Realized Volatility. *Journal of Financial Econometrics*, 7:174–196.
- Diebold, F., Hickman, A., Inoue, A., and Schuermann, T. (1998). Converting 1-Day Volatility to h-Day Volatility: Scaling by \sqrt{h} Is Worse than You Think. Wharton Financial Institutions Center, Working Paper 97-34.

- Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(1):253–263.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50:987–1008.
- Findley, D. (1983). On the Use of Multiple Models of Multi-period Forecasting. *Proceedings of the Business and Statistics Section, American Statistical Association*, 1:528–531.
- Findley, D. (1985). Model selection for multi-step-ahead forecasting. In Ghosh, S. and Young, P., editors, *Proceedings of the 7th Symposium on Identification and System Parameter Estimation*, pages 1039–1044. Pergamon, Oxford.
- Forsberg, L. and Ghysels, E. (2006). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, 5:31–67.
- French, K. R., Schwert, W., and Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, 19(1):3–29.
- Ghysels, E. and Hall, A. (1990). A test for structural stability of Euler conditions parameters estimated via the generalized method of moments estimator. *International Economic Review*, 31(2):355–364.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2002). The MIDAS touch: Mixed data sampling regression models. Working paper, UNC and UCLA.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return tradeoff after all. *Journal of Financial Economics*, 76:509–548.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006a). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131:59–95.
- Ghysels, E., Sinko, A., and Valkanov, R. (2006b). MIDAS Regressions: Further Results and New Directions. *Econometric Reviews*, 26:53–90.
- Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 15(1):549–582.

- Hansen, P. and Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(1):873–889.
- Hansen, P. R. and Timmermann, A. (2015). Equivalence between out-of-sample forecast comparisons and Wald Statistics. *Econometrica*, 83:2485–2505.
- Hoffman, D. and Pagan, A. (1989). Post-sample prediction tests for generalized method of moments estimators. *Oxford Bulletin of Economics and Statistics*, 51(3):333–343.
- Ing, C.-K. (2003). Multistep Prediction in Autoregressive Processes. *Econometric Theory*, 19(1):254–279.
- J.P.Morgan/Reuters (1996). Riskmetrics—technical document. J.P. Morgan/Reuters, New York.
- Lin, J.-L. and Granger, C. (1994). Model selection for multiperiod forecasts. *Journal of Forecasting*, 13(1):1–9.
- Marcellino, M., Stock, J., and Watson, M. (2006). A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series. *Journal of Econometrics*, 135(1):499–526.
- McCracken, M. W. and McGillicuddy, J. T. (2018). An empirical investigation of direct and iterated multistep conditional forecasts. *Journal of Applied Econometrics*, forthcoming.
- Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361.
- Patton, A. (2011). Volatility Forecast Comparison Using Imperfect Volatility Proxies. *Journal of Econometrics*, 160:246–256.
- Sheppard, K. (2018). Direct volatility modeling. University of Oxford working paper.
- West, K. (1996). Asymptotic Inference about Predictive Ability. *Econometrica*, 64:1067–1084.
- West, K. and Cho, D. (1995). The Predictive Ability of Several Models of Exchange Rate Volatility. *Journal of Econometrics*, 69:367–391.

Appendix

A MIDAS Weights: Various Specifications

We summarize the various weights specifications discussed in the paper.

1. Exponential:

$$b_k(j, \theta_{1,k}, \theta_{2,k}) = \frac{\exp\{\theta_{1,k}j + \theta_{2,k}j^2\}}{\sum_{i=0}^{\infty} \exp\{\theta_{1,k}i + \theta_{2,k}i^2\}}. \quad (6)$$

This function, first proposed by Ghysels et al. (2005), guarantees that the weights are positive (which in turn ensures that the forecasted volatility is also positive) and that they add up to one. Also, the functional form in equation (6) can produce a wide variety of shapes for different values of the two parameters, and it is parsimonious, with only two parameters to estimate. Finally, as long as the coefficient $\theta_{2,k}$ is negative, the weights go to zero as the lag length increases. The speed with which the weights decay controls the effective number of observations used to estimate the conditional volatility. We do not impose this constraint in the estimation.

2. Beta:

$$b_k(j, \theta_{1,k}, \theta_{2,k}) = \frac{f(\frac{j}{j^{max}}, \theta_{1,k}; \theta_{2,k})}{\sum_{i=1}^{j^{max}} f(\frac{i}{j^{max}}, \theta_{1,k}; \theta_{2,k})} \quad (7)$$

where: $f(z, a, b) = z^{a-1}(1-z)^{b-1}/\beta(a, b)$ and $\beta(a, b)$ is based on the Gamma function, or $\beta(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$. Specification (7) was introduced in Ghysels et al. (2002) and further explored in Ghysels et al. (2006b). One appealing feature is positivity of the coefficients, which is necessary for *a.s.* positive definiteness of the forecasted volatility. For $\theta_{1,k} = 1$ and $\theta_{2,k} > 1$ one has a slowly decaying pattern typical of volatility filters, which means that only one parameter is left to determine the shape, whereas in the case of $\theta_{1,k} = \theta_{2,k} = 1$ we obtain equal weights, which corresponds to a rolling estimator of the volatility (French et al. (1987), Ghysels et al. (2005)). The flexibility of the Beta function is well known and is often used in Bayesian econometrics to impose flexible, yet parsimonious prior distributions. The function can take many shapes, including flat weights, gradually declining weights as well as hump-shaped patterns. We impose no constraints when estimating the Beta specification.

3. HAR [Corsi \(2009\)](#):

$$b(j, \theta_1, \theta_2, \theta_3) = \frac{\theta_1 1_{\{j=1\}} + \theta_2 1_{\{j=2, \dots, 5\}} + \theta_3 1_{\{j=6, \dots, 22\}}}{\sum_{i=1}^3 \theta_i} \quad (8)$$

where $1_{\{j\}}$ is a dummy variable that equals to one if j is within a given range, and zero, otherwise. This is a step function, with steps at 1 day, two to five days, and six to 22 days. Beyond 22 daily lags, the weights are zero. Notice that the shape of the function does not depend on the horizon, k . The MIDAS HAR model can be written as a linear regression, where the three regressors are the realized variance of the last day, the realized variance over days 2 to 5, and the realized variance over days 6 to 22 as in the original [Corsi \(2009\)](#) paper. Hence, the parameters can be estimated by OLS. While there are technically four parameters, ϕ_k , θ_1 , θ_2 , and θ_3 , the constraint that the θ s sum up to unity insures identification.

Table 1: **Direct vs Iterated vs Scaled: In-Sample Results**

This table reports the results of the Direct vs Iterated vs Scaled test for the group of three GARCH models (Panel A) and three RV models (Panel B) for the 30 assets and 5 horizons. The Direct GARCH is 1, the Iterated GARCH is 2, and the Scaled GARCH is 3. The Direct RV is 4, the iterated RV is 5, and the Scaled RV is 6. The exact definition of these models is in section 2. For a given asset-horizon combination, the number refers to the best model based on the QLIKE criterion. All tests are in-sample. The bottom rows report the list of models and the fraction of assets a given model is the best across assets for a given horizon, and across All horizons. Entries in bold indicate that the best model's QLIKE is significantly lower (at the 5% c.l.) than the second best model's QLIKE, using the one-sided test outlined in Section 3.3.

| <i>Best model</i> | | | | | | <i>Best model</i> | | | | | |
|-----------------------|----------------|----------|----------|----------|----------|-----------------------|----------------|----------|----------|----------|----------|
| Panel A: GARCH models | | | | | | Panel B: RV models | | | | | |
| Asset | Horizon (days) | | | | | Asset | Horizon (days) | | | | |
| | 5 | 10 | 22 | 44 | 66 | | 5 | 10 | 22 | 44 | 66 |
| Currencies | | | | | | Currencies | | | | | |
| Australian Dollar | 2 | 2 | 2 | 2 | 2 | Australian Dollar | 4 | 4 | 4 | 6 | 6 |
| British Pound | 2 | 2 | 2 | 2 | 2 | British Pound | 4 | 4 | 4 | 6 | 6 |
| Canadian Dollar | 2 | 2 | 2 | 2 | 2 | Canadian Dollar | 4 | 4 | 4 | 6 | 6 |
| Euro | 2 | 2 | 2 | 2 | 2 | Euro | 4 | 4 | 4 | 4 | 4 |
| Japanese Yen | 2 | 2 | 2 | 2 | 2 | Japanese Yen | 4 | 4 | 4 | 6 | 4 |
| Swiss Franc | 1 | 3 | 1 | 3 | 3 | Swiss Franc | 4 | 4 | 4 | 4 | 4 |
| Interest Rates | | | | | | Interest Rates | | | | | |
| 2Y Treasury | 1 | 1 | 1 | 1 | 2 | 2Y Treasury | 4 | 4 | 6 | 4 | 4 |
| 10Y Treasury | 2 | 2 | 2 | 2 | 2 | 10Y Treasury | 4 | 4 | 4 | 4 | 4 |
| 30Y Treasury | 2 | 2 | 2 | 2 | 2 | 30Y Treasury | 4 | 4 | 4 | 4 | 4 |
| Stocks | | | | | | Stocks | | | | | |
| S&P 500 | 3 | 2 | 2 | 2 | 3 | S&P 500 | 6 | 6 | 6 | 4 | 4 |
| Russell 2000 | 2 | 2 | 2 | 2 | 3 | Russell 2000 | 4 | 6 | 6 | 4 | 6 |
| Nasdaq | 3 | 3 | 2 | 2 | 3 | Nasdaq | 6 | 6 | 6 | 4 | 4 |
| Nikkei 225 | 2 | 2 | 2 | 2 | 2 | Nikkei 225 | 6 | 4 | 6 | 6 | 4 |
| Energies | | | | | | Energies | | | | | |
| Crude Oil | 3 | 3 | 2 | 2 | 3 | Crude Oil | 4 | 4 | 4 | 4 | 6 |
| Heating Oil | 3 | 3 | 1 | 3 | 3 | Heating Oil | 4 | 4 | 4 | 4 | 4 |
| Natural Gas | 2 | 2 | 2 | 2 | 2 | Natural Gas | 4 | 4 | 4 | 4 | 4 |
| Metals | | | | | | Metals | | | | | |
| Gold | 3 | 2 | 2 | 2 | 2 | Gold | 4 | 4 | 4 | 4 | 4 |
| Copper | 3 | 3 | 2 | 2 | 2 | Copper | 4 | 4 | 4 | 6 | 4 |
| Silver | 2 | 2 | 2 | 2 | 2 | Silver | 4 | 4 | 4 | 4 | 4 |
| Food | | | | | | Food | | | | | |
| Cocoa | 2 | 2 | 2 | 2 | 2 | Cocoa | 4 | 4 | 4 | 4 | 4 |
| Coffee | 2 | 2 | 2 | 2 | 2 | Coffee | 4 | 4 | 4 | 4 | 6 |
| Cotton | 2 | 2 | 2 | 1 | 3 | Cotton | 4 | 6 | 6 | 6 | 6 |
| Orange Juice | 1 | 1 | 2 | 2 | 2 | Orange Juice | 4 | 4 | 4 | 4 | 4 |
| Sugar | 1 | 2 | 2 | 2 | 2 | Sugar | 4 | 4 | 4 | 4 | 4 |
| Grains | | | | | | Grains | | | | | |
| Corn | 1 | 1 | 3 | 1 | 1 | Corn | 4 | 4 | 4 | 4 | 4 |
| Oats | 2 | 2 | 2 | 2 | 1 | Oats | 4 | 4 | 4 | 4 | 6 |
| Soybeans | 2 | 2 | 2 | 2 | 2 | Soybeans | 4 | 6 | 6 | 6 | 6 |
| Wheat | 2 | 2 | 2 | 2 | 2 | Wheat | 4 | 4 | 4 | 4 | 4 |
| Livestock | | | | | | Livestock | | | | | |
| Feeder Cattle | 1 | 2 | 2 | 2 | 1 | Feeder Cattle | 4 | 4 | 4 | 4 | 4 |
| Live Cattle | 1 | 2 | 2 | 2 | 2 | Live Cattle | 4 | 4 | 4 | 4 | 4 |

| <i>Fraction of times a model is the best</i> | | | | | | | <i>Fraction of times a model is the best</i> | | | | | | |
|--|----------------|-----|-----|-----|-----|-----|--|----------------|-----|-----|-----|-----|-----|
| Model | Horizon (days) | | | | | | Model | Horizon (days) | | | | | |
| | 5 | 10 | 22 | 44 | 66 | All | | 5 | 10 | 22 | 44 | 66 | All |
| 1: GARCH-D | 23% | 10% | 10% | 10% | 10% | 13% | 4: RV-D | 90% | 83% | 77% | 73% | 70% | 79% |
| 2: GARCH-I | 57% | 73% | 87% | 83% | 67% | 73% | 5: RV-I | 0% | 0% | 0% | 0% | 0% | 0% |
| 3: GARCH-S | 20% | 17% | 3% | 7% | 23% | 14% | 6: RV-S | 10% | 17% | 23% | 27% | 30% | 21% |

Table 2: **Best model against Best of other groups: In-Sample Results**

This table reports the results of the test for the best model against the best model of the other two groups, for the 30 assets and 5 horizons. The Direct GARCH is 1, the Iterated GARCH is 2, the Scaled GARCH is 3; the Direct RV is 4; the Iterated RV is 5; the Scaled RV is 6; the Beta MIDAS is 7; the exponential Almon MIDAS is 8; and the HAR MIDAS is 9. The exact definition of these models is in section 2. For a given asset-horizon combination, the numbers in the leftmost panel report the best model based on the QLIKE criterion. All tests are in-sample. The rightmost panel reports the list of models by group, and the fraction of assets a given model is the best for a given horizon, and across all horizons. In bold, we report the sum of these fractions across the models of the same group (GARCH, RV, or MIDAS). Entries in bold indicate that the best model's QLIKE is significantly lower (at the 5% c.l.) than the QLIKE of the best model of the other two groups, using the one-sided test outlined in Section 3.3.

| <i>Best model</i> | | | | | |
|-----------------------|----------------|----------|----------|----------|----------|
| Asset | Horizon (days) | | | | |
| | 5 | 10 | 22 | 44 | 66 |
| Currencies | | | | | |
| Australian Dollar | 9 | 8 | 8 | 6 | 6 |
| British Pound | 9 | 8 | 9 | 9 | 7 |
| Canadian Dollar | 9 | 7 | 8 | 8 | 7 |
| Euro | 8 | 8 | 9 | 9 | 7 |
| Japanese Yen | 9 | 8 | 7 | 8 | 8 |
| Swiss Franc | 9 | 9 | 7 | 4 | 9 |
| Interest Rates | | | | | |
| 2Y Treasury | 8 | 8 | 9 | 8 | 7 |
| 10Y Treasury | 9 | 8 | 8 | 8 | 8 |
| 30Y Treasury | 8 | 8 | 8 | 8 | 7 |
| Stocks | | | | | |
| S&P 500 | 9 | 8 | 2 | 2 | 3 |
| Russell 2000 | 9 | 8 | 6 | 7 | 3 |
| Nasdaq | 8 | 6 | 9 | 2 | 8 |
| Nikkei 225 | 2 | 8 | 7 | 2 | 8 |
| Energies | | | | | |
| Crude Oil | 9 | 7 | 9 | 9 | 3 |
| Heating Oil | 8 | 8 | 9 | 9 | 8 |
| Natural Gas | 8 | 8 | 9 | 9 | 9 |
| Metals | | | | | |
| Gold | 8 | 7 | 8 | 7 | 2 |
| Copper | 9 | 7 | 9 | 9 | 7 |
| Silver | 8 | 9 | 9 | 7 | 2 |
| Food | | | | | |
| Cocoa | 8 | 9 | 8 | 7 | 7 |
| Coffee | 8 | 8 | 8 | 8 | 9 |
| Cotton | 2 | 2 | 2 | 1 | 3 |
| Orange Juice | 8 | 9 | 7 | 8 | 8 |
| Sugar | 1 | 7 | 2 | 4 | 7 |
| Grains | | | | | |
| Corn | 8 | 8 | 7 | 4 | 7 |
| Oats | 8 | 8 | 7 | 8 | 9 |
| Soybeans | 2 | 2 | 2 | 2 | 2 |
| Wheat | 8 | 8 | 7 | 4 | 8 |
| Livestock | | | | | |
| Feeder Cattle | 8 | 9 | 8 | 8 | 7 |
| Live Cattle | 8 | 8 | 9 | 7 | 7 |

| <i>Fraction of times a model is the best</i> | | | | | | |
|--|----------------|------------|------------|------------|------------|------------|
| Model | Horizon (days) | | | | | |
| | 5 | 10 | 22 | 44 | 66 | All |
| 1: GARCH-D | 3% | 0% | 0% | 3% | 0% | 1% |
| 2: GARCH-I | 10% | 7% | 13% | 13% | 10% | 11% |
| 3: GARCH-S | 0% | 0% | 0% | 0% | 13% | 3% |
| GARCH 1–3 | 13% | 7% | 13% | 17% | 23% | 15% |
| 4: RV-D | 0% | 0% | 0% | 13% | 0% | 3% |
| 5: RV-I | 0% | 0% | 0% | 0% | 0% | 0% |
| 6: RV-S | 0% | 3% | 3% | 3% | 3% | 3% |
| RV 4–6 | 0% | 3% | 3% | 17% | 3% | 5% |
| 7: MIDAS-B | 0% | 17% | 23% | 17% | 37% | 19% |
| 8: MIDAS-E | 53% | 57% | 27% | 30% | 23% | 38% |
| 9: MIDAS-H | 33% | 17% | 33% | 20% | 13% | 23% |
| Midas 7–9 | 87% | 90% | 83% | 67% | 73% | 80% |

Table 3: **Direct vs Iterated vs Scaled: Out-of-Sample Results**

This table reports the results of the Direct vs Iterated vs Scaled test for the group of three GARCH models (Panel A) and three RV models (Panel B) based on the out-of-sample exercise outlined in Section 4.3, for the 30 assets and 5 horizons. The Direct GARCH is 1, the Iterated GARCH is 2, and the Scaled GARCH is 3. The Direct RV is 4, the iterated RV is 5, and the Scaled RV is 6. The exact definition of these models is in section 2. For a given asset-horizon combination, the number refers to the winning model based on the QLIKE criterion. The bottom rows report the list of models and the fraction of assets a given model is the winner across assets for a given horizon, and across All horizons. Entries in bold indicate that the best model's QLIKE is significantly lower (at the 5% c.l.) than the second best model's QLIKE, using the one-sided test outlined in Section 3.3.

| <i>Best model</i> | | | | | <i>Best model</i> | | | | |
|---|----------------|----------|----------|-----|---|----------------|------|------|------|
| Panel A: GARCH models | | | | | Panel B: RV models | | | | |
| Asset | Horizon (days) | | | | Asset | Horizon (days) | | | |
| | 5 | 10 | 22 | | | 5 | 10 | 22 | |
| Currencies | | | | | Currencies | | | | |
| Australian Dollar | 2 | 2 | 2 | | Australian Dollar | 4 | 4 | 4 | |
| British Pound | 2 | 3 | 3 | | British Pound | 4 | 4 | 4 | |
| Canadian Dollar | 1 | 3 | 3 | | Canadian Dollar | 4 | 4 | 4 | |
| Euro | 3 | 1 | 2 | | Euro | 4 | 4 | 4 | |
| Japanese Yen | 1 | 1 | 1 | | Japanese Yen | 4 | 4 | 4 | |
| Swiss Franc | 3 | 1 | 1 | | Swiss Franc | 4 | 4 | 4 | |
| Interest Rates | | | | | Interest Rates | | | | |
| 2Y Treasury | 1 | 2 | 2 | | 2Y Treasury | 4 | 4 | 4 | |
| 10Y Treasury | 3 | 3 | 3 | | 10Y Treasury | 4 | 4 | 4 | |
| 30Y Treasury | 2 | 2 | 2 | | 30Y Treasury | 4 | 4 | 4 | |
| Stocks | | | | | Stocks | | | | |
| S&P 500 | 3 | 1 | 3 | | S&P 500 | 4 | 4 | 4 | |
| Russell 2000 | 3 | 3 | 3 | | Russell 2000 | 4 | 4 | 4 | |
| Nasdaq | 3 | 1 | 3 | | Nasdaq | 4 | 4 | 4 | |
| Nikkei 225 | 1 | 1 | 1 | | Nikkei 225 | 4 | 4 | 4 | |
| Energies | | | | | Energies | | | | |
| Crude Oil | 3 | 3 | 2 | | Crude Oil | 4 | 4 | 4 | |
| Heating Oil | 1 | 1 | 3 | | Heating Oil | 4 | 4 | 4 | |
| Natural Gas | 1 | 3 | 3 | | Natural Gas | 4 | 4 | 4 | |
| Metals | | | | | Metals | | | | |
| Gold | 3 | 3 | 3 | | Gold | 4 | 4 | 4 | |
| Copper | 1 | 1 | 1 | | Copper | 4 | 4 | 4 | |
| Silver | 3 | 3 | 3 | | Silver | 4 | 4 | 4 | |
| Food | | | | | Food | | | | |
| Cocoa | 3 | 3 | 2 | | Cocoa | 4 | 4 | 4 | |
| Coffee | 3 | 2 | 2 | | Coffee | 4 | 4 | 4 | |
| Cotton | 3 | 1 | 3 | | Cotton | 4 | 4 | 4 | |
| Orange Juice | 2 | 2 | 2 | | Orange Juice | 4 | 4 | 4 | |
| Sugar | 3 | 1 | 1 | | Sugar | 4 | 4 | 4 | |
| Grains | | | | | Grains | | | | |
| Corn | 1 | 3 | 3 | | Corn | 4 | 4 | 4 | |
| Oats | 2 | 1 | 1 | | Oats | 4 | 4 | 4 | |
| Soybeans | 3 | 3 | 3 | | Soybeans | 4 | 4 | 4 | |
| Wheat | 3 | 3 | 3 | | Wheat | 4 | 4 | 4 | |
| Livestock | | | | | Livestock | | | | |
| Feeder Cattle | 2 | 3 | 2 | | Feeder Cattle | 4 | 4 | 4 | |
| Live Cattle | 2 | 3 | 1 | | Live Cattle | 4 | 4 | 4 | |
| <i>Fraction of assets a model is the best</i> | | | | | <i>Fraction of assets a model is the best</i> | | | | |
| Model | 5 | 10 | 22 | All | Model | 5 | 10 | 22 | All |
| 1: GARCH-D | 27% | 37% | 23% | 29% | 4: RV-D | 100% | 100% | 100% | 100% |
| 2: GARCH-I | 23% | 17% | 30% | 23% | 5: RV-I | 0% | 0% | 0% | 0% |
| 3: GARCH-S | 50% | 47% | 47% | 48% | 6: RV-S | 0% | 0% | 0% | 0% |

Table 4: **Best model against Best of other groups: Out-of-Sample Results**

This table reports the results of the test for the best model against the best model of the other two groups based on the out-of-sample exercise outlined in Section 4.3, for the 30 assets and 5 horizons. The Direct GARCH is 1, the Iterated GARCH is 2, the Scaled GARCH is 3; the Direct RV is 4; the Iterated RV is 5; the Scaled RV is 6; the Beta MIDAS is 7; the exponential Almon MIDAS is 8; and the HAR MIDAS is 9. The exact definition of these models is in section 2. For a given asset-horizon combination, the numbers in the leftmost panel report the best model based on the QLIKE criterion. The rightmost panel reports the list of models by group, and the fraction of assets a given model is the best for a given horizon, and across all horizons. In bold, we report the sum of these fractions across the models of the same group (GARCH, RV, or MIDAS). Entries in bold indicate that the best model's QLIKE is significantly lower (at the 5% c.l.) than the QLIKE of the best model of the other two groups, using the one-sided test outlined in Section 3.3.

| <i>Best Model</i> | | | | | | | | |
|-----------------------|----------------|----------|----------|---|--|--|--|--|
| Asset | Horizon (days) | | | | | | | |
| | 5 | 10 | 22 | | | | | |
| Currencies | | | | | | | | |
| Australian Dollar | 2 | 7 | 2 | | | | | |
| British Pound | 8 | 9 | 3 | | | | | |
| Canadian Dollar | 8 | 8 | 8 | | | | | |
| Euro | 8 | 9 | 9 | | | | | |
| Japanese Yen | 7 | 9 | 8 | | | | | |
| Swiss Franc | 9 | 9 | 9 | | | | | |
| Interest Rates | | | | <i>Fraction of assets a model is the best</i> | | | | |
| 2Y Treasury | 9 | 9 | 8 | | | | | |
| 10Y Treasury | 8 | 8 | 8 | | | | | |
| 30Y Treasury | 9 | 2 | 2 | | | | | |
| Stocks | | | | | | | | |
| S&P 500 | 8 | 9 | 3 | | | | | |
| Russell 2000 | 8 | 8 | 3 | | | | | |
| Nasdaq | 8 | 9 | 7 | | | | | |
| Nikkei 225 | 9 | 9 | 7 | | | | | |
| Energies | | | | | | | | |
| Crude Oil | 9 | 8 | 9 | | | | | |
| Heating Oil | 8 | 7 | 7 | | | | | |
| Natural Gas | 1 | 9 | 9 | | | | | |
| Metals | | | | | | | | |
| Gold | 8 | 9 | 3 | | | | | |
| Copper | 4 | 8 | 9 | | | | | |
| Silver | 9 | 8 | 3 | | | | | |
| Food | | | | | | | | |
| Cocoa | 8 | 8 | 8 | | | | | |
| Coffee | 9 | 9 | 2 | | | | | |
| Cotton | 3 | 1 | 3 | | | | | |
| Orange Juice | 8 | 2 | 2 | | | | | |
| Sugar | 8 | 9 | 8 | | | | | |
| Grains | | | | | | | | |
| Corn | 1 | 8 | 3 | | | | | |
| Oats | 8 | 8 | 8 | | | | | |
| Soybeans | 3 | 9 | 3 | | | | | |
| Wheat | 3 | 8 | 8 | | | | | |
| Livestock | | | | | | | | |
| Feeder Cattle | 9 | 9 | 9 | | | | | |
| Live Cattle | 9 | 8 | 7 | | | | | |

| Model | Horizon (days) | | | |
|------------------|----------------|------------|------------|------------|
| | 5 | 10 | 22 | All |
| 1: GARCH-D | 7% | 3% | 0% | 3% |
| 2: GARCH-I | 3% | 7% | 13% | 8% |
| 3: GARCH-S | 10% | 0% | 27% | 12% |
| GARCH 1–3 | 20% | 10% | 40% | 23% |
| 4: RV-D | 3% | 0% | 0% | 1% |
| 5: RV-I | 0% | 0% | 0% | 0% |
| 6: RV-S | 0% | 0% | 0% | 0% |
| RV 4–6 | 3% | 0% | 0% | 1% |
| 7: MIDAS-B | 3% | 7% | 13% | 8% |
| 8: MIDAS-E | 43% | 37% | 27% | 36% |
| 9: MIDAS-H | 30% | 47% | 20% | 32% |
| MIDAS 7–9 | 77% | 90% | 60% | 76% |

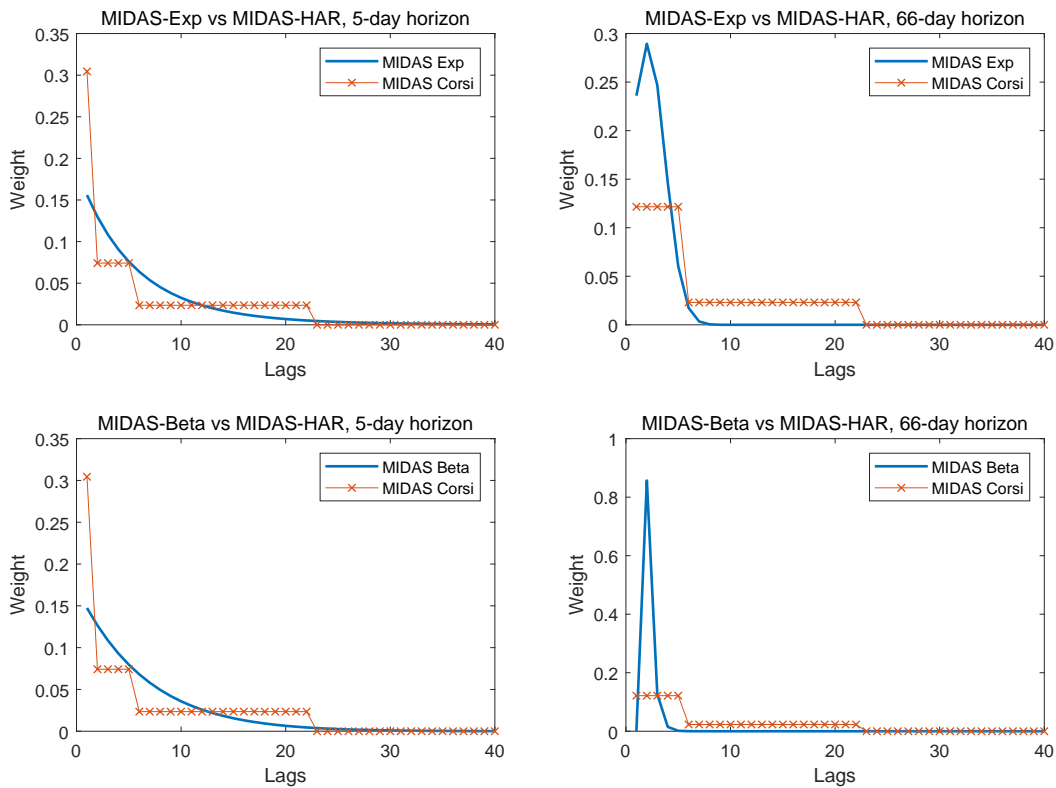


Figure 1: Optimal MIDAS weights, Euro

This figure displays the optimal weights from Table 2 for the MIDAS models of predicting the realized variance of the USD/Euro exchange rate future contract at the 5-day (left plots) and 66-day (right plots). The top plots display the optimal weights of the MIDAS-Exponential and MIDAS-HAR models, whereas the bottom plots display the optimal weights of the MIDAS-Beta and MIDAS-HAR models.

Table A.1: **Futures Data**

The first four columns indicate the asset class, contract, symbol, and exchange for each series. The relevant exchanges are Chicago Mercantile Exchange (CME), Intercontinental Exchange (ICE), Chicago Board of Trade (CBOT), New York Mercantile Exchange (NYMEX), and Commodity Exchange (COMEX). The start date for each series is in the format (MM/YY). All the data end on October 31, 2018. Roll days is the number of trading days before the expiration date that a contract is rolled. Values with a * indicate the number of trading days from the end of the month prior to expiration that a contract is rolled. N is the average daily number of 5-minute returns.

| Asset class | Contract | Symbol | Exchange | Start Date | Roll Days | N |
|----------------|-------------------|--------|----------|------------|-----------|-----|
| Currencies | Australian Dollar | AD | CME | 7-Sep-05 | 2 | 272 |
| Currencies | British Pound | BP | CME | 21-Aug-07 | 2 | 233 |
| Currencies | Canadian Dollar | CD | CME | 17-Apr-06 | 3 | 266 |
| Currencies | Euro | EU | CME | 27-Feb-07 | 2 | 220 |
| Currencies | Japanese Yen | JY | CME | 4-Jan-07 | 2 | 222 |
| Currencies | Swiss Franc | SF | CME | 9-Jun-06 | 2 | 267 |
| Interest Rates | 2Y Treasury | TU | CBOT | 9-May-06 | 1* | 262 |
| Interest Rates | 10Y Treasury | TY | CBOT | 27-Mar-06 | 1* | 271 |
| Interest Rates | 30Y Treasury | US | CBOT | 5-May-06 | 1* | 270 |
| Stocks | S&P 500 | ES | CME | 24-Apr-06 | 7 | 268 |
| Stocks | Russell 2000 | TF | ICE | 17-Nov-05 | 7 | 274 |
| Stocks | Nasdaq 100 | NQ | CME | 23-Nov-05 | 7 | 273 |
| Stocks | Nikkei 225 | NKD | CME | 7-Mar-06 | 5 | 260 |
| Energies | Crude Oil | CL | NYMEX | 18-Feb-08 | 3 | 276 |
| Energies | Heating Oil | HO | NYMEX | 28-Feb-08 | 3 | 218 |
| Energies | Natural Gas | NG | NYMEX | 26-Feb-08 | 3 | 254 |
| Metals | Gold | GC | COMEX | 24-May-07 | 2* | 269 |
| Metals | Copper | HG | COMEX | 17-May-07 | 2* | 259 |
| Metals | Silver | SI | COMEX | 31-Aug-07 | 2* | 270 |
| Food | Cocoa | CC | ICE | 17-May-07 | 19* | 107 |
| Food | Coffee | KC | ICE | 17-May-07 | 16* | 116 |
| Food | Cotton | CT | ICE | 17-May-07 | 14* | 163 |
| Food | Orange Juice | OJ | ICE | 17-May-07 | 7* | 59 |
| Food | Sugar | SB | ICE | 17-May-07 | 10 | 121 |
| Grains | Corn | C | CBOT | 5-Mar-07 | 8* | 190 |
| Grains | Oats | O | CBOT | 17-May-07 | 8* | 54 |
| Grains | Soybeans | S | CBOT | 12-Apr-07 | 8* | 187 |
| Grains | Wheat | W | CBOT | 17-Apr-07 | 8* | 186 |
| Livestock | Feeder Cattle | GF | CME | 17-May-07 | 15 | 70 |
| Livestock | Live Cattle | LE | CME | 17-May-07 | 15 | 97 |