Forecasting the Realized Variance in the Presence of Intraday Periodicity

Ana Maria H. Dumitru* [a], Rodrigo Hizmeri†[b], Marwan Izzeldin‡[c]

[a] *Conning*
[b] *University of Liverpool*
[c] *Lancaster University*

November 14, 2024

Abstract

This paper examines the impact of intraday periodicity on forecasting realized volatility using a heterogeneous autoregressive model (HAR) framework. We show that periodicity inflates the variance of the realized volatility and biases jump estimators. This combined effect adversely affects forecasting. To account for this, we propose a periodicity-adjusted HAR model, HARP, where predictors are constructed from the periodicity-filtered data. We demonstrate empirically (using 30 stocks from various business sectors and the SPY for the period 2000–2020) and via Monte Carlo simulations that the HARP models produce significantly better forecasts across all forecasting horizons. We also show that adjusting for periodicity when estimating the variance risk premium improves return predictability.

Keywords: realized volatility, heterogeneous autoregressive models, intraday periodicity, forecast, variance risk-premium.
JEL codes: C14, C22, C58, G12, G17.

◇ *We are grateful to the editor (Geert Bekaert) and two anonymous referees for their extremely helpful comments. We also thank Torben Andersen, Lars Forsberg, Aleksey Kolokolov, Alex Kostakis, Sébastien Laurent, Ingmar Nolte, Roberto Renò, participants at the 13th International Conference on Computational and Financial Econometrics, the 72nd European Meeting of the Econometric Society, the Workshop on Asset Pricing and Econometrics at Manchester University, the Infiniti Conference on International Finance 2019 at the University of Glasgow, and the 2nd International Conference on Quantitative Finance and Financial Econometrics at the Aix-Marseille School of Economics for helpful comments.*

*Corresponding author: Conning, Gertrudenstr. 30-36, 50667 Köln, Nordrhein-Westfalen, Germany. E-mail: ana.dumitru@conning.com.

†Accounting and Finance Group, University of Liverpool Management School, Liverpool, L69 7ZH, United Kingdom. Email: r.hizmeri@liverpool.ac.uk.

‡Department of Economics, Lancaster University Management School, Bailrigg, LA1 4YX, United Kingdom. Email: m.izzeldin@lancaster.ac.uk.

Introduction

With the availability of high-frequency data in the late 1990s, realized volatility (RV) and related measures were developed as proxies for the daily observed volatility of all financial securities for which intraday price observations were available (Andersen and Bollerslev, 1998a). The shift in volatility from latent to quasi-observable[1] meant forecasting could now rely on simple autoregressive models. Corsi (2009)'s heterogenous autoregressive model (HAR) emerged as the standard in forecasting univariate realized volatility.

In this paper, we show that the periodicity of intraday volatility impacts realized volatility forecasts based on autoregressive models through two channels. The first and most important channel is by distorting the variance of the realized volatility, which in turn contributes to biasing the coefficients of the forecasting models. The second channel is via the realized jumps regressors that appear in some predictive models and can also be biased in the presence of intraday periodicity.

To address the observed impact of periodicity, we propose forecasting RV with predictors based on data from which periodicity is filtered out. We refer to these autoregressive models as ARP or HARP, with "P" from periodicity-filtered. When the integrated volatility follows a simple AR(1) model, we show that filtering out periodicity lowers the RV variance when periodicity is estimated on a big sample. We compare the forecasting performance of the HARP models with several HAR models existing in the literature. To this end, we perform a simulation exercise, followed by an empirical application based on high frequency data for the SPDR S&P 500 ETF (SPY) and 30 S&P 500 constituents, observed over the period 2000-2020.

Our analysis attests the superiority of HARP models across all forecasting horizons, with greater gains for the 1-day to 1-week ahead forecasts. For SPY, we observe improvements of up to 7% in the forecast losses using HARP models. For the stock average, depending on the model specification, filtering data reduces the forecast losses by approximately 6.6% to 7.3% at the 1-day and 1-week horizons. These results are consistent when using different time-varying windows to estimate periodicity. Moreover, at the 1-month horizon, HARP models still outperform their counterparts for both SPY and individual stocks. However, we observe a decrease in the number of models significantly outperforming standard HAR models, suggesting that the impact of intraday periodicity diminishes as the horizon lengthens. Nevertheless, when comparing the eight HAR models consid-

[1]The use of the term quasi here is due to the fact that all realized measures are estimates.

ered throughout the paper using the model confidence set approach (Hansen et al., 2011), the filtered models consistently rank first. Finally, and to demonstrate the usefulness of our approach, we use monthly RV predictions for SPY and individual stocks to estimate the variance risk premium as the difference between the Risk-Neutral Variance (RNV) of Bakshi et al. (2003) and the RV forecast. When predicting excess returns with the variance risk premium, we find that VRP measures derived from HARP models lead to improvements in forecasting future monthly excess returns. While the improvement for the aggregate market is substantial, yielding an increase in the R^2 of up to 50% relative to their counterparts, for individual stocks, the R^2 improvement is marginal, and the overall significance is modest.

Andersen et al. (2003) document the presence of long memory in the time series of logarithmic realized volatilities and suggest a fractionally integrated autoregressive approach in modelling. Inspired by the heterogeneous autoregressive conditional heteroskedasticity (HARCH) model featured in Müller et al. (1997) and Dacorogna et al. (1997), Corsi (2009) proposes the HAR model which regresses realized volatilities on past daily, weekly and monthly realized volatilities. This model can replicate the high levels of persistence observed in the series of daily realized volatilities, without relying on fractional integration. Given its simple linear structure and ease of estimation, the HAR has become the most popular framework in forecasting realized volatilities.

The daily quadratic variation includes a continuous component and a jump part,[2] with the former component featuring a high level of persistence, while the jump component shows little or no persistence. To account for the different levels of persistence in the two components, Andersen et al. (2007a) propose adding the lagged realized daily squared jump as an extra explanatory variable to the HAR regression, leading to the HAR-J model. They also propose the HAR-CJ model, which uses as predictors daily, weekly and monthly estimates of the integrated variance and integrated squared jumps. They find that accounting for jumps generally leads to an increase in the explanatory power. This finding is also confirmed by Corsi et al. (2010), who perform a more exhaustive forecasting exercise.

Corsi and Renò (2012) add negative returns to the previous HAR specifications, in order to account for a potential leverage effect. They show improved accuracy in forecasting the S&P 500. Bollerslev et al. (2016) argue that all realized measures used in HAR models are bound to include measurement errors, which should be taken into account

[2]See, for instance, Barndorff-Nielsen and Shephard (2004, 2006b), Mancini (2009), Christensen and Podolskij (2007), Corsi et al. (2010), Andersen et al. (2012) and Bu et al. (2023).

in modelling. The model, abbreviated HAR-Q,[3] performs well in environments of high variability of the measurement error.

The impact of periodicity on the dynamic properties of high frequency returns was first examined by Andersen and Bollerslev (1997). They model intraday volatility as a product of two components: a deterministic periodic component and the actual volatility, i.e. a stochastic component reflecting variability in the fundamental value of the financial security. Such specification has become the literature standard and is also considered in our analysis. Andersen et al. (2001) and Bollerslev et al. (2000) employ similar specifications in modelling intraday volatility in the FX and the US Treasury bond markets.

While the periodicity component does not impact the realized variance, by integrating to 1 over the trading day,[4] little is known of its impact on other realized measures. Andersen et al. (2018) propose a statistical test for time-varying intraday periodicity in high frequency data and associated realized measures. Christensen et al. (2018) develop a test for the hypothesis that time-variation in intraday volatility is caused solely by intraday periodicity. Dette et al. (2023) examine the effect of periodicity on the realized bi-power variation, its variance and covariance with the realized variance, as well as on the realized quarticity under a constant volatility data generating process (DGP hereafter).

Intraday periodicity has also been shown to impact the jump detection ability of the intraday jump tests proposed by Andersen et al. (2007b) and Lee and Mykland (2008), where high levels of periodicity can increase the probability of type I error (Andersen et al., 2007b). This highlights the confounding impact of jumps and periodicity on the price process and related functions. Boudt et al. (2011, 2012) recommend applying intraday jump tests on returns from which periodicity is filtered out. They propose non-parametric and parametric methods to estimate periodicity that are robust to jumps in prices and time-varying volatility.

The rest of the paper is structured as follows. Section 1 provides the theoretical background on realized-type volatility measures and intraday periodicity estimation. Section 2 discusses the impact of intraday periodicity on RV forecasting and proposes periodicity filtering. Sections 3 and 4 present simulation results and empirical applications, respectively. Section 5 concludes. Appendices A and B provide AR(1) model derivations and intraday periodicity estimator details.

[3] "Q" comes from the fact that the realized quarticity, as the estimated asymptotic variance of the realized variance, is included in the specification.

[4] Intraday periodicity has no impact on the realized variance when the spot variance is assumed constant throughout the day. However, with a time-varying spot variance, intraday periodicity may also impact the realized variance.

1 Theoretical background

Let $p(t)$ denote a logarithmic asset price at time t belonging to a special class of semimartingales with jumps:

$$\mathrm{d}p(t) = \mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t) + \mathrm{d}L(t), \quad t \in [0, T] \tag{1}$$

where $\mu(t)$ is a continuous and locally bounded drift term, $\sigma(t)$ is the spot volatility which is adapted and càdlàg. $W(t)$ is a one-dimensional standard Brownian motion, while $L(t)$ is a jump process. Without loss of generality, we assume that T is an integer, representing the number of trading days. All integers in $[0, T]$ mark the end of a trading day. The volatility at time t over the past day is given by the integrated variance, $IV_t = \int_{t-1}^{t} \sigma^2(u)\,\mathrm{d}u$.

Within each trading day, there are M observations, which, for now, we assume equally spaced, with the time interval between any two consecutive observations equal to $\Delta = \frac{1}{M}$.[5] Let $r_{t,i}$, $i = 1, ..., M$, be the i-th intraday return during the one-day interval $(t-1, t]$, such that $r_{t,i} = p(t-1+i\Delta) - p(t-1+(i-1)\Delta)$. In the absence of jumps, the integrated volatility is consistently estimated by the realized variance (Andersen and Bollerslev, 1998a, Andersen et al., 2003), defined as:

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2.$$

If the price contains jumps, RV_t is no longer consistent for the integrated variance, converging to the quadratic variance of the price process, $\int_{t-1}^{t} [\sigma^2(u) + L^2(u)]\,\mathrm{d}u$. To estimate the integrated variance, it becomes necessary to rely on a robust to jumps estimator, such as the realized bipower variation of Barndorff-Nielsen and Shephard (2004) given by:

$$BV_t = \frac{M}{M-1} 1.57 \sum_{i=2}^{M} |r_{t,i}||r_{t,i-1}|. \tag{2}$$

Barndorff-Nielsen and Shephard (2006a) build a statistical test comparing RV_t and the jump robust BV_t to infer whether jumps occur in the interval $(t-1, t]$. The test statistic is given in Appendix B.2, while Section IA.2 in the internet appendix confirms all our results for an alternative jump test.

Kolokolov and Renò (2024) show that price staleness (zero returns) has a great impact on realized power variations, and on the tests for jumps in prices using such measures.

[5] Note that defining realized volatility does not require equally spaced observations. We make this assumption here for simplicity.

They develop staleness-robust estimators, where the above equidistant sampling scheme is replaced by a stochastic scheme $i_1, \ldots, i_{N_M}$, such that returns r_{t,i_j}, $j = 1, \ldots, N_M$ are non-zero (Kolokolov and Renò, 2024).[6] Whenever our analysis involves identifying and estimating jump components, we implement the Kolokolov and Renò (2024) estimators.[7] In formulas, derivations and the simulation exercise, we use the classical equidistant sampling scheme throughout.

We define the intraday volatility periodicity, $f(t)$, as a deterministic function multiplying the actual spot volatility stochastic process, $s(t)$, as in Andersen and Bollerslev (1997, 1998b), Andersen et al. (2001), Boudt et al. (2011):

$$\sigma(t) = s(t)f(t), \tag{3a}$$

$$\text{such that} \int_{t-1}^{t} f^2(u)\,\mathrm{d}u = 1, \tag{3b}$$

so that intraday periodicity has no impact on the integrated variance when $s(t)$ is constant, i.e., $\int_{t-1}^{t} \sigma^2(u)\,\mathrm{d}u = \int_{t-1}^{t} s^2(u)\,\mathrm{d}u$. However, in a more realistic scenario where time-varying spot variance is considered, the latter equality does not always hold. In practice, as we observe a discrete number of observations, the condition in equation (3b) can be written using the following Riemann sum:

$$\Delta \sum_{i=1}^{M} f_i^2 = 1, \tag{4}$$

where f_i is the i-th value of the function $f(\cdot)$ observed during a trading day. Clearly, when Δ approaches 0, the Riemann sum converges to the integral in (3b).

The two components of spot volatility defined in (3a) differ greatly. The periodic component is a deterministic function of intraday time and reflects intraday trading patterns. The actual spot volatility $s(t)$ is a stochastic process which varies over time reflecting the available information on the asset.

To estimate the intraday periodicity, we use the nonparametric approach proposed by Boudt et al. (2011),[8] which is robust to the presence of jumps in the price process. Let $\bar{r}_{t,i} = \frac{r_{t,i}}{\sqrt{\Delta BV_t}}$, $i = 1, \ldots, M$, $t = 1, \ldots, T$ be the standardized intraday returns, with BV_t

[6] Kolokolov and Renò (2024) redefine the realized bipower variation as $\overline{\Delta_M} 1.57 \sum_{j=2}^{N_M} |\Delta_j{}^{-1/2} r_{t,i_j}||\Delta_{j-1}{}^{-1/2} r_{t,i_{j-1}}|$, where $\Delta_{j-1} = i_j - i_{j-1}$ are the durations between price changes, and $\overline{\Delta_M}$ is the average over these durations. Similar definitions apply to other realized power variations.

[7] We are grateful to Roberto Renò for providing us with the implementation code.

[8] Among the various applications of periodicity-filtered returns, the authors demonstrate that these filtered returns can also be employed to formulate a jump-robust estimate of the autocorrelation and long-memory in the absolute or squared filtered high-frequency return series.

given in equation (2). For a certain intraday time, i, we observe T standardized intraday returns. The intraday periodicity estimator is defined as:

$$\hat{f}_i^{WSD} = \frac{WSD_i}{\sqrt{\Delta \sum_{j=1}^{M} WSD_j^2}} \tag{5}$$
$$WSD_i = \sqrt{1.081 \frac{\sum_{l=1}^{T} \chi_{l,i} \bar{r}_{l,i}^2}{\sum_{l=1}^{T} \chi_{l,i}}},$$

for all $i = 1, \ldots, M$, where WSD_i is the weighted standard deviation (WSD) and $\chi_{l,i}$, $l = 1, \ldots, T$ are weights computed using the shortest half scale periodicity estimator and defined in appendix B.1. The periodicity-filtered returns are defined as:

$$r_{t,i}^f = \frac{r_{t,i}}{\hat{f}_i^{WSD}}. \tag{6}$$

We can further define periodicity-filtered realized measures, such as the filtered realized variance $RV_t^f = \sum_{i=1}^{M} (r_{t,i}^f)^2$, the filtered realized bipower variation $BV_t^f = \frac{M}{M-1} 1.57$ $\sum_{i=2}^{M} |r_{t,i}^f||r_{t,i-1}^f|$ and so on.

2 Forecasting realized volatility in the presence of periodicity

For a simple DGP, we consider the impact of intraday periodicity on the forecasting regression, comparing the resulting coefficients with those obtained in the absence of periodicity.

2.1 The simple AR(1) model

We assume the daily integrated variance (IV) evolves according to an AR(1) process.

$$IV_t = \Theta + \Phi IV_{t-1} + \epsilon_t, \tag{7}$$

where $t \in \{1, 2, ..., T\}$, $\Theta > 0$, $|\Phi| < 1$, and ϵ_t is i.i.d. with $\text{Var}(\epsilon_t) = \sigma_\epsilon^2$. In addition, within each trading day, the actual spot volatility remains constant at a level equal to a fraction of the daily integrated variance, ΔIV_t. If we also account for intraday periodicity, the spot volatility for the i-th Δ-length window during a trading day equals $\Delta IV_t f_i^2$. Assuming no drift, the i-th return is $r_{t,i} = \sqrt{\Delta IV_t} f_i w_i$, where w_i is i.i.d. $\mathcal{N}(0,1)$ and independent of present and past values of $s(\cdot)$. Suppose one attempts to forecast volatility using the following AR(1) model for the realized variance:

$$RV_t = \theta + \phi RV_{t-1} + e_t, \tag{8}$$

with ϕ equal to the well-known formula:

$$\phi = \frac{\text{Cov}\left(RV_t, RV_{t-1}\right)}{\text{Var}\left(RV_t\right)}. \tag{9}$$

In the above equations, as RV_t is only a proxy for the integrated variance, it is subject to measurement error, leading to an attenuation bias in the estimate of ϕ (Bollerslev et al., 2016). Below, we show that periodicity further increases this bias, resulting in a further reduction –in absolute value– in the ϕ estimate.

To assess the impact of periodicity on the value of ϕ, we compute the numerator and denominator in equation (9) in the presence/absence of periodicity. The required derivations are enclosed in section A.1 of the appendix. While the auto-covariance remains unaffected by periodicity, we obtain the following variance formulae for the case in which periodicity is present (equation (10a)), compared to the case when it is absent (equation (10b)):

$$\text{Var}(RV_t) = Var(IV_t) + 2\Delta E(IV_t^2)\Delta \sum_{i=1}^{M} f_i^4, \tag{10a}$$

$$\text{Var}(RV_t)^{NP} = Var(IV_t) + 2\Delta E(IV_t^2), \tag{10b}$$

where the superscript NP above stands for "no periodicity".

The main difference between the above formulae resides in the term $\sum_{i=1}^{M} f_i^4$. From equation (4), we know that $\sum_{i=1}^{M} f_i^2 = \Delta^{-1} > 1$. Then, $\sum_{i=1}^{M} f_i^4 > \sum_{i=1}^{M} f_i^2 = \Delta^{-1}$, so that $\text{Var}(RV_t) > \text{Var}(RV_t)^{NP}$.

The biased coefficient, $\phi = \Phi\left(1 + \frac{2\Delta E(IV_t^2)}{Var(IV_t)}\Delta \sum_{i=1}^{M} f_i^4\right)^{-1}$ (for derivations, see Appendix A.1), is lower –in absolute value– than the corresponding coefficient for the case of no periodicity. ϕ understates the true correlation coefficient, Φ, for two reasons. First, the presence of measurement error leads to the variance distortion in (10b), pushing ϕ downwards from Φ. Second, as shown in equation (10a), the presence of periodicity generates a further increase in the variance of realized volatility, further reducing ϕ.

A bias correcting solution in the spirit of Bollerslev et al. (2016) would be running a regression with a time varying AR(1) parameter: $RV_t = \theta + \left[\phi + \phi_1 \Delta \sqrt{\widehat{E(IV_t^2)} \sum_{i=1}^{M} \hat{f}_i^4}\right] \cdot RV_{t-1} + u_t$. This correction would require pre-estimating periodicity, $\hat{f}_i$, $i = 1, \ldots, M$, in addition to estimating the integrated quarticity $IQ_t = \Delta E(IV_t^2)$, which featured in the above-mentioned work. If we wish to extend to an AR(p) model (p$>$ 1), the issue of correcting some or all regressors adds additional layers of complications.

Instead of applying a correction for periodicity, we propose using regressors based on returns pre-filtered for periodicity. We thus define an ARP(1) model ("P" stands for

periodicity-filtered):

$$RV_t = \theta^f + \phi^f {RV^f}_{t-1} + \epsilon_t^f, \tag{11}$$

where RV_{t-1}^f is the lagged RV based on $r_{t,i}^f$, specified in (6). Appendix A.2 shows that if the number of days in the sample, T, is high enough, the variance of the regressor decreases, leading to a decrease in the variance of the error. In the above regression, the dependent variable is unfiltered, while the regressor is filtered. As a result, the model in equation (11) is not strictly an autoregression in RV. It is an autoregression in integrated volatility proxies, where the proxies used on the two sides of the equal sign are different.

2.2 HAR models

As integrated volatility is likely a long memory process, HAR models are widely preferred for its forecast when using high frequency data. Such models rely not only on recent RV lags, but also on cumulated weekly and monthly realized variances. Let h be the forecasting horizon, measured in days. Then, $RV_{t,t+h-1}$ is the forecasted realized variance over the next h days (starting from day t). The HAR model is defined as (Corsi, 2009):

$$\textbf{HAR} \quad RV_{t,t+h-1} = \beta_0 + \beta_d RV_{t-1} + \beta_w RV_{t-5,t-1} + \beta_m RV_{t-22,t-1} + \epsilon_{t,t+h-1}, \tag{12}$$

where RV_{t-1} is the first lag of the (daily) realized variance, $RV_{t-5,t-1}$ is the average realized variance over the past week and $RV_{t-22,t-1}$ the average realized variance over the past month. In addition, $\epsilon_{t,t+h-1}$ is the forecasting error, β_0 the regression constant term, while β_d, β_w and β_m are the coefficients corresponding to the one-day, one-week and one-month lagged values of the realized variance.

Defining RV_{t-1}^f, $RV_{t-5,t-1}^f$ and $RV_{t-22,t-1}^f$ as the periodicity-filtered one-day, one-week and one-month lagged realized variances, we can extend the ARP model to HARP:[9]

$$\textbf{HARP} \quad RV_{t,t+h-1} = \beta_0 + \beta_d RV_{t-1}^f + \beta_w RV_{t-5,t-1}^f + \beta_m RV_{t-22,t-1}^f + \epsilon_{t,t+h-1}^f, \tag{13}$$

with $\epsilon_{t,t+h-1}^f$ the forecasting error.

In addition to the HAR/HARP models above, we also analyse three other HAR-type models and their periodicity-filtered counterparts: the HAR-J and HAR-CJ models by Andersen et al. (2007a), and the HAR-Q model by Bollerslev et al. (2016). Below, we present the forecasting regressions for these three HAR models and their HARP counterparts.

[9] It is debatable whether the ARP/HARP specifications can be considered new models or a sub-class of generic AR/HAR models for measures of integrated volatility. We take the liberty to refer to HARP as models for simplicity and to the extent the literature generally refers to all HAR variations as models.

HAR-J and HARP-J

Andersen et al. (2007a) define the contribution of jumps to the daily quadratic variation of the price as $J_t = \max(RV_t - C_t, 0)$, for $t = 1, \ldots, T$, where C_t is a consistent estimator of the integrated variance. Similarly, we can define a jump regressor based on periodicity-filtered returns, J_t^f. The HAR-J and HARP-J models are obtained by including J_{t-1} and J_{t-1}^f, respectively, in the forecasting regression:

$$\text{HAR-J} \quad RV_{t,t+h-1} = \beta_0 + \beta_d RV_{t-1} + \beta_w RV_{t-5,t-1} + \beta_m RV_{t-22,t-1} + \beta_J J_{t-1} + \epsilon_{t,t+h-1}, \tag{14}$$

$$\text{HARP-J} \quad RV_{t,t+h-1} = \beta_0 + \beta_d RV_{t-1}^f + \beta_w RV_{t-5,t-1}^f + \beta_m RV_{t-22,t-1}^f + \beta_J J_{t-1}^f + \epsilon_{t,t+h-1}^f. \tag{15}$$

HAR-CJ and HARP-CJ

In this model, past lags of the estimated continuous and discontinuous components of the quadratic variation are considered in the forecasting regression, as follows:

$$\begin{aligned} \text{HAR-CJ} \quad RV_{t,t+h-1} &= \beta_0 + \beta_{C_d} C_{t-1} + \beta_{C_w} C_{t-5,t-1} + \beta_{C_m} C_{t-22,t-1} + \beta_{J_d} J_{t-1} + \\ &\quad \beta_{J_w} J_{t-5,t-1} + \beta_{J_m} J_{t-22,t-1} + \epsilon_{t,t+h-1}, \end{aligned} \tag{16}$$

$$\begin{aligned} \text{HARP-CJ} \quad RV_{t,t+h-1} &= \beta_0 + \beta_{C_d} C_{t-1}^f + \beta_{C_w} C_{t-5,t-1}^f + \beta_{C_m} C_{t-22,t-1}^f + \beta_{J_d} J_{t-1}^f + \\ &\quad \beta_{J_w} J_{t-5,t-1}^f + \beta_{J_m} J_{t-22,t-1}^f + \epsilon_{t,t+h-1}^f, \end{aligned} \tag{17}$$

where C_{t-1}, $C_{t-5,t-1}$ and $C_{t-22,t-1}$ are the one-day, one-week and one-month lagged estimates of the integrated variance, and J_{t-1}, $J_{t-5,t-1}$ and $J_{t-22,t-1}$ are the one-day, one-week and one-month lagged estimates of the jumps' contribution to the quadratic variation. In equation (17), all these regressors are computed on periodicity-filtered returns, hence the ^{f} superscript. In computing C_t and C_t^f, we employ the method in Andersen et al. (2007a): $C_t = RV_t \cdot \mathbb{I}_t(\text{no jumps}) + BV_t \cdot \mathbb{I}_t(\text{jumps})$, for $t = 1, \ldots, T$, where $\mathbb{I}_t(\cdot)$ is the indicator function for whether jumps were identified on day t or not. To test for jumps we rely on the BNS test of Barndorff-Nielsen and Shephard (2006a),[10] at the 1% significance level.

HAR-Q and HARP-Q

As Bollerslev et al. (2016) indicate, the variance of the realized volatility measurement error is a function of the integrated quarticity, $\int_{t-1}^{t} \sigma^4(u)\,\mathrm{d}u$, $t = 1, \ldots, T$. Their main forecasting model accounts for the error in measuring the one-day lagged realized

[10] This test is outlined in Appendix B.2.

variance,[11] as follows:

$$\text{HAR-Q} \quad RV_{t,t+h-1} = \beta_0 + (\beta_d + \beta_{dQ} RQ_{t-1}^{1/2}) RV_{t-1} + \beta_w RV_{t-5,t-1} + \beta_m RV_{t-22,t-1} + \epsilon_{t,t+h-1}, \tag{18}$$

$$\text{HARP-Q} \quad RV_{t,t+h-1} = \beta_0 + (\beta_d + \beta_{dQ} (RQ_{t-1}^{f})^{1/2}) RV_{t-1}^{f} + \beta_w RV_{t-5,t-1}^{f} + \beta_m RV_{t-22,t-1}^{f} + \epsilon_{t,t+h-1}^{f}, \tag{19}$$

where $RQ_{t-1} = \frac{M}{3} \sum_{i=1}^{M} r_{t-1,i}^4$ estimates the integrated quarticity using unfiltered data, while RQ_{t-1}^{f} is its counterpart estimate based on periodicity-filtered data.

3 Simulation evidence

3.1 Simulation set-up

The starting point for our simulation is the one-factor stochastic volatility model previously analyzed by Huang and Tauchen (2005) and given in equation (20a).[12] To this model, we add compound Poisson jumps in the price, accompanied by co-jumps in volatility.[13] For the intraday periodicity function, $f(t)$, we employ the specification in Andersen et al. (2012) and Hasbrouck (1999):

$$\begin{aligned}
\mathrm{d}p(t) &= 0.03\,\mathrm{d}t + f(t)\nu(t)\left(-0.62\,\mathrm{d}W_{\nu_1}(t) + \sqrt{0.6156}\,\mathrm{d}W_p(t)\right) + z_p(t)\,\mathrm{d}N(t), \\
\nu^2(t) &= \exp\{0.125\nu_1^2(t)\}, \\
\mathrm{d}\nu_1^2(t) &= -0.1\nu_1^2(t)\,\mathrm{d}t + \mathrm{d}W_{\nu_1}(t) + z_{\nu_1}(t)\,\mathrm{d}N(t), \\
N(t) &\sim \text{Poisson}(0.4\text{t}), \\
z_p(t), z_{\nu_1}(t) &\sim \mathcal{N}(0, 1.284).
\end{aligned} \tag{20a}$$

$$f(t) = 0.88929198 + 0.75e^{-10t} + 0.25e^{-10(1-t)}, \tag{20b}$$

where W's are correlated standard Brownian motions, $N(t)$ a counting process, while $z_p(t)$ and $z_{\nu_1}(t)$ are the respective jump sizes for price and volatility, independent of each

[11] Authors explain that measurement errors for the one week and one month realized volatilities do not have a significant impact on forecasting.

[12] Section IB. , in the internet appendix, provides some results for some alternative DGPs. Section IB.5 uses some of these alternative DGPs to perform some robustness checks for our analysis.

[13] The variance of both the price and volatility jump sizes is set to 1.284, accounting for approximately 30% of the quadratic variation in the model when no jumps are present. This number corresponds to $\exp(0.125^2)$, where the value 0.125 is the coefficient multiplying the volatility factor in equation (20a).

other. The parameters in $f(t)$ are calibrated such that the integrated squared periodicity function equals 1.

In the absence of jumps, the model volatility is predictable and moderate, not leading to a large number of extreme returns. When jumps are added to the price, it generates unpredictable extreme returns that are independent of the volatility dynamics. However, when volatility is allowed to co-jump with the price, we generate both extreme and persistent volatility along with extreme returns. This occurs because jumps in volatility propagate forward through the autoregressive specification in ν_1^2. This model adequately characterizes both tranquil periods in the absence of jumps, as well as turbulent periods due to the co-jump feature.

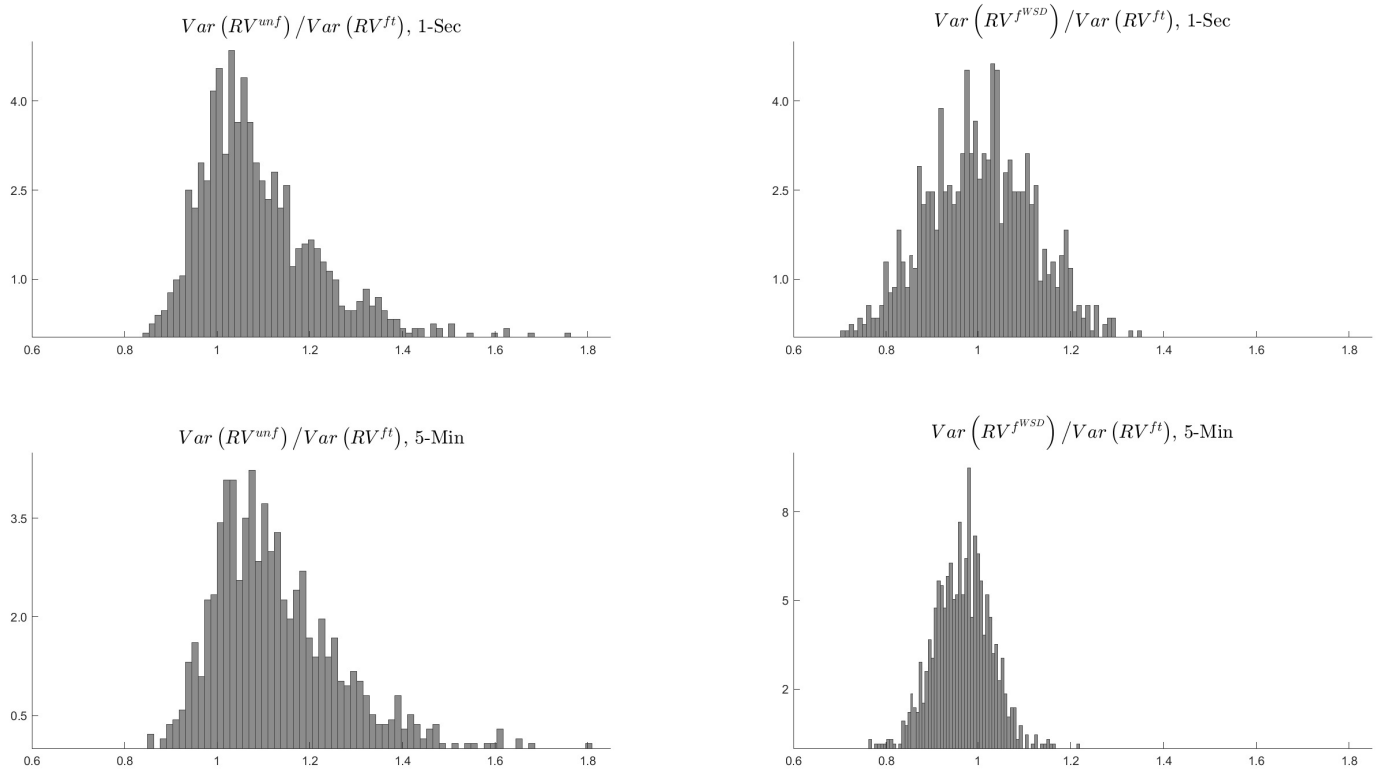
Simulations are generated using an Euler scheme based on 23,400 initial data points (corresponding to seconds). We further aggregate data up to the following lower sampling frequencies: 5 seconds (4680 observations), 30 seconds (780 observations), 1 minute (390 observations), 1.5 minutes (260 observations), 2 minutes (195 observations), 2.5 minutes (156 observations), 5 minutes (78 observations), 10 minutes (39 observations), 15 minutes (26 observations) and 30 minutes (13 observations). We simulate a total of 1,000 sample paths of length 2,000 days.

3.2 Impact of filtering on the realized volatility

The simple AR(1) exercise in Section 2.1 showed that the presence of periodicity inflated the variance of the realized volatility. Here, we investigate this effect, along with the impact of filtering out periodicity, using the more complex simulated price process described in Equation (20a). Let $\text{Var}(RV_t^{unf})$, $\text{Var}(RV_t^{ft})$ and $\text{Var}(RV_t^{WSD})$ be the variances of the realized volatility estimators respectively based on: unfiltered returns, returns filtered by the true periodicity, and returns filtered with the weighted standard deviation method as shown in Section 1. Figure 1 shows the histograms of the ratios $\text{Var}(RV_t^{unf})/\text{Var}(RV_t^{ft})$ and $\text{Var}(RV_t^{WSD})/\text{Var}(RV_t^{ft})$ computed on simulated returns. We consider two sampling frequencies: 1-second, the frequency at which data is generated, and 5-minute, the standard sampling frequency used in applications.

Both plots on the left show that when periodicity is present, the distribution of $\text{Var}(RV_t^{unf})/\text{Var}(RV_t^{ft})$ is almost entirely shifted to the right of 1, suggesting that the realized volatility variance increases substantially when periodicity is present. The plots on the right show that filtering out periodicity using the weighted standard deviation method is on average beneficial, as the distributions of $\text{Var}(RV_t^{WSD})/\text{Var}(RV_t^{ft})$ for both sampling frequencies are centered close to 1.

Figure 1: Simulated impact of periodicity on the unconditional variance of RV

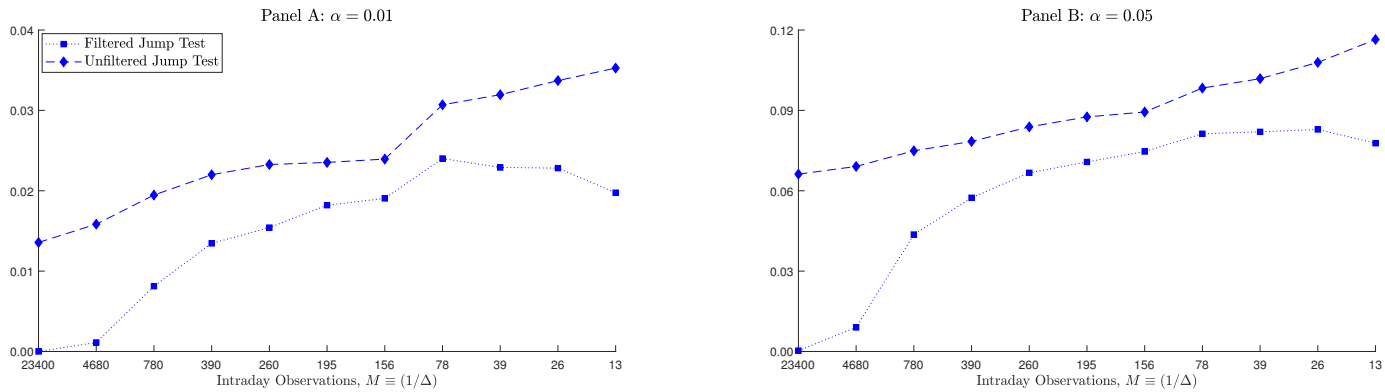


Note: This figure plots the distributions of the ratios $\text{Var}(RV_t^{unf})/\text{Var}(RV_t^{ft})$ and $\text{Var}(RV_t^{WSD})/\text{Var}(RV_t^{ft})$ for simulated returns sampled every second and every 5 minutes. $\text{Var}(RV_t^{unf})$, $\text{Var}(RV_t^{ft})$ and $\text{Var}(RV_t^{WSD})$ are the variances of the realized volatility estimators based on, respectively, unfiltered returns, returns filtered by the true periodicity, and returns filtered with the weighted standard deviation method as shown in Section 1.

3.3 Impact via jump regressors

Two of the most popular HAR models, HAR-J (equation (14)) and HAR-CJ (equation (16)), use the estimated daily squared jumps as predictors. These estimates depend on the outcome of jump tests that decide whether jumps in prices have occurred during a particular trading day. Previous literature has already documented that the presence of periodicity is likely to interfere with jump detection (Boudt et al., 2011, 2012, Dette et al., 2023). Figure 2 plots the proportion of spurious jumps identified for simulated returns and their filtered counterparts, where we relied on 1% and 5% significance levels in jump testing.

Figure 2: Simulated proportion of spurious jumps by sampling frequency for filtered and unfiltered data



Note: This figure depicts the proportion of spurious jumps across sampling frequencies. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

The figure illustrates that a higher number of spurious jumps are detected in unfiltered returns, a finding that holds true across all sampling frequencies. This suggests that jump regressors in HAR-J and HAR-CJ models are likely affected by periodicity-related estimation errors, which can further impact the accuracy of the realized variance forecasts.

3.4 HARP forecasting performance

In this section, we use simulated data to compare the forecasting performance of the HARP models to that of the HAR models. To evaluate the forecasting performance of the two classes of models, we use two distinct loss functions, the mean squared error (MSE) and the quasi-likelihood (QLIKE) loss, defined in equation (21) below:

$$\begin{aligned} MSE(RV_t, F_t) &= (RV_t - F_t)^2 \\ QLIKE(RV_t, F_t) &= \frac{RV_t}{F_t} - \log \frac{RV_t}{F_t} - 1, \end{aligned} \tag{21}$$

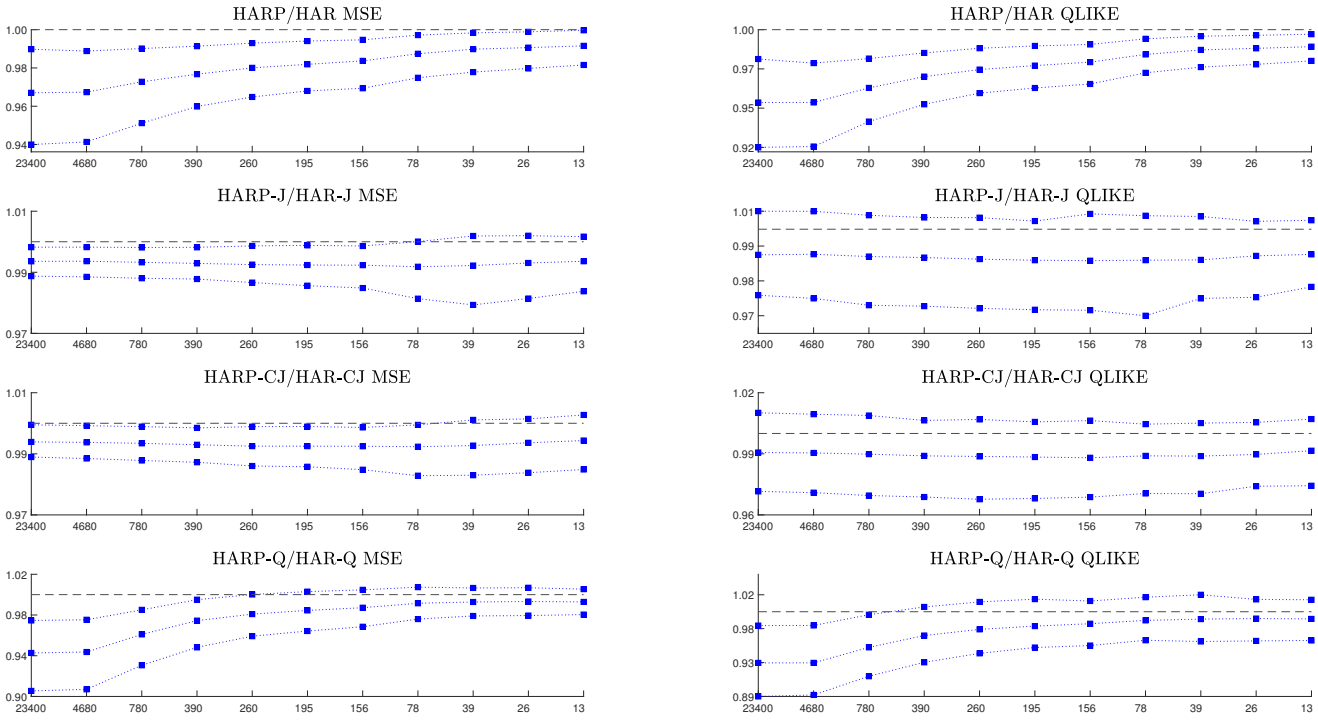
where F_t denotes the out-of-sample forecast of the realized variance.

For forecast horizons beyond 1-day, both HARP and HAR models are adapted to the new time scale via direct projection, replacing the daily RVs on the left-hand-side with the weekly and monthly RVs. Separate models are fitted for each forecasting horizon. We compute out-of-sample forecasts, re-estimating the models daily using a rolling window of 1000 days. For each forecasting model introduced in Section 1 and each forecast horizon, we calculate the ratio of forecast losses for the HARP version of the model versus the

HAR model.[14] A ratio below one signals the superiority of the model based on filtered returns.

Figure 3 illustrates the median and the 5% and 95% quantiles of the one-day ahead forecast loss ratios from the HARP models versus the HAR models against various sampling frequencies. In most cases, all quantiles consistently fall below 1 across the range of sampling frequencies. Furthermore, regardless of the forecast horizon, model, or sampling frequency, at least 70% of the loss ration distribution remains below 1. These findings are also corroborated for forecast loss ratios based on one-week and one-month horizons. We present these additional figures in Section IB.4 of the internet appendix.

Figure 3: One-day ahead loss ratio for the simulated model



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

[14]Days with jumps are estimated using the BNS test proposed by Barndorff-Nielsen and Shephard (2006a) at the 1% significance level. This test is outlined in Appendix B.2.

4 Empirical evidence

4.1 Data

We use intraday price data from the TickData database for the SPDR S&P 500 ETF (SPY) and 30 individual stocks in the S&P 500 basket. The sample period spans from January 2000 to December 2020 (5,284 trading days). The selection criteria for individual stocks are as follows: (i) we only consider stocks that are continuously traded over the sample period; (ii) we require these stocks to be very liquid to mitigate biases arising from price staleness; to mitigate potential biases from price staleness, in each sector we select stocks ranking in the top 15th percentile by trading volume. This ensures that the selected stocks also fall within the bottom 20th percentile in terms of the proportion of zero returns.[15] (iii) to account for scenarios where intraday periodicity could exhibit sector-specific patterns, we select at least one representative stock from each sector as classified by the Global Industry Classification Standard (GICS), ensuring they meet conditions (i) and (ii). This yields the following sectorial composition: Industrials (6 stocks), Materials (2 stocks), Consumer Staples (3 stocks), Financials (2 stocks), Energy (2 stocks), Real Estate (1 stock), Consumer Discretionary (4 stocks), Communication Services (2 stocks), Health Care (3 stocks), Utilities (2 stocks), and Information Technology (3 stocks).

Data is aggregated from tick level using previous tick interpolation and then sampled every 5 minutes. This sampling frequency is standard in the literature, motivated by the trade-off between bias and variance (for more details, see Aït-Sahalia et al., 2005, Hansen and Lunde, 2006).[16]

As recent literature suggests, intraday periodicity, which is likely time-varying (Andersen et al., 2018), is computed for each day t based on data from the previous 20 trading days.[17] This approach covers a reasonable degree of variation in the periodicity function from one day to another, with the results below showing improved performance of HARP models with respect to HAR models. It is unclear, though, how our results would change

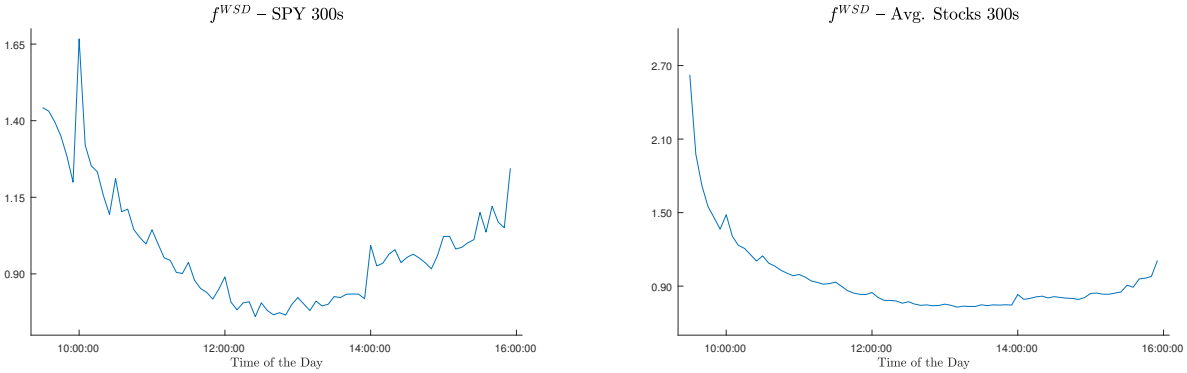
[15] These stocks are usually Dow Jones or S&P 100 constituents. For more details regarding the stocks, please see Table 1. From those stocks that satisfy the criteria, we have implemented a random selection process. However, it is important to bear in mind that stocks from the same sector, which meet the criteria, typically exhibit a return correlation of around 0.35 to 0.50, and the correlation among their variances oscillates around 0.65 to 0.90. This suggests that the results should hold if one replaces one of our stocks with another stock that satisfies the criteria and belongs to the same sector.

[16] Liu et al. (2015) and Bu et al. (2023) show that 5-min RV is generally very difficult to beat, corroborating previous findings in the literature regarding the good variance-bias trade-off afforded by 5-minute RV.

[17] This approach also avoids any potential forward-looking bias. In addition, for simplicity, the in-sample exercise is carried out using the whole sample intraday periodicity estimate.

if the true periodicity function varied drastically from day to day.[18] Figure 4 plots the estimated periodicity for SPY and the average estimated periodicity for the 30 S&P 500 stocks considered. Both plots reveal the characteristic U-shape for the estimated curve.

Figure 4: Intraday estimated periodicity for SPY (left) and average periodicity for all stocks (right).



Note: The panel on the left shows the estimated periodicity for SPY, while the panel on the right shows the average estimated periodicity for the 30 S&P 500 stocks considered. Periodicity was estimated using all available data and a 5-minute sampling frequency.

Table 1 reports, for each ticker in our sample, the minimum, maximum and median values of the realized variance, the number of jumps detected and the estimated proportion of the continuous component relative to the total RV. Days with jumps, in both panels, are estimated at the 1% significance level using the staleness-robust Barndorff-Nielsen and Shephard (2006a) procedure proposed by Kolokolov and Renò (2024). The left (right) panel of the table reports these statistics for the unfiltered (filtered) return data.

For SPY, we detect 324 jumps for unfiltered data, meaning that we identify jumps on 6.13% of days. When data is filtered, the number of jumps drops to 271, suggesting that 5.12% days had jumps.[19] Results for individual stocks show high variability in the number of jumps identified for both filtered and unfiltered data. On average, we observe 504 jumps for the unfiltered data, which decreases substantially after filtering to 183. As shown in Section 3.3 above, the presence of intraday periodicity can lead to spurious jump detection.

[18] In the internet appendix, we also report results for intraday periodicity estimated using sampling windows' sizes equivalent to 40, 125, 250, 500, and 1,000 trading days.

[19] The percentage of days with jumps are very close to the total number of FOMC meetings and related macroeconomic announcements. In addition, we found that the Covid-19 period concentrates about 15% of these jumps.

Table 1: Realized variance minimum, maximum and median, realized number of jumps and the estimated proportion of integrated variance in the quadratic variation for SPY and 30 stocks

			Unfiltered					Filtered				
Company	Ticker	Sector	Min RV	Max RV	Med RV	# Jumps	%QV	Min RV	Max RV	Med RV	# Jumps	%QV
SPDR ETF	SPY	ETF Index	0.010	59.863	0.433	324	98.773	0.010	53.325	0.432	271	98.984
3M	MMM	IND	0.082	91.955	0.993	529	96.532	0.084	88.370	0.981	178	98.152
Air Product&Chem	APD	MAT	0.134	97.709	1.301	555	95.664	0.126	113.489	1.291	209	97.908
American Tower	AMT	REIT	0.165	1048.657	1.684	559	95.639	0.165	1304.046	1.637	204	97.037
Brown-Forman Corp	BFB	CS	0.074	998.842	1.182	599	92.861	0.102	254.205	1.168	316	96.971
Citigroup	C	FIN	0.130	975.858	1.901	409	97.165	0.093	997.884	1.896	150	97.818
Coca-Cola	KO	CS	0.046	73.789	0.770	548	95.932	0.064	66.469	0.752	193	98.056
Conoco Phillips	COP	EN	0.165	200.297	1.751	436	96.862	0.171	151.894	1.738	163	98.251
Duke Energy	DUK	UT	0.051	189.935	1.061	572	96.315	0.057	197.351	1.039	185	98.275
eBay	EBAY	CD	0.134	236.419	2.432	528	97.383	0.123	361.374	2.421	186	98.283
General Dynamics	GD	IND	0.081	63.282	1.225	535	95.873	0.064	75.786	1.201	197	97.948
General Electric	GE	IND	0.108	180.389	1.516	445	97.054	0.101	139.008	1.494	198	98.004
Halliburton	HAL	EN	0.229	372.677	3.443	438	96.103	0.207	379.735	3.414	134	97.661
Home Depot	HD	CD	0.156	103.477	1.407	450	97.041	0.142	100.545	1.382	148	98.358
Honeywell	HON	IND	0.082	268.331	1.419	504	96.432	0.102	161.226	1.403	164	97.537
Howmet Aerospace	HWM	IND	0.280	291.089	2.959	443	97.081	0.191	209.633	2.910	215	98.151
Humana	HUM	HC	0.174	194.898	2.442	541	96.400	0.170	309.298	2.353	166	98.171
Intel	INTC	IT	0.154	107.011	1.886	497	97.871	0.153	101.833	1.833	204	98.135
Interpublic Group	IPG	MAT	0.229	615.148	2.412	501	93.063	0.190	155.507	2.411	181	97.689
McDonald's	MCD	CD	0.087	161.156	0.968	509	94.988	0.087	106.360	0.956	139	98.281
Microsoft	MSFT	IT	0.078	62.386	1.328	453	97.606	0.053	80.630	1.300	191	98.200
Pfizer	PFE	HC	0.150	64.478	1.271	496	95.664	0.137	70.559	1.272	165	98.183
Procter&Gamble	PG	CS	0.101	82.119	0.742	559	96.129	0.090	127.868	0.736	150	98.351
Southern Co.	SO	UT	0.092	97.041	0.917	482	96.438	0.109	84.671	0.881	144	98.250
Starbucks	SBUX	CD	0.169	89.444	1.909	541	96.689	0.155	99.210	1.898	181	98.154
Travelers C Inc	TRV	FIN	0.102	263.929	1.108	582	95.474	0.106	229.683	1.107	181	98.827
United Health	UNH	HC	0.129	225.956	1.637	556	95.460	0.145	135.383	1.656	183	98.838
UPS	UPS	IND	0.081	216.939	0.896	533	95.732	0.059	118.149	0.888	189	97.850
Verizon	VZ	COMS	0.122	102.221	1.066	534	95.843	0.106	102.236	1.048	166	98.955
Vodafone	VOD	COMS	0.110	70.936	0.844	268	98.747	0.075	81.745	0.838	207	99.196
Xerox	XRX	IT	0.230	412.153	2.684	506	95.018	0.182	352.620	2.657	217	98.109
		Avg. Stocks	0.131	265.284	1.572	504	96.169	0.120	225.226	1.552	183	98.120

Note: The table reports the descriptive statistics for the RV of the 30 individual stocks and SPY estimated at the 300 second frequency. The %QV is estimated as $\%QV = \frac{\sum_{t=1}^{T} C_t}{\sum_{t=1}^{T} (C_t + J_t)}$, while # jumps indicates the total number of days with jumps estimated at the 1% significance level using the staleness-robust Barndorff-Nielsen and Shephard (2006a) procedure (Kolokolov and Renò, 2024). The sector column shows, for each stock, the corresponding industry group based on the Global Industry Classification Standard (GICS). Industrials (IND), Materials (MAT), Real Estate (REIT), Consumer Staples (CS), Financials (FIN), Energy (EN), Utilities (UT), Consumer Discretionary (CD), Health Care (HC), Information Technology (IT), and Communication Services (COMS).

4.2 In-sample forecasting results

Tables 2, 3, 4 and 5 report the regression results for all HAR and HARP models, estimated on the entire sample, for SPY and a stock average.[20] Estimated standard errors are robust to heteroscedasticity and autocorrelation, as we allow for serial correlation of up to orders 5, 10 and 44 for the 1-day, 5-days and 22-days models, respectively. We compute both in-sample and out-of-sample R-squared coefficients, reported as R^2_{is} and R^2_{oos}, where the computation of R^2_{oos} is based on Campbell and Thompson (2007) and

[20]Section IA.4 of the internet appendix also reports results for the Rivers and Vuong (2002) test, comparing the in-sample performance of HARP and HAR models for SPY. Results indicate HARP models generally outperform their HAR counterparts.

uses over 3,000 observations.[21]

All tables show that for both SPY and the average stock, R^2_{oos} from HARP models is uniformly higher than for the HAR models, while R^2_{is} is higher for all but one case (1-day forecast with the HARPQ model). For SPY, the majority of the coefficients' standard errors are lower following filtering. At the same time, the levels of persistence in the residuals of the estimated HARP models are much lower than for the HAR models. These results confirm that HARP models are generally better specified and have a greater goodness of fit in comparison to HAR models.

Across all models, β_d (β_{C_d}) decreases as a result of filtering, while β_w (β_{C_w}) increases. Instead, the behaviour of the monthly coefficients varies with the model considered. All tables report for SPY $\beta_d+\beta_w+\beta_m$ ($\beta_{C_d}+\beta_{C_w}+\beta_{C_m}$ for the HAR-CJ model), representing the level of persistence. While all models show a very high degree of persistence, we observe lower levels for the HARP and HARP-Q models over all horizons compared to their unfiltered counterparts. On the contrary, the models including jump regressors see an increase in persistence after filtering. For the average stock, persistence decreases following filtering.

At first glance, a decrease in persistence alongside an increase in both in-sample and out-of-sample R^2 may appear puzzling. Moreover, results for the simple AR(1) in Section 2.1 above indicated that periodicity decreased persistence and filtering it out was likely to restore it. The puzzling effect observed for lots of HARP models is simply the result of having more than one explanatory variable. Consider the variance of the error for the HAR model in equation (12) and its HARP counterpart:

$$\begin{aligned}\mathrm{Var}(\epsilon_{t,t+h-1}) = \mathrm{Var}(RV_{t,t+h-1}) - \beta_d^2\,\mathrm{Var}(RV_{t-1}) - \beta_w^2\,\mathrm{Var}(RV_{t-5,t-1}) - \beta_m^2 Var(RV_{t-22,t-1}) - \\ 2\beta_d\beta_w\,\mathrm{Cov}(RV_{t-1}, RV_{t-5,t-1}) - 2\beta_d\beta_m\,\mathrm{Cov}(RV_{t-1}, RV_{t-22,t-1}) - \\ 2\beta_m\beta_w\,\mathrm{Cov}(RV_{t-22,t-1}, RV_{t-5,t-1})\end{aligned}$$

Assuming $\mathrm{Var}(RV_t)$ does not change much with time, we have:

$$\mathrm{Var}(\epsilon_{t+h-1}) = \mathrm{Var}(RV_t) - \mathrm{Var}(X)(\beta_d^2 + \beta_w^2 + \beta_m^2) - C,$$

where $\mathrm{Var}(X) = \mathrm{Var}(RV_t)$ for HAR and $\mathrm{Var}(X) = \mathrm{Var}(RV_t^f)$, while C denotes the sum over all covariance terms. C does not change greatly in the presence or absence of periodicity for two reasons: first, covariances are either not impacted or less impacted by periodicity; second, products of coefficients with values in $(0, 1)$ are unlikely to vary

[21] Section IA.2 in the internet appendix reports, as a robustness check, the results obtained for the HAR(P)-J and HAR(P)-CJ models where the Andersen et al. (2012) test is used to identify jumps.

greatly with periodicity. Given that $\text{Var}(RV_t^f) \leq \text{Var}(RV_t)$ for big T, the decrease in the variance of the error after filtering can only be due to an increase in $\beta_d^2 + \beta_w^2 + \beta_m^2$. Such an increase happens even if the persistence diminishes. Take for example the HAR and HARP models in Table 2 below with $h = 1$. We observe a decrease in persistence following filtering from 0.887 to 0.875. The sum of the squared coefficients is $0.268^2 + 0.522^2 + 0.097^2 = 0.354$ for HAR and $0.231^2 + 0.549^2 + 0.095^2 = 0.364$ for HARP, ensuring an overall improved performance. This is possible because filtering in a HAR model increases the coefficient with the highest value, β_w, which has the biggest impact in the sum of squares. At the same time, this coefficient is also estimated more precisely after filtering, with its standard error decreasing for all forecasting models and horizons.

Table 2: Estimated 1-, 5-, and 22-day ahead HAR(P) models for SPY and stocks average

	HAR			HARP		
	$h = 1$	$h = 5$	$h = 22$	$h = 1$	$h = 5$	$h = 22$
β_0	0.112***	0.183***	0.375***	0.110***	0.180***	0.372***
s.e.	(0.042)	(0.049)	(0.092)	(0.041)	(0.048)	(0.092)
β_d	0.268***	0.242***	0.147***	0.231***	0.200***	0.127***
s.e.	(0.091)	(0.069)	(0.038)	(0.083)	(0.062)	(0.039)
β_w	0.522***	0.427***	0.316***	0.549***	0.473***	0.336***
s.e.	(0.114)	(0.099)	(0.090)	(0.103)	(0.097)	(0.086)
β_m	0.097	0.147	0.159	0.095	0.131	0.152
s.e.	(0.072)	(0.097)	(0.108)	(0.075)	(0.094)	(0.103)
R^2_{is}	0.546	0.619	0.452	0.550	0.625	0.465
R^2_{oos}	0.479	0.555	0.440	0.502	0.581	0.450
$\beta_d + \beta_w + \beta_m$	0.887	0.816	0.622	0.875	0.805	0.615
Average Stocks						
$\overline{R}^2_{is}$	0.464	0.560	0.478	0.479	0.576	0.487
$\overline{R}^2_{oos}$	0.410	0.493	0.409	0.425	0.513	0.420
$\beta_d + \beta_w + \beta_m$	0.861	0.798	0.649	0.810	0.747	0.602

Note: This table reports the regression coefficients, standard errors in parentheses, and in- and out-of-sample R-squared for the HAR and HARP models based on various horizons, estimated on SPY data. The standard errors are estimated using the Newey-West HAC estimator. The bottom panel shows the stock average in- and out-of-sample R-squared obtained for HAR and HARP models of various horizons. *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

For the HARP model, standard errors for the constant and the daily coefficient are lower for the 1-day ahead and 1-week ahead forecasting horizons, while standard errors for the weekly coefficients are uniformly smaller. β_m decreases following filtering and so

does its standard error for $h = 5$ and $h = 22$.

Table 3: Estimated 1-, 5-, and 22-day ahead HAR(P)-J models for SPY and stocks average

	HAR-J			HARP-J		
	$h = 1$	$h = 5$	$h = 22$	$h = 1$	$h = 5$	$h = 22$
β_0	0.115***	0.181***	0.358***	0.092**	0.163***	0.358***
s.e.	(0.039)	(0.047)	(0.088)	(0.040)	(0.046)	(0.093)
β_d	0.223**	0.231***	0.140***	0.222**	0.201***	0.129***
s.e.	(0.111)	(0.066)	(0.041)	(0.091)	(0.066)	(0.042)
β_w	0.568***	0.439***	0.297***	0.586***	0.503***	0.347***
s.e.	(0.135)	(0.102)	(0.090)	(0.125)	(0.099)	(0.089)
β_m	0.093	0.146	0.190*	0.114*	0.146	0.172
s.e.	(0.072)	(0.094)	(0.113)	(0.068)	(0.095)	(0.109)
β_{J_d}	−0.440**	−0.489***	−0.292**	−0.480***	−0.484***	−0.256**
s.e.	(0.217)	(0.152)	(0.142)	(0.184)	(0.138)	(0.123)
R^2_{is}	0.527	0.616	0.457	0.546	0.630	0.468
R^2_{oos}	0.458	0.554	0.460	0.510	0.582	0.461
$\beta_d + \beta_w + \beta_m$	0.885	0.816	0.627	0.922	0.850	0.649
Average Stocks						
$\overline{R}^2_{is}$	0.492	0.597	0.510	0.506	0.606	0.514
$\overline{R}^2_{oos}$	0.385	0.488	0.407	0.418	0.515	0.418
$\beta_d + \beta_w + \beta_m$	0.881	0.822	0.675	0.827	0.766	0.616

Note: The table reports the regression coefficients, standard errors in parentheses, and in- and out-of-sample R-squared for the HAR-J and HARP-J models based on various horizons, estimated on SPY data. Days with jumps are estimated using the BNS test of Barndorff-Nielsen and Shephard (2006a) at the 1% significance level. The test is implemented using the staleness-robust adjustment of Kolokolov and Renò (2024). The standard errors are estimated using the Newey-West HAC estimator. The bottom panel shows the stock average in- and out-of-sample R-squared obtained for HAR and HARP models of various horizons. *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

In the case of the HAR-J and HARP-J models, β_{J_d} is always negative, in line with the existing literature (see Andersen et al., 2007a), and has smaller standard errors for the filtered model. As filtering out periodicity reduces the number of detected spurious jumps (see Section 3.3), the jump predictor for the HARP-J model is less affected by measurement error and, as a result, it is more informative. Persistence increases following filtering and standard errors are generally smaller for the HARP-J model.

Table 4: Estimated 1-, 5-, and 22-day ahead HAR(P)-CJ models for SPY and stocks average

	HAR-CJ			HARP-CJ		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
β_0	0.120***	0.183***	0.356***	0.097***	0.163***	0.353***
s.e.	(0.036)	(0.043)	(0.086)	(0.037)	(0.043)	(0.091)
β_{C_d}	0.220**	0.229**	0.139***	0.220**	0.199***	0.129***
s.e.	(0.111)	(0.066)	(0.041)	(0.091)	(0.066)	(0.042)
β_{C_w}	0.573***	0.443***	0.299***	0.589***	0.507***	0.349***
s.e.	(0.135)	(0.102)	(0.091)	(0.126)	(0.100)	(0.087)
β_{C_m}	0.093	0.145	0.189***	0.115	0.144	0.170
s.e.	(0.073)	(0.095)	(0.114)	(0.078)	(0.096)	(0.110)
β_{J_d}	−0.028	−0.123	−0.080	−0.064	−0.117	−0.104*
s.e.	(0.161)	(0.095)	(0.053)	(0.138)	(0.074)	(0.054)
β_{J_w}	−0.403	−0.303	−0.201	−0.407	−0.511	−0.032
s.e.	(0.278)	(0.380)	(0.316)	(0.344)	(0.369)	(0.464)
β_{J_m}	0.306	0.484	0.798	0.283	0.934	1.214
s.e.	(0.588)	(0.916)	(1.471)	(0.542)	(0.861)	(1.423)
R^2_{is}	0.527	0.616	0.457	0.547	0.631	0.468
R^2_{oos}	0.457	0.551	0.457	0.503	0.580	0.462
$\beta_d+\beta_w+\beta_m$	0.886	0.817	0.627	0.924	0.850	0.647
Average Stocks						
$\overline{R}^2_{is}$	0.495	0.603	0.518	0.508	0.609	0.523
$\overline{R}^2_{oos}$	0.367	0.501	0.425	0.412	0.514	0.424
$\beta_d+\beta_w+\beta_m$	0.870	0.805	0.658	0.819	0.754	0.603

Note: The table reports the regression coefficients, standard errors in parentheses, and in- and out-of-sample R-squared for the HAR-CJ and HARP-CJ models based on various horizons, estimated on SPY data. Days with jumps are estimated using the BNS test of Barndorff-Nielsen and Shephard (2006a) at the 1% significance level. The test is implemented using the staleness-robust adjustment of Kolokolov and Renò (2024). The standard errors are estimated using the Newey-West HAC estimator. The bottom panel shows the stock average in- and out-of-sample R-squared obtained for HAR and HARP models of various horizons. *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

In line with our findings for the HARP-J model, for the HARP-CJ model, we notice an important reduction in the standard errors for most coefficients of the realized jumps regressors in comparison to the unfiltered model. For the 1-week ahead model, the standard errors of β_{C_w} decrease uniformly following filtering. While standard errors for some coefficients and some horizons go up following filtering, probably impacted by the use of a high number of explanatory variables, both in-sample and out-of-sample R^2 indicate

an improvement in performance for all forecasting horizons.

Table 5: Estimated 1-, 5-, and 22-day ahead HAR(P)-Q models for SPY and stocks average

	HAR-Q			HARP-Q		
	$h=1$	$h=5$	$h=22$	$h=1$	$h=5$	$h=22$
β_0	0.005	0.111**	0.319***	0.059	0.139***	0.342***
s.e.	(0.045)	(0.050)	(0.083)	(0.039)	(0.047)	(0.087)
β_d	0.669***	0.512***	0.359***	0.476***	0.396***	0.272***
s.e.	(0.081)	(0.097)	(0.087)	(0.071)	(0.075)	(0.061)
β_w	0.363***	0.320***	0.232***	0.420***	0.371***	0.260***
s.e.	(0.082)	(0.093)	(0.092)	(0.081)	(0.091)	(0.091)
β_m	0.023	0.097	0.120	0.060	0.103	0.131
s.e.	(0.072)	(0.093)	(0.106)	(0.068)	(0.089)	(0.100)
β_Q	−0.007***	−0.005***	−0.004***	−0.003***	−0.003***	−0.002***
s.e.	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R^2_{is}	0.584	0.642	0.479	0.571	0.647	0.482
R^2_{oos}	0.553	0.545	0.424	0.560	0.555	0.435
$\beta_d+\beta_w+\beta_m$	1.055	0.929	0.711	0.957	0.870	0.663
Average Stocks						
$\overline{R}^2_{is}$	0.490	0.589	0.502	0.498	0.600	0.510
$\overline{R}^2_{oos}$	0.411	0.477	0.407	0.418	0.504	0.409
$\beta_d+\beta_w+\beta_m$	0.981	0.895	0.730	0.905	0.834	0.686

Note: The table reports the regression coefficients, standard errors in parentheses, and in- and out-of-sample R-squared for the HAR-Q and HARP-Q models based on various horizons, estimated on SPY data. The standard errors are estimated using the Newey-West HAC estimator. The bottom panel shows the stock average in- and out-of-sample R-squared obtained for HAR and HARP models of various horizons. *, ** and *** denote significance at 10%, 5% and 1% level, respectively.

For the HARP-Q model, standard errors for the daily, weekly and monthly coefficients are lower across the board than for HAR-Q, while the standard error of β_Q, the estimated coefficient for $RQ_t^{1/2}RV_t$, stays approximately constant. While all out-of-sample R^2 indicate a better performance following filtering, the in-sample 1-day ahead model for SPY is lower for the HARP-Q model. Bollerslev et al. (2016) explain that the adjustment to account for measurement error has a bigger impact at shorter horizons, while measurement errors tend to average out at longer horizons. It is thus plausible that for the 1-day model, the HAR-Q correction trims "just enough" of the RV_{t-1} regressor, while further trimming to account for periodicity increases the in-sample standard error.

4.3 Out-of-sample forecasting results

In addition to the MSE and QLIKE forecast loss functions defined in equation (21), we compute the GonzÃ¡lez-Rivera et al. (2004) Value-at-Risk-based smoothed loss function, given by:

$$\text{VaR} = \left[\alpha - \frac{1}{1 + e^{\delta(r_t - \text{CVaR}_t^\alpha)}}\right](r_t - \text{CVaR}_t^\alpha),$$

where

$$\text{CVaR}_t^\alpha = \mu_t + \Phi^{-1}(\alpha)\sqrt{F_t},$$

where CVaR is the conditional value at risk, μ_t the forecasted conditional mean, assumed time-invariant here, and F_t the forecasted conditional variance. Φ^{-1} is the inverse of the standard normal cdf, $\delta > 0$ a parameter controlling the smoothness. Following Clements and Preve (2021), we set $\delta = 25$ and $\alpha = 0.05$. We will further refer to the this smoothed loss function simply as value-at-risk.

For SPY, as well as for the 30 stocks in our sample, we assess the significance of the forecasting gains attained for the HARP models relative to the HAR models by applying the Diebold and Mariano (1995) test. Let $\epsilon_{t,t+h-1}$ be the errors from one of the HAR models in equations (12), (14),(16) and (18) and $\epsilon^f_{t,t+h-1}$ the errors from these models' HARP counterparts. Further, let $L(\cdot)$ denote one of the above loss functions and $d_t = L(\epsilon^f_{t,t+h-1}) - L(\epsilon_{t,t+h-1})$. Then, the Diebold and Mariano (1995) test statistic is defined as:

$$DM = \frac{\frac{1}{T}\sum_{t=1}^{T} d_t}{\sqrt{\widehat{\text{Var}}\left(\frac{1}{T}\sum_{t=1}^{T} d_t\right)}} \to \mathcal{N}(0,1), \tag{23}$$

where $\widehat{\text{Var}}\left(\frac{1}{T}\sum_{t=1}^{T} d_t\right)$ is a consistent estimator for the variance of the d_t sample mean. We run a two-tailed test, where rejection when $DM < 0$ ($DM > 0$) means that the average loss from HARP (HAR) models is lower than the average loss from HAR (HARP) models.

We remind the reader that the intraday periodicity of day t is estimated using a rolling window of 20 trading days. In addition, we re-estimate all HAR and HARP models on rolling windows of 1000 days and compute out-of-sample forecast losses.[22] The ratios of the losses from HARP versus HAR models for the 1-day, 1-week and 1-month horizons are reported in Table 6. For each forecasting horizon, the top panel shows results for the SPY and the average across all stocks.

[22]The rolling window length is equivalent to approximately 20% of the length of the dataset.

Table 6: Out-of-sample forecast losses

		(a) $h = 1$				(b) $h = 5$				(c) $h = 22$			
		HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ
SPY	MSE	0.956*	0.985	0.957*	0.955*	0.934*	0.978	0.934*	0.933*	0.984	0.966*	0.985	0.976*
	QLIKE	0.979*	1.355$^{\blacklozenge}$	0.985	1.018	0.932*	0.936*	0.928*	0.950*	0.957*	0.930*	0.956*	0.945*
	VaR	0.993	1.009	0.992*	0.991*	0.999	1.001	1.000	0.998	0.999	1.003	0.999	0.994*
Avg. Stocks	MSE	0.971	0.991	0.966	0.950	0.951	0.947	0.945	0.958	0.969	0.985	0.975	0.983
	QLIKE	0.934	0.984	0.936	0.940	0.927	0.962	0.931	0.946	0.937	0.952	0.942	0.941
	VaR	0.992	0.994	0.993	0.993	0.997	0.998	0.997	0.998	0.997	0.998	0.997	0.999
Diebold & Mariano Test – Individual Stocks													
	MSE	14 : 0	8 : 0	12 : 0	14 : 0	16 : 1	15 : 0	17 : 0	18 : 1	13 : 0	14 : 2	9 : 1	12 : 2
	QLIKE	23 : 0	15 : 0	24 : 2	21 : 0	22 : 2	18 : 1	20 : 1	21 : 1	16 : 1	14 : 1	16 : 1	15 : 2
	VaR	15 : 5	14 : 4	11 : 6	12 : 5	13 : 5	11 : 4	13 : 6	13 : 5	15 : 4	14 : 6	15 : 7	9 : 5

Note: This table reports the ratio of the losses from HARP versus HAR models for various forecasting horizons. $^{\star(\blacklozenge)}$ indicates that the losses of the HARP (HAR) models are significantly lower compared to the HAR (HARP) model at the 5% significance level based on the Diebold and Mariano (1995) test. The entries of type "$xx : yy$" summarize the results of the Diebold and Mariano test for the 30 stocks considered. The first number, "xx", shows the number of stocks for which the HARP model significantly outperforms the HAR model, while the second number, "yy", indicates the number of stocks for which the opposite is true. Filtered measures used in the HARP models are constructed using a rolling window of 20 trading days to estimate the intraday periodicity of day t.

Starred numbers indicate HARP outperforms HAR at a 5% significance level, while numbers with a diamond superscript indicate that HAR significantly outperforms HARP. For each forecasting horizon, the last three rows show the results of the Diebold and Mariano (1995) test for the 30 stocks in our sample. For each entry, the first value indicates the number of stocks for which HARP models outperform HAR models at 5% significance level, while the second value shows the number of stocks for which the opposite is true.

In the case of the 1-day ahead forecasts ($h = 1$), all except two loss ratios take values below 1 for SPY. The Diebold and Mariano (1995) test applied to the MSE loss function indicates significant gains (at 5% significance level) from forecasting the SPY RV based on filtered data for the HARP, HARP-J, and HARP-CJ. For the QLIKE and VaR loss functions, we find significant gains after filtering for respectively the HARP and the models with realized jumps in their specifications (last two columns).

The average loss ratios for all considered stocks and all models are below 1. Moreover, independently of the loss function, forecasting model or horizon, we generally find significant gains from filtering for more than one-third of the stocks considered using the MSE and VaR, while for the QLIKE we find significant gains for at least half of the stocks. In contrast, for QLIKE, the number of stocks for which the opposite is true is at most 2, while for the VaR loss function, it is 6, and for MSE, it is 0.

For the 1-week ahead forecasts ($h = 5$), all except two loss ratios are below 1 in the case of SPY. Furthermore, for the first two loss functions, the Diebold and Mariano (1995) test shows significant losses for at least 3 out of the 4 HARP models considered. Across all models, the average loss ratios for all stocks under consideration are below 1. For 9 out of the 12 loss-function - model pairs, more than half of the stocks feature significantly lower HARP losses. By comparison, HAR models outperform HARP models, based on the MSE and QLIKE, for a number of stocks between 0 and 2.

For the 1-month ahead forecasts ($h = 22$), all but one ratios are below 1 for SPY. The exception occurs for the VaR loss ratio HARP-Q/HAR-Q, but even in this case, the ratio remains close to 1. For the HARP-CJ model, the Diebold and Mariano (1995) test applied to all loss functions indicates significant gains over the HAR model. The remaining HARP models always significantly outperform their HAR counterpart based on the QLIKE loss function, whereas for the HARP-Q model we also find significantly lower MSE.

The stock average loss ratios are always lower than 1. Furthermore, MSE is significantly lower in the case of HARP models for a number of stocks ranging between 9 and

14, while the HARP QLIKE loss is significantly lower for a number of stocks ranging between 14 and 16. For the VaR loss function, the number of stocks showing significant forecast gains following filtering is between 9 and 15. By comparison, the number of stocks for which the HAR models outperform the HARP models oscillates between 0 and 2 for the MSE and QLIKE. Although this number oscillates between 4 and 7 for the VaR loss function, we find that the number of significant stocks for the HARP models generally doubles that of the HAR models.

We further compare the performance of the filtered models with that of the unfiltered ones by ranking the eight models considered throughout the paper (HAR,HAR-J, HAR-CJ, HAR-Q and their filtered counterparts) based on the model confidence set approach of Hansen et al. (2011). For every possible pair of models p, q, where $p, q = 1, \ldots, 8$, we compute the difference between the values for a given loss function $d_{pq,t} = L(\epsilon_{p,[t,t+h-1]}) - L(\epsilon_{q,[t,t+h-1]})$. Models with the highest losses are then sequentially eliminated until the differences between the remaining models are not significant at a 5% significance level. The remaining models are then ranked from the one exhibiting the lowest losses to the one with the highest. The covariances for the employed tests statistics are estimated using a block bootstrap with length equal to 20 days and 5,000 replications.

Table 7 summarizes the results on the model confidence set for SPY, as well as the 30 stocks in our sample, using the three loss functions we have used throughout this section. For SPY, conditional on the model not having been eliminated, we report its ranking. For all forecasting horizons, filtered models generally occupy at least 3 out of the 4 first places for the MSE, while for the QLIKE and VaR losses, filtered models generally take at least 2 of the 4 first places. When $h = 5$, the periodicity filtered models are ranked 1–4 out of 8 based on the MSE losses, taking the following order: HARP-J, HARP, HARP-CJ, HARP-Q.

For the stocks, for each model, forecasting horizon and loss function, we report the number of stocks for which the model was retained in the model confidence set. In the majority of cases, this number increases following filtering. For instance, for $h = 1$, the HAR model was not eliminated based on the QLIKE criterion for 10 out of the 30 stocks in the sample. At the same time, the HARP model was retained for 26 out of the 30 stocks.

In summary, empirical evidence supports the notion that filtering periodicity results in superior forecasting performance across all considered models and forecasting horizons.[23] Further validation of these findings is provided in the internet appendix. For

[23]Section IA.5 in the internet Appendix presents the results of a nested HAR model that incorporates

instance, Section IA.1 presents out-of-sample results for log-transformed models, Section IA.2 demonstrates out-of-sample results using an alternative test for jumps by Andersen et al. (2012), and Section IA.3 reports out-of-sample results using alternative window sizes to compute intraday periodicity –specifically, windows equivalent to 40, 125, 250, 500, and 1,000 trading days.

Table 7: Model confidence set

		Unfiltered Measures				Filtered Measures			
		HAR	HAR-Q	HAR-J	HAR-CJ	HARP	HARP-Q	HARP-J	HARP-CJ
Panel A: $h = 1$									
	MSE	6	2	7	8	3	1	4	5
SPY	QLIKE	5	3	4		1		2	
	VaR		4			3		1	2
	MSE	28	26	29	28	30	30	30	29
Stocks	QLIKE	10	15	11	9	26	22	28	21
	VaR	22	26	23	25	26	28	27	28
Panel B: $h = 5$									
	MSE	6	7	5	8	2	4	1	3
SPY	QLIKE		3				1	2	
	VaR	4	5	2	8	3	7	1	6
	MSE	26	28	27	27	29	29	29	29
Stocks	QLIKE	13	19	12	10	21	28	22	19
	VaR	26	28	25	27	28	30	27	28
Panel C: $h = 22$									
	MSE	6	8	4	5	3	7	2	1
SPY	QLIKE						1	2	3
	VaR		1	4		5	2	3	
	MSE	27	29	28	26	29	29	29	29
Stocks	QLIKE	18	26	19	21	26	30	26	25
	VaR	25	26	27	26	26	30	28	28

Note: The table reports results on the model confidence set across the three forecasting horizons: one-day ($h = 1$), one-week ($h = 5$), and one-month ($h = 22$). For SPY, we report the ranking of each model in the model confidence set. Empty entries indicate that the model was excluded from the set. For the stocks, we report the number of stocks for which a model was retained in the model confidence set. The covariances for the employed tests statistics are estimated using a block bootstrap with length equal to 20 days and 5,000 replications.

all unfiltered and filtered regressors. We employ a LASSO approach to regularize the model. The results confirm the advantages of using regressors from which periodicity has been filtered out. Specifically, we find that the daily and weekly filtered coefficients are selected in at least 70% (and up to 91%) of the days, while the same coefficients for unfiltered models are selected in less than 50% of the days at $h = 1$.

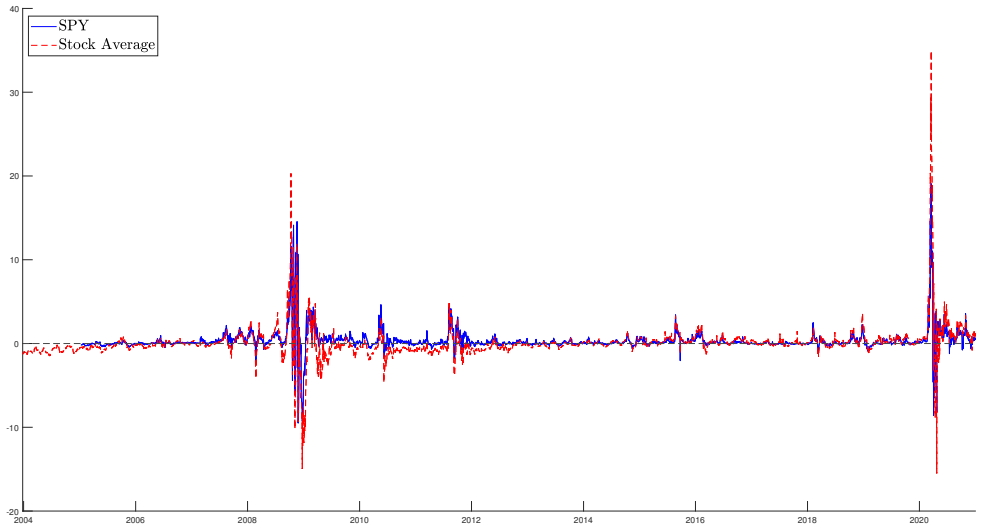
4.4 An analysis of return predictability using periodicity-filtered variance risk premium measures

Bollerslev et al. (2009) show that the variance risk-premium, measured as the difference between the implied and the realized variance, dominates other predictors in explaining the variation in quarterly stock returns.[24] Bekaert and Hoerova (2014) revisit this question by employing a wide range of competing models to forecast realized variance. In this section, we assess whether the superior forecasting performance of HARP models translates into improved predictive power for stock returns. Specifically, we propose using the HARP methodology to construct variance risk premium measures that may better predict stock returns. We measure the variance risk-premium (VRP) as follows:

$$VRP_t = \mathbb{E}_t^{\mathbb{Q}}[RV_{t+1}] - \mathbb{E}_t^{\mathbb{P}}[RV_{t+1}], \tag{24}$$

where $\mathbb{E}_t^{\mathbb{Q}}[RV_{t+1}]$ represents the risk-neutral variance with a 30-day maturity, proxied by the model-free implied variance of Bakshi et al. (2003),[25] and $\mathbb{E}_t^{\mathbb{P}}[RV_{t+1}]$ denotes the monthly forecast derived from either the HAR or HARP models.

Figure 5: Variance risk-premium



Note: The plot depicts the SPY variance risk-premium (blue line), and the average of variance risk-premia over individual stocks (red dashed line). $\mathbb{E}_t^{\mathbb{P}}[RV_{t+1}]$ in equation (24) is derived with the HARP model. Options written on SPY are available from January 10, 2005, so the SPY VRP starts on the latter date, while the stocks' VRP is available from December 26, 2003.

[24] Some of these predictors, typically available at the monthly or quarterly frequencies, are the Treasury bill rate (Campbell, 1987), valuation ratios (Kothari and Shanken, 1997, Hodrick, 1992, Ang and Bekaert, 2007), and information derived from corporate payout and financing activities (Lamont, 1998, Baker and Wurgler, 2000), among others.

[25] For details on the construction of the model-free implied variance, we direct the reader to Section IC. of the internet appendix.

Figure 5 illustrates the VRP for SPY (solid blue line) and the average VRP across all individual stocks (dashed red line). Our findings align with those reported by Carr and Wu (2008). The VRPs are statistically significant and consistently positive, indicating a persistent premium for variance risk in the market. The magnitude of the SPY VRP is larger than the average VRP of individual stocks.

Table 8: SPY monthly predictive regression

	HAR	HAR-Q	HAR-J	HAR-CJ	HARP	HARP-Q	HARP-J	HARP-CJ
Panel A: Univariate Predictive Regressions								
VRP	1.044	0.850	1.044	0.987	1.442	1.275	1.436	1.299
t-stat^{NW}	[2.542]**	[2.449]**	[2.638]***	[2.535]**	[2.816]***	[2.603]***	[2.894]***	[2.719]***
t-stat^{HD}	[2.205]**	[1.876]*	[2.218]**	[2.122]**	[2.386]**	[2.129]**	[2.398]**	[2.210]**
W^{IVX}	[6.310]**	[3.273]*	[8.043]***	[5.815]**	[7.903]***	[5.386]**	[8.196]***	[5.500]**
R^2 (%)	0.582	0.597	0.853	0.783	**0.999**	**0.791**	**0.991**	**0.801**
Panel B: Predictive Regressions with Controls								
VRP	1.530	1.494	1.537	1.512	1.880	1.800	1.876	1.756
t-stat^{NW}	[2.601]***	[2.508]**	[2.600]***	[2.672]***	[2.944]***	[2.912]***	[3.036]***	[2.961]***
t-stat^{HD}	[2.465]**	[2.468]**	[2.640]***	[2.491]**	[2.800]***	[2.728]***	[2.784]***	[2.819]***
W^{IVX}	[5.165]**	[3.662]*	[6.566]**	[5.767]**	[6.689]***	[6.651]***	[6.654]***	[6.664]***
SPRD	0.806	0.798	0.805	0.794	0.767	0.790	0.767	0.772
t-stat^{NW}	[3.113]***	[3.058]***	[3.109]***	[3.044]***	[2.966]***	[3.079]***	[2.965]***	[2.981]***
t-stat^{HD}	[2.971]***	[2.865]***	[2.965]***	[2.847]***	[2.741]***	[3.008]***	[2.791]***	[2.868]***
W^{IVX}	[7.214]***	[6.961]***	[7.182]***	[6.875]***	[6.765]***	[7.386]***	[6.688]***	[6.798]***
CV	2.333	2.354	2.323	2.365	2.127	2.075	2.125	2.178
t-stat^{NW}	[3.574]***	[3.539]***	[3.557]***	[3.655]***	[3.672]***	[3.563]***	[3.561]***	[3.637]***
t-stat^{HD}	[2.756]***	[2.704]***	[2.753]***	[2.708]***	[2.892]***	[2.795]***	[2.790]***	[2.662]***
W^{IVX}	[7.593]***	[6.162]**	[6.909]***	[6.592]**	[7.696]***	[6.562]**	[6.737]***	[8.158]***
adj R^2 (%)	5.101	5.011	5.096	5.135	**5.307**	**5.043**	**5.255**	**5.350**

Note: The table shows results for the monthly predictive regressions $r_{t+1} = \alpha + \beta VRP_t + \boldsymbol{\theta}'\boldsymbol{X}_t + e_{t+1}$, where r_{t+1} is the next month's SPY return in excess of the 3-month T-bill rate, $\boldsymbol{X}$ is the matrix with controls, and $\boldsymbol{\theta}$ is the vector of coefficients. SPRD is the implied volatility spread (e.g., Cremers and Weinbaum, 2010), and CV is the conditional variance (e.g., Bekaert and Hoerova, 2014). t-stat^{NW} is calculated using Newey-West standard errors with a lag length of 44-days. t-stat^{HD} is calculated using Hodrick (1992)'s standard errors, while W^{IVX} is the robust Wald test of Kostakis et al. (2015). R-squares are reported as a percentage and are displayed in bold font whenever they exceed those of the HAR (HARP) counterparts. Symbols *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Our primary objective is to forecast the monthly market excess return, defined as the SPY return in excess of the 3-month T-bill rate. To achieve this, we employ overlapping predictive regressions. Table 8 presents the main findings in two panels. Panel A reports the results for univariate regressions, where the subsequent monthly excess return is re-

gressed on the VRP. Panel B shows the results for regressions with additional control variables: conditional variance (CV) and the implied volatility spread (SPRD). While VRP has been interpreted as an indicator of the representative agent's risk aversion (e.g., Bollerslev et al., 2011, Bekaert et al., 2022), CV serves as a proxy for uncertainty over consumption or dividend growth (e.g., Bekaert et al., 2009). These represent two important state variables driving time variation in the equity premium. SPRD, calculated as the difference between at-the-money (ATM) call and put implied volatilities, measures price pressure related to market sentiment and risk perceptions regarding future economic states. Such factors bolster the disparity between the risk-neutral and the physical variances.[26] Extant literature has identified SPRD as a robust predictor of both the market equity premium (e.g., Atilgan et al., 2015, Cao et al., 2020, Han and Li, 2021) and the cross-section of stock returns (e.g., Cremers and Weinbaum, 2010, Bali and Hovakimian, 2009).

Given our use of overlapping excess returns, we employ multiple approaches to ensure a robust statistical inference. We calculate Newey-West standard errors with a lag length of 44 days,[27] as well as Hodrick (1992) standard errors. In addition, we implement the robust Wald test proposed by Kostakis et al. (2015), which addresses the issue of persistence in predictors.

Results for the univariate regressions suggest a positive relationship between VRP and future excess returns, consistent with previous findings in the literature. If investors anticipate higher future volatility, a discount is built into prices, leading to higher future returns. This positive relationship remains statistically significant regardless of whether we employ HAR or HARP models, and across all the test statistics.VRP measures derived from HARP models exhibit superior explanatory power, as evidenced by higher R-squared values. Specifically, the HARP, HARP-J, and HARP-CJ models yield the highest R^2 values, respectively. These improvements are sizeable; for instance, the predictive regression using the HARP model for VRP construction achieves an R^2 of 0.999%, which is almost twofold higher than that obtained by the VRP derived from the HAR model. The findings in Panel B further support the superiority of HARP-based VRP measures in predicting excess returns. They display higher adjusted R^2 and highly significant VRP regressors for all models considered.[28]

[26] We find that SPRD has a correlation of 13% with VRP, which is in line with the magnitude reported by Atilgan et al. (2015). This indicates that these measures likely capture different types of information. For instance, Bali and Hovakimian (2009) find that SPRD is more closely related with jump risk.

[27] The use of a relatively large number of lags (2×horizon) improves the power of the Newey-West estimator (e.g., Bekaert and Hoerova, 2014, Sun et al., 2008).

[28] Table IA.6, of the internet appendix, shows that similar conclusions are drawn if we only control

We further investigate the predictive power of variance risk premia for future stock excess returns. Consistent with previous literature (e.g., Carr and Wu, 2008, Vilkov, 2008), we identify non-zero VRPs for individual stocks, which are an order of magnitude smaller than the market index VRP. This finding naturally leads us to examine whether these individual VRPs contain valuable predictive information for future stock returns. To address this question, we employ a panel regression framework with firm fixed effects and clustered robust standard errors (Petersen, 2008). Our findings are presented in Table 9.[29]

Table 9: Stocks monthly predictive regression

	HAR	HAR-Q	HAR-J	HAR-CJ	HARP	HARP-Q	HARP-J	HARP-CJ
Panel A: Univariate Predictive Regressions								
VRP	0.103	0.082	0.098	0.097	0.110	0.080	0.113	0.107
t-stat	[1.853]*	[1.386]	[1.661]*	[1.775]*	[1.976]**	[1.433]	[2.000]**	[2.098]**
R^2 (%)	0.054	0.033	0.050	0.047	**0.060**	**0.037**	**0.063**	**0.056**
Panel B: Predictive Regressions with Controls								
VRP	0.183	0.165	0.181	0.180	0.170	0.147	0.172	0.168
t-stat	[1.762]*	[1.629]	[1.728]*	[1.630]	[1.790]*	[1.519]	[1.771]*	[1.817]*
SPRD	0.133	0.138	0.134	0.134	0.137	0.144	0.136	0.138
t-stat	[2.100]**	[2.079]**	[2.116]**	[2.164]**	[2.065]**	[2.156]**	[2.064]**	[2.090]**
CV	0.115	0.126	0.116	0.119	0.123	0.137	0.122	0.124
t-stat	[0.983]	[1.060]	[0.996]	[1.056]	[1.032]	[1.260]	[1.126]	[1.123]
adj R^2 (%)	0.176	0.168	0.178	0.178	**0.187**	**0.174**	**0.185**	**0.189**

Note: The table shows results for the monthly predictive regressions $r_{i,t+1} = \alpha + \beta VRP_{i,t} + \boldsymbol{\theta}' \boldsymbol{X}_{i,t} + e_{i,t+1}$, where $r_{i,t+1}$ denotes the next month's return in excess of the 3-month T-bill rate for the ith stock, $\boldsymbol{X}$ is the matrix with controls, and $\boldsymbol{\theta}$ is the vector of coefficients. SPRD is the implied volatility spread (e.g., Cremers and Weinbaum, 2010), and CV is the conditional variance (e.g. Bekaert and Hoerova, 2014). The model is estimated within a panel framework with firm fixed effects. The t-statistics are estimated using clustered robust standard errors (e.g., Petersen, 2008). R-squares are reported as a percentage and are displayed in bold font whenever they exceed those of the HAR (HARP) counterparts. Symbols *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Results show that individual variance risk premia positively predict future monthly excess returns, consistent with findings at the aggregate level. However, the coefficients of all individual VRPs are less significant compared to the market VRP, indicating that while individual VRPs possess some predictive power for future returns, their information content is less rich. This result is unsurprising given our observation of a weakly positive

for the conditional variance of the models.

[29] Table IA.7 in the Internet Appendix presents qualitatively similar results when t-statistics are computed using clustered robust standard errors with Newey-West correction and 44 lags.

average VRP. Carr and Wu (2008) also note that individual stock VRP averages are typically smaller than market VRP and can even be negative for some stocks.[30]

Despite the weaker predictive power of individual VRPs, we consistently observe that HARP-based VRP measures outperform their HAR-based counterparts. With the sole exception of the HARP-Q-based VRP, all HARP-based VRPs are significant at the 5% level. In contrast, the HAR-based VRPs, are only significant at the 10% level. These results are robust to controlling for the implied volatility spread and the conditional variance, albeit weaker compared to the univariate results. While the HARP-based VRPs are now significant at the 10% level, only the HAR-based and HAR-CJ-based VRPs remain significant at the 10% level.

In summary, we document that VRP measures derived from HARP models are likely more informative than their HAR counterparts in predicting future monthly excess returns. For the aggregate market, the improvement is sizeable, yielding an increase in the R^2 of up to 50% relative to HAR-based measures. However, for individual stocks, the R^2 improvement is marginal, and the overall significance is modest.

4.5 An alternative approach to dealing with periodicity

The HARP methodology proposed in this paper addresses the issue of forecasting with periodicity-contaminated regressors by substituting them with filtered-out estimates. An alternative approach is to explicitly incorporate periodicity into the forecasting regression, akin to the HARQ model introduced by Bollerslev et al. (2016). This model, referred to as HAR-QP, is constructed as a HAR model in which the coefficient of RV_{t-1} is corrected for both measurement error and periodicity, as follows:

$$RV_{t,t+h-1} = \beta_0 + \left(\beta_d + \beta_{d_Q}\left[RQ_{t-1}\Delta\hat{f}^4_{t-1}\right]^{1/2}\right) RV_{t-1} + \beta_w RV_{t-5,t-1} + \beta_m RV_{t-22,t-1} + \epsilon_{t,t+h-1}. \tag{25}$$

While this model, like the HARP models, depends on a pre-estimate of periodicity, its coefficient correction is predicated on the assumption that IV_t follows an AR(1) process. More complex DGPs would require more involved corrections.

We compare the out of sample performance of this approach to the performances of HARQ and HARP-Q models. Results are included in Table 10. Panels A and B present, respectively, the ratio of losses from the HAR-P and HAR-QP models versus

[30] Please note that Carr and Wu (2008) define their VRP as realized minus implied variances, while we define it as implied minus realized variances.

the HAR-Q model for the 1-day, 1-week and 1-month horizons. In addition, the numbers in parentheses, included in the superscript of each loss ratio, represent results from the model confidence set. For Panel A (SPY), we provide the ranking of each model within the model confidence set. Empty entries indicate that the model was excluded from the set. For Panel B (stock average), we report the number of stocks for which a model was retained by the model confidence set. The covariances for the employed test statistics are estimated using a block bootstrap with a length of 20 days and 5,000 replications.

Table 10: out-of-sample forecast losses

	HAR-Q	HARP-Q	HAR-QP	HAR-Q	HARP-Q	HAR-QP
	Panel A: SPY			Panel B: Stock Average		
	$h = 1$					
MSE	$1.000^{(3)}$	$0.985^{(2)}$	$0.947^{(1)}$	$1.000^{(30)}$	$0.991^{(30)}$	$0.993^{(30)}$
QLIKE	$1.000^{(2)}$	1.355	$0.994^{(1)}$	$1.000^{(21)}$	$0.984^{(29)}$	$0.996^{(23)}$
VaR	$1.000^{(2)}$	$1.009^{(3)}$	$0.997^{(1)}$	$1.000^{(17)}$	$0.994^{(24)}$	$1.000^{(18)}$
	$h = 5$					
MSE	$1.000^{(2)}$	$0.978^{(1)}$	$1.002^{(3)}$	$1.000^{(26)}$	$0.947^{(30)}$	$0.979^{(28)}$
QLIKE	1.000	$0.936^{(1)}$	1.002	$1.000^{(24)}$	$0.962^{(30)}$	$0.999^{(27)}$
VaR	$1.000^{(2)}$	$1.001^{(3)}$	$0.991^{(1)}$	$1.000^{(20)}$	$0.998^{(23)}$	$0.996^{(25)}$
	$h = 22$					
MSE	1.000	$0.966^{(1)}$	$0.973^{(2)}$	$1.000^{(28)}$	$0.985^{(28)}$	$0.968^{(30)}$
QLIKE	1.000	$0.930^{(2)}$	$0.923^{(1)}$	$1.000^{(22)}$	$0.952^{(30)}$	$0.965^{(28)}$
VaR	$1.000^{(2)}$	$1.003^{(3)}$	$0.999^{(1)}$	$1.000^{(21)}$	$0.998^{(25)}$	$0.998^{(25)}$

Note: This table presents the ratio of losses from the HARP-Q and HAR-QP models to the HAR-Q model for various forecasting horizons. The numbers in parentheses, included in the superscript of each loss ratio, represent results in the model confidence set. For SPY, we provide the ranking of each model within the model confidence set. Empty entries indicate that the model was excluded from the set. For the stocks, we report the number of stocks for which a model was retained in the model confidence set. The covariances for the employed test statistics are estimated using a block bootstrap with a length of 20 days and 5,000 replications.

As can be seen, the new HAR-QP model generally outperforms the standard HAR-Q model, confirming the importance of accounting for intraday periodicity when forecasting future RV. Additionally, the HAR-QP model performs well against the HARP-Q model. Specifically, for SPY, improvements of up to 5% and 4% at $h = 1$, and up to 8% and 0.7% at $h = 22$ are observed relative to the HAR-Q and HARP-Q models, respectively. The Model Confidence Set corroborates these improvements by consistently ranking the HAR-QP model first at $h = 1$ and $h = 22$ across all loss functions, except for the MSE

at $h = 22$. Notably, the HAR-Q model is never ranked first and is often excluded from the set, affirming the superior performance of our proposed models. Turning our attention to the results for the stock average (Panel B), both the HAR-QP and HARP-Q models outperform the HAR-Q model across all forecasting horizons and loss functions. Furthermore, the number of stocks retained by the MCS approach is consistently higher for models that account for intraday periodicity, further confirming their superior performance over unfiltered approaches. In sum, our results indicate that the HARP-Q and HAR-QP models should be preferred over the standard HAR-Q model.

5 Conclusion

The contribution of this paper is twofold. Firstly, we document the impact of volatility intraday periodicity on forecasting the realized variance using heterogenous autoregressive (HAR) models. While periodicity has no impact on the realized volatility itself, it distorts its variance, leading to biases in the coefficients of the forecasting models. We derive the variance and the 1-lag auto-correlation coefficient for the realized variance in the case of a very simple DGP and show that periodicity artificially inflates the variance and has a decreasing impact on the autocorrelation. We further show that filtering periodicity out shrinks back the variance, provided we estimate periodicity over a big sample. In addition, we document that periodicity leads to spurious jumps detection.

Secondly, we propose a HAR model where predictors rely on data from which periodicity is filtered out and denote this model "HARP". We provide a thorough set of in-sample and out-of-sample forecasting comparisons between the HARP and HAR models, relying on both simulated and empirical data. Our analysis encompasses the HARP versions of the most common HAR models in the literature, the HAR model by Corsi (2009), the HAR-J and HAR-CJ models by Andersen et al. (2007a), and the HAR-Q model by Bollerslev et al. (2016). Our dataset includes intraday observations for the SPDR ETF and 30 S&P500 constituents for the period 2000 to 2020. The simulation and empirical evidence indicates that pre-filtering the data for periodicity leads to forecasting gains for all model specifications and all forecasting horizons. The MCS also ranks the filtered models above the unfiltered. Finally, an empirical application examining return predictability through the RV-based variance risk premium demonstrates that filtering for periodicity to construct the VRP results in a more accurate predictor of excess returns for both the index and individual stocks, albeit with weaker results for the latter. These findings remain robust even after controlling for the implied volatility spread and conditional variance.

References

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

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

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2):115–158.

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

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

Andersen, T. G., Bollerslev, T., and Das, A. (2001). Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns. *The Journal of Finance*, 56(1):305–327.

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

Andersen, T. G., Bollerslev, T., and Dobrev, D. (2007b). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138:125–180.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Andersen, T. G., Thyrsgaard, M., and Todorov, V. (2018). Time-varying periodicity in intraday volatility. *Journal of the American Statistical Association*, 0(ja):1–39.

Ang, A. and Bekaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies*, 20(3):651–707.

Atilgan, Y., Bali, T. G., and Demirtas, K. O. (2015). Implied volatility spreads and expected market returns. *Journal of Business & Economic Statistics*, 33(1):87–101.

Baker, M. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *the Journal of Finance*, 55(5):2219–2257.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *The Review of Financial Studies*, 16(1):101–143.

Bali, T. G. and Hovakimian, A. (2009). Volatility spreads and expected stock returns. *Management Science*, 55(11):1797–1812.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

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

Barndorff-Nielsen, O. E. and Shephard, N. (2006b). Multipower variation and stochastic volatility. In *Stochastic Finance*, pages 73–82. Springer.

Bekaert, G., Engstrom, E., and Xing, Y. (2009). Risk, uncertainty, and asset prices. *Journal of Financial Economics*, 91(1):59–82.

Bekaert, G., Engstrom, E. C., and Xu, N. R. (2022). The time variation in risk appetite and uncertainty. *Management Science*, 68(6):3975–4004.

Bekaert, G. and Hoerova, M. (2014). The vix, the variance premium and stock market volatility. *Journal of Econometrics*, 183(2):181–192. Analysis of Financial Data.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Bollerslev, T., Cai, J., and Song, F. M. (2000). Intraday periodicity, long memory volatility, and macroeconomic announcement effects in the us treasury bond market. *Journal of Empirical Finance*, 7(1):37–55.

Bollerslev, T., Gibson, M., and Zhou, H. (2011). Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. *Journal of Econometrics*, 160(1):235–245.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). Expected Stock Returns and Variance Risk Premia. *The Review of Financial Studies*, 22(11):4463–4492.

Boudt, K., Cornelissen, J., Croux, C., and Laurent, S. (2012). *Nonparametric Tests for Intraday Jumps: Impact of Periodicity and Microstructure Noise*, chapter Eighteen, pages 447–463. John Wiley & Sons, Ltd.

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

Bu, R., Hizmeri, R., Izzeldin, M., Murphy, A., and Tsionas, M. (2023). The contribution of jump signs and activity to forecasting stock price volatility. *Journal of Empirical Finance*, 70:144–164.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics*, 18(2):373–399.

Campbell, J. Y. and Thompson, S. B. (2007). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, 21(4):1509–1531.

Cao, C., Simin, T., and Xiao, H. (2020). Predicting the equity premium with the implied volatility spread. *Journal of Financial Markets*, 51:100531.

Carr, P. and Wu, L. (2008). Variance Risk Premiums. *The Review of Financial Studies*, 22(3):1311–1341.

Christensen, K., Hounyo, U., and Podolskij, M. (2018). Is the diurnal pattern sufficient to explain intraday variation in volatility? a nonparametric assessment. *Journal of Econometrics*, 205(2):336–362.

Christensen, K. and Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349.

Clements, A. and Preve, D. P. (2021). A practical guide to harnessing the har volatility model. *Journal of Banking & Finance*, 133:106285.

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

Corsi, F., Pirino, D., and Renò, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288.

Corsi, F. and Renò, R. (2012). Discretetime volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, 30(3):368–380.

Cremers, M. and Weinbaum, D. (2010). Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(2):335–367.

Dacorogna, M., Müller, A., Pictet, V., and Olsen, R. (1997). Modelling short-term volatility with garch and harch models. *Nonlinear Modelling of High Frequency Financial Time Series*.

Dette, H., Golosnoy, V., and Kellermann, J. (2023). The effect of intraday periodicity on realized volatility measures. *Metrika*, 86(3):315–342.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Dumitru, A.-M. and Urga, G. (2012). Identifying jumps in financial assets: a comparison between nonparametric jump tests. *Journal of Business & Economic Statistics*, 30(2):242–255.

GonzÃ¡lez-Rivera, G., Lee, T.-H., and Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, 20(4):629–645.

Han, B. and Li, G. (2021). Information content of aggregate implied volatility spread. *Management Science*, 67(2):1249–1269.

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

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Hasbrouck, J. (1999). The dynamics of discrete bid and ask quotes. *The Journal of Finance*, 54(6):2109–2142.

Hodrick, R. J. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *The Review of Financial Studies*, 5(3):357–386.

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4):456–499.

Kolokolov, A. and Renò, R. (2024). Jumps or staleness? *Journal of Business & Economic Statistics*, 42(2):516–532.

Kostakis, A., Magdalinos, T., and Stamatogiannis, M. P. (2015). Robust econometric inference for stock return predictability. *The Review of Financial Studies*, 28(5):1506–1553.

Kothari, S. P. and Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics*, 44(2):169–203.

Lamont, O. (1998). Earnings and expected returns. *The Journal of Finance*, 53(5):1563–1587.

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

Liu, L. Y., Patton, A. J., and Sheppard, K. (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1):293–311.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36:270–296.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., and von Weizsäcker, J. E. (1997). Volatilities of different time resolutions - analyzing the dynamics of market components. *Journal of Empirical Finance*, 4(2):213–239.

Petersen, M. A. (2008). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480.

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *The Econometrics Journal*, 5(1):1–39.

Sun, Y., Phillips, P. C., and Jin, S. (2008). Optimal bandwidth selection in heteroskedasticity–autocorrelation robust testing. *Econometrica*, 76(1):175–194.

Vilkov, G. (2008). Variance risk premium demystified. *Available at SSRN 891360*.

A Some Proofs for the Simple AR(1) Model

A.1 AR(1) with periodicity

Under the assumptions of Section 2.1, $RV_t = \sum_{i=1}^{M} r_{t,i}^2 = \Delta IV_t \sum_{i=1}^{M} f_i^2 w_i^2$.

$$\mathbb{E}(RV_t) = \Delta\, \mathbb{E}(IV_t) \sum_{i=1}^{M} f_i^2 = \mathbb{E}(IV_t),$$

where we used the fact that $\mathbb{E}(w_i^2) = 1$.

Proof of equation (10a).

$$\begin{aligned}
\mathbb{E}\left(RV_t^2\right) &= \Delta^2\, \mathbb{E}(IV_t^2)\, \mathbb{E}\left(\sum_{i=1}^{M} f_i^4 w_i^4 + \sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M} f_i^2 w_i^2 f_j^2 w_j^2\right) \\
&= \Delta^2\, \mathbb{E}(IV_t^2)\left(3\sum_{i=1}^{M} f_i^4 + \sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M} f_i^2 f_j^2\right) \\
&= \Delta^2\, \mathbb{E}(IV_t^2)\left[3\sum_{i=1}^{M} f_i^4 + \sum_{i=1}^{M} f_i^2\left(\frac{1}{\Delta} - f_i^2\right)\right] \\
&= \mathbb{E}(IV_t^2) + \mathbb{E}(IV_t^2)\Delta^2\left(2\sum_{i=1}^{M} f_i^4\right).
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(RV_t) &= \mathbb{E}\left(RV_t^2\right) - \left[\mathbb{E}\left(RV_t\right)\right]^2 = \mathbb{E}(IV_t^2) + \mathbb{E}(IV_t^2)\Delta^2\left(2\sum_{i=1}^{M} f_i^4\right) - \left[\mathbb{E}(IV_t)\right]^2 \\
&= \mathrm{Var}(IV_t) + 2\Delta^2 \sum_{i=1}^{M} f_i^4\, \mathbb{E}(IV_t^2). \qquad \square
\end{aligned}$$

For comparison purposes, we compute the same variance in the absence of periodicity, where the superscript NP below stands for "no periodicity".:

Proof of equation (10b).

$$\begin{aligned}
\mathrm{Var}(RV_t)^{NP} &= \Delta^2\, \mathbb{E}(IV_t^2)\, \mathbb{E}\left(\sum_{i=1}^{M} w_i^4 + \sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M} w_i^2 w_j^2\right) - \left[\mathbb{E}(IV_t)\right]^2 \\
&= \Delta^2\, \mathbb{E}(IV_t^2)\left[\frac{3}{\Delta} + \frac{1}{\Delta}\left(\frac{1}{\Delta} - 1\right)\right] - \left[\mathbb{E}(IV_t)\right]^2 \\
&= \Delta^2\, \mathbb{E}(IV_t^2)\left(\frac{2}{\Delta} + \frac{1}{\Delta^2}\right) - \left[\mathbb{E}(IV_t)\right]^2 \\
&= \mathrm{Var}(IV_t) + 2\Delta\, \mathbb{E}(IV_t^2). \qquad \square
\end{aligned}$$

Let w_i, $i = 1, \ldots, M$ be a sequence of i.i.d. standard normal variables entering the intraday returns on day t and w_i^*, $i = 1, \ldots, M$ another sequence of i.i.d. standard normals, independent of w_i, entering returns on day $t-h$, $h \geq 1$. The auto-covariance of lag h is obtained below.

Auto-covariance derivation.

$$\begin{aligned}\operatorname{Cov}(RV_t, RV_{t-h}) &= \Delta^2 \operatorname{Cov}(IV_t, IV_{t-1})\, \mathbb{E}\left(\sum_{i=1}^{M} f_i^2 {w_{i,t}}^2\right)\, \mathbb{E}\left(\sum_{i=1}^{M} f_i^2 {w_{i,t-1}}^2\right) \\ &= \operatorname{Cov}(IV_t, IV_{t-1}),\end{aligned}$$ □

where we used the fact that $w_{i,t} \perp w_{i,t-1}$ and $\sum_{i=1}^{M} f_i^2 = \frac{1}{\Delta}$

Proof of the formula for ϕ.

$$\begin{aligned}\phi &= \frac{\operatorname{Cov}(RV_t, RV_{t-1})}{\operatorname{Var}(RV_t)} \\ &= \frac{\operatorname{Cov}(IV_t, IV_{t-1})}{\operatorname{Var}(IV_t) + 2\Delta^2 \sum_{i=1}^{M} f_i^4\, \mathbb{E}(IV_t^2)} \\ &= \frac{\Phi \operatorname{Var}(IV_t)}{\operatorname{Var}(IV_t) + 2\Delta^2 \sum_{i=1}^{M} f_i^4\, \mathbb{E}(IV_t^2)} \\ &= \frac{\Phi}{1 + 2\Delta^2 \sum_{i=1}^{M} f_i^4 \frac{\mathbb{E}(IV_t^2)}{\operatorname{Var}(IV_t)}}.\end{aligned}$$ □

A.2 AR(1) with filtered data

The filtered returns are $r_{i,t}^f = \frac{r_{i,t}}{\widehat{f}_i}$, leading to the filtered RV: $RV_t^f = \sum \frac{\Delta IV_t f_i^2 w_{i,t}^2}{\widehat{f}_i^{\,2}} = \Delta IV_t \sum \frac{f_i^2}{\widehat{f}_i^{\,2}} w_{i,t}^2$.

Let us first define a simpler version of the estimator $\widehat{f}_i$, which is more efficient, but less robust to jumps. This is not a problem as in this set-up we do not have jumps. Let $\overline{r_{i,t}}$, $t = 1 \ldots T$, assumed standardized with $\sqrt{\Delta IV_t}$ for simplicity: $\overline{r_{i,t}} = f_i w_{i,t} \sim \mathcal{N}(0, f_i^2)$.

We define an estimator for the variance of $\overline{r_{i,t}}$ as $\widehat{f}_i^2 = \frac{\sum_{t=1}^{T} \overline{r_{i,t}}^2}{T}$, with the following mean and variance:

$$\begin{aligned}\mathbb{E}\left(\widehat{f}_i^2\right) &= \mathbb{E}\left(f_i^2 \frac{\sum_{t=1}^{T} {w_{i,t}}^2}{T}\right) = f_i^2 \\ \mathbb{E}\left(\widehat{f}_i^2 - f_i^2\right)^2 &= \mathbb{E}\left(\widehat{f}_i^4\right) - f_i^4 = \mathbb{E}\left(\frac{\sum_{t=1}^{T} \overline{r_{i,t}}^2}{T}\right)^2 - f_i^4 \\ &= \frac{1}{T^2}\left[\sum_{t=1}^{T} \mathbb{E}\left(\overline{r_{i,t}}^4\right) + \sum\sum_{t \neq \tau} \mathbb{E}\left(\overline{r_{i,t}}^2\right)\, \mathbb{E}\left(\overline{r_{i,\tau}}^2\right)\right] - f_i^4 \\ &= \frac{1}{T^2}\left[3T f_i^4 + T(T-1) f_i^4\right] - f_i^4 = \frac{f_i^4}{T^2}\left[3T + T^2 - T - T^2\right] = \frac{2 f_i^4}{T}.\end{aligned}$$

Note that $\left(\frac{2f_i^4}{T}\right)^{-\frac{1}{2}} \widehat{f_i}^2 = O_p(1)$, so $\frac{f_i^2}{\widehat{f_i}^2} \xrightarrow{p} 1$ based on the continuous mapping theorem.

As $\widehat{f_i}^2$ enters RV_t^f as a ratio, we will approximate the variance of this ratio. Let us assume $\frac{1}{x} = \frac{1}{\mu_x + \epsilon}$, for x a random variable with expected value $\mu_x > 0$. The second order Taylor series approximation of $\frac{1}{x}$ around μ_x is:

$$\frac{1}{x} \approx \frac{1}{\mu_x} - \frac{1}{\mu_x^2}(x - \mu_x) + \frac{2}{2\mu_x^3}(x - \mu_x)^2 + \ldots = \frac{1}{\mu_x} - \frac{1}{\mu_x^2}(x - \mu_x) + \frac{1}{\mu_x^3}(x - \mu_x)^2$$

Applying the expectation and variance operators:

$$\mathbb{E}\left(\frac{1}{x}\right) \approx \frac{1}{\mu_x} + \frac{1}{\mu_x^3}\mathbb{E}(x - \mu_x)^2 = \frac{1}{\mu_x} + \frac{1}{\mu_x^3}\operatorname{Var}(x)$$

$$\begin{aligned}\operatorname{Var}\left(\frac{1}{x}\right) &\approx \frac{1}{\mu_x^4}\operatorname{Var}(x) + \frac{1}{\mu_x^6}\operatorname{Var}(x - \mu_x)^2 = \frac{1}{\mu_x^4}\operatorname{Var}(x) + \frac{1}{\mu_x^6}\operatorname{Var}\left(x^2 - 2\mu_x x + \mu_x^2\right) \\ &= \frac{1}{\mu_x^4}\operatorname{Var}(x) + \frac{1}{\mu_x^6}\left[\operatorname{Var}\left(x^2\right) + 4\mu_x^2\operatorname{Var}(x) - 4\mu_x\operatorname{Cov}\left(x, x^2\right)\right]\end{aligned}$$

In our case, we have: $x = \widehat{f_i}^2; \frac{1}{x} = \frac{1}{\widehat{f_i}^2} = \frac{1}{f_i^2 + \epsilon}$ *so* $\mu_x = f_i^2,\ \mu_x^2 = f_i^4,\ \mu_x^4 = f_i^8,\ \mu_x^6 = f_i^{12}$.

$$\mathbb{E}\left(\frac{1}{\widehat{f_i^2}}\right) \approx \frac{1}{f_i^2} + \frac{2}{f_i^2 T}$$

$$\begin{aligned}\operatorname{Var}\left(\frac{1}{\widehat{f_i^2}}\right) &\approx \frac{2}{f_i^4 T} + \frac{1}{f_i^{12}}\left[\frac{\left(8T^3 + 40T^2 + 48T\right)f_i^8}{T^4} + 4f_i^4\frac{2f_i^4}{T} - 4f_i^2\frac{f_i^6}{T^2}(8 + 4T)\right] \\ &= \frac{2}{f_i^4 T} + \frac{1}{f_i^4}\left[\frac{8T^3 + 40T^2 + 48T}{T^4} + \frac{8}{T} - \frac{4}{T^2}(8 + 4T)\right] \\ &= \frac{2}{f_i^4 T} + \frac{1}{f_i^4}\left[\frac{8T^2 + 40T + 48}{T^3} - \frac{8}{T} - \frac{32}{T^2}\right] = \frac{2}{f_i^4 T} + \frac{1}{f_i^4}\left(\frac{8T^2 + 40T + 48 - 8T^2 - 32T}{T^3}\right) \\ &= \frac{2}{f_i^4 T} + \frac{1}{f_i^4}\left(\frac{8T + 48}{T^3}\right) = \frac{2}{f_i^4 T} + O_p(T^{-2})\end{aligned}$$

To compute the above variance, we needed higher order moments for $\widehat{f_i^2}$. To determine them, we used the fact that $z = \frac{\sum_{t=1}^T \overline{r_{i,t}}^2}{f_i^2} \sim \chi^2(T)$ as a sum of squared i.i.d. standard normals:

$$\widehat{f_i^2} = \frac{\sum_{t=1}^T \overline{r_{i,t}}^2}{T} = \frac{f_i^2}{T}z \implies z = T\widehat{f_i^2}/f_i^2 \sim \chi^2(T)$$

The moments for $\widehat{f_i^2}$ are given to the right below, as a function of corresponding moments for z:

$$\begin{aligned}\operatorname{Var}(z) &= \operatorname{Var}\left(\frac{T\widehat{f_i^2}}{f_i^2}\right) = 2T \\ \operatorname{Var}\left(z^2\right) &= \operatorname{Var}\left(\frac{T\widehat{f_i^2}}{f_i^2}\right)^2 = 8T^3 + 40T^2 + 48T \\ \operatorname{Cov}\left(z, z^2\right) &= \operatorname{Cov}\left[\left(\frac{T\widehat{f_i^2}}{f_i^2}\right)^2, \frac{T\widehat{f_i^2}}{f_i^2}\right] = 4T^2 + 8T\end{aligned} \implies \begin{aligned}\operatorname{Var}\left(\widehat{f_i^2}\right) &= \tfrac{2f_i^4}{T} \\ \operatorname{Var}\left(\widehat{f_i^2}\right)^2 &= \tfrac{\left(8T^3 + 40T^2 + 48T\right)f_i^8}{T^4} \\ \operatorname{Cov}\left[\left(\widehat{f_i^2}\right)^2, \widehat{f_i^2}\right] &= \tfrac{\left(4T^2 + 8T\right)f_i^6}{T^3}\end{aligned}$$

The variance of the filtered RV can be thus approximated:

$$\begin{aligned}\mathrm{Var}\left(RV_t^f\right) \approx\ & \mathrm{Var}\left(IV_t\right) + 2\Delta\,\mathbb{E}\left(IV_t^2\right) + \mathrm{Var}\left(IV_t\right)\frac{4}{T}\left(1+\frac{1}{T}\right) + \Delta\,\mathbb{E}\left(IV_t^2\right)\left(\frac{14}{T}+\frac{32}{T^2}+\frac{144}{T^3}\right)\\ =\ & \mathrm{Var}\left(RV_t^{NP}\right) + \mathrm{Var}\left(IV_t\right)\frac{4}{T}\left(1+\frac{1}{T}\right) + \Delta\,\mathbb{E}\left(IV_t^2\right)\left(\frac{14}{T}+\frac{32}{T^2}+\frac{144}{T^3}\right),\end{aligned}$$

where $\mathrm{Var}\left(RV_t^{NP}\right)$ is the variance when no periodicity is present.

By comparison, the variance of the unfiltered RV is:

$$\begin{aligned}\mathrm{Var}\left(RV_t\right) =\ & \mathrm{Var}\left(IV_t\right) + 2\Delta\,\mathbb{E}\left(IV_t^2\right)\Delta\sum_{i=1}^{M} f_i^4\\ =\ & \mathrm{Var}\left(IV_t\right) + 2\Delta\,\mathbb{E}\left(IV_t^2\right) + 2\Delta\,\mathbb{E}\left(IV_t^2\right)\left(\Delta\sum_{i=1}^{M} f_i^4 - 1\right)\\ =\ & \mathrm{Var}\left(RV_t^{NP}\right) + 2\Delta\,\mathbb{E}\left(IV_t^2\right)\left(\Delta\sum_{i=1}^{M} f_i^4 - 1\right)\end{aligned}$$

To understand the difference between the two variances, we will compare them for the case when $T = 1000$ and for $\mathrm{Var}(IV_t)$, $\mathbb{E}\left(IV_t^2\right)$ and $\Delta\sum f_i^4$ calibrated to their values for our simulated data: $\mathrm{Var}\left(IV_t\right) \approx 0.5$, $\mathbb{E}\left(IV_t^2\right) = 1.615$ and $\Delta\sum f_i^4 = 1.0815$. This gives us bias for $\mathrm{Var}\left(RV_t\right)$ of $1.615 \times \frac{2}{78} \times 0.0815 = 0.0034$. For $\mathrm{Var}\left(RV_t^f\right)$, the biggest part of the bias is $\mathrm{Var}\left(IV_t\right)\frac{4}{T} + \Delta\,\mathbb{E}\left(IV_t^2\right)\frac{14}{T} = \frac{4}{1000}0.5 + \frac{14}{78\times1000}1.615 = 0.0023$.

B More on Methodology

B.1 Weighted Standard Deviation

For each $i = 1, \ldots, M$, we observe T standardized returns, $\bar{r}_{t,i}$, which we sort in increasing order, as follows:

$$\bar{r}_{(1),i} \leq \bar{r}_{(2),i} \leq \cdots \leq \bar{r}_{(T),i}$$

Given the above ordered set, we define the sub-sets containing half ($\kappa = \lfloor T/2 \rfloor + 1$) contiguous observations: $\{\bar{r}_{(1),i}, \ldots, \bar{r}_{(\kappa),i}\}, \ldots, \{\bar{r}_{(T-\kappa+1),i}, \ldots, \bar{r}_{(T),i}\}$. The shortest half scale estimator is the shortest length of these subsets:

$$ShortH_i = 0.741 \min(\bar{r}_{(\kappa),i} - \bar{r}_{(1),i}, \ldots, \bar{r}_{(T),i} - \bar{r}_{(T-\kappa+1),i})$$

The shortest half scale periodicity estimator is given by:

$$\hat{f}_i^{ShortH} = \frac{ShortH_i}{\Delta\sum_{j=1}^{M} ShortH_j^2}, \quad \Delta \equiv 1/M$$

The weights used to compute the weighted standard deviation in equation (5) are defined, for all $l = 1, \ldots, T$ and all $i = 1, \ldots, M$, as:

$$\chi_{l,i} = \chi(\bar{r}_{l,i}/\hat{f}_i^{ShortH}),$$
$$\chi(z) = \begin{cases} 1 & \text{if } z^2 \leq 6.635 \\ 0 & \text{otherwise.} \end{cases}$$

B.2 Jump Tests

In this paper, we identify jumps relying mostly on the test proposed by Barndorff-Nielsen and Shephard (2006a) and further developed by Huang and Tauchen (2005). The test statistic, Z_t^{BV}, is given by:

$$Z_t^{BV} = \frac{1 - BV_t/RV_t}{\sqrt{0.61 M^{-1} \max(1, TPQ_t/BV_t^2)}} \sim \mathcal{N}(0,1)$$

where TPQ_t is the realized tripower quarticity, that consistently estimates the integrated quarticity in the presence of jumps and is defined as:

$$TPQ_t = M 1.74 \frac{M}{M-2} \sum_{i=3}^{M} |r_{t,i}|^{4/3} |r_{t,i-1}|^{4/3} |r_{t,i-2}|^{4/3} \xrightarrow{p} \int_{t-1}^{t} \sigma^4(u)\, \mathrm{d}u.$$

The above test is widely used in empirical work due to its simplicity and reasonable size and power properties under various scenarios (see Dumitru and Urga, 2012).

Internet Appendix

Forecasting the Realized Variance in the Presence of Intraday Periodicity

Abstract

This appendix collects additional simulation and empirical results supporting the main paper.

IA. Additional Empirical Results

IA.1 Results for log-transformed HAR and HARP models

Table IA.1: Out-of-sample forecast losses for log-transformed models

		(a) $h = 1$				(b) $h = 5$				(c) $h = 22$			
		HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP/ HAR	HARP-Q/ HAR-Q	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ
SPY	MSE	0.998	0.903*	0.991	0.994	0.990	0.969*	0.997	0.998	1.002	0.985*	1.003	0.988*
	QLIKE	0.987*	0.993	0.990	0.993	0.984*	0.979*	0.987	0.983*	0.994	0.957*	0.988*	0.993
	VaR	0.999	0.997	0.999	0.999	0.990*	0.998	0.995*	0.996	0.991*	0.989*	0.982*	0.997
Avg.	MSE	0.962	0.985	0.972	0.983	0.948	0.952	0.957	0.965	0.976	0.970	0.980	0.971
	QLIKE	0.955	0.942	0.961	0.964	0.955	0.945	0.959	0.966	0.965	0.952	0.969	0.965
Stocks	VaR	0.996	0.994	0.997	0.997	0.999	0.998	0.998	0.998	0.999	0.998	0.999	0.997
		Diebold & Mariano Test – Individual Stocks											
	MSE	24 : 0	7 : 0	15 : 0	10 : 0	25 : 0	9 : 0	21 : 0	13 : 1	11 : 1	14 : 1	16 : 0	12 : 1
	QLIKE	30 : 0	17 : 0	28 : 0	25 : 0	29 : 0	26 : 2	29 : 0	23 : 0	24 : 0	18 : 1	21 : 0	17 : 1
	VaR	11 : 0	10 : 0	10 : 0	9 : 0	8 : 1	5 : 0	6 : 0	6 : 2	5 : 1	3 : 1	6 : 3	4 : 0

Note: This table reports the ratio of the losses from HARP versus HAR models for various forecasting horizons. $^{\star(\blacklozenge)}$ indicates that the losses of the HARP (HAR) models are significantly lower compared to the HAR (HARP) model at the 5% significance level based on the Diebold and Mariano (1995) test. The entries of type "$xx : yy$" summarize the results of the Diebold and Mariano test for the 30 stocks considered. The first number, "xx", shows the number of stocks for which the HARP model significantly outperforms the HAR model, while the second number, "yy", indicates the number of stocks for which the opposite is true.

IA.2 Results based on the jump test by Andersen et al. (2012)

In what follows, we report some out-of-sample empirical findings for an alternative test for jumps, i.e. the test by Andersen et al. (2012). This test relies on the median realized variance to estimate the integrated variation and is shown to have better finite sample properties than the original test by Barndorff-Nielsen and Shephard (2006a). The test statistic is given below:

$$Z_t^{MedRV} = \frac{1 - MedRV_t/RV_t}{\sqrt{0.96M^{-1}\max(1, MedRQ_t/MedRV_t^2)}} \sim \mathcal{N}(0,1), \tag{IA.1}$$

with

$$MedRV_t = \frac{M}{M-2}1.42\sum_{i=2}^{M-1}\text{med}(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}|)^2 \xrightarrow{p} \int_{t-1}^{t}\sigma^2(u)\,\mathrm{d}u$$

and

$$MedRQ_t = \frac{M^2}{M-2}0.92\sum_{i=2}^{M-1}\text{med}(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}|)^4 \xrightarrow{p} \int_{t-1}^{t}\sigma^4(u)\,\mathrm{d}u.$$

As with the Barndorff-Nielsen and Shephard (2006a) test, we use staleness-robust estimators as proposed by Kolokolov and Renò (2024).

Table IA.2: Out-of-sample forecast losses

		(a) $h = 1$		(b) $h = 5$		(c) $h = 22$	
		HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ
SPY	MSE	0.952*	0.932*	0.882*	0.852*	0.982*	0.984
	QLIKE	0.966*	0.904*	0.922*	0.920*	0.997	0.972*
	VaR	0.993	0.992*	0.997	0.998	0.999	0.998
Avg. Stocks	MSE	0.957	0.956	0.961	0.972	0.988	0.988
	QLIKE	0.931	0.930	0.950	0.945	0.973	0.985
	VaR	0.994	0.993	0.999	0.999	0.997	0.996
Diebold & Mariano Test – Individual Stocks							
	MSE	16 : 0	12 : 0	15 : 0	12 : 1	15 : 3	12 : 0
	QLIKE	25 : 0	21 : 1	19 : 1	22 : 1	22 : 0	18 : 0
	VaR	14 : 0	13 : 0	7 : 0	5 : 1	11 : 0	9 : 2

Note: This table reports the ratio of the losses from HARP versus HAR models for various forecasting horizons. Days with jumps are estimated using the ADS test of Andersen et al. (2012) at the 1% significance level. $^{\star(\blacklozenge)}$ indicates that the losses of the HARP (HAR) models are significantly lower compared to the HAR (HARP) model at the 5% significance level based on the Diebold and Mariano (1995) test. The entries of type "$xx : yy$" summarize the results of the Diebold and Mariano test for the 30 stocks considered. The first number, "xx", shows the number of stocks for which the HARP model significantly outperforms the HAR model, while the second number, "yy", indicates the number of stocks for which the opposite is true.

IA.3 Estimating intraday periodicity by varying the rolling window size

In what follows, we estimate intraday periodicity using different size windows, which are equivalent to 40, 125, 250, 500, and 1,000 trading days. Then, we re-run the out-of-sample forecast exercise across all horizons. As in the main results in the paper, we assess the performance using the QLIKE, MSE and the VaR loss functions. We also evaluate the predictive ability of the HARP models using the test proposed by Diebold and Mariano (1995) at the 5% significance level. The model incorporating jump regressors estimate days with significance jumps using the BNS test of Barndorff-Nielsen and Shephard (2006a) at the 1% level. The test is implemented using the staleness-robust adjustment of Kolokolov and Renò (2024).

Table IA.3: Forecasts results for SPY when periodicity is estimated across different size windows

		$W=20$	$W=40$	$W=125$	$W=250$	$W=500$	$W=1000$	$W=20$	$W=40$	$W=125$	$W=250$	$W=500$	$W=1000$
		HARP/HAR						HARP-Q/HAR-Q					
MSE	$h=1$	0.956*	0.922*	0.924*	0.924*	0.918*	0.922*	0.985	0.984	0.975	0.972*	0.974*	0.971*
QLIKE		0.979*	0.988	0.972*	0.968*	0.970*	0.965*	1.355$^\blacklozenge$	1.334$^\blacklozenge$	1.450$^\blacklozenge$	1.166$^\blacklozenge$	1.235$^\blacklozenge$	1.436$^\blacklozenge$
VaR		0.993	0.992*	0.997	0.995	0.997	0.997	1.009	1.004	1.005	1.002	1.003	1.005
MSE	$h=5$	0.934*	0.998	0.989	0.977*	0.998	0.989	0.978	0.966*	0.949*	0.945*	0.925*	0.904*
QLIKE		0.932*	0.970	0.990	0.995	0.958*	0.949*	0.936*	0.939*	0.911*	0.936*	0.987	0.922*
VaR		0.999	0.998	0.995*	0.996	0.996	0.995*	1.001	1.001	0.998	0.996	0.997	0.998
MSE	$h=22$	0.984	0.997	0.963*	0.971	0.970*	0.933*	0.966*	0.976	0.992	1.017	1.012	0.958*
QLIKE		0.957*	0.978	0.965*	0.978	0.984	0.930*	0.930*	0.958*	0.951*	1.001	0.996	0.908*
VaR		0.999	0.999	0.999	0.998	0.997	0.996*	1.003	1.002	0.998	0.996*	0.997	0.997
		HARP-J/HAR-J						HARP-CJ/HAR-CJ					
MSE	$h=1$	0.957*	0.924*	0.923*	0.923*	0.918*	0.921*	0.955*	0.948*	0.921*	0.923*	0.918*	0.921*
QLIKE		0.985	0.981	0.974	0.970*	0.972*	0.967*	1.018	0.975	0.964*	0.961*	0.960*	0.957*
VaR		0.992*	0.991*	0.998	0.997	0.999	0.998	0.991*	0.989*	0.999	0.997	0.999	0.999
MSE	$h=5$	0.934*	0.998	0.989	0.978	0.998	0.989	0.933*	0.993	0.985	0.974*	1.010	0.985
QLIKE		0.928*	0.951*	0.992	0.998	0.960	0.951*	0.950*	0.951*	0.964*	0.974	0.934*	0.924*
VaR		1.000	0.998	0.995*	0.997	0.997	0.995*	0.998	0.997	0.993*	0.995	0.992*	0.993*
MSE	$h=22$	0.985	1.020	0.965*	0.973	0.972*	0.934*	0.976*	0.969*	0.955*	0.962*	0.961*	0.924*
QLIKE		0.956*	0.989	0.966*	0.980	0.985	0.931*	0.945*	0.935*	0.959*	0.970*	0.975	0.926*
VaR		0.999	1.008	0.999	0.998	0.997	0.997	0.994*	0.983*	0.993*	0.990*	0.990*	0.989*

Note: This table reports the out-of-sample forecasts loss ratio across different size windows used to estimate intraday periodicity. $^{\star(\blacklozenge)}$ indicates that the losses of the HARP (HAR) models are significantly lower compared to the HAR (HARP) models at the 5% significance level using the test proposed by Diebold and Mariano (1995). The first row shows the length of the window used to estimate the intraday periodicity, given in number of trading days.

IA.4 Forecast diagnostics based on the Rivers and Vuong (2002) test

Table IA.4: Results for the Rivers and Vuong (2002) test applied to SPY data

			BV		Med-RV	
	HARP/ HAR	HARPQ/ HARQ	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ	HARP-J/ HAR-J	HARP-CJ/ HAR-CJ
$h = 1$	0.811	0.129	0.850	0.674	0.003*	0.000*
$h = 5$	0.039*	0.000*	0.474	0.439	0.000*	0.000*
$h = 22$	0.000*	0.000*	0.040*	0.026*	0.025*	0.006*

Note: The table reports the p-values for the two-sided Rivers and Vuong (2002) test. The test evaluates at the 5% significance level whether the in-sample performance of the HARP (HAR) models is significantly better relative to the performance of the HAR (HARP) models. Starred values indicate significance for the HARP models, while a diamond superscript indicates significance in favour of the HAR models. Results in the "BV" column are based on the Barndorff-Nielsen and Shephard (2006a) test, while those in the "Med-RV" column are based on the Andersen et al. (2012) test.

IA.5 Nesting the HAR and HARP models on a LASSO Approach

Throughout the paper, we have compared the performance of the HAR model with that of our proposed HARP models. These comparisons have been conducted using statistical and economic loss functions, as well as tests for equal predictive ability. So far, we have consistently demonstrated that HARP models significantly outperform their unfiltered counterparts across all forecasting horizons, regardless of whether we consider the SPY or individual stocks. Additionally, a return predictability exercise, which regresses the next monthly excess return on the VRP constructed using the forecasts of the HAR(P) models, corroborates the relevance of filtering out intraday periodicity.

However, in this section, we take a further step in understanding the benefits of filtering out intraday periodicity. To achieve this, we create a nested HAR model that

considers all unfiltered and filtered regressors. This model is outlined below:

$$RV_{t,t+h-1} = \beta_0 + \beta_d RV_{t-1} + \beta_w RV_{t-5,t-1} + \beta_m RV_{t-22,t-1} + \beta_d^f RV_{t-1}^f + \beta_w^f RV_{t-5,t-1}^f + \beta_m^f RV_{t-22,t-1}^f + e_{t,t+h-1}. \quad \text{(IA.2)}$$

We rely on a LASSO approach to eliminate the non-informative or redundant regressors. A priori, we should expect that if filtering intraday periodicity results in more informative realized measures, then unfiltered measures should be removed from the model, and vice versa for the reverse case. We compute these regressions out-of-sample using a rolling window of size equal to 1,000 days, and the intraday periodicity of day t is computed using a time-varying window equivalent to 20 trading days. The regularization parameter is selected using a K-fold cross-validation with $K = 10$.[31]

The proportion of days per year in which the filtered and/or unfiltered coefficients were selected is plotted in Figure IA.1. As can be seen, the daily and weekly filtered coefficients are selected in at least 70% (and up to 91%) of the days, while the same coefficients for unfiltered models are selected less than 50% of the time at $h = 1$ and up to 76% at $h = 22$. Interestingly, at $h = 1$, the daily and weekly unfiltered coefficients are never selected during the Global Financial Crisis (GFC). A plausible rationale for this finding could be explained by the fact that during periods of financial turmoil, investors may rebalance their portfolios more often to capture the best returns, which may further increase activity at the beginning and end of the market, making intraday periodicity more vivid. As a result, unfiltered measures are more prone to be estimated with errors, making them less informative.

Finally, we assess the performance of the HAR-LASSO against the HAR model, and for completeness, we also report the performance of the HARP model. These results are reported in Table IA.5.[32] Looking at the MSE loss function, we find that the HAR-LASSO always outperforms the HAR model, while it outperforms the HARP model at $h = 5$ and

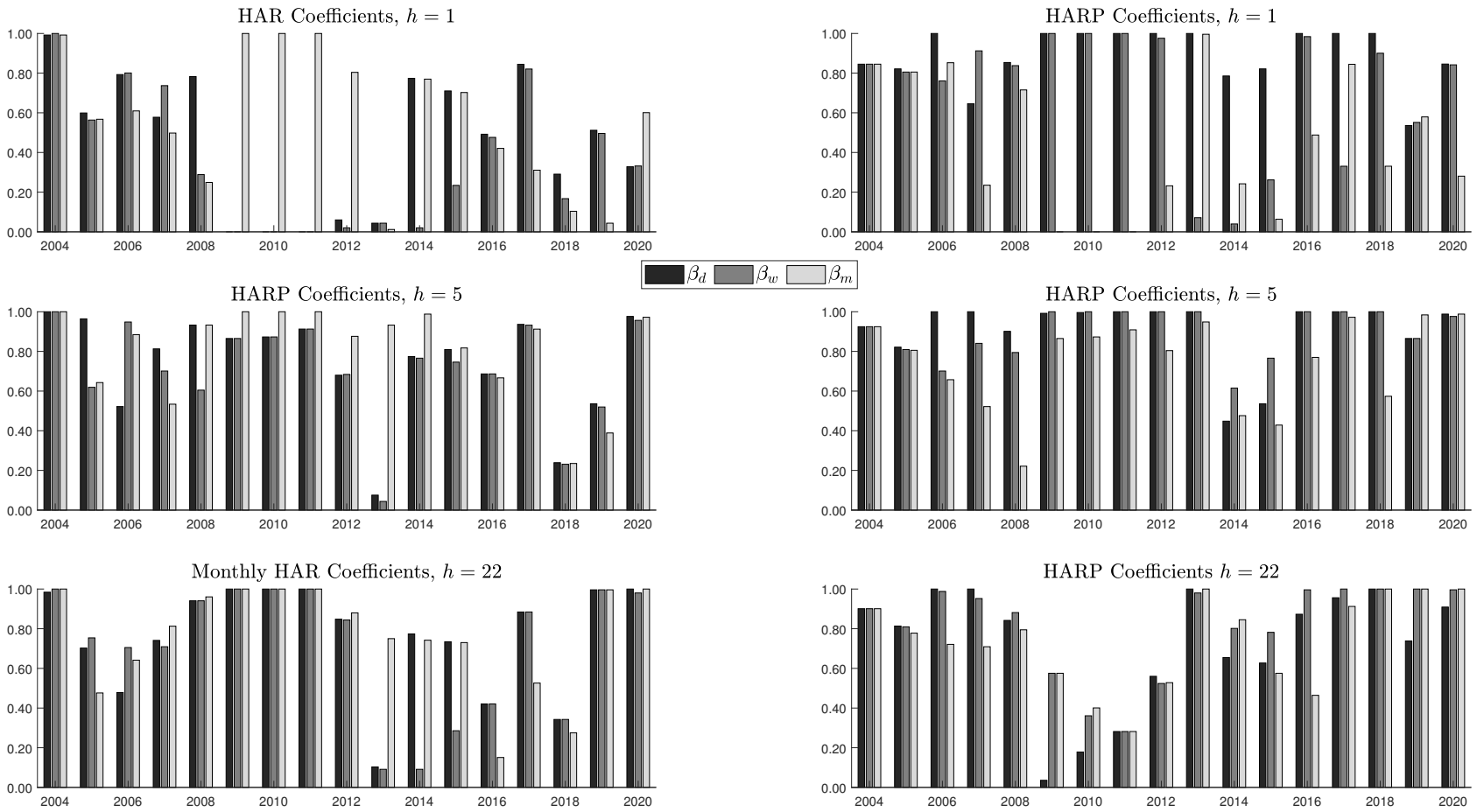
[31] Different values of K render qualitatively similar conclusions.

[32] Please note that the performance is based on a combination of the HAR and HARP regressors, which are selected dynamically throughout the sample period by optimizing the regularization parameter used to shrink the model.

$h = 22$. A very intriguing finding is observed when looking at the QLIKE loss function, where the QLIKE loss of the HAR-LASSO is twice as large as that of the HAR model. To understand this result, Figure IA.2 plots in the left panel the out-of-sample fitted values for the HARP, HAR-LASSO, and the RV, while the right panel depicts the image of the MSE and QLIKE loss functions. As can be seen, the QLIKE is an asymmetric loss that penalizes more heavily underestimates of the true value than overestimates, and as shown in the left panel, the forecasts of the HAR-LASSO –although very close to the RV– generally underestimates the future RV, explaining why the QLIKE provides such a large value.

In sum, the HAR-LASSO approach corroborates our main results, as unfiltered measures are more often excluded from the model, indicating that filtering for periodicity enriches the information set of the realized regressors.

Figure IA.1: HAR-LASSO selection of coefficients



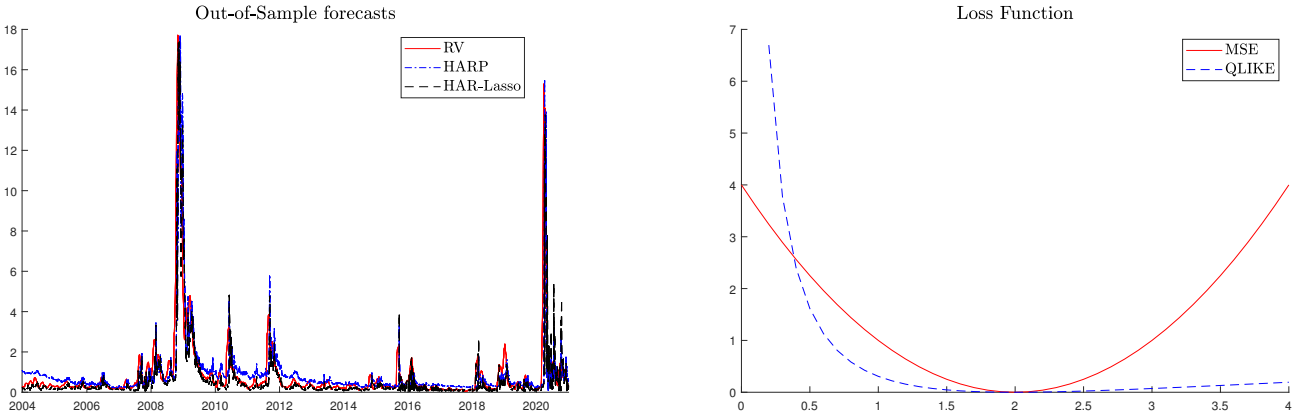
Note: The plot depicts the proportion of days, per year, in which the HAR and HARP coefficients were selected by a HAR-LASSO estimation that includes the filtered and unfiltered measures. From top to bottom, the left (right) panels show the results for HAR (HARP) coefficients for daily, weekly, and monthly out-of-sample regressions. Intraday periodicity was estimated using a time-varying window of length $W = 20$ days.

Table IA.5: HAR-LASSO out-of-sample forecast losses

	HAR/ HARP	HAR/ HAR-LASSO
	$h = 1$	
MSE	0.956	0.965
QLIKE	0.979	1.825
	$h = 5$	
MSE	0.934	0.841
QLIKE	0.932	2.239
	$h = 22$	
MSE	0.984	0.918
QLIKE	0.957	2.578

Note: This table reports the ratio of the losses from HAR-LASSO versus HAR models for various forecasting horizons.

Figure IA.2: Out-of-sample fitted values and loss-function



Note: The plot depicts in two panels the SPY RV together with the out-of-sample fitted values of the HARP and HAR-LASSO models, and the image of the MSE and QLIKE loss functions where the true forecast value is 2.

IA.6 Additional Predictive Regression Results

Table IA.6 reports the predictive regression results controlling only for the conditional variance proxied by the out-of-sample forecasts of the HAR(P) models. The upper (lower) panel contains respectively the results for SPY and individual stocks. As can be seen, the results remain quantitatively and qualitatively similar after removing the implied volatil-

ity spread (SPRD) as a control. That is, the VRP is a strong predictor of future monthly excess returns, and VRP measures derived from HARP models moderately outperform their HAR counterparts as measured by a higher adjusted R^2.

Moreover, our results indicate

Table IA.6: VRP Predictive regressions controlling for CV

	HAR	HAR-Q	HAR-J	HAR-CJ	HAR-P	HARP-Q	HARP-J	HARP-CJ
Panel A: SPY Predictive Regressions								
VRP	1.547	1.505	1.554	1.522	1.896	1.759	1.906	1.817
t-stat^{NW}	[2.458]**	[2.255]**	[2.725]***	[2.563]**	[2.630]***	[2.588]***	[3.032]***	[2.966]***
t-stat^{HD}	[2.157]**	[1.854]*	[2.420]**	[2.381]**	[2.496]**	[2.305]**	[2.688]***	[2.609]***
W^{IVX}	[6.260]**	[4.590]**	[6.432]**	[5.165]**	[7.081]***	[6.102]**	[7.142]***	[6.740]***
CV	2.312	2.340	2.302	2.351	2.157	2.120	2.145	2.188
t-stat^{NW}	[3.267]***	[3.388]***	[3.266]***	[3.250]***	[3.223]***	[3.243]***	[3.223]***	[3.134]***
t-stat^{HD}	[2.422]**	[2.279]**	[2.201]**	[2.258]**	[2.217]**	[2.304]**	[2.206]**	[2.268]**
W^{IVX}	[6.850]***	[5.840]**	[6.904]***	[7.172]***	[6.807]***	[6.972]***	[6.705]***	[7.541]***
adj R^2 (%)	4.887	4.833	4.886	4.976	**4.987**	**4.876**	**5.046**	**5.087**
Panel B: Stocks Predictive Regressions								
VRP	0.162	0.146	0.159	0.160	0.149	0.127	0.151	0.145
t-stat	[1.686]**	[1.490]	[1.726]*	[1.635]	[1.760]*	[1.635]	[1.774]*	[1.800]*
CV	0.066	0.074	0.068	0.071	0.073	0.084	0.072	0.074
t-stat	[0.921]	[1.024]	[0.938]	[1.016]	[0.998]	[1.143]	[0.987]	[0.987]
adj R^2 (%)	0.091	0.077	0.094	0.096	**0.104**	**0.091**	**0.106**	**0.108**

Note: The table shows, in two panels, results for predictive regressions, where the monthly excess return is regressed on the VRP and CV (e.g., Bekaert and Hoerova, 2014). For SPY (Panel A), we report t-statistics computed using the Newey-West statistics (t-stat^{NW}) with 44-lags, Hodrick (1992) standard errors (t-stat^{HD}), and the robust Wald test (W^{IVX}) of Kostakis et al. (2015). For stocks, we employ a panel regression framework with firm fixed effects and compute t-statistics using clustered robust standard errors (e.g. Petersen, 2008). Adjusted R-squares are reported as a percentage and are displayed in bold font whenever they exceed those of the HAR (HARP) counterparts. Symbols *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Table IA.7: Stocks monthly Predictive Regressions

	HAR	HAR-Q	HAR-J	HAR-CJ	HAR-P	HARP-Q	HARP-J	HARP-CJ
Panel A: Univariate Predictive Regressions								
VRP	0.103	0.082	0.098	0.097	0.110	0.080	0.113	0.107
t-stat^{NW}	[1.969]**	[1.555]	[1.911]*	[1.945]*	[2.174]**	[1.687]*	[2.221]**	[2.052]**
R^2 (%)	0.054	0.033	0.050	0.047	**0.060**	**0.037**	**0.063**	**0.056**
Panel B: Predictive Regressions with Controls								
VRP	0.183	0.165	0.181	0.180	0.170	0.147	0.172	0.168
t-stat^{NW}	[1.685]*	[1.551]	[1.660]*	[1.663]*	[1.769]*	[1.629]	[1.743]*	[1.746]*
SPRD	0.133	0.138	0.134	0.134	0.137	0.144	0.136	0.138
t-stat^{NW}	[2.076]**	[2.101]**	[2.086]**	[2.106]**	[2.095]**	[2.125]**	[2.095]**	[2.188]**
CV	0.115	0.126	0.116	0.119	0.123	0.137	0.122	0.124
t-stat^{NW}	[1.088]	[1.195]	[1.103]	[1.163]	[1.161]	[1.307]	[1.152]	[1.146]
adj R^2 (%)	0.176	0.168	0.178	0.178	**0.187**	**0.174**	**0.185**	**0.189**

Note: The table shows results for the monthly predictive regressions $r_{i,t+1} = \alpha + \beta VRP_{i,t} + \boldsymbol{\theta}' \boldsymbol{X}_{i,t} + e_{i,t+1}$, where $r_{i,t+1}$ denotes the next month's return in excess of the 3-month T-bill rate for the ith stock, $\boldsymbol{X}$ is the matrix with controls, and $\boldsymbol{\theta}$ is the vector of coefficients. SPRD is the implied volatility spread (e.g., Cremers and Weinbaum, 2010), and CV is the conditional variance (e.g. Bekaert and Hoerova, 2014). The model is estimated within a panel framework with firm fixed effects. The t-statistics are estimated using clustered robust standard errors with Newey-West correction with 44 lags. R-squares are reported as a percentage and are displayed in bold font whenever they exceed those of the HAR (HARP) counterparts. Symbols *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

IB. Additional Results on Simulated Data

For all simulated models, we use an Euler scheme based on 23,400 initial data points (corresponding to seconds). We simulate a total of 1,000 sample paths of length 1,000 days.

IB.1 One-factor volatility model plus jumps

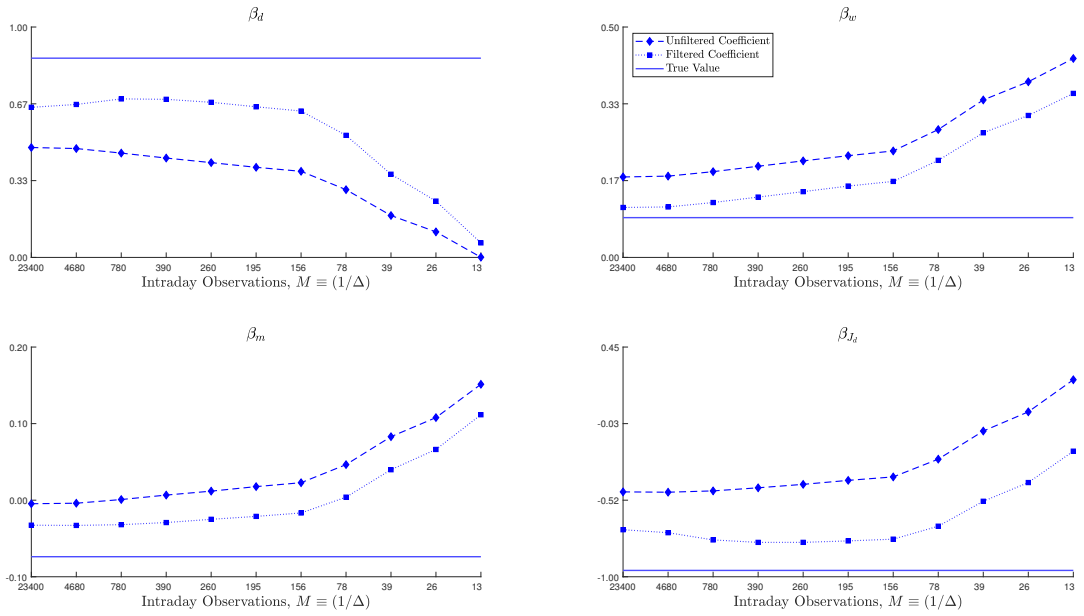
$$\mathrm{d}p(t) = 0.03\,\mathrm{d}t + f(t)\nu(t)\left(-0.62\,\mathrm{d}W_{\nu_1}(t) + \sqrt{0.6156}\,\mathrm{d}W_p(t)\right) + z_p(t)\,\mathrm{d}N(t), \quad \text{(IB.3a)}$$

$$\nu^2(t) = \exp\{0.125\nu_1^2(t)\}, \quad \mathrm{d}\nu_1^2(t) = -0.1\nu_1^2(t)\,\mathrm{d}t + \mathrm{d}W_{\nu_1}(t),$$

$$N(t) \sim \text{Poisson}(0.4\text{t}), \quad z_p(t) \sim \mathcal{N}(0, 1.284).$$

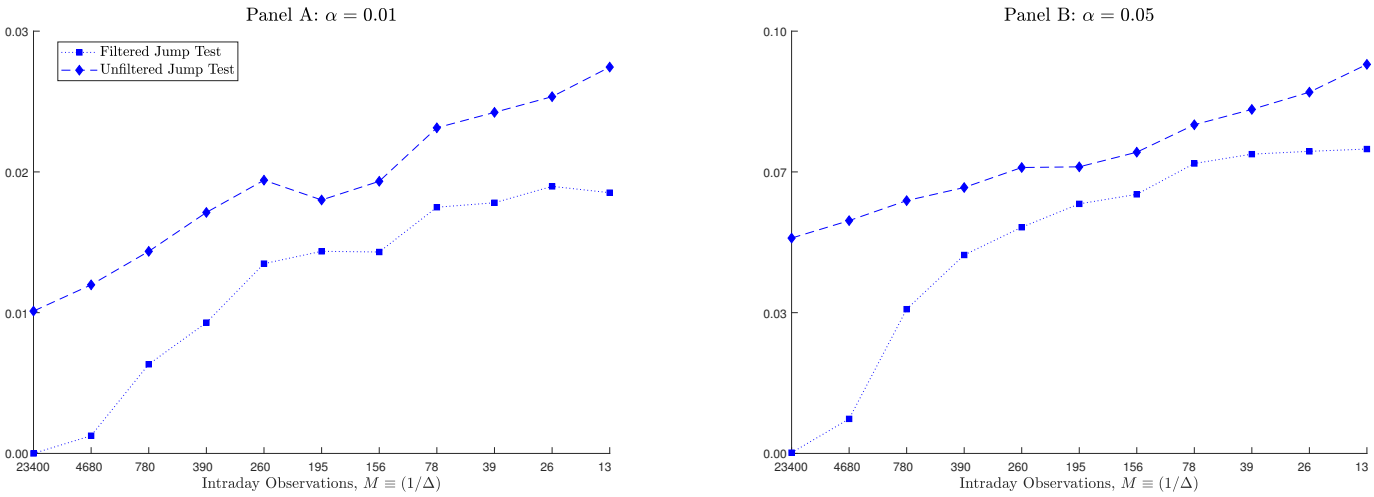
$$f(t) = 0.88929198 + 0.75e^{-10t} + 0.25e^{-10(1-t)},$$

Figure IB.3: HARP-J and HAR-J coefficients for the SV1F model with jumps



Note: This figure compares the estimates of the HARP-J (squared marker) and HAR-J (diamond marker) models against the true estimates across different sampling frequencies. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

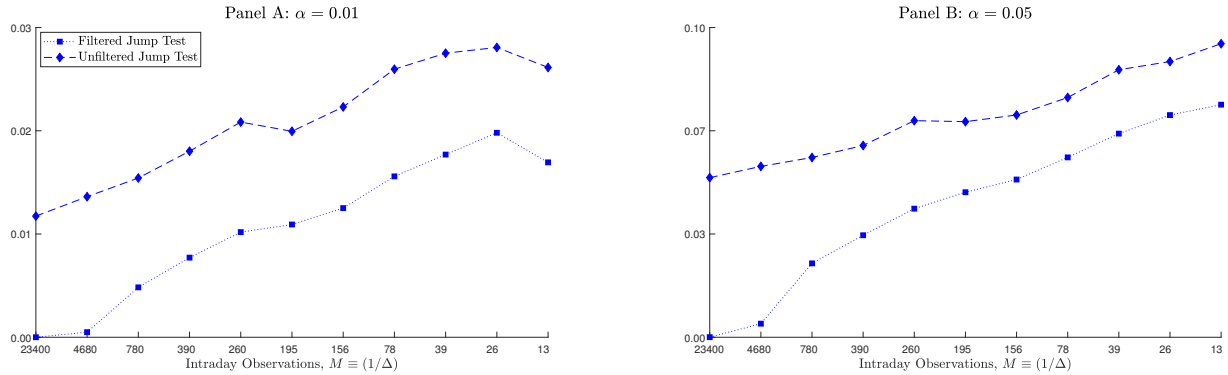
Figure IB.4: Proportion of spurious jumps by sampling frequency for filtered and unfiltered data



Note: This plot graphs the proportion of spurious jumps across sampling frequencies. Jumps were detected using the Barndorff-Nielsen and Shephard (2006a) jump test evaluated at the 1% and 5% significance level. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

IB.2 Proportion of spurious jumps

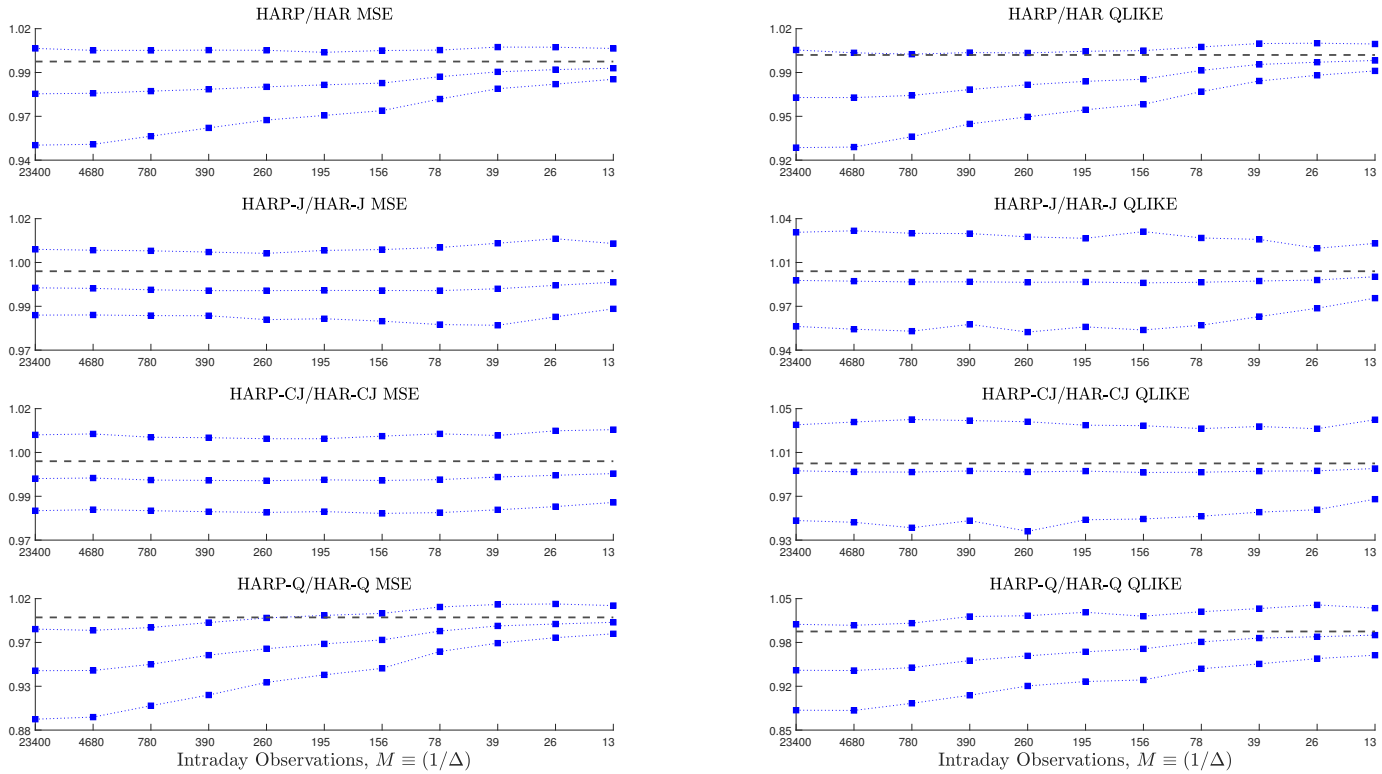
Figure IB.5: Proportion of spurious jumps by sampling frequency for filtered and unfiltered data based on the jump test by Andersen et al. (2012)



Note: This plot graphs the proportion of spuriously detected jumps across sampling frequencies using the jump test given in (IA.1), evaluated at the 1% and 5% significance level. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

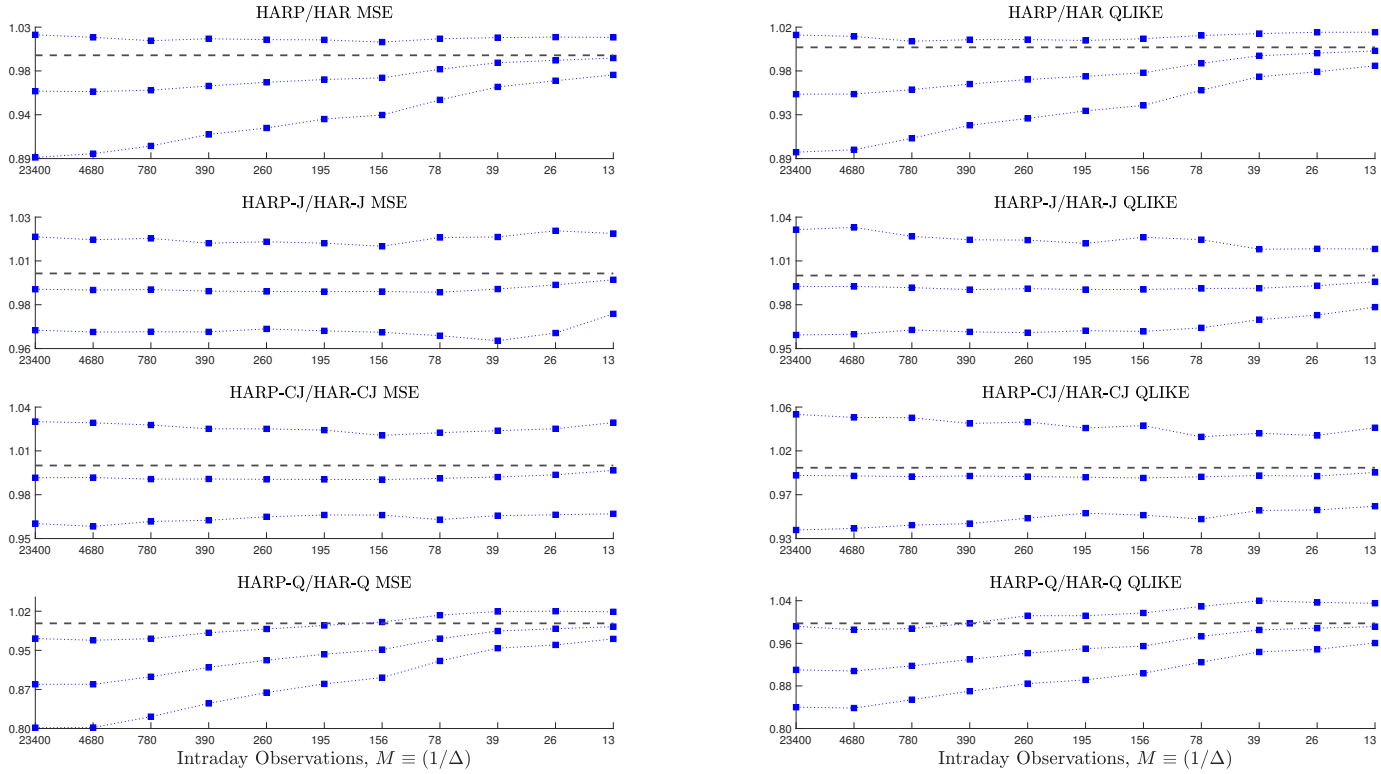
IB.2.1 Loss ratios for the one-factor model plus jumps

Figure IB.6: One-day ahead loss ratio



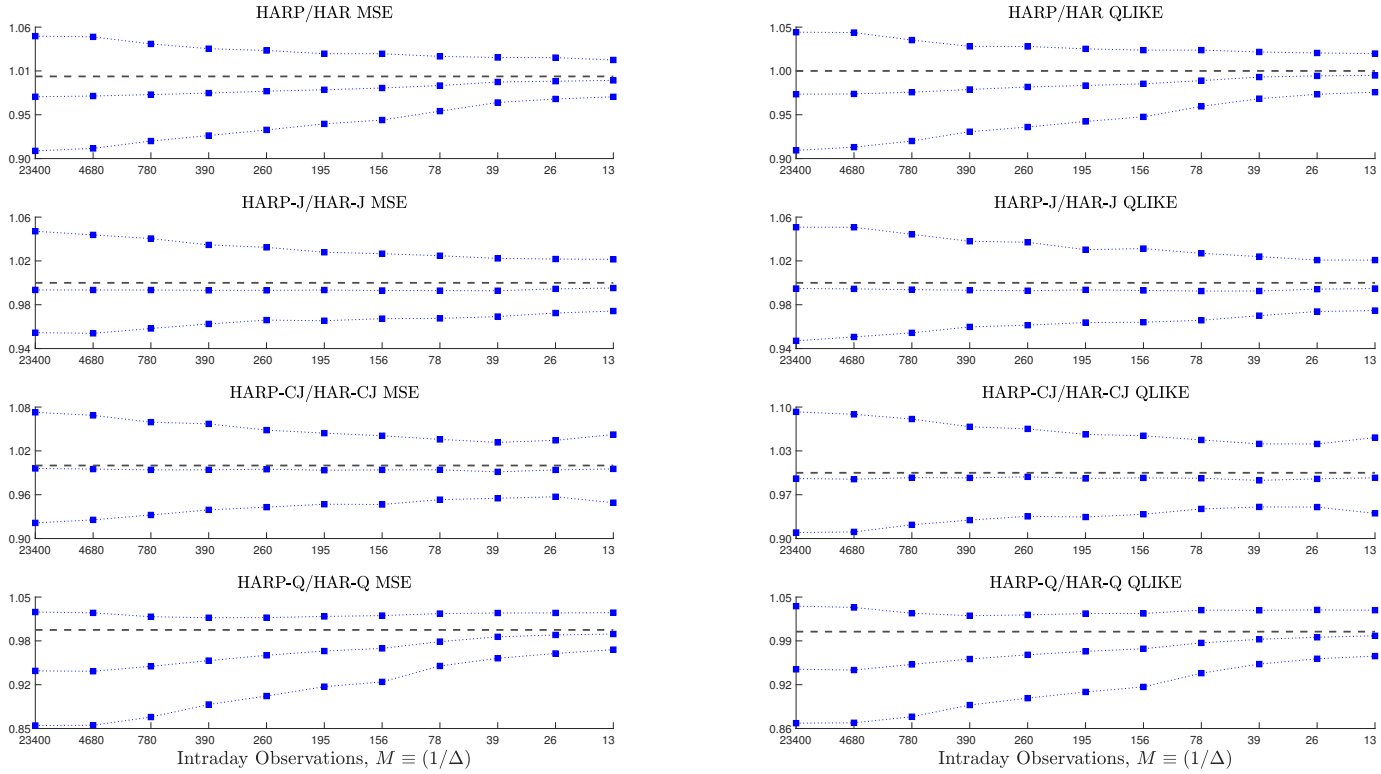
Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

Figure IB.7: One-week ahead loss ratio



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

Figure IB.8: One-month ahead loss ratio



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

IB.3 Two-factor volatility model

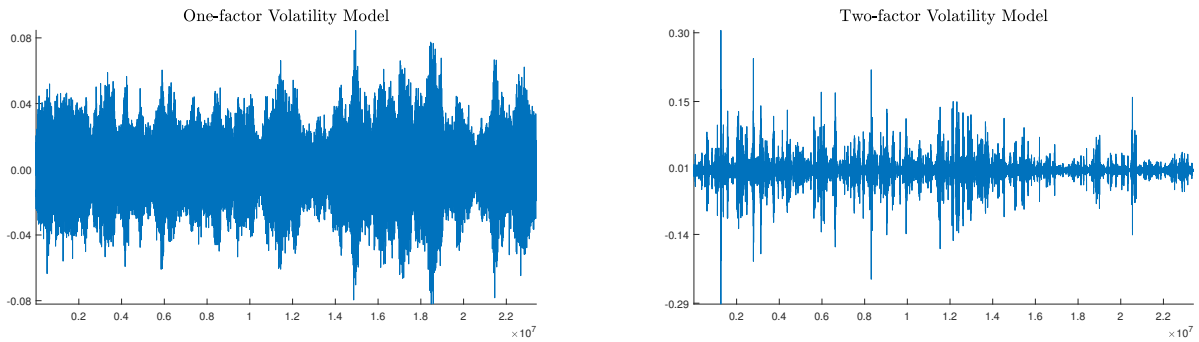
Two volatility factors (SV2F) (IB.4a)

$$
\begin{aligned}
\mathrm{d}p(t) &= 0.03\,\mathrm{d}t + f(t)\nu(t)\left(-0.3\,\mathrm{d}W_{\nu_1}(t) - 0.3\,\mathrm{d}W_{\nu_2}(t) + \sqrt{0.82}\,\mathrm{d}W_p(t)\right), \\
\nu^2(t) &= \text{s-exp}\{-1.2 + 0.04\nu_1^2(t) + 1.5\nu_2^2(t)\}, \\
\mathrm{d}\nu_1^2(t) &= -0.00137\nu_1^2(t)\,\mathrm{d}t + \mathrm{d}W_{\nu_1}(t), \\
\mathrm{d}\nu_2^2(t) &= -1.386\nu_2^2(t)\,\mathrm{d}t + \left(1 + 0.25\nu_2^2(t)\right)\mathrm{d}W_{\nu_2}(t).
\end{aligned}
$$

$$
f(t) = 0.88929198 + 0.75e^{-10t} + 0.25e^{-10(1-t)}, \tag{IB.4b}
$$

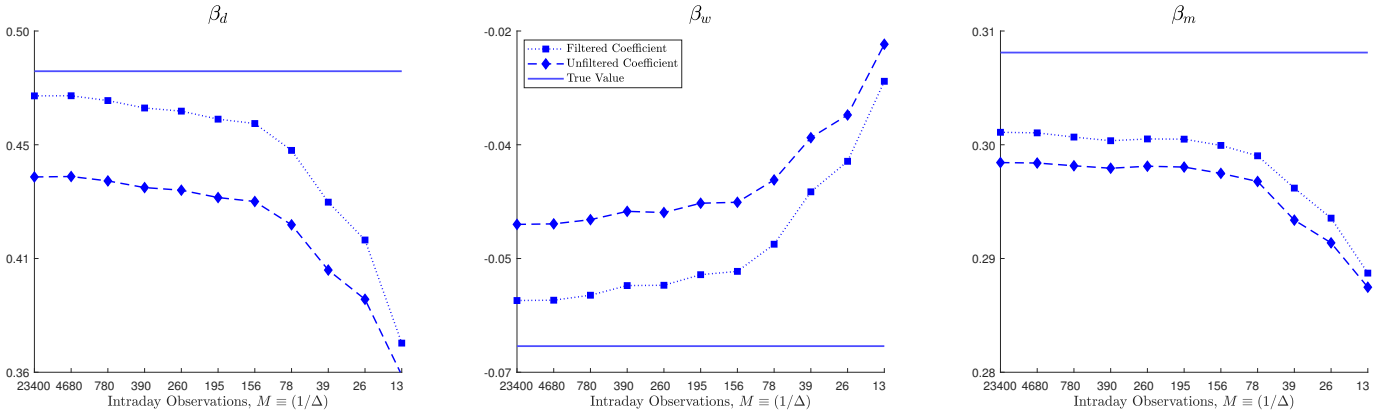
where s-exp denotes the exponential function with a polynomial spline at high values to avoid explosive behaviour.

Figure IB.9: Simulated returns for the one-factor stochastic volatility model (no jumps) and the two-factor stochastic volatility model



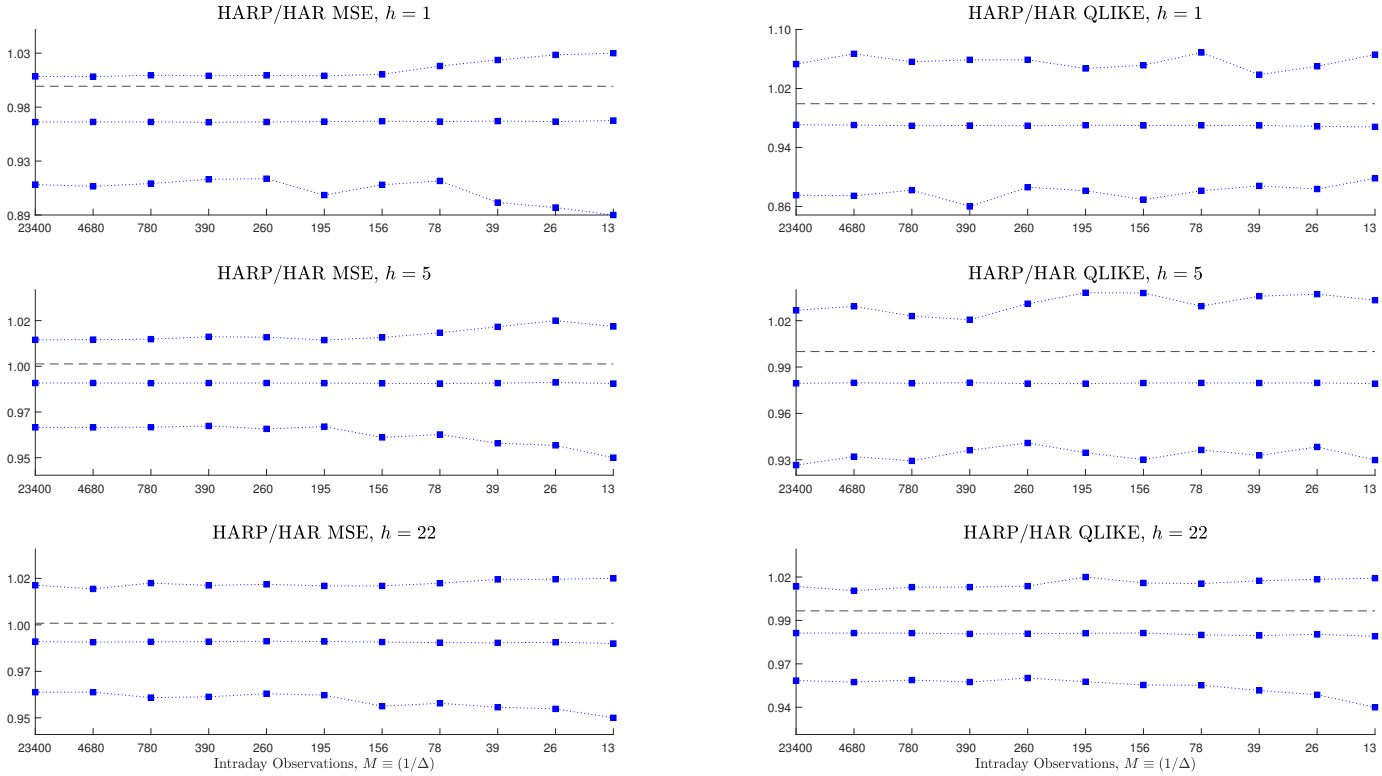
Note: The plot shows the dynamics of simulated 1-second returns over 1000 trading days.

Figure IB.10: HARP and HAR coefficients



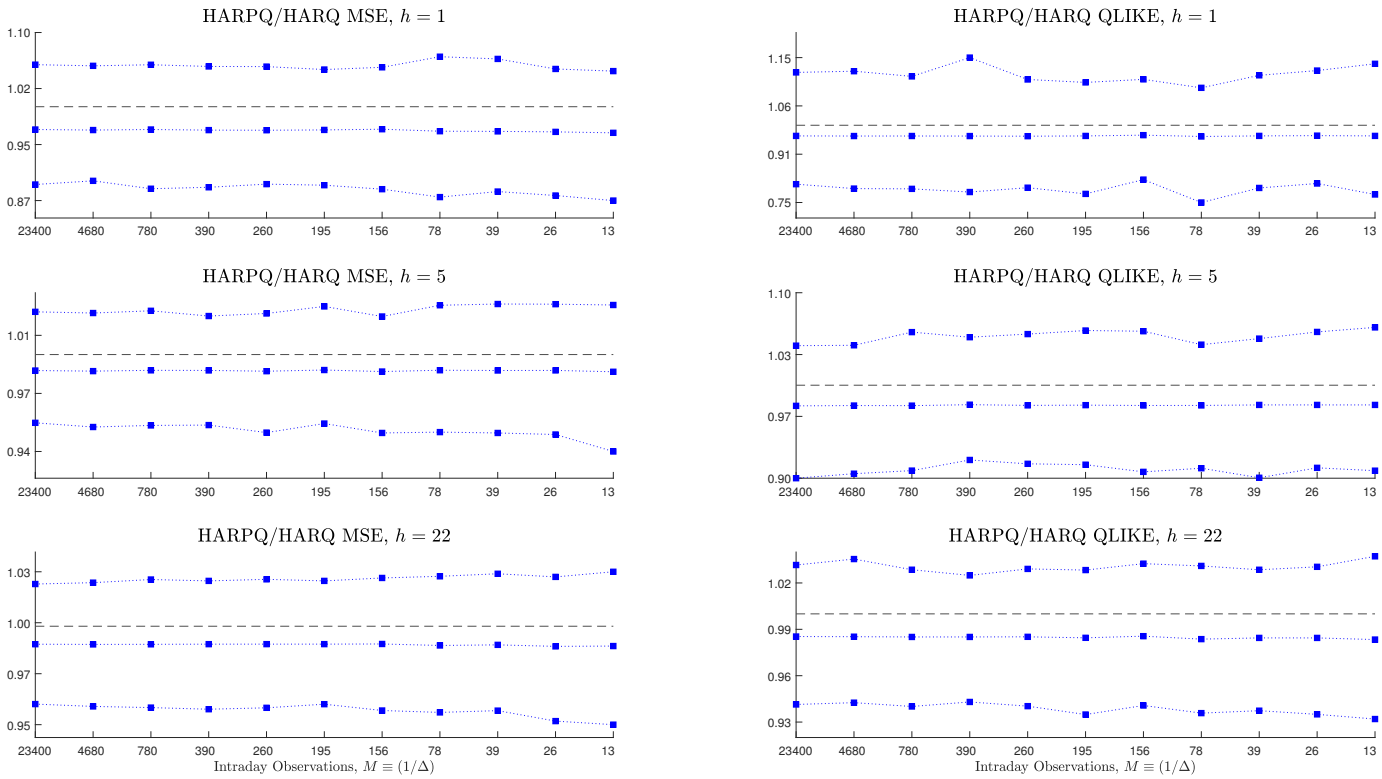
Note: This figure compares the estimates of the HARP (squared marker) and HAR (diamond marker) models against the true estimates across different sampling frequencies. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13). From left to right, the first panel of the figure corresponds to β_d from (12), the middle panel to β_w, while the last panel to β_m. The straight line in each panel represents the corresponding estimates on the daily quadratic variance. These are referred to as "true" coefficients.

Figure IB.11: HARP/HAR loss ratio for the two-factor volatility model



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for the HARP versus the HAR model. All forecasting horizons are included: one-day ($h = 1$), one-week ($h = 5$) and one-month ($h = 22$). The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

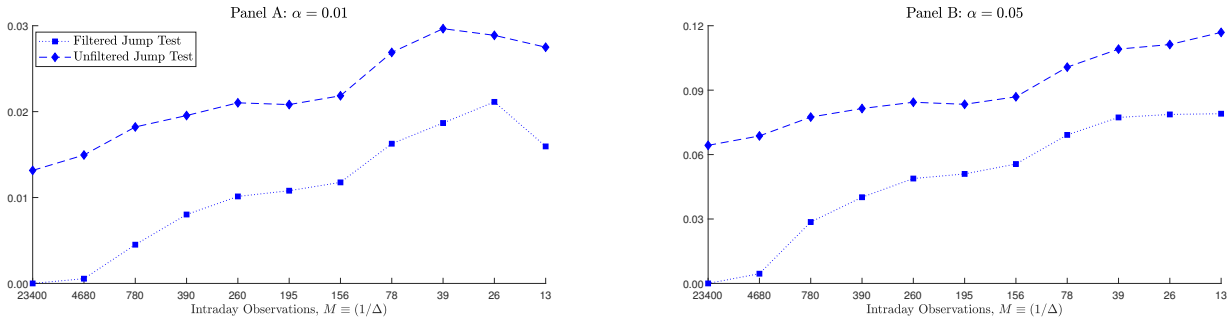
Figure IB.12: HARP-Q/HAR-Q loss ratio for the two-factor volatility model



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for the HARP versus the HAR model. All forecasting horizons are included: one-day ($h = 1$), one-week ($h = 5$) and one-month ($h = 22$). The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

IB.4 One-factor volatility model with jumps and co-jumps in volatility

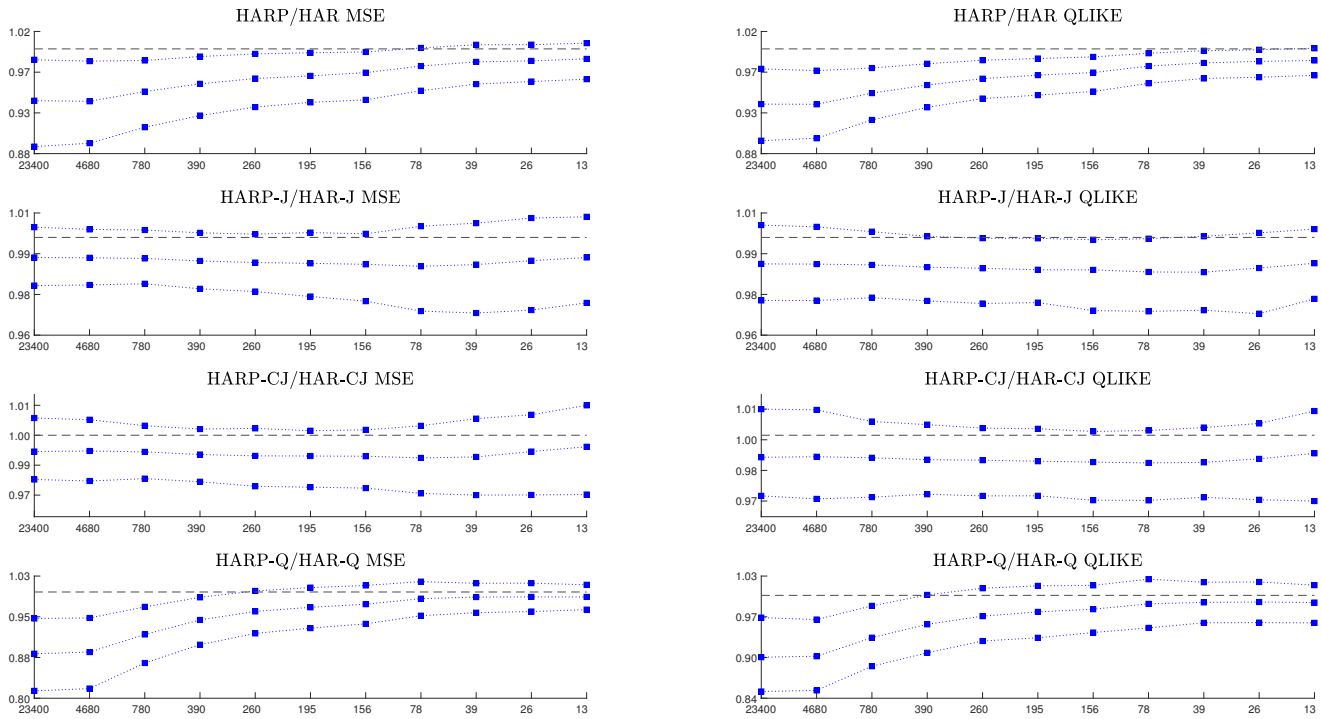
Figure IB.13: Proportion of spurious jumps by sampling frequency for filtered and unfiltered data using the jump test by Andersen et al. (2012)



Note: This plot graphs the proportion of spuriously detected jumps across sampling frequencies using the jump test given in (IA.1), evaluated at the 1% and 5% significance level. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

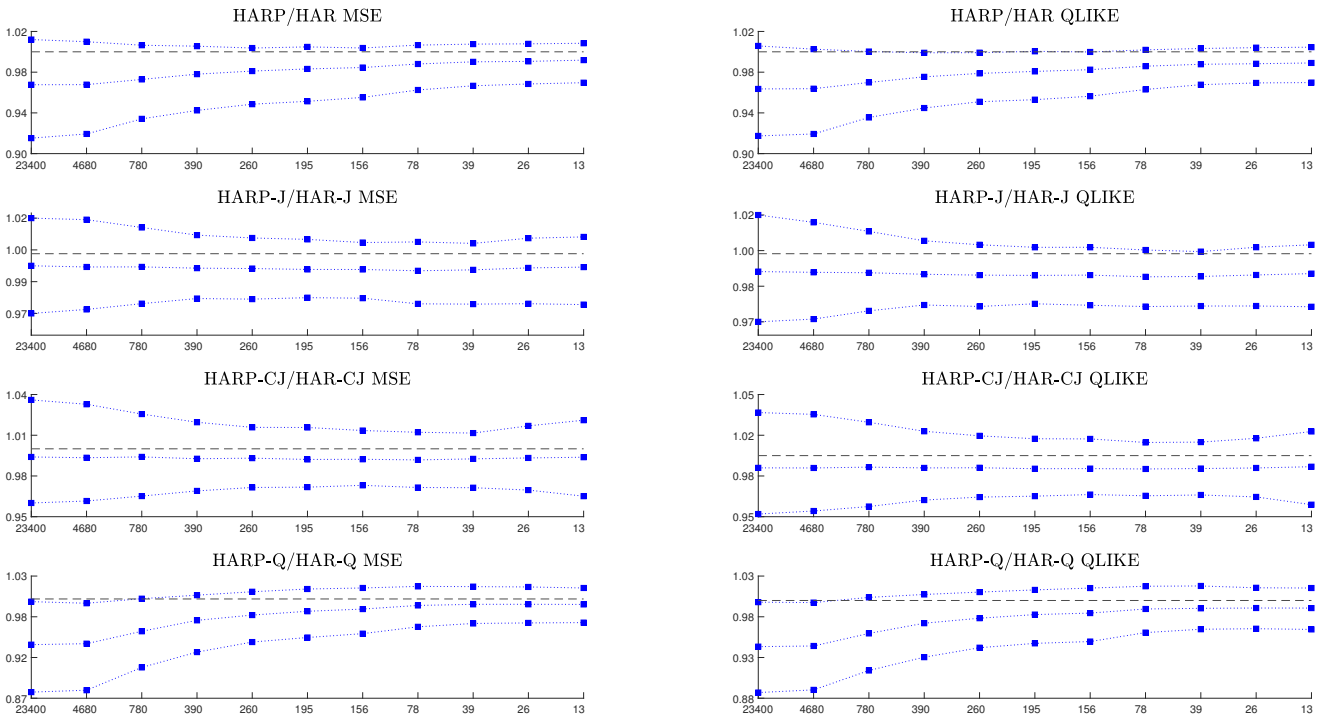
The figure also illustrates that, for the jump test proposed byAndersen et al. (2012), more spurious jumps are detected for unfiltered returns, a result that holds true across all sampling frequencies. This result corroborates our main results based on the Barndorff-Nielsen and Shephard (2006a) test, thereby suggesting that jump regressors in models HAR-J and HAR-CJ are likely to be affected by periodicity-related estimation error, which can further impact the forecast of the realized variance.

Figure IB.14: One-week ahead loss ratio for the simulated model



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

Figure IB.15: One-month ahead loss ratio for the simulated model



Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for HARP versus HAR models. The dashed horizontal line corresponds to the value 1. The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

Figures IB.14 and IB.15 plot the median and the 5% and 95% quantiles of the forecast loss ratios from the HARP models versus the HAR models against the sampling frequency for one-week and one-month forecast horizons. In most cases, all quantiles are consistently below 1 across sampling frequencies. These results corroborate those reported for the one-day-ahead forecasts, thereby confirming that HARP models provide more accurate forecasts than models that do not account for intraday periodicity.

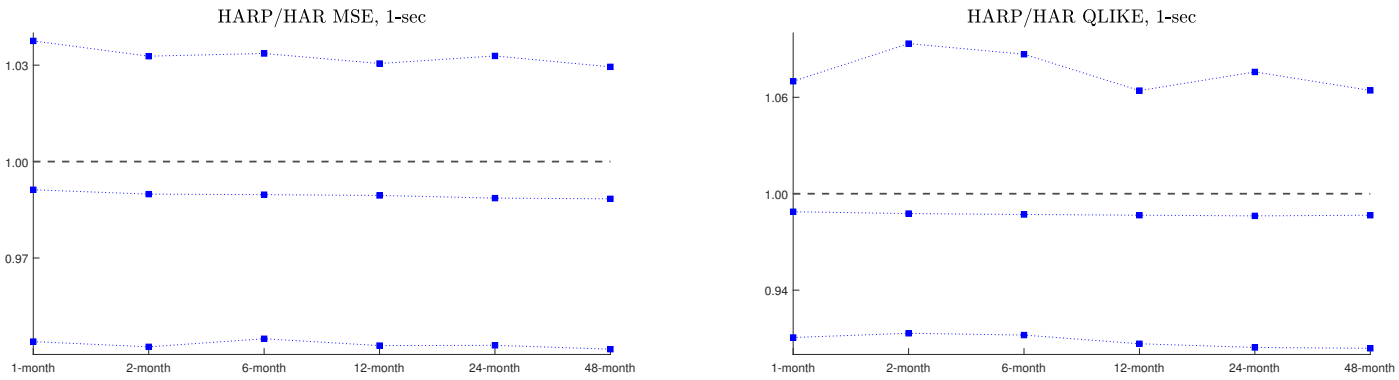
IB.5 Sensitivity and validity analysis

This section acts as a robustness check for our results. It explores how various sources of estimation error in the intraday periodicity estimates can impact our findings. We ultimately show that our results hold even in the presence of such errors. The estimation

error in disentangling periodicity has two main sources: the number of days used to estimate periodicity is too short and the jumps in the price interfere with periodicity estimation, especially at higher frequencies.

A time-varying periodicity function (see Andersen et al., 2018) calls for estimating periodicity over shorter windows of time. As shorter estimation windows can lead to less reliable periodicity estimates, figure IB.16 explores the sensitivity of our results on forecasting RV to the length of the periodicity estimation window. We plot the distribution of the HARP/HAR loss ratios obtained at the highest sampling frequency for the SV2F model against the length of the periodicity estimation window. At very high sampling frequencies and in the absence of jumps, the impact of measurement error emanating from any other source than the length of the estimation window is insignificant.

Figure IB.16: The impact of the length of the periodicity estimation window on the performance of HARP models



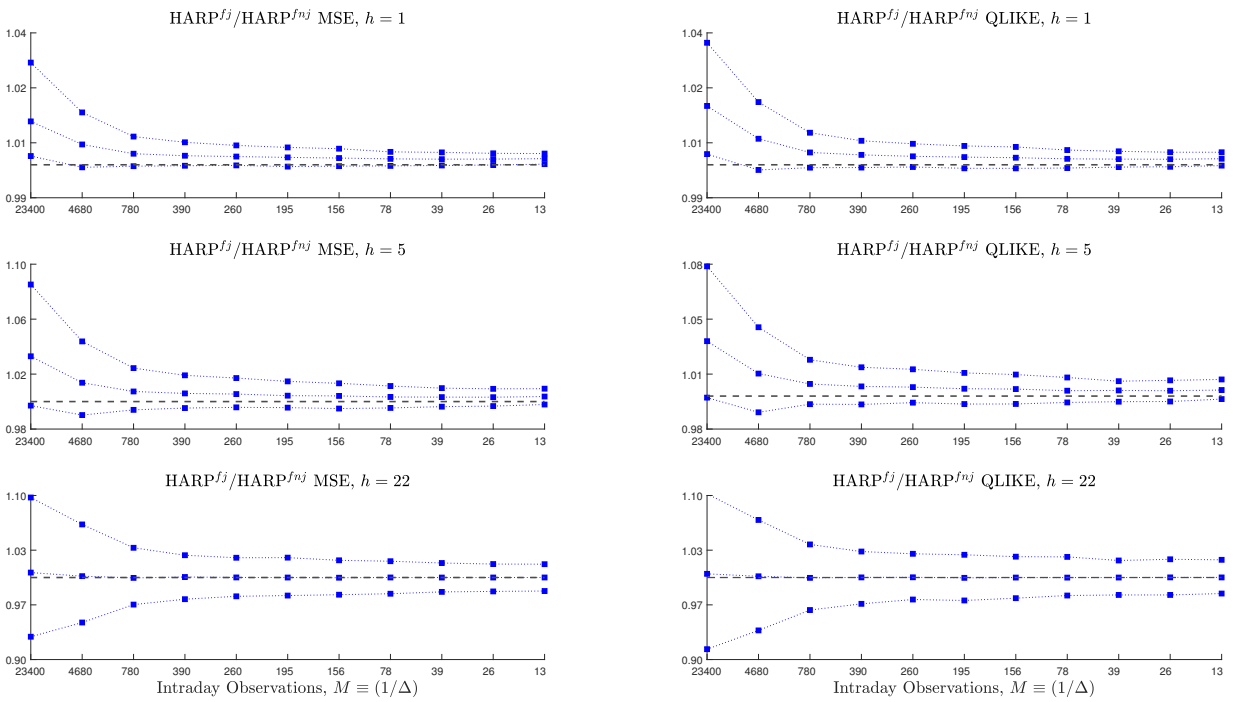
Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for the HARP versus the HAR model against the length of the time window over which periodicity was estimated. Simulations are based on the SV2F model and shown results are for a sampling frequency equal to 1 sec.

The distribution of the loss ratios does not change much with the length of the estimation window for periodicity. The median is always below 1, confirming that filtering improves the forecasting performance. The distribution of the QLIKE ratios is slightly more dispersed than the distribution of the MSE ratios for shorter estimation windows.

We further consider the impact of the jump-related periodicity estimation error on our analysis. To this end, we compare the HARP forecast loss for the filtered SV1F process with jumps to the forecast loss for the filtered SV1F model to which we add jumps only

after applying the periodicity filter. Specifically, for the latter forecast loss, we apply the periodicity filter at different sampling frequencies before adding the jumps also sampled correspondingly. The first forecast loss is impacted by jump-related periodicity estimation error, while the second loss is not. The distributions of the ratios of the two losses for different forecasting horizons are plotted against the sampling frequencies in figure IB.17. The "$HARP^{fj}$" notation indicates the HARP model where filtering ("f") occurs on data with jumps ("j"), while the "$HARP^{fnj}$" denotes the forecasting model where filtering ("f") occurs on data with no jumps ("nj").

Figure IB.17: The impact of the jump-related periodicity estimation error on the performance of HARP models



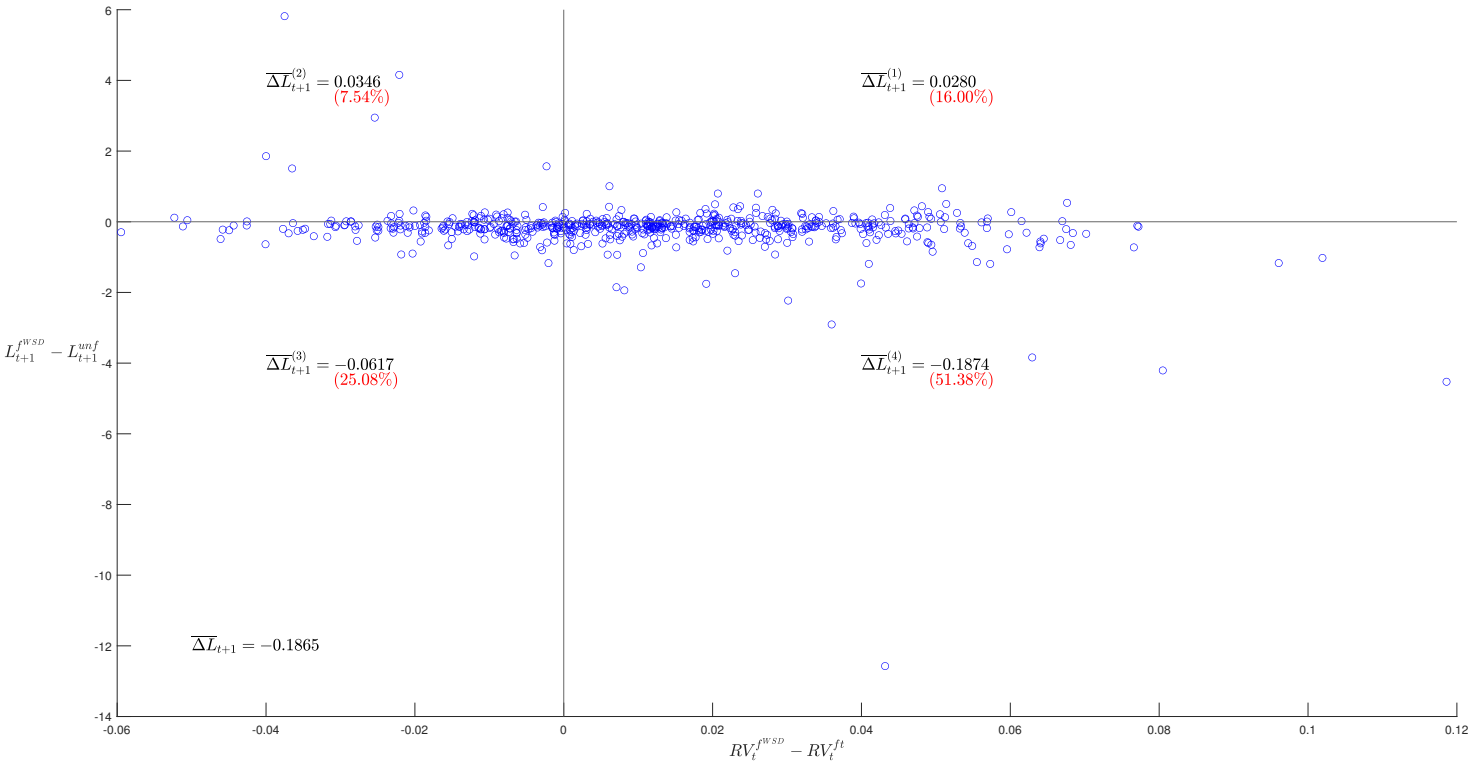
Note: The figure plots the median and the 5% and 95% quantiles for the MSE and QLIKE loss ratios, for the $HARP^{fj}$ model, for which filtering is applied to data containing jumps, versus the $HARP^{fnj}$ model, for which filtering is performed before adding jumps to the data. Simulations are based on the SV1F model plus jumps. All forecasting horizons are included: one-day ($h = 1$), one-week ($h = 5$) and one-month ($h = 22$). The number of intraday observations on the x axis corresponds to the following sampling frequencies: 1 second (23400), 5 seconds (4680), 30 seconds (780), 1 minute (390), 1.5 minutes (260), 2 minutes (195), 2.5 minutes (156), 5 minutes (78), 10 minutes (39), 15 minutes (26) and 30 minutes (13).

At high frequencies, the distribution of the loss ratio shifts above 1 and is more dis-

persed than at lower frequencies. This shift is mostly visible for the 1-day and 1-week ahead forecasts, where all three quantiles are located above 1 for sampling frequencies higher than 30 seconds. For the one-month ahead forecasts, the median and the 95% quantile at high frequency are above 1, indicating an upwards shift, but the distribution is a lot more dispersed, with the 5% quantile well below 1. In this case, aggregation of data over long horizons makes the impact of jump-related estimation error less clear in terms of direction, but still very much visible in terms of dispersion. For all forecasting horizons, the impact of jump-related estimation error gradually decreases with the sampling frequency.

Finally, we examine whether excessive or insufficient filtering impacts our results. We rely on the SV2F 5-minute data and employ the forecast loss definition used in presenting the Diebold and Mariano (1995) test in section 4.3. Let $L(\epsilon_{t+1}^{WSD})$ and $L(\epsilon_{t+1}^{unf})$ be, respectively, the HARP and HAR model forecast losses computed using the MSE loss function. Let RV_t^{ft} and RV_t^{WSD} be the realized variance estimators based on, respectively, returns filtered by the true periodicity, and returns filtered with the weighted standard deviation method outlined in section 1. We define excessive filtering the situation for which $RV_t^{WSD} < RV_t^{ft}$ and insufficient filtering when $RV_t^{WSD} > RV_t^{ft}$. In figure IB.18, we plot the loss differential $L(\epsilon_{t+1}^{WSD}) - L(\epsilon_{t+1}^{unf})$ against $RV_t^{WSD} - RV_t^{ft}$. The surface of the plot is split in four quadrants based on the criteria: $RV_t^{WSD} - RV_t^{ft} \lessgtr 0$ and $L(\epsilon_{t+1}^{WSD}) - L(\epsilon_{t+1}^{unf}) \lessgtr 0$. In each quadrant, we also report the average loss difference per quadrant, $\overline{\Delta L}_{t+1}$, as well as the percentage of points.

Figure IB.18: One-day ahead loss differential as a function of the amount of filtering



Note: The figure depicts the loss differential for HARP versus HAR models, $L(\epsilon^{WSD}_{t+1}) - L(\epsilon^{unf}_{t+1})$, against the previous day difference between the filtered and no periodicity realized volatilities, $RV^{WSD}_t - RV^{ft}_t$. The loss function considered is the MSE. Each quadrant reports the average loss difference, $\overline{\Delta L}_{t+1}$, as well as the percentage number of points in that quadrant (red). Data is generated from the SV2F model.

Overall, filtering leads to forecast gains, as more than 75% of the points in the scatter plot are situated below the line $L(\epsilon^{WSD}_{t+1}) = L(\epsilon^{unf}_{t+1})$, where loss differentials are also higher in absolute value. Over 67% of the points are to the right of the $RV^{WSD}_t = RV^{ft}_t$ line, out of which 51% are in the fourth quadrant, where $L(\epsilon^{WSD}_{t+1}) - L(\epsilon^{unf}_{t+1}) < 0$ and $RV^{WSD}_t - RV^{ft}_t > 0$. This quadrant also features the most extreme points of the scatter plot and the highest loss difference in absolute value. The second quadrant, defined by $L(\epsilon^{WSD}_{t+1}) - L(\epsilon^{unf}_{t+1}) > 0$ and $RV^{WSD}_t - RV^{ft}_t < 0$, also contains some extreme points, showing that excessive filtering can have adverse effects. This is not worrisome though, as this quadrant has the lowest percentage of points and a relatively low loss differential in absolute value.

IC. Estimation of Risk-Neutral Measures

Option data are downloaded from OptionMetrics. We used standardized option data from the volatility surface file to have the same maturity for our options every day. We compute the Risk-Neutral Variance (RNV) of the option-implied stock return distribution using the model-free methodology of Bakshi et al. (2003). Using the time t prices of out-of-the-money (OTM) call, $C_t(\tau; K)$, and put, $P_t(\tau; K)$, options with strike price K and time-to-expiration τ, the RNV is defined as:

$$RNV_t(\tau) = \exp(r_f \tau) V_t(\tau) - \mu_t(\tau)^2, \tag{IC.5}$$

where r_f is the risk-free rate, and $\mu_t(\tau)$ is given by:

$$\mu_t(\tau) = \exp(r_f \tau) - 1 - \frac{\exp(r_f \tau)}{2} V_t(\tau) - \frac{\exp(r_f \tau)}{6} W_t(\tau) - \frac{\exp(r_f \tau)}{24} X_t(\tau), \tag{IC.6}$$

and $V_t(\tau)$, $W_t(\tau)$, and $X_t(\tau)$ are the time t prices of τ-maturity quadratic, cubic, and quartic contracts, defined as contingent claims with payoffs equal to the second, third, and fourth power of the asset log-return, respectively. The corresponding prices of these contracts are given by:

$$V_t(\tau) = \int_{S_t}^{\infty} \frac{2\left(1 - \log\left(\frac{K}{S_t}\right)\right)}{K^2} C_t(\tau; K) dK + \int_0^{S_t} \frac{2\left(1 + \log\left(\frac{S_t}{K}\right)\right)}{K^2} P_t(\tau; K) dK, \tag{IC.7}$$

$$\begin{aligned} W_t(\tau) = &\int_{S_t}^{\infty} \frac{6\log\left(\frac{K}{S_t}\right) - 3\left(\log\left(\frac{K}{S_t}\right)\right)^2}{K^2} C_t(\tau; K) dK \\ &- \int_0^{S_t} \frac{6\log\left(\frac{S_t}{K}\right) + 3\left(\log\left(\frac{S_t}{K}\right)\right)^2}{K^2} P_t(\tau; K) dK, \end{aligned} \tag{IC.8}$$

$$X_t(\tau) = \int_{S_t}^{\infty} \frac{6\log\left(\frac{K}{S_t}\right) - 3\left(\log\left(\frac{K}{S_t}\right)\right)^2}{K^2} C_t(\tau; K)dK$$
$$- \int_{0}^{S_t} \frac{6\log\left(\frac{S_t}{K}\right) + 3\left(\log\left(\frac{S_t}{K}\right)\right)^2}{K^2} P_t(\tau; K)dK, \tag{IC.9}$$

where S_t is the price of the underlying stock price. To compute the integrals, we interpolate implied volatilities between the lowest and highest available moneyness using cubic splines and perform constant extrapolation with 1% and 300% moneyness as bounds, resulting in 1000 grid points. Then, these implied volatility are converted to option prices using the Black and Scholes (1973) formula and use those prices to numerically calculate the above integrals.[33]

[33]It is important to note that this approach does not assume that the Black and Scholes (1973) model is valid. Rather, the Black and Scholes (1973) formula is only used as a one-to-one mapping between the option prices and IVs. This is done because fitting the IV curve is much easier than fitting option prices as IVs are comparable across strikes.