

# Exploiting the Information of Signed Jumps and Jump Activity in Forecasting Stocks Returns Volatility

Rodrigo Hizmeri<sup>†</sup>, Marwan Izzeldin<sup>†</sup>, Anthony Murphy<sup>\*‡</sup>, and Mike G. Tsionas<sup>†</sup>

<sup>†</sup>*Lancaster University Management School*

<sup>‡</sup>*Federal Reserve Bank of Dallas*

First version: February 14, 2019

## Abstract

We document the forecasting gains achieved by incorporating measures of signed, finite, and infinite jumps in forecasting the volatility of equity prices, using high-frequency data from 2000 to 2016. We consider the SPY and 20 stocks that vary by sector, volume and degree of jump activity. We use extended HAR-RV models, and consider different frequencies (5, 60, and 300 seconds), forecast horizons (1, 5, 22, and 66 days) and the use of standard and robust-to-noise volatility and threshold bipower variation measures. Incorporating signed finite and infinite jumps generates significantly better real-time forecasts than the HAR-RV model, although no single extended model dominates. In general, standard volatility measures at the 300 second frequency generate the smallest real-time mean squared forecast errors. Finally, the forecasts from simple model averages generally outperform forecasts from the single best model.

**Keywords:** Realized volatility, Signed Jumps, Finite Jumps, Infinite Jumps, Volatility Forecasts, Noise-Robust Volatility, Model Averaging.

**JEL classification:** C22, C51, C53, C58.

---

<sup>◊</sup>We would like to thank Ingmar and Sandra Nolte, Gerald Steele, seminar participants at the University of Warwick and participants at the 11<sup>th</sup> International Conference in Computational Statistics and Financial Econometrics in London, the 2018 QuantFest at the Federal Reserve Bank of Chicago and the 12<sup>th</sup> International Conference in Computational Statistics and Financial Econometrics in PISA for helpful comments. The views expressed are those of the authors, and do not represent the views of the Federal Reserve Bank of Dallas, or the Federal Reserve System.

\*Corresponding address: Federal Reserve Bank of Dallas, 2200 N. Pearl St., Dallas, Texas 75201. Email address: anthony.murphy@dal.frb.org

# 1 Introduction

Modeling and forecasting asset return volatility is central to asset pricing, portfolio optimization and risk management. As volatility is an unobservable variable, its estimation is a challenge for both researchers and practitioners. Popular approaches to modeling and forecasting volatility include the (G)ARCH family of models proposed by [Engle \(1982\)](#) and [Bollerslev \(1986\)](#), and the stochastic volatility (SV) models proposed by [Taylor \(1986\)](#). However, these models fail to capture certain stylized facts, including the long-memory aspect of volatility.

The introduction and use of high-frequency data provided a framework for directly measuring and capturing the main stylized facts of volatility. This enables us to calculate a non-parametric measure of volatility, realized volatility (RV), given as the sum of the intra-day squared returns. Early adoption of RV in modeling and forecasting was featured in the work of [Andersen and Bollerslev \(1998\)](#), [Andersen et al. \(2001, 2003, 1999, 2005\)](#), [Barndorff-Nielsen and Shephard \(2002a,b\)](#), and [Meddahi \(2002\)](#) inter alios.

In these early studies, the implicit or explicit assumption was that the underlying asset return process was best described by a continuous, diffusion process, which cannot generate the fat tails observed in the distribution of returns. For that reason, the literature has advocated the use of jump diffusion processes to explain the behavior of stocks returns (see [Chernov et al., 2003](#); [Christoffersen et al., 2008](#); [Duan et al., 2006](#); [Eraker et al., 2003](#), among others). Recent developments using threshold and multipower variation estimators dissect the components of realized variance into its continuous and jump parts (see, for instance, [Aït-Sahalia, 2004](#); [Barndorff-Nielsen and Shephard, 2002a, 2004](#); [Corsi et al., 2010](#); [Mancini, 2001, 2009](#)). These estimators were further utilized to build daily and intra-daily jump tests that are capable of detecting significant jumps (see [Andersen et al., 2007b](#); [Huang and Tauchen, 2005](#)). Moreover, the work of [Aït-Sahalia and Jacod \(2012\)](#) provide a unified framework that allows the finer characterizations of jumps into finite and infinite activity jumps.

As a consequence of these developments, the literature has explored the use of jumps to forecast RV. These studies have generally employed a Heterogeneous Autoregressive (HAR) framework, as proposed by [Müller et al. \(1997\)](#), and subsequently popularized by [Corsi \(2009\)](#) as the HAR-RV. For instance, [Andersen et al. \(2007a\)](#) introduce the HAR-RV-J and HAR-RV-CJ specifications that extend the HAR-RV model using a daily measure of the jump component, and lagged values of the continuous and discontinuous components. Whilst jumps based on the bi-power variation (BPV) of [Barndorff-Nielsen and Shephard \(2004\)](#) improve the out-of-sample performance, the threshold bi-power variation (TBPV) of [Corsi et al. \(2010\)](#) generates the largest out-of-sample forecasting gains. The reason is that TBPV is a more robust estimator of the integrated variance than the BPV, which tends to be upwardly biased.

Further developments were subsequently made using signed jumps. [Barndorff-Nielsen et al.](#)

(2010) propose the estimation of realized semivariances, which are constructed as the sum of the squares of positive/negative returns in a fixed interval of time. The difference between the two semi-variances is the so-called signed jumps, which are only present when the distribution of returns is asymmetric. Several studies have considered the use of jumps and signed jumps to forecast realized volatility (see, for instance, Andersen et al., 2007a; Busch et al., 2011; Corsi et al., 2010; Duong and Swanson, 2015; Forsberg and Ghysels, 2007; Ghysels and Sohn, 2009; Giot and Laurent, 2007; Martens et al., 2009; Patton and Sheppard, 2015; Prokopczuk et al., 2016; Sévi, 2014, and references therein).

Whether we consider jumps or signed jumps, the literature generally provides mixed evidence regarding their value added in forecasting. There are two strands here. The first reports gains in forecasting from incorporating jumps. Andersen et al. (2007a) find that separating the jump component from volatility improves out-of-sample volatility forecasts, and that jumps are closely related to macroeconomic news. Corsi et al. (2010) show that the use of a threshold bipower estimator to obtain the jump component leads to greater out-of-sample performance. Patton and Sheppard (2015) argue that volatility is strongly related to the volatility of past negative returns, and report that negative jumps produce better out-of-sample performance. Duong and Swanson (2015) study the importance of large and small jumps using higher-order power variations, and find that small jumps are more important for forecasting volatility than large jumps.

In contrast to these positive findings, many researchers report that jumps do not significantly improve volatility forecasts. For instance, Forsberg and Ghysels (2007), Giot and Laurent (2007), Martens et al. (2009), Busch et al. (2011), Sévi (2014), and Prokopczuk et al. (2016) consider the use of both total and signed jumps to forecast future volatility. Their results suggest that the inclusion of jumps produces a better fitting in-sample model, but does not generate any significant out-of-sample forecasting gains.

This paper contributes to the literature in a number of ways. First, we address the general question of whether jumps add to forecasting. We then consider the contribution of jumps in volatility forecasts by studying whether dissecting jumps by activity (finite/infinite) and by sign (positive/negative) helps to predict future volatility. Finite jumps are generally linked to macroeconomic announcements and so almost certainly play an important role, whereas infinite jumps may or may not be relevant to financial decision making and asset allocation. To the best of our knowledge only Duong and Swanson (2015) have attempted to use jump components in the form of large and small jumps using higher-order power variations. Here, we have two reservations. First, as shown by Ghysels and Sohn (2009), this class of models works best with volatility measures that are not based on squared returns, which is the type of volatility that researchers and practitioners care about the most. Second, a drawback in disentangling jumps

using higher-order power variation is the choice of the power value. High power values can generate inaccurate estimates.<sup>1</sup>

Our approach in this paper is to implement novel theoretical advances in the areas of jump testing and the characterization of continuous time processes with jumps in order to isolate and dissect jumps by their activity (finite/infinite) and sign (positive/negative). We use the so-called ABD test in Andersen et al. (2007b), to test for jumps in the asset returns process. The idea underlying this method is to normalize the intraday returns and then evaluate the normalized returns using the Šidák approach. A day is further classified as a jump day if at least one of the intraday returns exceeds the Šidák approach.<sup>2,3</sup> Additionally, we use the work of Aït-Sahalia and Jacod (2011), SFA test hereafter, to test whether a day is driven mainly by finite or infinite activity jumps.<sup>4</sup> The SFA test has a null of finite activity. Hence, when the null is rejected, we do not know whether the rejection is driven by the presence of finite jumps and/or Brownian motion increments. To address this issue, we combine the ABD and SFA tests, thereby identifying days with infinite jumps (when both tests are rejected) and days with finite jumps (when only the ABD test is rejected). To obtain signed jumps, we intersect the ABD and SFA test with an indicator function for the sign of the jumps.

It is well established in the literature that the characterization of jumps varies by sampling frequency. At higher frequencies, say 5 seconds, infinite jumps are more abundant than at lower frequencies, say 300 seconds. However, the presence of market microstructure at higher frequencies may cast doubt on this characterization making the analysis using the previous methodology irrelevant since we do not observe the true price. Instead, we observe a contaminated price, so making the realized volatility a very noisy estimator (see Bandi and Russell, 2006; Zhang et al., 2005). In order to correctly capture the characterization of jumps at higher frequencies, we use the work of Christensen et al. (2014) and Jacod et al. (2009) to estimate noise-robust estimators of quadratic and integrated variation. We introduce two methodological innovations. First, we modify the ABD test to obtain a version that is noise-robust. Second, we use a two-time scale realized semivariance, which is a modification of the work of Zhang

---

<sup>1</sup>Aït-Sahalia and Jacod (2014) show that, when using power values ( $p$ ) below(above) 2, the limit of the power variation converges to the continuous(jump) component, while for  $p = 2$ , we obtain the RV estimator. In the absence of a clear choice for  $p$ , the user must arbitrarily select a value of  $p > 2$ . For values of  $p$  close to 2, say  $p = 2.5$ , the jump component will retain some Brownian increments and infinite jumps. On the contrary, finite jumps are isolated when the value of  $p$  is big enough, say  $p \geq 4$ . However, this method exaggerates big jumps and when they are used to predict volatility using linear regression, the future volatility is either over or under-estimated depending on the sign of the coefficient.

<sup>2</sup>For other examples of work in this area see Aït-Sahalia and Jacod (2009b), Barndorff-Nielsen and Shephard (2006), Jiang and Oomen (2008), Huang and Tauchen (2005). For a survey of jump tests, see Dumitru and Urga (2012) and the references therein.

<sup>3</sup>We choose the ABD test after testing the power of several jump tests under a DGP that allows for infinite jumps (alpha-stable process).

<sup>4</sup>Using individual stocks, Aït-Sahalia and Jacod (2012, 2014) show that the degree of jump activity at higher-frequencies is in the range of [1.4, 1.6], which implies the existence of both finite and infinite jumps activity.

[et al. \(2005\)](#).

Our application uses high-frequency, TickData data from 2000 to 2016. We forecast the volatility of the SPY and 20 stocks, which vary by sector and volume, using extended HAR-RV models. We consider different frequencies (5, 60, and 300 seconds), forecast horizons (1, 5, 22, and 66 days) and the use of noise-robust measures. We find evidence that jumps characterize the structure of both SPY and the 20 individual stocks under examination.

Our results suggest improvements in performance, in both in- and out-of-sample, when jumps are used as additional predictors in the extended HAR-RV models. We focus on the mean squared prediction error (MSPE) results from real time, pseudo out-of-sample forecasts using rolling window regressions. The classification of jumps by activity shows that infinite jumps are relatively more important at shorter horizons, whereas finite jumps are superior at longer horizons. We find little difference in jump asymmetry. Incorporating signed finite and infinite jumps generates significantly better real-time forecasts than the HAR-RV model, although no single extended model dominates.

The use of noise-robust estimators substantially improves the out-of-sample performance of our extended HAR-RV models, especially at very high frequencies. The gains are greater for the individual stocks than for SPY. However, the greatest gains in real-time forecast performance are generally found using returns sampled at 300 seconds intervals, rather than at 5 or 60 seconds intervals. At the 300 second frequency, the forecasting performance of standard and noise-robust measures is similar. In line with [Ghysels and Sinko \(2011\)](#), robust-to-noise measures only improve the forecasting performance at higher frequencies when market microstructure noise is important.

Finally, since no single model dominates in terms of forecasting performance, we show that simple model averaging generally results in economically significant out-of-sample forecasting performance. These gains are documented using both the SPY and individual stocks across horizons. The gains are greatest using returns sampled every 300 seconds. We assess the predictive accuracy of the model averaging using the pair-wise DM test of [Diebold and Mariano \(2002\)](#). The results show that the model averaging produces significantly smaller MSPEs, even at long horizons of 66 days / three months.

The remainder of the paper is as follows: Section 2 introduces the theoretical background. Section 3 describes the estimation of jumps and their decomposition. Robust-to-noise volatility measures are also discussed. Section 4 highlights the forecasting framework, and sets out the forecasting models and forecast evaluation criteria. The data used in this study are described in Section 5, where the incidence of various types of jumps are detailed. The real-time forecasting results are discussed in Section 6, where the gains from adding the different types of jumps to extended HAR-RV models are documented. A range of model averaging results are presented

in Section 7. Finally, Section 8 summarizes the paper and presents our conclusions.

## 2 Theoretical Background

Let  $X$  be an Itô-semimartingale defined on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , with the following representation

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + J_t, \quad t \in [0, T] \quad (1)$$

where  $a$  is a locally bounded and predictable drift term,  $\sigma$  is the spot volatility, being both adapted and càdlàg. In the current context,  $X_t$  is the log of the price of an equity or an equity index. Additionally,  $W$  is a standard Brownian motion and  $J_t$  is a pure jump process given by  $J_t^F + J_t^I$ , where  $J_t^F$  has finite activity and  $J_t^I$  has infinite activity and is a Lévy process.

$$J_t^F := \int_0^t \int_{|x| \geq \epsilon} x \mu(dx, ds) = \sum_{l=1}^{N_t} \gamma_l \quad (2)$$

$$J_t^I := \int_0^t \int_{|x| \leq \epsilon} x (\mu(dx, ds) - \nu(dx) ds) \quad (3)$$

where  $\mu$  is a Poisson random measure with compensator  $\nu$ , and  $\epsilon > 0$  is an arbitrary number.  $N_t$  is a Poisson process with constant intensity  $\lambda$ , jumping at times denoted by  $(\tau_l)_{l=1, \dots, N_t}$ , and where each  $\gamma_l$  denotes the size of the jump occurring at  $\tau_l$ . All  $\gamma_l$  are i.i.d. and independent from  $N_t$ . For more details on Itô-semimartingale processes, see [Aït-Sahalia and Jacod \(2014\)](#) and the references therein.

Since volatility is a latent variable, realized measures are often employed to give consistent estimates of the quadratic variation of the process on a fixed interval  $[0, T]$ , using high-frequency data. The Quadratic variation of the price process is defined as

$$QV_t = \underbrace{\int_0^t \sigma_s^2 ds}_{\text{Integrated Variation (IV)}} + \underbrace{\sum_{0 < s \leq t} (\Delta_s X)^2}_{\text{Contribution from Jumps}} \quad (4)$$

where  $\Delta_s X := X_s - X_{s-} \neq 0$ , if and only if  $X$  jumps at time  $s$ .  $\sum_{0 < s \leq t} (\Delta_s X)^2$  is the total number of discontinuities on the interval  $[0, t]$ . A widely known realized measure is the realized volatility (RV), which converges in probability to the quadratic variation as  $\Delta_n \rightarrow 0$ . The RV is defined as

$$RV_t = \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} (\Delta_i^n X)^2 \xrightarrow{P} QV_t \quad (5)$$

where  $\lfloor \cdot \rfloor$  is the integer part. The log-returns are defined as  $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$ ,  $i = 1, 2, \dots, 1/\Delta_n (\equiv M)$ , where  $\Delta_n$  is the equally-spaced sampling interval. To separate the integrated variation from the jump part, we use the threshold bi-power variation (TBPV) – which

is a modified version of the so-called bi-power variation, [Barndorff-Nielsen and Shephard \(2004\)](#) – proposed by [Corsi et al. \(2010\)](#). The TBPV is robust to jumps in both the stochastic limit and the asymptotic distribution defined as follows

$$TBPV_t = \mu_1^{-2} \frac{M}{M-1} \sum_{i=2}^{\lfloor 1/\Delta_n \rfloor} |\Delta_i^n X| \mathbf{1}_{\{(\Delta_i^n X)^2 \leq \vartheta_i\}} |\Delta_{i-1}^n X| \mathbf{1}_{\{(\Delta_{i-1}^n X)^2 \leq \vartheta_{i-1}\}} \xrightarrow{p} \int_0^t \sigma_s^2 ds \quad (6)$$

where  $\mu_1 \equiv \mathbb{E}|Z| = \sqrt{2/\pi}$ ,  $Z \sim \mathcal{N}(0, 1)$ ,  $M/(M-1)$  is a small sample correction,  $\vartheta = c_\vartheta^2 \widehat{V}_t$ ,  $c_\vartheta^2 = 3$ , and  $\widehat{V}_t$  is an auxiliary estimator of  $\sigma^2$ .

[Barndorff-Nielsen et al. \(2010\)](#) introduce the realized semivariance (RSV) estimator. The RSV captures the variation of only positive or negative returns and is defined as

$$RSV_t^+ = \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} (\Delta_i^n X)^2 \mathbf{1}_{\{\Delta_i^n X > 0\}} \xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 ds + \sum_{0 < s \leq t} (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s > 0\}} \quad (7)$$

$$RSV_t^- = \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} (\Delta_i^n X)^2 \mathbf{1}_{\{\Delta_i^n X < 0\}} \xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 ds + \sum_{0 < s \leq t} (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s < 0\}} \quad (8)$$

[Barndorff-Nielsen et al. \(2010\)](#) show that the difference of the two semivariances is equal to the signed jumps, obtained by subtracting the positive and negative semivariances; i.e. the halves integrated variations canceled out.

## 3 Realized Jumps

### 3.1 Jump test statistics

To identify days with significant jumps, we employ the intraday jump test proposed by [Andersen et al. \(2007b\)](#), and select the highest intra-daily value of the test within a day. If this value is greater than the cut-off, we classify this day as a jump day. The jump test is defined as

$$ABD_{t,i} = |\Delta_i^n X| \mathbf{1}_{\{|\Delta_i^n X| > \Phi_{1-\beta/2}^{-1} \sqrt{\Delta_n TBPV_t}\}} \quad (9)$$

where  $\Phi_{(\cdot)}^{-1}$  is the inverse of the standard normal cumulative distribution function,  $\beta = 1 - (1 - \alpha)^{\Delta_n}$  is the Šidák approach, and  $\alpha$  is the significance level. The indicator for a day with a significant jumps is estimated as follows

$$\mathcal{K}_t = \begin{cases} 1, & \text{if } \max \left( \frac{|\Delta_i^n X|}{\sqrt{\Delta_n TBPV_t}} \right) > \Phi_{1-\beta/2}^{-1} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

### 3.2 Infinite/Finite Jump Test

To identify whether the days with significant jumps are mainly driven by infinite or finite activity, we employ the  $SFA_t$  test proposed by [Aït-Sahalia and Jacod \(2011\)](#). The  $SFA_t$  test

statistic, with  $k = 2$ , and  $p = 4$  is defined as

$$SFA_t = \frac{B(p, v_n, k\Delta_n)_t}{B(p, v_n, \Delta_n)_t} \xrightarrow{p} \begin{cases} k^{p/2-1}, & \text{if } X_t \text{ has finitely many jumps on } [0, t] \\ 1, & \text{if } X_t \text{ has infinitely many jumps on } [0, t] \end{cases} \quad (11)$$

$$\frac{(SFA_t - k^{p/2-1})}{\sqrt{\widehat{V}_t}} \xrightarrow{L} \mathcal{N}(0, 1) \quad (12)$$

where  $B(p, v_n, \Delta_n) = \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} |\Delta_i^n X|^p \mathbf{1}_{\{|\Delta_i^n X| \leq v_n\}}$ ,  $v_n = \varrho \Delta_n^\varpi$ ,  $\varrho > 0$ ,  $\varpi \in (0, 1/2)$ . The  $SFA_t$  test statistic uses truncated returns to eliminate large jumps. The variance of the test is defined as

$$\widehat{V}_t = N(p, k) \frac{B(2p, v_n, \Delta_n)_t}{(B(p, v_n, \Delta_n)_t)^2} \quad (13)$$

$$N(p, k) = \frac{1}{m_{2p}} (k^{p/2}(1+k)m_{2p} + k^{p-2}(k-1)m_p^2 - 2k^{p/2-1}m_{k,p})$$

$$m_p \equiv \mathbb{E}[|Z|^p] = \frac{2^{p/2}}{\Gamma(1/2)} \Gamma\left(\frac{p+1}{2}\right)$$

$$m_{k,p} = \mathbb{E}[|Z|^p | Z + \sqrt{k+1}U|^p]$$

$$= \frac{2p}{\Gamma(1/2)} (k-1)^{p/2} \Gamma\left(\frac{1+r}{2}\right)^2 F_{2,1}\left(-\frac{p}{2}; \frac{p+1}{2}; \frac{1}{2}; \frac{-1}{k-1}\right)$$

For more details and settings, see [Aït-Sahalia and Jacod \(2011\)](#). The null hypothesis is that jumps are finite. The indicator  $\mathcal{C}_t = \left\{ SFA_t < k^{p/2-1} - \Phi_{1-\alpha}^{-1} \sqrt{\widehat{V}_t} \right\}$  is used to identify days with finite/infinite activity jumps.  $\Phi_{(\cdot)}^{-1}$  is the standard normal quantile, and  $\alpha$  is the significance level.

### 3.3 Jumps Classification and Identification

The partitioning of variation due to continuous and jump components can be done by taking the difference between measures of realized volatility and integrated variation. [Andersen et al. \(2007a\)](#) use this approach to construct measures of the variation of daily jump and continuous component. In this paper, we follow their method to disentangle the jump from the diffusive component, where we extend this methodology to construct measures for infinite and finite jumps. Following [Andersen et al. \(2007a\)](#) we disentangle the jump and continuous component as follows

$$\widehat{C}_t = RV_t \cdot \mathcal{K}_t + TBPV_t \cdot (1 - \mathcal{K}_t)$$

$$\widehat{J}_t = \max(RV_t - TBPV_t, 0) \cdot \mathcal{K}_t \quad (14)$$

The intersection of the ABD and the SFA tests enable jumps to be classified by activity. Infinite jumps are obtained when both tests reject their null of no jumps and finite activity of jumps, respectively. When only the ABD test rejects its null, we classify this jump as a finite jump.



This new jump classification is outlined as follows

$$IJ_t = \begin{cases} \widehat{J}_t, & \text{if } \mathcal{K}_t = 1 \cap \mathcal{C}_t = 1 \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

$$FJ_t = \begin{cases} \widehat{J}_t, & \text{if } \mathcal{K}_t = 1 \cap \mathcal{C}_t = 0 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

We classify a jump day as an infinite jump day when the ABD and SFA test respectively reject the nulls of no jumps and finite activity jumps. In the same fashion, we classify a day as finite activity jump day when the ABD test identifies a jump and the SFA test fails to reject its null of finite activity of jumps.

Barndorff-Nielsen et al. (2010) show that by taking the difference between the positive and negative semivariance it is possible to obtain the so-called signed jumps,

$$\Delta RJ_t = RSV_t^+ - RSV_t^- \xrightarrow{p} (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s > 0\}} - (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s < 0\}} \quad (17)$$

Using the intersection of the ABD and SFA tests along with an indicator function highlighting the sign of the signed jumps provide a further classification of jump by activity and sign as follows

$$\begin{aligned} \widehat{J}_t^+ &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t > 0 \cap \mathcal{K}_t = 1\}}, & \widehat{J}_t^- &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t < 0 \cap \mathcal{K}_t = 1\}} \\ FJ_t^+ &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t > 0 \cap \mathcal{K}_t = 1 \cap \mathcal{C}_t = 0\}}, & FJ_t^- &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t < 0 \cap \mathcal{K}_t = 1 \cap \mathcal{C}_t = 0\}} \\ IJ_t^+ &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t > 0 \cap \mathcal{K}_t = 1 \cap \mathcal{C}_t = 1\}}, & IJ_t^- &= \Delta RJ_t \mathbf{1}_{\{\Delta RJ_t < 0 \cap \mathcal{K}_t = 1 \cap \mathcal{C}_t = 1\}} \end{aligned} \quad (18)$$

### 3.4 Contribution of Jumps to Quadratic Variation

Following Aït-Sahalia and Jacod (2012), we estimate the relative contribution of jumps to the quadratic variation (QV) as follows

$$CV(\%) = \frac{\widehat{C}_T}{\widehat{C}_T + \widehat{J}_T} \times 100, \quad JV(\%) = \frac{\widehat{J}_T}{\widehat{C}_T + \widehat{J}_T} \times 100 \quad (19)$$

We can also obtain the relative contribution of the jumps dissected by activity and sign to the total jump component as follows

$$\begin{aligned} \%JV^+ &= \frac{\widehat{J}_T^+}{\widehat{J}_T} \times 100, & \%JV^- &= \frac{\widehat{J}_T^-}{\widehat{J}_T} \times 100 \\ \%FJV &= \frac{FJ_T}{\widehat{J}_T} \times 100, & \%IJV &= \frac{IJ_T}{\widehat{J}_T} \times 100 \\ \%FJV^+ &= \frac{FJ_T^+}{\widehat{J}_T} \times 100, & \%FJV^- &= \frac{FJ_T^-}{\widehat{J}_T} \times 100 \\ \%IJV^+ &= \frac{IJ_T^+}{\widehat{J}_T} \times 100, & \%IJV^- &= \frac{IJ_T^-}{\widehat{J}_T} \times 100 \end{aligned} \quad (20)$$

where  $x_T = \sum_{t=1}^T x_t$  and  $x = \{\widehat{C}_t, \widehat{J}_t, FJ_t, IJ_t, \widehat{J}_t^+, \widehat{J}_t^-, FJ_t^+, FJ_t^-, IJ_t^+, IJ_t^-\}$ . Using those measures, we may obtain a better understanding of the contribution of jumps to quadratic variation, as well as the composition of these jumps in terms of level of activity and sign. For instance, our prior is that infinite jumps are more important at higher frequencies than at lower frequencies. In addition, using these measures, we can also see if the level of activity and sign of the jumps matter in modeling and forecasting volatility.

### 3.5 Market Microstructure Noise

We now examine how the market microstructure distorts the estimation of realized volatility measures, and hence the identification of jumps. From Table 2, we know that the contribution of jumps varies across frequencies, and as  $\Delta_n \rightarrow 0$  the level of market microstructure noise increases. As a result, standard realized volatility measures tend to be biased upward, distorting the jump test statistics.<sup>5</sup> Robust-to-noise volatility measures should be used instead. We assume that the observed prices are contaminated by additive, microstructure noise

$$Y_t = X_t + u_t \quad (21)$$

where  $X_t$  is the process described in equation (1),  $u_t$  is an i.i.d. noise process with  $\mathbb{E}[u_t] = 0$  and  $\mathbb{E}[u_t^2] = \omega^2$ , and  $u_t \perp\!\!\!\perp X_t$ . Jacod et al. (2009) and Christensen et al. (2014) propose pre-averaging the returns and using the following robust-to-noise estimators for the realized variance and the bipower variation in this case.<sup>6</sup>

$$RV_t^* = \frac{M}{M-L+2} \frac{1}{L\psi_2^L} \sum_{i=0}^{M-L+1} |\Delta_i^n X^*|^2 - \frac{\psi_1^L \hat{\omega}_{AC}^2}{\theta^2 \psi_2^L} \quad (22)$$

$$BPV_t^* = \frac{M}{M-2L+2} \frac{1}{L\psi_2^L \mu_1^2} \sum_{i=0}^{M-2L+1} |\Delta_i^n X^*| |\Delta_{i+L}^n X^*| - \frac{\psi_1^L \hat{\omega}_{AC}^2}{\theta^2 \psi_2^L} \quad (23)$$

where  $L = \theta\sqrt{M} + o(M^{-1/4})$ ,  $M/(M-L+2)$ , and  $M/(M-2L+2)$  are small sample corrections,<sup>7</sup> while  $\frac{\psi_1^L \omega^2}{\theta^2 \psi_2^L}$  is a bias-correction removing a leftover effect of the noise that is not completely wiped out by the pre-averaging estimator. The unknown noise variance  $\omega^2$  can be approximated using either the Bandi and Russell (2006) estimator  $\hat{\omega}_{RV}^2 = \frac{1}{2M} \sum_{i=1}^M (\Delta_i^n Y)^2$ , or Oomen (2006) estimator  $\hat{\omega}_{AC}^2 = -\frac{1}{M-1} \sum_{i=2}^M \Delta_{i-1}^n Y \Delta_i^n Y$ . In this paper we use the latter procedure.

The pre-averaging returns are estimated as follows

$$\Delta_i^n X^* = \sum_{j=1}^{L-1} g\left(\frac{j}{L}\right) \Delta_{i+j}^n Y \quad (24)$$

where  $g = (x \wedge 1 - x)$ , and  $\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n}$ . The constants associated with  $g$  are defined as

$$\psi_1^L = L \sum_{j=1}^L \left[ g\left(\frac{j}{L}\right) - g\left(\frac{j-1}{L}\right) \right]^2, \quad \psi_2^L = \frac{1}{L} \sum_{j=1}^{L-1} g^2\left(\frac{j}{L}\right) \quad (25)$$

For estimating the up and downside volatilities in a noise-robust framework, we modify the two-time scale realized variance of Zhang et al. (2005) to obtain the two-time scale realized semivariance as follows.

<sup>5</sup>Since  $\mathbb{E}[|\Delta_i^n X|] \leq \mathbb{E}[|\Delta_i^n X + \eta_i|]$  where  $\eta_i = u_i - u_{i-1}$ , (see Hansen and Lunde, 2006; Huang and Tauchen, 2005, for more details).

<sup>6</sup>We also tried the threshold bi-power variation proposed by Christensen et al. (2018) but the differences were negligible.

<sup>7</sup>We try  $\theta = \{1/3, 1\}$  following Christensen et al. (2014).

**Proposition 1** (Robust Semivariances). *Assume that the true log-price process is contaminated by an additive noise, and hence we only observe a log-price as in (21), then the two-time scale realized semivariances are described as*

$$TSRV_t^+ = \frac{1}{M_k} \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor - k + 1} (Y_{(k+i\Delta_n)} - Y_{i\Delta_n})^2 \mathbf{1}_{\{(Y_{(k+i\Delta_n)} - Y_{i\Delta_n}) > 0\}} - \frac{\bar{M}}{M} \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} (\Delta_i^n Y)^2 \mathbf{1}_{\{\Delta_i^n Y > 0\}}$$

$$\xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 ds + \sum_{0 < s \leq t} (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s > 0\}} \quad (26)$$

$$TSRV_t^- = \frac{1}{M_k} \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor - k + 1} (Y_{(k+i\Delta_n)} - Y_{i\Delta_n})^2 \mathbf{1}_{\{(Y_{(k+i\Delta_n)} - Y_{i\Delta_n}) < 0\}} - \frac{\bar{M}}{M} \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} (\Delta_i^n Y)^2 \mathbf{1}_{\{\Delta_i^n Y < 0\}}$$

$$\xrightarrow{p} \frac{1}{2} \int_0^t \sigma_s^2 ds + \sum_{0 < s \leq t} (\Delta X_s)^2 \mathbf{1}_{\{\Delta X_s < 0\}} \quad (27)$$

where  $\bar{M} = (M - k + 1)/k$ ,  $M_k = \frac{N_k}{N/k} \approx k$ , where  $N_k$  represents the number of observations after using the scale  $k$ .  $k = \lceil cM^{2/3} \rceil$ , and  $c$  is the bandwidth stated as in Zhang et al. (2005).

Table 11 presents the finite sample performance exercise we carry out using the simulation design with finite jumps described in appendix B. The entries in Table 11 are MSEs. The results indicate that the realized semivariance is very sensitive to market microstructure noise, resulting in large MSEs even when the level of noise-to-signal ratio is moderate. On the other hand, the performance of the two-scale realized semivariance is excellent when the noise-to-signal ratio is moderate, even though the performance deteriorates a little when both the frequency and noise-to-signal level are higher.

### 3.5.1 Robust-to-Noise Jump Test Statistics

The ABD test in Andersen et al. (2007b) can be modified into a test that is robust to the presence of additive market microstructure noise by replacing the raw returns by pre-averaged returns and using the pre-averaging bipower variation without the bias correction. The bias correction is not required since the remaining noise from the pre-averaged returns and pre-averaged bipower variation cancel out.

**Proposition 2** (Robust ABD test). *Assume that the observed log-price process is (21), then using robust measures of volatility and pre-averaged returns we obtain the following daily-jump test statistics*

$$\mathcal{K}_t^* = \begin{cases} 1, & \text{if } \max \left( \frac{|\Delta_i^n X^*|}{\sqrt{\Delta_n^* BV_t^*}} \right) > \Phi_{1-\beta/2}^{-1} \\ 0, & \text{otherwise} \end{cases} \quad (28)$$

$$BV_t^* = \frac{M}{M - 2L + 2} \frac{1}{L\psi_2^L \mu_1^2} \sum_{i=0}^{M-2L+1} |\Delta_i^n X^*| |\Delta_{i+L}^n X^*|$$

Finally, we employ the methodology outlined in Section 3.3 to decompose the noise-robust RV into its continuous and discontinuous component, and dissect the jump component by activity (finite/infinite) and sign (positive/negative).

Table 12 shows the results of a small Monte Carlo exercise that explores the size and power of the two version of the ABD test under finite and infinite jumps, with a moderate and higher level of noise-to-signal ratio. Our simulation design is described in Appendix B. As expected the size and power of the noise-robust ABD test is much closer to the theoretical values when the magnitude of noise is greater, i.e. at higher frequencies. At lower frequencies the ABD test provides more accurate size and power. When testing for infinite jumps in the presence of moderate noise levels, the power of both tests is good, but the power of standard ABD test is badly affected by noise-to-signal levels.

## 4 Forecasting Models and Methodology

The basic HAR-RV in Corsi (2009) models current and future RV as a linear function of lagged values of the RV at the daily, weekly and monthly levels. Our forecasting models extend the HAR-RV models by including finite and infinite jumps and signed jumps. Andersen et al. (2007a) and Corsi et al. (2010) originally added jumps to the HAR-RV model. Our baseline HAR-RV model is

$$\widehat{RV}_{t,t+h} = \beta_0 + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \epsilon_{t+h} \quad (29)$$

where  $RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \dots + RV_{t+h}]$  aggregates information between  $\{t+1, t+h\}$  in order to present the coefficients values at the same scale with daily measures.  $h$  is an integer specifying the forecasting horizon. Our benchmark HAR-RV takes the original specification proposed by Corsi (2009) where future RV depends only on its lagged values. We extend the HAR-RV framework using a family of 9 different HAR models.

---

**Total Jumps Models:**

$$\begin{aligned}\widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d} \widehat{J}_t + \beta_{J_w} \widehat{J}_{t-5,t} + \beta_{J_m} \widehat{J}_{t-22,t} + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d^+} \widehat{J}_t^+ + \beta_{J_w^+} \widehat{J}_{t-5,t}^+ + \beta_{J_m^+} \widehat{J}_{t-22,t}^+ + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d^-} \widehat{J}_t^- + \beta_{J_w^-} \widehat{J}_{t-5,t}^- + \beta_{J_m^-} \widehat{J}_{t-22,t}^- + \epsilon_{t+h}\end{aligned}$$

**Finite Jumps Models:**

$$\begin{aligned}\widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{FJ_d} FJ_t + \beta_{FJ_w} FJ_{t-5,t} + \beta_{FJ_m} FJ_{t-22,t} + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{FJ_d^+} FJ_t^+ + \beta_{FJ_w^+} FJ_{t-5,t}^+ + \beta_{FJ_m^+} FJ_{t-22,t}^+ + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{FJ_d^-} FJ_t^- + \beta_{FJ_w^-} FJ_{t-5,t}^- + \beta_{FJ_m^-} FJ_{t-22,t}^- + \epsilon_{t+h}\end{aligned}$$

**Infinite Jumps Models:**

$$\begin{aligned}\widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{IJ_d} IJ_t + \beta_{IJ_w} IJ_{t-5,t} + \beta_{IJ_m} IJ_{t-22,t} + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{IJ_d^+} IJ_t^+ + \beta_{IJ_w^+} IJ_{t-5,t}^+ + \beta_{IJ_m^+} IJ_{t-22,t}^+ + \epsilon_{t+h} \\ \widehat{RV}_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{IJ_d^-} IJ_t^- + \beta_{IJ_w^-} IJ_{t-5,t}^- + \beta_{IJ_m^-} IJ_{t-22,t}^- + \epsilon_{t+h}\end{aligned}$$

---

where the continuous and jump component were estimated as in equation (14). We respectively estimate finite and infinite jumps as in equation (16) and (15), while the estimation of signed jump activity follows from equation (18). We also have an additional 9 models where all the right-hand volatility measures are the noise-robust measures discussed in Section 3.5.

## 4.1 Forecasting Evaluation

We are primarily interested in the performance of real-time, pseudo out-of-sample forecasts. We consider horizons  $h = 1, 5, , 22,$  and  $66,$  corresponding to one day, one week, one month, and one quarter ahead. We also use rolling window regressions of size 1000, or approximately four years, to estimate the models.

The out-of-sample performance is evaluated using the mean squared prediction error (MSPE) loss function and the out-of-sample  $R^2$ . The MSPE has been shown to be robust to noise in the proxy for volatility in Patton (2011). We estimate the  $R_{oos}^2$  as in Andersen et al. (1999); Patton and Sheppard (2015)

$$MSPE : L(RV_{t,t+h}, \widehat{RV}_{t,t+h}) = (RV_{t,t+h} - \widehat{RV}_{t,t+h})^2 \quad (30)$$

$$R_{oos}^2 = 1.0 - \frac{L(RV_{t,t+h}, \widehat{RV}_{t,t+h})}{L(RV_{t,t+h}, \overline{RV}_{t,t+h})} \quad (31)$$

where the numerator of the  $R_{oos}^2$  represents the out-of-sample MSPE from the models, and the denominator is the MSPE from a forecast that includes only an intercept. Additionally, we carry out pair-wise DM test, (see Diebold and Mariano, 2002), which have a null of equal predictive ability based on the MSPE loss criterion. We use robust HAC standard errors for the DM test. The Model Confidence Set (MCS) procedure of Hansen et al. (2011) is used to identify the subset of models that significantly outperform the others. We denote by  $\mathcal{M}$  the

set of all the HAR models. We define  $d_{h,i,j} = L(RV_{t,t+h}, \widehat{RV}_{t,t+h}^{(i)}) - L(RV_{t,t+h}, \widehat{RV}_{t,t+h}^{(j)})$  as the difference in the loss of model  $i$  and model  $j$ . We use a quadratic loss function as  $L$ . Finally, we construct the average loss difference,  $\bar{d}_{h,i,j}$ , and define the test statistics as follows

$$t_{i,j}^h = \frac{\bar{d}_{h,i,j}}{\sqrt{\widehat{\text{Var}}(\bar{d}_{h,i,j})}}, \quad \forall i, j \in \mathcal{M} \quad (32)$$

The MCS test statistics is given by  $T_{\mathcal{M}} = \max_{i,j \in \mathcal{M}} |t_{i,j}^h|$  and has the null hypothesis,  $H_0$  that all models have the same expected loss. The alternative hypotheses is that there is some model  $i$  with a MSPE that is greater than the MSPE's of all the other models  $j \in \mathcal{M} \setminus i$ . When the null is rejected the worst performing model is eliminated, and this process is iterated until no further model can be eliminated. The surviving models denoted by  $\mathcal{M}_{MCS}$  are retained with a confidence level  $\alpha = 0.05$ . We implement the MCS via a block bootstrap using a block length of 10 days and 5000 bootstrap replications.<sup>8</sup>

## 5 Data

We use the SPDR S&P 500 ETF (SPY) and 20 individual stocks constituents of the S&P 500 for the period 2000 – 2016, comprising 4277 trading days. The individual stocks were chosen based on their jump activity index, allowing for different levels of both finite and infinite jumps. The data source is TickData database. We employ the previous tick interpolation for aggregating down the data to the required sampling frequency.<sup>9</sup> Table 1 reports the descriptive statistics and average volume of the stocks under examination. This table highlights the heterogeneous composition of our dataset, with great variation in the trading volume, yet still active enough to avoid higher levels of price staleness, and different levels of volatility which ensures a great variety of jump activity across our stocks.

As this paper studies the role of the different types of jumps and the effect of market microstructure noise in forecasting realized volatility, we sample our data every 5, 60, and 300 seconds. The choice of 300 seconds is standard in high-frequency finance studies, and is motivated by the trade-off between bias and variance (see [Ait-Sahalia et al., 2005](#); [Bandi and Russell, 2006](#); [Zhang et al., 2005](#), for a more detailed discussion).

---

<sup>8</sup>We tried different block sizes representing 20 and 50 days, and we also use 10000 and 20000 bootstrap replications, however, the results are qualitatively similar.

<sup>9</sup>TickData provides pre-cleaned and filtered price series which usually undergo a series of algorithmic data filters to identify bad prints, decimal errors, transposition errors and other data irregularities. These filters take advantage of the fact that since we are not producing data in real time, we have the capacity to look at the tick following a suspected bad tick before we decide whether or not the tick is valid. The filters are proprietary and are based upon recent tick volatility, moving standard deviation windows, and time day. For a more detailed explanation refer to [TickData](#). We use the previous tick interpolation following [Hansen and Lunde \(2006\)](#) who suggest that in applications using quadratic variation, the interpolation method leads to realized volatilities with value 0.

Table 2 reports the contribution of different types of jumps to the quadratic variation across frequencies estimated as in equation (19). The contribution of total jumps decreases as the sampling interval increases, i.e. from 5 to 300 seconds. This means that aggregating the data reduces the identification of jumps, since longer intervals of time make the data look more continuous. This is why the continuous component in the SPY increases from 56% to 85% as the sampling interval increases from 5 to 300 seconds. At higher frequencies, most jumps in SPY are infinite activity jumps. For the 20 individual stocks, the share of infinite jumps is somewhat higher. The contribution of infinite jumps decreases in line with the total jumps as the sampling interval increases, suggesting that data aggregation reduces the estimated contribution of infinite jumps faster than that of finite jumps. This result is expected since, at longer sampling intervals, the small variations that characterize infinite jumps are closer to Brownian increments (or completely disappear).

Figure 1 plots the RV of the SPY and three individual stocks – AMZN, HD and KO – over our sample period 2000 to 2016. Days with jumps are shown in red and other days in blue, so the contribution of the jump component of RV can be identified. Note the different scales of the time series plots. The highest spikes in volatility occur around the dot-com and subprime crises (shaded areas), but many other spikes in volatility do not always coincide across stocks, so there is considerable heterogeneity in the level and timing of volatility. The autocorrelation function of the SPY 5 (left panel) and 300 (right panel) second realized measures are shown in Figure 2. In the right panel the autocorrelation of the continuous component start at 0.6, fall steady for about two months and then levels out at about 0.3, whereas the autocorrelation of the jump component are low throughout. By contrast, the left panel indicates that the RV is more persistence than the continuous component, and also infinite jumps present signs of persistence throughout the 70 days period. The greater level of persistence in the RV is due to infinite jumps, and implies that jumps at higher frequencies might contain predictive information power.

The index of jump activity,  $\hat{\beta}_{IJA}$ , measures the activity of small increments. The estimated values are in line up with the estimated contribution of finite and infinite jump components. At higher frequencies, the index is 1.45 for both the SPY and the average of the stocks. This implies that infinite jumps determine most of the jump component at higher frequencies. At 300 seconds, the estimated  $\hat{\beta}_{IJA}$  is 0.78 for SPY and 0.72 for the 20 stocks, suggesting that both types of jumps are present, but finite jumps are more important. Finally, little evidence of asymmetry is observed in signed jumps, even though the percentage contribution of positive jumps is marginally higher. Similarly, we find that positive/negative jumps are marginally more abundant in finite/infinite jumps.

Table 3 presents the correlations of the realized measures for SPY in the entries below the

diagonal, and the average correlations for the 20 stocks in the entries above the diagonal. We find that jumps, finite jumps, and infinite jumps are relatively highly correlated with both  $RV_t$  and  $C_t$ . However, the level of correlation is somewhat greater for the average of the stocks. Interestingly, correlations of finite and infinite jumps with total jumps are 0.939 and 0.293, respectively. The Correlation between the signed jumps is very low, around 0.05. The difference levels of correlations observed in jumps dissected by sign and activity suggest that there is novel information in this decomposition that might help in predicting future realized volatility.

## 6 Empirical Findings

### 6.1 RV Prediction: SPY and Individual Stocks

Table 4–7 presents the in-sample coefficients, as well as the in- and out-of-sample  $R^2$ , of the baseline HAR-RV and extended HAR-RV models for SPY for four forecast horizons  $-h = 1$  (day),  $h = 5$  (week),  $h = 22$  (month),  $h = 66$  (three months) using returns sampled every 300 seconds. The baseline HAR-RV is presented in Table 4, while the HAR-CJ (jumps), HAR-CFJ (finite jumps), and the HAR-CIJ (infinite jumps) models are presented in the subsequent tables. The significance of the coefficients is evaluated using Newey-West HAC robust standard errors, allowing for serial correlation of up to order 5 ( $h = 1$ ), 10 ( $h = 5$ ), 44 ( $h = 22$ ), and 132 ( $h = 66$ ), since the random error term in the models is serially correlated at least up to order  $h - 1$ . The  $R^2_{(oos)}$  is estimated as 1 minus the ratio of the out-of-sample models-based MSPE to the out-of-sample MSPE from a forecast including only a constant.

Table 4 reports the baseline HAR-RV estimates, which are all significant at even  $h = 66$ , confirming the existence of highly persistent volatility dependence. The magnitude of the daily and weekly coefficients decrease as we lengthen the forecast horizon. The magnitude of the monthly coefficient changes little with the horizons, although its relative importance increases at longer horizons.<sup>10</sup>

Table 5–7 report the coefficients for various specifications including the total, finite and infinite jumps, respectively. The results are in line with those found in the previous literature: substantial persistence, with the sum of the coefficients on the daily, weekly and monthly integrated variation (continuous volatility) measures  $\beta_{C_d} + \beta_{C_w} + \beta_{C_m}$  close to 1. The importance of recent information also decreases with the forecast horizon. The daily and weekly continuous estimates are significant for all of the specifications, across the four forecast horizons, suggesting that the daily and weekly continuous components are important to forecast future volatility. On the other hand, the estimated coefficient on the monthly continuous volatility is generally

---

<sup>10</sup>These results have been well-documented in the literature, (see Andersen et al., 2007a; Corsi, 2009; Corsi et al., 2010, among others).



insignificant in Tables 5 (jumps) and 6 (finite jumps), even though the estimates increase with the forecast horizons. However, in the specification using infinite jumps (Table 7), the estimated coefficients on the monthly continuous component is highly significant and quite large for all horizons.

If the level of activity of the jumps does not matter when forecasting volatility, then we should expect that  $\beta_J = \beta_{FJ} = \beta_{IJ}$ . We easily reject this restriction for  $\beta_J = \beta_{IJ}$ . Looking at signed jumps, we also reject the hypothesis that  $\beta_J = \beta_{J+} = \beta_{J-}$  at the 5% for all horizons. This means that dissecting jumps by their nature and sign do help to predict future volatility. Of course, we ideally need to show that the real-time, pseudo out-of-sample results from HAR type models with finite and infinite jumps, and/or signed and unsigned jumps generate significantly better out of sample forecasts using a standard loss function such as the MSPE.

Negative/positive jumps usually increase/decrease the level of future volatility, whereas unsigned total jumps also reduce the estimation of future volatility. These results are line with those of Andersen et al. (2007a) and Patton and Sheppard (2015). Turning to the nature of jumps we find that infinite jumps increase the level of future volatility, whereas finite jumps act like unsigned jumps reducing the value of the expected future volatility. The daily jump component is potentially important in predicting future volatility when total and positive jumps are used. On the other hand, the weekly and monthly jump components are more important for positive and negative jumps, respectively. Infinite jump estimates are usually insignificant, except for the daily negative and weekly positive infinite jumps which are significant across horizons.

The in-sample R-squared,  $R_{(in)}^2$ , suggests that incorporating jumps as predictors results in a better fit for our models outperforming the baseline HAR-RV across the four horizons we have examined. The out-of-sample R-squared,  $R_{(oos)}^2$ , shows that all models based on finite and total jumps outperform the baseline model across horizons, whereas specifications based on infinite jumps outperform the HAR-RV for 1- and 5-day ahead predictions. Taking the nature of jumps results, the  $R_{(oos)}^2$  indicates that specifications based on total and finite jumps perform better at longer horizons. For instance, the HAR-CFJ<sup>+</sup> has an increment of 11.89% and 3.82% in terms of  $R_{(oos)}^2$  for  $h = 5$  and  $h = 22$ , respectively. Infinite jump specifications perform best at shorter horizons only. For instance, adding infinite jumps at  $h = 1$  and  $h = 5$  result in an improvement in the  $R_{(oos)}^2$  of 16% and 5%, whilst at  $h = 22$  the  $R_{(oos)}^2$  reduces by 8%, suggesting that the information power of infinite jumps is only important at shorter horizons. Infinite jumps contribute very little to the QV(%) at lower frequencies, yet its information is very rich. However, the use of moving averages to construct the measures for  $h > 1$  presumably reduces the importance of infinite jumps at longer horizons.

## 6.2 Rolling Regression Forecast Results

Table 8 and 9 report the relative MSPE and the Model Confidence Set (MCS) for the SPY and our selection of 20 stocks. Average results for all the 20 stocks are tabulated in Table 9. The results are based on real-time, pseudo out-of-sample rolling regression forecast using a 1000 day window. We use the Diebold and Mariano test to identify extended HAR jump models with significantly lower MSPE's than the benchmark HAR-RV models. We identify the set of retained models – models with MSPE's that are not significantly higher than the MSPE of other retained models – using the MCS procedure.

The bottom row of the two tables are the MSPEs for the baseline HAR-RV model using standard or raw volatility measures. The remaining MSPE entries are relative MSPEs, i.e. the ratio of the proposed models to the MSPE of the benchmark model. Values below one generate better forecasts than baseline HAR-RV, and vice versa for values above one. The entries in the top panel are based on forecasts using standard volatility measures as explanatory variables; the bottom panel entries using robust-to-noise volatility measures. The MCS entries in Table 8 are the MCS rankings while entries in Table 9 are the number of times (out of 20 stocks) each extended HAR model is retained in the MCS.

Taking the SPY results (Table 8) first, the starred MSPE entries and the bold MCS rank entries indicate that many of the extended HAR models with jumps forecast as well as, or better, than the baseline HAR-RV models. Based on the DM and MCS tests, the MSPEs of these models are similar or lower than the MSPEs of the baseline models. However, there is no consistent pattern across time horizons, or the use of raw versus robust volatility measures. The noise-robust measures, Panel B, generally outperform their raw measure counterparts at the 5 second frequency and sometimes at the 60 second frequency. As expected, raw measure models are usually excluded for the MCS at longer horizons, confirming the importance of accounting for microstructure noise at higher frequencies (Ghysels and Sinko, 2011). Nevertheless, the MSPE numbers for the baseline HAR-RV model in the final row suggest that models using 300 second volatility measures tend to forecast better than models using 5 or 60 second returns, irrespective of whether standard or robust-to-noise volatility measures are used.

Turning now to Table 9, recall that the MSPE entries are averages across the 20 stocks, while the MCS entries are counts of the number of times (out of 20 stocks) each baseline HAR-RV or extended HAR model with jumps is retained in the MCS. The relative MSPE entries are clustered around the value one, but less dispersed than in Table 8, in part because we are reporting averages. In addition, many of the MCS entries, including entries for the baseline HAR-RV models, are close or equal to 20. Thus, the improvement in forecasting performance is less obvious for the 20 stocks than for the SPY.

Consistent with the results for SPY, we show that forecasts using 300 second volatility

measures are generally better than the forecast using 5 or 60 second measures, and that the relative MSPEs of the raw volatility measures tend to be somewhat lower than the relative MSPEs of the robust-to-noise measures. Since no single extended HAR model with jumps dominates the others, and given that the gains obtained are small at the aggregate level, we now examine whether or not model averaging does generate forecasting gains.

## 7 The Gains from Model Averaging

So far we have shown that a variety of HAR type volatility models that account for the nature and sign of jumps generate significant improvements in forecasting performance. However, there is little to choose when it comes to selecting the best approach, since no specification consistently outperforms the other models across horizons and frequencies.

This section employs four simple approaches to assigning model averaging weights. We experimented using more complicated averaging procedures, but the results were similar to those we present below. The aim of model averaging is to capture the different information embedded in the various jump specifications, hopefully producing an ensemble model that outperforms the benchmark HAR-RV model and, more importantly, the best single jump model. Our approaches follow the literature closely (see, for instance [Aiolfi et al., 2011](#); [Aiolfi and Timmermann, 2006](#); [Bates and Granger, 1969](#); [Elliott and Timmermann, 2016](#), and references therein). We present model averaging results for the four sets of weights tabulated below – weights minimizing the estimated variance of the prediction errors, inverse MSPE weights, inverse MSPE rank weights and equal weights. In the first three cases, the weights are recalculated every time a new set of rolling window forecasts are generated, and we prune the set of models under consideration by only averaging models which are retained in the Model Confidence Set.

Weight	Formula	Models
Min. Prediction Error Variance	$w_t^h = \underset{w}{\operatorname{argmin}} w' \Sigma_t^h w \text{ s.t. } \iota' w = 1$	MCS
Inverse MSPE	$w_{t,i}^h = \frac{(MSPE_{t,i}^h)^{-1}}{\sum_{i \in \mathcal{M}_J} (MSPE_{t,i}^h)^{-1}}$	MCS
Inverse Rank	$w_{t,i}^h = \frac{(Rank_{t,i}^h)^{-1}}{\sum_{i \in \mathcal{M}_J} (Rank_{t,i}^h)^{-1}}$	MCS
Equal Weights	$w_{i,t}^h = \frac{1}{N}$	All

Note:  $\Sigma_t^h$  is the estimated, rolling window variance-covariance matrix of the set of MCS retained horizon  $h$  volatility forecasting models at time  $t$ .  $\iota$  is a vector of ones representing each retained model.  $MSPE_{t,i}^h$  and  $Rank_{t,i}^h$  are the rolling window MSPEs and MCS Ranks for the MCS retained horizon  $h$  forecasting model at time  $t$ . Finally,  $N$  represents all the jump specifications used in this study.

We present model averaging results for SPY and four individual stocks that were chosen based on their level of jump activity. All the stocks have estimated Blumenthal-Gettoor indices

in the range 0 to 1, so their returns include both types of jump activity with finite jumps dominating. BA and KO with jump activity of 0.58 and 0.91 are the extreme cases.

Table 10 reports the relative MSPE for the four different model averaging approaches, the baseline model, and the best jump specification for each stock and horizon. Model averages with significantly lower MSPE's than the baseline HAR-RV models are denoted by the superscript \*. Model averages with lower MSPE's than the baseline HAR-RV and the best extended HAR model are shown in bold, while the superscript \*\* denotes model averages with significantly lower MSPEs than the extended HAR models. We find that the four model averaging schemes generate forecasts that significantly outperform the benchmark model in most cases, for the four forecast horizons examined –  $h = 1$  (one day),  $h = 5$  (one week),  $h = 22$  (one month) and  $h = 66$  (three months). For example, in the case of SPY with 300 second returns, the one-week relative MSPE of the best extended HAR model is 0.753 versus 0.693 to 0.715 for the four model averages.

The largest MSPE reductions are generally found at the one-week horizons, followed by the one-month horizon. We present model averaging results for SPY using 60 and 300 second returns. The SPY model averaging forecasts using 300 second returns dominate the forecasts using 60 second returns, generating lower MSPEs. The model averaging forecasts using 300 second SPY returns also dominate the unreported model averaging based on 5 second returns. This result also holds for the other stocks. The 300 second model average MSPEs are mostly lower than the MSPEs of both the baseline HAR-RV and the best extended HAR type model with jumps. In about a quarter of the cases, the 300 second model average forecasts significantly outperform the best extended HAR type model with jumps.

Model averaging forecasts from the extended HAR models generally result in lower MSPE forecasts, although no single model averaging procedure stands out. Pruning the models by only averaging the model in MCS, appears to help and the weights from minimizing the prediction error variance do not predominate the other weighting schemes. Overall, pruning dominated models and the use of inverse MSPE weights or inverse MSPE ranks for model averaging work quite well.

## 8 Conclusion

We examine the gains in forecasting the volatility of equity prices by decomposing jumps by activity (finite/infinite) and by sign (positive/negative) using high-frequency data for SPY and 20 individual stocks which vary sector and volume. Our key findings are as follows. Quadratic variation contains a significant jump component even at the 300 second frequency. However, the share of infinite jumps at this frequency is relatively small. Extending the HAR-RV style models with a variety of jump measures helps to predict future volatility, generating significant

in- and out-of-sample improvements for both SPY and the 20 individual stocks.

We find that noise-robust estimators substantially improve the estimation of future volatility at higher frequencies. However, since the level of market microstructure declines as the sampling interval increases, the forecasting advantage of the noise-robust jump volatility wanes.

The real-time, pseudo out-of-sample forecasting results using a rolling window suggest that the lowest MSPE forecasts are obtained using the returns sampled every 300 seconds, as opposed to every 5 or 60 seconds. This result holds for the horizons we examined – one day, one week, one month and three months – irrespective of whether robust-to-noise volatility measures were or were not used. In terms of MSPE's, there is little to choose between using standard or robust-to-noise measure at this frequency.<sup>11</sup>

Finally, since no single forecasting model dominates, we investigated whether various model-averaging procedures generated further real-time forecasting gains. We report results for simple model averaging procedures, since the forecasting performance of more complicated model averaging procedures were similar. In some cases, we pruned the set of models using the model confidence set procedure of [Hansen et al. \(2011\)](#) to eliminate dominated models. We find that simple model averaging procedures generally result in significant gains in forecasting performance vis-à-vis the single best forecasting model, which in turn outperforms our HARV-RV baseline. For example, model averaged results using equal weights, or the rather ad hoc, normalized inverse MSE weights in [Bates and Granger \(1969\)](#) perform as well as the model averaged results where the weights minimize the variance of the prediction error. We believe that these findings will prove helpful for practitioners.

---

<sup>11</sup>In future research, we will check whether our findings carry over when returns are sampled in business time as opposed to clock time

## References

- Aiolfi, M., Capistrán, C., and Timmermann, A. (2011). Forecasts combinations. In Clements, M. P. and Hendry, D. F., editors, *The Oxford Handbook of Economic Forecasting*. OUP.
- Aiolfi, M. and Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, 135(1-2):31–53.
- Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics*, 74(3):487–528.
- Aït-Sahalia, Y. and Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37(5A):2202–2244.
- Aït-Sahalia, Y. and Jacod, J. (2009b). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184–222.
- Aït-Sahalia, Y. and Jacod, J. (2011). Testing whether jumps have finite or infinite activity. *The Annals of Statistics*, pages 1689–1719.
- Aït-Sahalia, Y. and Jacod, J. (2012). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature*, 50(4):1007–50.
- Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.
- 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. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, pages 885–905.
- Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007a). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4):701–720.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.
- 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 iid noise: Theory and testable distributional implications. *Journal of Econometrics*, 138(1):125–180.
- Andersen, T. G., Bollerslev, T., and Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-a-vis forecast horizon. *Journal of Empirical Finance*, 6(5):457–477.
- Andersen, T. G., Bollerslev, T., and Meddahi, N. (2005). Correcting the errors: Volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica*, 73(1):279–296.
- Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, 79(3):655–692.
- Barndorff-Nielsen, O., Kinnebrock, S., and Shephard, N. (2010). Volatility and time series econometrics: Essays in honor of robert f. engle, chapter measuring downside risk-realised semivariance.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.
- Barndorff-Nielsen, O. E. and Shephard, N. (2002a). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.
- Barndorff-Nielsen, O. E. and Shephard, N. (2002b). Estimating quadratic variation using realized volatility. *Journal of Applied Econometrics*.
- 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. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.
- Bates, J. M. and Granger, C. W. (1969). The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.
- Busch, T., Christensen, B. J., and Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, 160(1):48–57.

- Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics*, 116(1-2):225–257.
- 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*.
- Christensen, K., Oomen, R. C., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114(3):576–599.
- Christoffersen, P., Jacobs, K., Ornathanalai, C., and Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, 90(3):272–297.
- 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 Reno, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288.
- Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.
- Duan, J.-C., Ritchken, P., and Sun, Z. (2006). Approximating garch-jump models, jump-diffusion processes, and option pricing. *Mathematical Finance*, 16(1):21–52.
- 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.
- Duong, D. and Swanson, N. R. (2015). Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction. *Journal of Econometrics*, 187(2):606–621.
- Elliott, G. and Timmermann, A. (2016). *Economic Forecasting*. Princeton University Press.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.
- Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, 58(3):1269–1300.
- Forsberg, L. and Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, 5(1):31–67.



- Ghysels, E. and Sinko, A. (2011). Volatility forecasting and microstructure noise. *Journal of Econometrics*, 160(1):257–271.
- Ghysels, E. and Sohn, B. (2009). Which power variation predicts volatility well? *Journal of Empirical Finance*, 16(4):686–700.
- Giot, P. and Laurent, S. (2007). The information content of implied volatility in light of the jump/continuous decomposition of realized volatility. *Journal of Futures Markets*, 27(4):337–359.
- 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.
- Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4):456–499.
- Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and Their Applications*, 119(7):2249–2276.
- Jiang, G. J. and Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise—a swap variance approach. *Journal of Econometrics*, 144(2):352–370.
- Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. (2012). On the jump activity index for semimartingales. *Journal of Econometrics*, 166(2):213–223.
- Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell’Istituto Italiano degli Attuari*, 64(19-47):44.
- Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.
- Martens, M., Van Dijk, D., and De Pooter, M. (2009). Forecasting s&p 500 volatility: Long memory, level shifts, leverage effects, day-of-the-week seasonality, and macroeconomic announcements. *International Journal of Forecasting*, 25(2):282–303.
- Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, 17(5):479–508.

- 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(2-3):213–239.
- Oomen, R. C. A. (2006). Comment on 2005 jbes invited address ”realized variance and market microstructure noise” by peter r. hansen and asger lunde. *Journal of Business & Economic Statistics*, 24(2):195–202.
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.
- Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3):683–697.
- Prokopczuk, M., Symeonidis, L., and Wese Simen, C. (2016). Do jumps matter for volatility forecasting? evidence from energy markets. *Journal of Futures Markets*, 36(8):758–792.
- Sévi, B. (2014). Forecasting the volatility of crude oil futures using intraday data. *European Journal of Operational Research*, 235(3):643–659.
- Taylor, S. J. (1986). *Modelling financial time series*. Wiley, New York, NY.
- Todorov, V. and Tauchen, G. (2010). Activity signature functions for high-frequency data analysis. *Journal of Econometrics*, 154(2):125–138.
- Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

## A The Index of Jump Activity, $\beta_{IJA}$

The index of jump activity denoted by  $\beta_{IJA}$  serves as an indicator of the activity of an Itô semimartingale processes, a rich class of stochastic processes which include diffusion, jump diffusion, Lévy process, and others. The Blumenthal-Gettoor index which is defined as

$$\beta_{IJA} := \inf \left\{ r \geq 0; \sum_{0 < s \leq t} |\Delta_s X|^r < \infty \right\} \quad (\text{A.1})$$

The values of the  $\beta_{IJA}$  range between 0 to 2. As shown by (Aït-Sahalia and Jacod, 2012, figure 7) when the  $\beta_{IJA}$  is close to zero, the jumps resemble Poisson jumps,<sup>12</sup> and when the  $\beta_{IJA}$  is close to two, the increments resemble a Brownian motion. Values in between will behave as a combination of finite and infinite activity. Aït-Sahalia and Jacod (2009a), Jing et al. (2012), and Todorov and Tauchen (2010) provide evidence using individual stocks, at higher frequencies, that the  $\beta_{IJA} \in [1.4, 1.6]$ , which indicates that both finite and infinite jumps are present in the data.

We follow Jing et al. (2012) and estimate the jump activity index as follows,

$$\hat{\beta}_{IJA} = \hat{\beta}_T(t, \varpi, \phi, \phi') =: \log \frac{U(\varpi, \phi, g)_t^n}{U(\varpi, \phi', g)_t^n} \bigg/ \log \left( \frac{\phi'}{\phi} \right) \quad (\text{A.2})$$

where

$$U(\varpi, \phi, g)_t^n =: \sum_{i=1}^{\lfloor 1/\Delta_n \rfloor} g \left( \frac{\Delta_i^n X}{\phi \Delta_n^\varpi} \right) \quad (\text{A.3})$$

$$g(x) = \begin{cases} c^{-1}|x|^p, & |x| \leq a \\ c^{-1} \left( a^p + \frac{pa^{p-1}}{2(b-a)} ((b-a)^2 - (|x|-b)^2) \right), & a \leq |x| \leq b \\ 1, & |x| \geq b \end{cases} \quad (\text{A.4})$$

where  $c = a^p + pa^{p-1}(b-a)/2$  and the following conditions must be satisfied:  $0 < \phi < \phi'$ ,  $\varpi \in (0, 1/2)$ ,  $0 < a < b$ . We set  $\varpi = 1/5$ ,  $\phi' = 2\phi$ ,  $p = 6$ ,  $a = 6/5$ ,  $b = 7/5$ . This estimator is a more efficient version of the jump activity index proposed by Aït-Sahalia and Jacod (2009a). Aït-Sahalia and Jacod (2009a)'s estimator uses only large increments ending with a very small effective sample size. By contrast, Jing et al. (2012)'s approach reduces the measurement error by using both small and large increments of the data. For more details, see Jing et al. (2012).

---

<sup>12</sup>Gamma and Variance gamma processes also generate  $\beta_{IJA}$  values of 0.

## B Simulation Design

Our simulation are based on the Heston model augmented with finite and infinite jumps as follows,

$$\begin{aligned} dX_t &= \sqrt{\nu_t}dW_t + \theta_L dL_t \\ d\nu_t &= \kappa(\eta_\nu - \nu_t)dt + \gamma_\nu \nu_t^{1/2}dB_t \end{aligned} \tag{B.1}$$

where  $\mathbb{E}[dW_t, dB_t] = \rho dt$ , and  $L_t$  is either a finite activity compound Poisson process or an infinite activity Cauchy process (a  $\beta$ -stable process with  $\beta = 1$ ).

We take  $\kappa = 5$ ,  $\eta_\nu = 1/16$ ,  $\rho = -0.5$  following [Aït-Sahalia and Jacod \(2011\)](#). The compound Poisson process has intensity  $\lambda$ , and jumps that are uniformly distributed on  $\sqrt{m}([-2, -1] \cup [1, 2])$ . We set  $m = 0.7$  and  $\lambda = 0.5$  such that there is on average one jump every two days. When jumps are of finite activity we set  $\theta_L = 1$ , while for infinite jumps we set  $\theta_L = 0.5$ . We follow [Barndorff-Nielsen et al. \(2008\)](#) and simulate the market microstructure as follows

$$\begin{aligned} u_{t,i} &\sim \mathcal{N}(0, \omega_t^2) \\ \omega_t^2 &= \xi^2 \int_0^t \nu_s ds \end{aligned} \tag{B.2}$$

This design enables the variance of the noise to be constant throughout the day, but change from day to day. This noise is then added to the  $X_{t,i}$  price process to obtain the time series of actual high-frequency simulate prices  $Y_{t,i} = X_{t,i} + u_{t,i}$ .  $\xi$  is the noise-to-signal ratio used to simulate the market microstructure noise.

Finally, we simulate  $T = 50$  days, consisting of 6.5 hours of trading per second, i.e.,  $n = 23400$ . We then sample the data every 5, 60, and 300 seconds. We use 3000 replications providing a total of 150,000 days, and the jump tests are evaluated at the 5% level.

## C Tables and Figures

Table 1: Summary Daily Statistics for RV and Trading Volume

Stock / Index	Mean	Std. Dev.	P25	RV Median	P75	Min	Max	Average Volume
SPY	1.037	2.259	0.240	0.485	1.030	0.013	59.863	98.972
AMZN	8.284	14.846	1.834	3.344	7.209	0.225	229.244	6.462
BA	2.812	3.900	0.908	1.602	3.232	0.086	55.570	4.347
BFB	1.920	4.861	0.752	1.152	1.912	0.074	240.414	0.875
CAT	3.239	4.889	1.116	1.873	3.486	0.185	105.908	5.826
CHL	2.052	3.608	0.511	1.063	2.245	0.082	65.965	1.078
COST	3.087	4.728	0.790	1.497	3.295	0.126	83.955	3.460
CVX	2.113	4.181	0.752	1.263	2.244	0.112	137.535	7.188
DOW	3.976	7.353	1.176	2.039	4.251	0.146	216.937	6.556
EXC	2.635	4.791	0.875	1.429	2.643	0.158	130.875	3.917
GILD	6.308	10.469	1.570	2.839	5.889	0.198	187.286	17.587
GS	4.186	11.978	1.062	1.757	3.683	0.153	400.346	5.731
HD	3.121	4.938	0.877	1.573	3.475	0.156	103.477	9.746
JNJ	1.385	3.482	0.408	0.692	1.488	0.076	179.016	8.559
JPM	4.615	10.848	0.869	1.770	4.466	0.114	252.877	20.538
KO	1.561	2.535	0.494	0.836	1.630	0.046	58.808	13.275
OKE	3.402	8.590	0.954	1.668	3.469	0.160	411.055	1.289
SO	1.744	2.773	0.555	0.937	1.947	0.092	97.041	3.203
UPS	1.649	4.140	0.526	0.851	1.673	0.081	216.939	3.045
WMT	2.045	3.277	0.580	0.976	2.117	0.090	71.485	10.569

Note: The table reports the summary RV statistics and average daily trading volume for SPY and the 20 stocks listed. The sample period is Jan 3, 2000 to Dec 30, 2016, which consists of 4277 trading days, RV is estimated using 300 second returns. P25 and P75 are the 25<sup>th</sup> and 75<sup>th</sup> percentiles of daily RV. Average daily trading volume is reported in millions.

Table 2: Contribution of Jumps to Total QV

	SPY			Avg. Stocks			AMZN	BA	BFB	CAT	CHL	COST	CVX
	5s	60s	300s	5s	60s	300s	300s	300s	300s	300s	300s	300s	300s
%CV	56.798	88.474	85.725	32.399	65.612	70.198	73.426	72.586	55.143	74.899	62.182	69.525	80.277
%JV	43.202	11.526	14.275	67.601	34.388	29.802	26.574	27.414	44.857	25.101	37.818	30.475	19.723
%JV+	21.847	6.450	8.257	33.946	16.535	14.992	15.208	14.362	22.474	12.574	17.978	15.963	9.849
%JV-	21.355	5.075	6.018	33.653	17.853	14.810	11.366	13.052	22.383	12.527	19.841	14.512	9.874
%FJV	10.602	10.419	14.156	33.394	32.417	29.597	26.410	27.228	44.649	24.852	37.314	30.357	19.576
%IJV	32.600	1.106	0.118	34.207	1.971	0.205	0.165	0.187	0.208	0.249	0.504	0.118	0.147
%FJV+	5.584	5.941	8.219	17.028	15.539	14.883	15.127	14.248	22.380	12.465	17.681	15.892	9.766
%FJV-	5.017	4.478	5.937	16.366	16.878	14.714	11.283	12.979	22.269	12.387	19.633	14.465	9.810
%IJV+	16.263	0.509	0.038	16.918	0.996	0.108	0.081	0.114	0.093	0.110	0.296	0.070	0.083
%IJV-	16.338	0.597	0.080	17.287	0.975	0.096	0.084	0.073	0.115	0.140	0.208	0.047	0.064
$\hat{\beta}_{IJA}$	1.454	1.056	0.778	1.455	1.040	0.723	0.461	0.576	0.802	0.621	0.763	0.697	0.748
	DOW	EXC	GILD	GS	HD	JNJ	JPM	KO	OKE	PG	SO	UPS	WMT
	300s	300s	300s	300s	300s	300s	300s	300s	300s	300s	300s	300s	300s
%CV	68.881	69.488	63.203	75.979	73.935	70.611	76.122	74.208	59.168	71.147	70.791	68.292	74.102
%JV	31.119	30.512	36.797	24.021	26.065	29.389	23.878	25.792	40.832	28.853	29.209	31.708	25.898
%JV+	15.029	15.506	18.911	12.311	13.875	12.919	12.926	12.498	19.059	15.416	14.486	15.477	13.013
%JV-	16.090	15.006	17.886	11.710	12.190	16.470	10.952	13.294	21.773	13.438	14.723	16.231	12.885
%FJV	30.849	30.400	36.458	23.941	25.940	29.279	23.822	25.519	40.602	28.777	28.642	31.527	25.802
%IJV	0.270	0.112	0.339	0.080	0.125	0.111	0.056	0.273	0.230	0.076	0.568	0.181	0.096
%FJV+	14.830	15.434	18.670	12.297	13.843	12.832	12.899	12.341	18.982	15.365	14.274	15.373	12.968
%FJV-	16.019	14.966	17.788	11.644	12.097	16.447	10.923	13.178	21.620	13.413	14.368	16.154	12.834
%IJV+	0.198	0.072	0.241	0.014	0.032	0.088	0.028	0.157	0.077	0.051	0.213	0.104	0.045
%IJV-	0.071	0.040	0.098	0.066	0.093	0.023	0.029	0.116	0.153	0.025	0.355	0.077	0.051
$\hat{\beta}_{IJA}$	0.579	0.725	0.522	0.610	0.665	0.971	0.606	0.913	0.645	0.955	0.878	0.895	0.824

Note: The table reports the contribution of the different jump measures to the total quadratic variation (QV) across frequencies for SPY and the average of all the stocks using 5 second, 60 second and 300 second returns. The results for the individual stocks were estimated using 300 second returns.  $\hat{\beta}_{IJA}$  is the estimated Blumenthal-Gettoor index of jump activity (Appendix A).

Table 3: Correlations of Volatility Measures

(Twenty stock average above and SPY below main diagonal)

	$RV_t$	$C_t$	$J_t$	$FJ_t$	$IJ_t$	$J_t^+$	$FJ_t^+$	$IJ_t^+$	$J_t^-$	$FJ_t^-$	$IJ_t^-$
$RV_t$		0.796	0.614	0.441	0.352	0.364	0.228	0.272	-0.491	-0.375	-0.273
$C_t$	0.713		0.564	0.433	0.301	0.354	0.263	0.242	-0.386	-0.312	-0.191
$J_t$	0.403	0.281		0.798	0.467	0.359	0.270	0.174	-0.513	-0.454	-0.174
$FJ_t$	0.301	0.217	0.939		-0.057	0.285	0.355	-0.044	-0.488	-0.561	0.041
$IJ_t$	0.332	0.212	0.293	-0.054		0.179	-0.048	0.400	-0.141	0.040	-0.448
$J_t^+$	0.683	0.555	0.431	0.370	0.222		0.817	0.464	0.057	0.043	0.047
$FJ_t^+$	0.413	0.605	0.495	0.536	-0.051	0.704		-0.037	0.043	0.033	0.036
$IJ_t^+$	0.544	0.169	0.105	-0.021	0.365	0.695	-0.020		0.043	0.032	0.035
$J_t^-$	-0.309	-0.334	-0.112	-0.085	-0.090	0.068	0.050	0.045		0.867	0.385
$FJ_t^-$	-0.176	-0.288	-0.107	-0.135	0.064	0.038	0.028	0.025	0.799		-0.031
$IJ_t^-$	-0.269	-0.157	-0.038	0.046	-0.238	0.061	0.044	0.040	0.556	-0.056	

Note: The table reports the correlation of the different realized measures estimated at the 300 second frequency. The entries below the diagonal are for SPY, and entries above the diagonal are average correlations for the 20 stocks.

Table 4: HAR-RV Benchmark

$$RV_{t,t+h} = \beta_0 + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \epsilon_{t,t+h}$$

	HAR-RV			
	h=1	h=5	h=22	h=66
$\beta_0$	0.095*	0.148**	0.288***	0.527***
$\beta_d$	0.246**	0.184***	0.103***	0.061***
$\beta_w$	0.422***	0.347***	0.322***	0.200***
$\beta_m$	0.238**	0.323***	0.290***	0.215***
$R^2_{(in)}$	0.512	0.629	0.562	0.337
$R^2_{(oos)}$	0.443	0.673	0.707	0.470
MSPE	3.102	1.322	0.944	1.262

Note: The table reports the OLS coefficient estimates and in- and out-of-sample R-squared for HAR-RV forecasting regressions for SPY RV at the daily ( $h = 1$ ), weekly ( $h = 5$ ), monthly ( $h = 22$ ) and quarterly ( $h = 66$ ) horizons. The RV measures are calculated using 300 second returns. The significance of the coefficients is based on Newey-West HAC standard errors, allowing for serial correlation up to order 5 ( $h = 1$ ), 10 ( $h = 5$ ), 44 ( $h = 22$ ), and 132 ( $h = 66$ ), respectively. The superscripts \*, \*\*, and \*\*\* indicate that the coefficient is statistically significant at the 10%, 5% or 1% level respectively. The out-of-sample R-squared,  $R^2_{oos}$ , is calculated as one minus the ratio of the out-of-sample model-based MSPE to the out-of-sample MSPE from a forecast that includes only a constant.

Table 5: HAR type regressions using unsigned and signed total jumps

$$\text{HAR-CJ: } RV_{t,t+h} = \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d} J_t + \beta_{J_w} J_{t-5,t} + \beta_{J_m} J_{t-22,t} + \epsilon_{t,t+h}$$

$$\text{HAR-CJ}^+: RV_{t,t+h} = \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d^+} J_t^+ + \beta_{J_w^+} J_{t-5,t}^+ + \beta_{J_m^+} J_{t-22,t}^+ + \epsilon_{t,t+h}$$

$$\text{HAR-CJ}^-: RV_{t,t+h} = \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{J_d^-} J_t^- + \beta_{J_w^-} J_{t-5,t}^- + \beta_{J_m^-} J_{t-22,t}^- + \epsilon_{t,t+h}$$

	HAR-CJ				HAR-CJ <sup>+</sup>				HAR-CJ <sup>-</sup>			
	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66
$\beta_0$	0.106***	0.148***	0.271***	0.527***	0.069	0.130**	0.291***	0.524***	0.003	-0.016	0.126**	0.431***
$\beta_{C_d}$	0.502***	0.340***	0.188***	0.109***	0.371***	0.232***	0.148***	0.086***	0.246**	0.184***	0.107***	0.063***
$\beta_{C_w}$	0.398**	0.382***	0.250**	0.159***	0.520**	0.562***	0.421***	0.240***	0.365***	0.245***	0.222**	0.141***
$\beta_{C_m}$	0.054	0.111	0.191**	0.214**	0.124	0.137	0.129	0.160	0.153	0.186*	0.170*	0.146*
$\beta_{J_d}$	-0.634***	-0.393**	-0.218***	-0.118**								
$\beta_{J_w}$	-0.614*	-0.725	0.004	0.083								
$\beta_{J_m}$	0.848	1.266*	0.934	-0.003								
$\beta_{J_d^+}$					-0.529**	-0.229*	-0.198***	-0.106***				
$\beta_{J_w^+}$					-0.922*	-1.500**	-0.781**	-0.326**				
$\beta_{J_m^+}$					0.203	0.844	1.225	0.333				
$\beta_{J_d^-}$									0.438	0.501*	0.242	0.118
$\beta_{J_w^-}$									0.980	0.151	-2.179	-1.376**
$\beta_{J_m^-}$									-6.741**	-9.834***	-6.705**	-3.789**
$R^2_{(in)}$	<b>0.555</b>	<b>0.666</b>	<b>0.572</b>	<b>0.338</b>	<b>0.541</b>	<b>0.668</b>	<b>0.578</b>	<b>0.341</b>	<b>0.523</b>	<b>0.664</b>	<b>0.612</b>	<b>0.362</b>
$R^2_{(oos)}$	<b>0.493</b>	<b>0.747</b>	<b>0.728</b>	0.465	<b>0.450</b>	<b>0.754</b>	<b>0.739</b>	<b>0.489</b>	<b>0.511</b>	<b>0.724</b>	0.690	0.445
MSPE	2.821*	1.017*	0.872*	1.274	3.059	0.995*	0.840*	1.218*	2.720*	1.110*	0.994	1.318

Note: See Notes to Table 4. Bold in-sample and out-of-sample R-squared entries indicate that the fit of the proposed models is better than that of the baseline HAR-RV model in Table 4. The \* indicates that the MSPE of the model is significantly lower than the MSPE of the benchmark HAR-RV model in Table 4.

Table 6: HAR type regressions using unsigned and signed finite jumps

$$\begin{aligned} \text{HAR-CFJ: } RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{F_{J_d}} FJ_t + \beta_{F_{J_w}} FJ_{t-5,t} + \beta_{F_{J_m}} FJ_{t-22,t} + \epsilon_{t,t+h} \\ \text{HAR-CFJ}^+: RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{F_{J_d}^+} FJ_t^+ + \beta_{F_{J_w}^+} FJ_{t-5,t}^+ + \beta_{F_{J_m}^+} FJ_{t-22,t}^+ + \epsilon_{t,t+h} \\ \text{HAR-CFJ}^-: RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{F_{J_d}^-} FJ_t^- + \beta_{F_{J_w}^-} FJ_{t-5,t}^- + \beta_{F_{J_m}^-} FJ_{t-22,t}^- + \epsilon_{t,t+h} \end{aligned}$$

	HAR-CFJ				HAR-CFJ <sup>+</sup>				HAR-CFJ <sup>-</sup>			
	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66
$\beta_0$	0.106**	0.148***	0.272***	0.527***	0.068	0.129**	0.290***	0.523***	0.007	-0.011	0.130**	0.432***
$\beta_{C_d}$	0.502***	0.340***	0.188***	0.109***	0.371***	0.232***	0.148***	0.086***	0.246**	0.184***	0.107***	0.063***
$\beta_{C_w}$	0.397**	0.382***	0.249**	0.159***	0.519**	0.562***	0.420***	0.241***	0.363***	0.242***	0.219**	0.139***
$\beta_{C_m}$	0.053	0.111	0.193**	0.216***	0.125	0.138	0.132	0.164	0.155	0.188*	0.171*	0.145*
$\beta_{F_{J_d}}$	-0.635***	-0.393**	-0.218***	-0.118***								
$\beta_{F_{J_w}}$	-0.613*	-0.723	0.008	0.084								
$\beta_{F_{J_m}}$	0.852	1.267*	0.923	-0.016								
$\beta_{F_{J_d}^+}$					-0.530**	-0.229*	-0.198***	-0.106***				
$\beta_{F_{J_w}^+}$					-0.918*	-1.496**	-0.776**	-0.326**				
$\beta_{F_{J_m}^+}$					0.196	0.831	1.183	0.278				
$\beta_{F_{J_d}^-}$									0.448*	0.507*	0.244	0.118
$\beta_{F_{J_w}^-}$									0.988	0.149	-2.188	-1.378**
$\beta_{F_{J_m}^-}$									-6.799**	-9.915***	-6.811**	-3.949**
$R_{(in)}^2$	<b>0.555</b>	<b>0.666</b>	<b>0.572</b>	<b>0.338</b>	<b>0.541</b>	<b>0.668</b>	<b>0.577</b>	<b>0.341</b>	<b>0.523</b>	<b>0.665</b>	<b>0.614</b>	<b>0.363</b>
$R_{(oos)}^2$	<b>0.493</b>	<b>0.747</b>	<b>0.728</b>	0.464	<b>0.449</b>	<b>0.753</b>	<b>0.734</b>	<b>0.478</b>	<b>0.511</b>	<b>0.724</b>	0.684	0.446
MSPE	2.822*	1.018*	0.874*	1.276	3.066	0.998*	0.857*	1.243	2.721*	1.112*	0.994	1.317

Note: See Notes to Table 4. Bold in-sample and out-of-sample R-squared entries indicate that the fit of the proposed models is better than that of the baseline HAR-RV model in Table 4. The \* indicates that the MSPE of the model is significantly lower than the MSPE of the benchmark HAR-RV model in Table 4.

Table 7: HAR type regression using unsigned and signed infinite jumps

$$\begin{aligned} \text{HAR-CIJ: } RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{I_{J_d}} IJ_t + \beta_{I_{J_w}} IJ_{t-5,t} + \beta_{I_{J_m}} IJ_{t-22,t} + \epsilon_{t,t+h} \\ \text{HAR-CIJ}^+: RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{I_{J_d}^+} IJ_t^+ + \beta_{I_{J_w}^+} IJ_{t-5,t}^+ + \beta_{I_{J_m}^+} IJ_{t-22,t}^+ + \epsilon_{t,t+h} \\ \text{HAR-CIJ}^-: RV_{t,t+h} &= \beta_0 + \beta_{C_d} C_t + \beta_{C_w} C_{t-5,t} + \beta_{C_m} C_{t-22,t} + \beta_{I_{J_d}^-} IJ_t^- + \beta_{I_{J_w}^-} IJ_{t-5,t}^- + \beta_{I_{J_m}^-} IJ_{t-22,t}^- + \epsilon_{t,t+h} \end{aligned}$$

	HAR-CIJ				HAR-CIJ <sup>+</sup>				HAR-CIJ <sup>-</sup>			
	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66	h=1	h=5	h=22	h=66
$\beta_0$	0.100**	0.154***	0.280***	0.506***	0.090*	0.139**	0.257***	0.478***	0.099*	0.156***	0.300***	0.544***
$\beta_{C_d}$	0.245**	0.183***	0.103***	0.061***	0.245**	0.183***	0.101***	0.059***	0.246**	0.184***	0.103***	0.061***
$\beta_{C_w}$	0.419***	0.342***	0.324***	0.209***	0.421***	0.345***	0.319***	0.198***	0.421***	0.345***	0.320***	0.197***
$\beta_{C_m}$	0.244**	0.332***	0.289***	0.200***	0.238**	0.324***	0.286***	0.204***	0.239**	0.327***	0.296***	0.223***
$\beta_{I_{J_d}}$	1.423	0.095	-0.266	-0.202								
$\beta_{I_{J_w}}$	-6.066*	-7.153	-8.157	-2.585								
$\beta_{I_{J_m}}$	-1.841	-1.708	15.370	26.391								
$\beta_{I_{J_d}^+}$					2.731*	0.069	-0.284	-0.097				
$\beta_{I_{J_w}^+}$					-10.826*	-11.283**	-13.030**	0.180**				
$\beta_{I_{J_m}^+}$					21.886	33.708	99.050	133.553				
$\beta_{I_{J_d}^-}$									-1.857*	-1.106**	-0.400*	-0.222*
$\beta_{I_{J_w}^-}$									0.827	2.731	1.190	0.484**
$\beta_{I_{J_m}^-}$									5.704	8.459	15.214	21.347
$R_{(in)}^2$	<b>0.512</b>	<b>0.630</b>	<b>0.563</b>	<b>0.340</b>	<b>0.512</b>	<b>0.630</b>	<b>0.576</b>	<b>0.381</b>	<b>0.512</b>	<b>0.629</b>	<b>0.563</b>	<b>0.339</b>
$R_{(oos)}^2$	<b>0.511</b>	<b>0.709</b>	0.644	0.452	<b>0.509</b>	<b>0.711</b>	0.652	<b>0.475</b>	<b>0.512</b>	<b>0.712</b>	0.651	0.454
MSPE	2.722*	1.173*	1.151	1.316	2.731*	1.168*	1.125	1.264	2.714*	1.162*	1.121	1.299

Note: See Notes to Table 4. Bold in-sample and out-of-sample R-squared entries indicate that the fit of the proposed models is better than that of the baseline HAR-RV model in Table 4. The \* indicates that the MSPE of the model is significantly lower than the MSPE of the benchmark HAR-RV model in Table 4.



Table 8: SPY Relative MSPEs and Model Confidence Set Rankings by Frequency – Raw vs. Noise-Robust Measures

	$h = 1$ (day)						$h = 5$ (week)						$h = 22$ (month)						$h = 66$ (quarter)					
	5s		60s		300s		5s		60s		300s		5s		60s		300s		5s		60s		300s	
	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank
Panel A: Raw Measures																								
HAR-RV	1.00	8	1.00	7	1.00	10	1.00	5	1.00	10	1.00	10	1.00	5	1.00	7	1.00	5	1.00	5	1.00	6	1.00	3
HAR-CJ	1.23	10	0.75*	2	0.90*	6	1.029	4	0.90	3	0.770*	3	0.98*	3	1.172	9	0.92*	3	0.968	4	1.167	9	1.010	5
HAR-CFJ	0.87*	3	0.752*	1	0.910*	7	1.181	9	0.92	4	0.770*	4	1.051	6	1.178	10	0.926*	4	1.010	6	1.171	10	1.011	6
HAR-CIJ	1.124	9	1.060	10	0.878*	4	1.022	3	1.034	9	0.888*	9	0.969*	2	1.001	8	1.220	10	0.940*	3	0.993	5	1.043	8
HAR-CJ <sup>+</sup>	0.903*	4	0.993	6	0.986	8	1.165	7	0.969	1	0.753*	1	1.147	9	0.894*	2	0.891*	1	1.074	8	0.977	3	0.965*	1
HAR-CJ <sup>-</sup>	0.848*	2	0.969	4	0.877*	2	1.124	5	1.017	6	0.840*	5	0.841*	1	0.936*	4	1.053	6	0.917*	1	1.020	7	1.045	10
HAR-CFJ <sup>+</sup>	0.925*	7	0.993	5	0.988	9	1.175	8	0.971	2	0.755*	2	1.198	10	0.877*	1	0.908*	2	1.096	9	0.959	2	0.985	2
HAR-CFJ <sup>-</sup>	0.915*	6	0.969	3	0.877*	3	1.215	10	1.035	10	0.841*	6	0.982	4	0.959*	5	1.054	7	1.035	7	1.020	8	1.044	9
HAR-CIJ <sup>+</sup>	0.910*	5	1.055	8	0.881*	5	1.151	6	1.020	7	0.884*	8	1.086	8	0.964*	6	1.192	9	1.136	10	0.940*	1	1.002	4
HAR-CIJ <sup>-</sup>	0.729*	1	1.059	9	0.875*	1	0.996	1	1.030	8	0.879*	7	1.054	7	0.921*	3	1.139	8	0.939*	2	0.977*	4	1.029	7
Panel B: Robust-to-Noise Measures																								
HAR-RV	0.843*	5	0.907*	5	1.000	1	0.882*	6	0.976	5	0.992	7	0.821*	5	1.031	6	1.154	8	0.893*	3	1.013	6	1.014	9
HAR-CJ	0.768*	1	0.966	8	1.015	5	0.895*	3	1.010	7	0.992	8	0.977	9	1.044	8	1.145	6	0.988	10	0.996	5	0.906*	2
HAR-CFJ	0.775*	2	0.960*	7	1.015	4	0.867*	4	1.060	10	0.953*	3	0.987	10	1.031	7	1.143	4	0.921*	6	0.925*	1	1.032	10
HAR-CIJ	0.791*	3	0.980	10	1.018	9	0.890*	9	1.025	9	0.965	9	0.803*	4	1.073	9	1.179	10	0.875*	2	1.016	7	0.998	7
HAR-CJ <sup>+</sup>	0.851*	6	0.684*	2	1.015	7	0.894*	8	0.930*	4	0.960	5	0.838*	6	0.907*	2	1.145	7	0.926*	7	1.037	8	0.991	4
HAR-CJ <sup>-</sup>	0.870*	8	0.852*	4	1.013	3	0.828*	2	0.889*	2	0.953*	1	0.772*	1	0.912	3	1.135	1	0.899*	4	0.968	4	0.997	6
HAR-CFJ <sup>+</sup>	0.866*	7	0.677*	1	1.015	6	0.895*	10	0.889*	1	0.960	4	0.861*	8	0.938*	4	1.145	5	0.919*	5	1.037	9	0.990	3
HAR-CFJ <sup>-</sup>	1.111	10	0.852*	3	1.013	2	0.882*	7	0.894*	3	0.953*	2	0.786*	2	0.902*	1	1.135	2	0.931*	9	0.953	3	0.753*	1
HAR-CIJ <sup>+</sup>	0.794*	4	0.972	9	1.026	10	0.875*	5	1.015	6	0.977	10	0.841*	7	1.166	10	1.164	9	0.930*	8	1.038	10	0.994	5
HAR-CIJ <sup>-</sup>	1.009	9	0.958	6	1.016	8	0.793*	1	1.015	8	0.961	6	0.794*	3	0.947*	5	1.137	3	0.852*	1	0.941*	2	1.000	8

Memo:

HAR-RV MSPE	3.364	4.550	3.102	1.553	1.350	1.322	1.443	1.025	0.944	1.778	1.344	1.262
-------------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Note: The relative MSPE ratios are the ratios of the MSPEs of the extended HAR models using standard volatility measures (top panel) or robust-to-noise measures (bottom panel) relative to the baseline HAR-RV models employing standard RV measures. The MSPEs are based on rolling regression using 1000 observations. The Relative MSPE ratio below one indicate better forecasting performance (lower MSPE) than the baseline HAR-RV model, and vice versa. The superscript \* denotes statistically significant reductions in MSPE. These are identified by the Diebold and Mariano test using a 5% level and robust standard errors. The bold MCS rank entries denote the models that are retained by the Hansen et al. (2011) MCS procedure using a 5% level and a 10 day block bootstrap with 5000 replications.

Table 9: Individual Stocks Relative MSPEs and Model Confidence Set by Frequency – Raw vs. Noise-Robust Measures

	$h = 1$ (daily)						$h = 5$ (week)						$h = 22$ (month)						$h = 66$ (quarter)					
	5s		60s		300s		5s		60s		300s		5s		60s		300s		5s		60s		300s	
	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank	Rel. MSPE	MCS Rank
Panel A – Standard Volatility Measures																								
HAR-RV	1.000	<b>18</b>	1.000	<b>19</b>	1.000	<b>20</b>	1.000	<b>20</b>	1.000	<b>20</b>	1.000	<b>16</b>	1.000	<b>19</b>	1.000	<b>18</b>	1.000	<b>15</b>	1.000	<b>9</b>	1.000	<b>15</b>	1.000	<b>19</b>
HAR-CJ	0.999	<b>18</b>	0.991	<b>20</b>	0.972	<b>20</b>	0.950	<b>20</b>	0.916	<b>20</b>	0.933	<b>20</b>	0.929	<b>19</b>	0.942	<b>19</b>	0.970	<b>20</b>	0.928	<b>17</b>	0.958	<b>17</b>	0.995	<b>20</b>
HAR-FJ	1.057	<b>17</b>	0.984	<b>20</b>	0.973	<b>20</b>	1.048	<b>14</b>	0.916	<b>20</b>	0.934	<b>20</b>	1.064	<b>13</b>	0.943	<b>18</b>	0.974	<b>19</b>	1.043	<b>11</b>	0.952	<b>16</b>	0.997	<b>20</b>
HAR-IJ	1.010	<b>20</b>	0.973	<b>19</b>	0.940	<b>20</b>	0.986	<b>19</b>	0.955	<b>19</b>	0.942	<b>20</b>	1.035	<b>15</b>	1.010	<b>16</b>	1.063	<b>16</b>	1.007	<b>13</b>	1.014	<b>10</b>	1.037	<b>15</b>
HAR-CJ <sup>+</sup>	1.044	<b>15</b>	1.000	<b>20</b>	0.968	<b>20</b>	1.098	<b>11</b>	0.939	<b>20</b>	0.945	<b>20</b>	1.203	<b>9</b>	0.994	<b>18</b>	1.038	<b>18</b>	1.127	<b>6</b>	1.004	<b>14</b>	1.033	<b>18</b>
HAR-CJ <sup>-</sup>	1.063	<b>18</b>	1.018	<b>19</b>	0.932	<b>20</b>	1.038	<b>15</b>	0.943	<b>20</b>	0.934	<b>20</b>	1.144	<b>11</b>	0.970	<b>18</b>	1.026	<b>17</b>	1.078	<b>7</b>	0.997	<b>10</b>	1.018	<b>18</b>
HAR-CFJ <sup>+</sup>	1.055	<b>13</b>	0.999	<b>20</b>	0.969	<b>20</b>	1.153	<b>9</b>	0.940	<b>20</b>	0.945	<b>20</b>	1.267	<b>7</b>	0.994	<b>17</b>	1.038	<b>18</b>	1.173	<b>7</b>	1.004	<b>15</b>	1.031	<b>17</b>
HAR-CFJ <sup>-</sup>	1.103	<b>11</b>	0.984	<b>19</b>	0.932	<b>20</b>	1.115	<b>10</b>	0.938	<b>20</b>	0.937	<b>20</b>	1.228	<b>9</b>	0.970	<b>18</b>	1.030	<b>18</b>	1.144	<b>7</b>	0.997	<b>11</b>	1.016	<b>19</b>
HAR-CIJ <sup>+</sup>	1.044	<b>15</b>	0.979	<b>19</b>	0.939	<b>20</b>	1.090	<b>12</b>	0.966	<b>19</b>	0.946	<b>20</b>	1.189	<b>7</b>	1.010	<b>16</b>	1.080	<b>17</b>	1.129	<b>7</b>	1.004	<b>10</b>	1.042	<b>15</b>
HAR-CIJ <sup>-</sup>	1.011	<b>18</b>	0.982	<b>19</b>	0.947	<b>20</b>	1.071	<b>12</b>	0.960	<b>20</b>	0.945	<b>20</b>	1.213	<b>9</b>	1.005	<b>17</b>	1.091	<b>16</b>	1.137	<b>6</b>	1.006	<b>13</b>	1.062	<b>15</b>
Panel B – Robust-to-Noise Volatility Measures																								
HAR-RV	0.966	<b>20</b>	0.916	<b>18</b>	0.969	<b>20</b>	0.975	<b>20</b>	1.017	<b>20</b>	0.998	<b>20</b>	0.975	<b>18</b>	1.081	<b>20</b>	1.138	<b>17</b>	0.956	<b>14</b>	1.050	<b>20</b>	1.032	<b>14</b>
HAR-CJ	0.958	<b>20</b>	0.935	<b>18</b>	0.975	<b>20</b>	0.934	<b>20</b>	0.975	<b>20</b>	0.990	<b>19</b>	0.958	<b>16</b>	1.077	<b>19</b>	1.135	<b>18</b>	0.949	<b>18</b>	1.040	<b>20</b>	0.962	<b>19</b>
HAR-FJ	0.980	<b>20</b>	0.939	<b>19</b>	0.976	<b>20</b>	0.962	<b>19</b>	1.003	<b>20</b>	0.996	<b>20</b>	0.966	<b>18</b>	1.082	<b>20</b>	1.075	<b>18</b>	0.882	<b>17</b>	0.963	<b>17</b>	0.994	<b>19</b>
HAR-IJ	0.969	<b>20</b>	0.926	<b>19</b>	0.970	<b>20</b>	0.956	<b>18</b>	1.022	<b>20</b>	0.985	<b>19</b>	0.943	<b>17</b>	1.064	<b>20</b>	1.122	<b>18</b>	0.905	<b>15</b>	1.042	<b>18</b>	1.021	<b>16</b>
HAR-CJ <sup>+</sup>	0.955	<b>20</b>	0.986	<b>19</b>	0.978	<b>20</b>	0.962	<b>19</b>	1.008	<b>19</b>	0.991	<b>19</b>	0.981	<b>18</b>	1.082	<b>20</b>	1.092	<b>16</b>	0.956	<b>15</b>	1.042	<b>19</b>	1.017	<b>14</b>
HAR-CJ <sup>-</sup>	0.973	<b>20</b>	0.943	<b>18</b>	0.961	<b>20</b>	0.950	<b>20</b>	0.984	<b>19</b>	0.994	<b>19</b>	0.938	<b>16</b>	1.043	<b>19</b>	1.126	<b>19</b>	0.936	<b>14</b>	1.030	<b>17</b>	1.019	<b>17</b>
HAR-CFJ <sup>+</sup>	0.947	<b>20</b>	0.987	<b>19</b>	0.980	<b>20</b>	0.952	<b>20</b>	1.010	<b>19</b>	0.993	<b>19</b>	0.967	<b>17</b>	1.086	<b>19</b>	1.091	<b>16</b>	0.924	<b>12</b>	1.044	<b>18</b>	1.014	<b>16</b>
HAR-CFJ <sup>-</sup>	0.963	<b>20</b>	0.938	<b>18</b>	0.962	<b>20</b>	0.962	<b>19</b>	0.984	<b>19</b>	0.994	<b>20</b>	0.948	<b>16</b>	1.047	<b>19</b>	1.107	<b>17</b>	0.945	<b>12</b>	1.031	<b>17</b>	1.024	<b>16</b>
HAR-CIJ <sup>+</sup>	0.972	<b>20</b>	0.926	<b>19</b>	0.950	<b>20</b>	0.957	<b>19</b>	1.022	<b>20</b>	0.994	<b>19</b>	0.966	<b>16</b>	1.073	<b>20</b>	1.091	<b>18</b>	0.952	<b>12</b>	1.045	<b>17</b>	1.008	<b>18</b>
HAR-CIJ <sup>-</sup>	0.964	<b>20</b>	0.935	<b>19</b>	0.948	<b>20</b>	0.948	<b>19</b>	1.025	<b>20</b>	0.986	<b>20</b>	0.969	<b>16</b>	1.061	<b>19</b>	1.116	<b>18</b>	0.943	<b>11</b>	1.037	<b>17</b>	1.033	<b>15</b>

Memo:

HAR-RV MSPE	373.1364	54.8865	22.7444	16.9684	85.5808	27.0842	8.8123	6.3931	17.2674	7.9011	6.2917
-------------	----------	---------	---------	---------	---------	---------	--------	--------	---------	--------	--------

Note: The relative MSPE entries are the 20 stock average ratio of the MSPEs of the extended HAR models using standard measures (top panel) or robust-to-noise measures (bottom panel), relative to the MSPE of the baseline HAR-RV models employing standard RV measures. The MSPEs are based on rolling regression using 1000 observations. Relative MSPE ratios below one indicate better forecast performance (lower MSPE) than the baseline HAR-RV model on average, and vice versa. The MCS entries are the number of stocks where the model is retained using the Hansen et al. (2011) MCS using a 5% level, a block bootstrap with 5000 replications and a 10 day window. The bold MCS entries denotes models which have been retained for 15 or more stocks.

Table 10: Model Averaging Relative MSPEs

	$h = 1$	$h = 5$	$h = 22$	$h = 66$	$h = 1$	$h = 5$	$h = 22$	$h = 66$
	SPY – 300 seconds				SPY – 60 seconds			
HAR-RV Baseline	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Best Extended HAR	0.875*	0.753*	0.891*	0.965*	0.752*	0.969	0.877	0.940*
Avg. – Min Var Weights	0.987	<b>0.693**</b>	0.895*	0.966*	0.812*	0.977	0.940*	0.971*
Avg. – MSPE Weights	0.879*	<b>0.706**</b>	<b>0.862**</b>	<b>0.919**</b>	0.875*	<b>0.914**</b>	<b>0.850*</b>	0.965*
Avg. – Rank Weights	0.910*	<b>0.715*</b>	<b>0.845**</b>	<b>0.873**</b>	0.880*	<b>0.923*</b>	<b>0.846*</b>	0.986
Avg. – Equal Weights	<b>0.873*</b>	<b>0.712*</b>	<b>0.876*</b>	<b>0.928*</b>	0.877*	<b>0.914**</b>	<b>0.852*</b>	0.964*
Memo:								
HAR-RV MSPE	3.102	1.322	0.944	1.262	4.550	1.350	1.025	1.344
	BA – 300 seconds				BFB – 300 seconds			
HAR-RV Baseline	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Best Extended HAR	0.981	0.937	0.993	0.864*	0.924*	0.836*	0.822*	0.876*
Avg. – Min Var Weights	0.992	<b>0.905**</b>	1.083	1.001	0.969*	0.845*	<b>0.751**</b>	<b>0.812**</b>
Avg. – MSPE Weights	<b>0.972*</b>	<b>0.906*</b>	<b>0.915**</b>	0.959*	0.926*	<b>0.823*</b>	<b>0.814*</b>	<b>0.856**</b>
Avg. – Rank Weights	<b>0.976*</b>	<b>0.923*</b>	<b>0.928**</b>	0.980	0.936*	<b>0.820*</b>	<b>0.810**</b>	<b>0.847**</b>
Avg. – Equal Weights	<b>0.972*</b>	<b>0.906*</b>	<b>0.919**</b>	0.961*	0.926*	<b>0.823*</b>	<b>0.816*</b>	0.878*
	COST – 300 seconds				KO – 300 seconds			
HAR-RV Baseline	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Best Extended HAR	0.958*	0.879*	0.925*	0.957*	0.814*	0.709*	0.882*	0.939*
Avg. – Min Var Weights	1.016	0.985	<b>0.881**</b>	<b>0.950*</b>	0.923*	<b>0.695**</b>	<b>0.837**</b>	<b>0.916*</b>
Avg. – MSPE Weights	0.962*	<b>0.871*</b>	<b>0.920*</b>	0.958*	0.817*	0.713*	0.888*	0.975*
Avg. – Rank Weights	0.969*	<b>0.856*</b>	<b>0.907**</b>	<b>0.945**</b>	<b>0.811*</b>	<b>0.686*</b>	<b>0.829**</b>	0.950*
Avg. – Equal Weights	0.962*	<b>0.873*</b>	<b>0.922*</b>	0.960*	0.817*	0.723*	0.914*	0.983*

Note: The table reports the relative MSPE, the ratio of MSPE of the model indicated in the first column to the MSPE of the baseline HAR-RV, in both cases using standard volatility measures as opposed to robust-to-noise measures. The best models refers to the min MSPE model from the set of jump models presented in Section 4. The superscript \* identifies models with significantly lower MSPEs than the benchmark HAR-RV. The bold entries highlight models where the MSPE of the model average is lower than the MSPEs of both the HAR-RV benchmark and the best model. The starred bold entries (superscript \*\*) identify models whose MSPEs are significantly lower than the MSPEs of both the benchmark HAR-RV and the best extended HAR model.

Table 11: Two-Scale Realized Semivariance – MSE Finite Sample Performance

	5s	60s	300s	5s	60s	300s
	$\xi = 0.01$			$\xi = 0.1$		
$RSV^+$	9.680	0.067	0.003	963.848	6.718	0.277
$RSV^-$	9.704	0.069	0.004	964.483	6.779	0.290
$TSRV^+$	0.001	0.001	0.002	0.113	0.014	0.008
$TSRV^-$	0.001	0.001	0.002	0.112	0.015	0.009

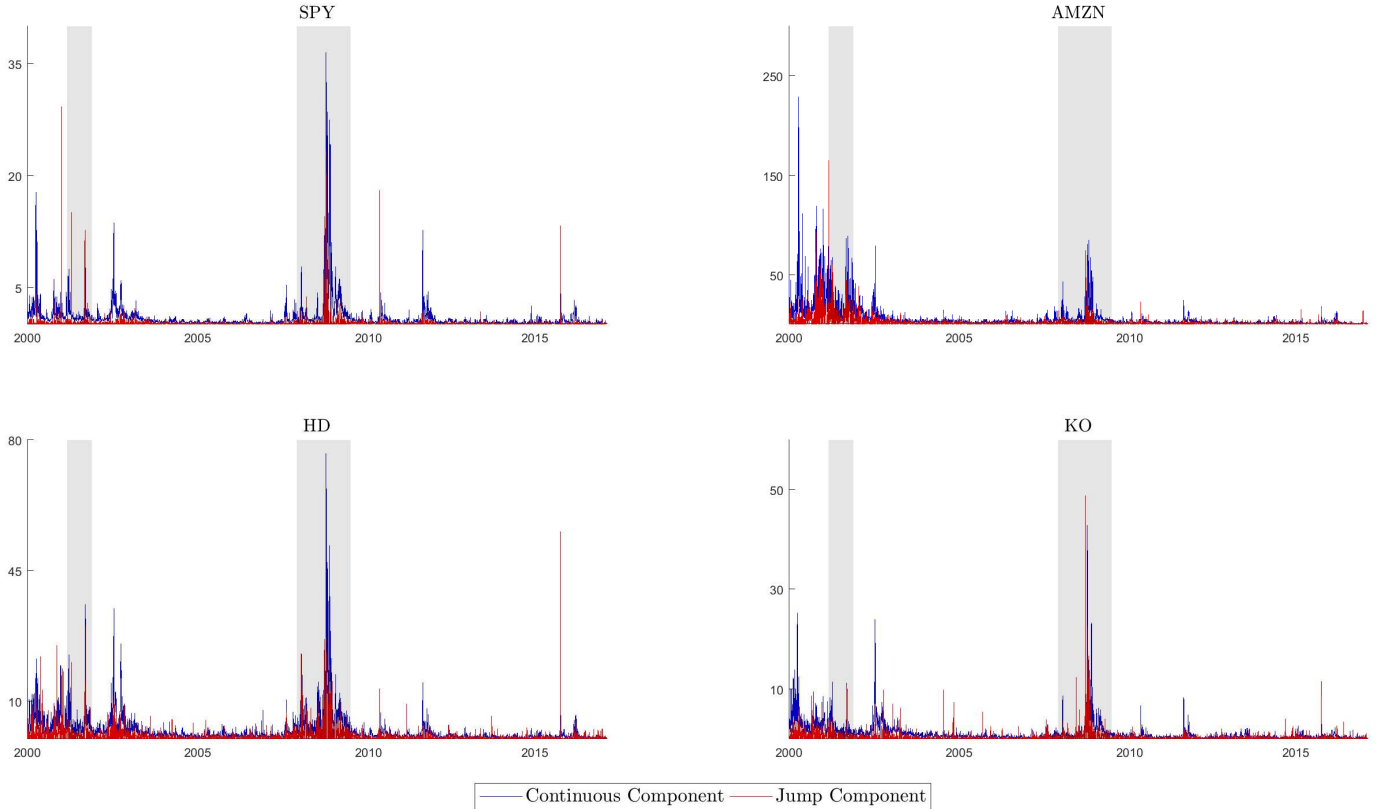
Note: The table reports the finite sample mean squared errors of the realized semivariances and two-scale realized semivariances in the simulation exercise described in Appendix B with finite activity compound Poisson jumps.  $\xi$  represents the noise-to-signal ratio used to simulate the market microstructure noise.

Table 12: Noise-Robust ABD Test – Size and Power

	5s	60s	300s	5s	60s	300s
	$\xi = 0.01$			$\xi = 0.1$		
	Size					
ABD Noise-robust	0.070	0.041	0.015	0.049	0.011	0.006
ABD	0.030	0.055	0.128	0.029	0.045	0.082
	Power – Compound Poisson (Finite Jumps)					
ABD Noise-robust	1.000	0.991	0.703	0.963	0.905	0.458
ABD	0.985	0.989	0.986	0.337	0.484	0.586
	Power – Cauchy Process (Infinite Jumps)					
ABD Noise-robust	0.948	0.784	0.410	0.670	0.572	0.342
ABD	0.732	0.774	0.764	0.361	0.415	0.463

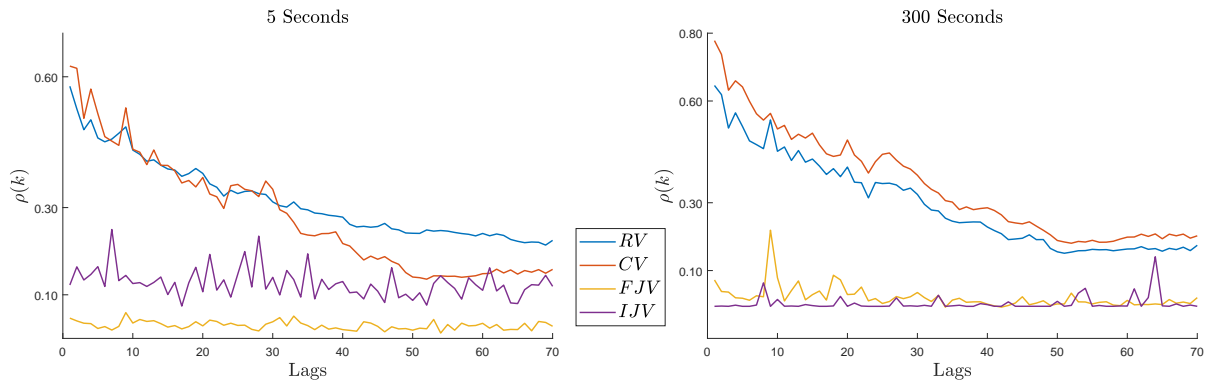
Note: This table report the empirical size and power of the ABD test of Andersen et al. (2007b) and our modified version that is noise-robust. The theoretical size is  $\alpha = 0.05$ , and the Monte Carlo settings and models are described in Appendix B.  $\xi$  represents the noise-to-signal ratio used to simulate the market microstructure noise.

Figure 1: Time Series of Realized Volatility – Jump and Continuous Components



Note: This figure depicts the elements of the realized volatility for SPY and three individual stocks estimated at the 300 second frequency. The three individual stocks have the largest, smallest and average RV. NBER dated U.S. recession are shaded grey.

Figure 2: Autocorrelation Function of SPY Realized Measures



Note: The figure graphs the autocorrelation of the realized variance and its elements. The autocorrelations at the 5 and 300 second frequencies were estimated using noise-robust and raw estimators, respectively.