Journal of
Risk and Financial Management

Article

Co-Jumps, Co-Jump Tests, and Volatility Forecasting: Monte Carlo and Empirical Evidence

Weijia Peng [1,*] and Chun Yao [2]

[1] Department of Finance, Jack Welch College of Business & Technology, Sacred Heart University, 5151 Park Avenue, Fairfield, CT 06825, USA
[2] US Economics Research, Barclays, 745 Seventh Avenue, New York, NY 10019, USA; chunyao.economics@gmail.com
* Correspondence: pengw@sacredheart.edu

Abstract: This study classifies jumps into idiosyncratic jumps and co-jumps to quantitatively identify systematic risk and idiosyncratic risk by utilizing high-frequency data. We found that systematic risk occurs more frequently and has larger magnitudes than the idiosyncratic risk in an individual asset, which indicates that volatilities from one sector are largely derived from the contagious effect of other sectors. We further investigated the importance of idiosyncratic jumps and co-jumps to predict the sector-level S&P500 exchange-traded fund (ETF) volatility. It was found that the predictive content of co-jumps is higher than that of idiosyncratic jumps, suggesting that systematic risk is more informative than idiosyncratic risk in volatility forecasting. Additionally, we carried out Monte Carlo experiments designed to examine the relative performances of the four co-jump tests. The findings indicate that the BLT test and the co-exceedance rule of the LM test outperform other tests, while the co-exceedance rule of the LM test has larger power and a smaller empirical size than that of the BLT test.

Keywords: volatility forecasting; co-jumps; co-jump tests; heterogeneous autoregressive model; high-frequency data

Citation: Peng, Peng, and Chun Yao. 2022. Co-Jumps, Co-Jump Tests, and Volatility Forecasting: Monte Carlo and Empirical Evidence. *Journal of Risk and Financial Management* 15: 334. https://doi.org/10.3390/jrfm15080334

Academic Editor: Svetlozar (Zari) Rachev

Received: 29 June 2022
Accepted: 25 July 2022
Published: 28 July 2022

1. Introduction

Effective stock volatility prediction remains a key goal for empirical finance researchers. For example, in portfolio management, agents utilize such predictions when allocating assets and managing risk. Various discrete-time models (e.g., ARCH-GARCH models) and continuous-time models (e.g., stochastic volatility and related models) have been developed to address this issue. Key papers in the discrete-time framework include Harvey et al. (1994), Andersen and Lund (1997), Hansen and Lunde (2005), Brandt and Jones (2006), and the references therein. With the availability of high-frequency intraday data, Andersen et al. (2001) developed a nonparametric volatility measure, which fostered the development of a number of time series volatility forecasting models, including mixed data sampling (MIDAS) models (see Ghysels et al. 2006; Ghysels and Sinko 2011) and long-memory ARMA models (e.g., Andersen et al. 2003; Koopman et al. 2005). In particular, this has led to a new class of parsimonious models, called heterogeneous autoregressive models of realized volatility (HAR-RV), including lag terms of RV over different time horizons. The original HAR forecasting model, introduced by Müller et al. (1997) and Corsi (2009), has been augmented by separating jump and diffusion components, including a variety of jump variation and intensity-related variables (e.g., Andersen et al. 2007; Corsi et al. 2010; Patton and Sheppard 2015; Lahaye and Neely 2020). These models have been successfully applied in various empirical applications.

In this study, we classified jumps into idiosyncratic jumps and co-jumps to decompose jump risk into idiosyncratic risk and systematic risk, and further investigated the importance of idiosyncratic jumps and co-jumps to predict sector-level S&P500 exchange-traded fund (ETF) volatility. Previous literature mostly focused on studying jumps in individual assets or a portfolio index and examining the predictability of jumps in forecasting volatilities of different target assets without decomposition. Our motivation for decomposing jumps into idiosyncratic jumps and co-jumps is to quantitatively identify idiosyncratic risk and systematic risk in the individual asset by utilizing 5-min price data and, thus, evaluate the predictability of idiosyncratic risk and systematic risk in forecasting volatilities of the target asset. Our impetus for this decomposition is the hypothesis that systematic jumps (risk) may be more informative than idiosyncratic jumps (risk) in volatility prediction. In addition, the identification of systematic jumps from idiosyncratic jumps is important on its own. In theory, idiosyncratic jumps can be alleviated in a weighted portfolio, whereas systematic jumps cannot. Hence, the failure to distinguish idiosyncratic jumps (idiosyncratic risk) from co-jumps (systematic risk) can lead to dire consequences when managing portfolios.

To evaluate this hypothesis, co-jumps and idiosyncratic jumps are included in predictive HAR-type regressions, which enables us to assess their relative marginal predictive content. Specifically, we compare the following alternative models: (i) the basic HAR model, which includes the continuous component of realized volatility; (ii) the HAR model that includes the continuous component and the co-jump component; (iii) the augmented HAR model that includes the continuous component and the idiosyncratic jump component; and (iv) the augmented HAR model that includes all three components. We predict the volatility of a variety of target assets, including the major S&P500 sector-level ETFs: financial (XLF), technology (XLK), consumer staple (XLP), utility (XLU), and consumer discretionary (XLY). Co-jumps and idiosyncratic jumps among sector-level ETFs are detected using the co-exceedance rule of the LM jump test in Lee and Mykland (2007) and Gilder et al. (2014).

Our experiment focused on one-day-ahead volatility predictions using rolling windows. Our sample period was from November 2008 to December 2019. For model evaluation, we used in-sample adjusted R^2, out-of-sample R^2, and mean absolute forecasting error (MAFE). In addition, we implement the model confidence set (MCS) tests of Hansen et al. (2011) to compare the predictive abilities of all candidate models. We summarize our main findings below.

First, among the four alternative HAR models in our forecasting experiment, models augmented with co-jumps and with both co-jumps and idiosyncratic jumps showed the largest predictive accuracy gains, in terms of in-sample adjusted R^2, out-of-sample R^2, and the mean absolute forecasting error (MAFE). This result was robust across different forecasting targets, i.e., S&P 500 sector ETF volatilities. For example, models augmented with co-jumps showed smaller MAFEs than the baseline HAR model and HAR model augmented with idiosyncratic jumps for all five target sectors, as listed in Table 10. Moreover, models combining both co-jumps and idiosyncratic jumps showed even smaller MAFEs, except for the financial sector. Moreover, the model confidence set (MCS) test, which assigns the value of 1 to the best model and a larger number for superior models, also suggests that models augmented with co-jumps or models augmented with both co-jumps and idiosyncratic jumps are superior to other models.

Second, models with co-jumps resulted in higher explanatory power than models with idiosyncratic jumps in both the in-sample and ex-ante forecasting experiments, in line with the hypothesis that systematic risks (as represented by co-jumps) potentially play a larger role than idiosyncratic risks in volatility forecasting.

Third, co-jumps occurred generally two to three times more frequently than idiosyncratic jumps during our sample period from 2007 to 2019, and the frequency of occurrences of both the co-jumps and idiosyncratic jumps increased especially during significant market turmoil, such as the 2008 financial crisis and the 2011 European debt crisis. This finding is consistent with the prevalence of systematic risk compared with the idiosyncratic risk in the financial market.

Fourth, we also found that co-jumps have a larger magnitude and variation compared to idiosyncratic jumps, suggesting systematic risk may pose a larger impact on the financial market than idiosyncratic risk.

Finally, given the importance of co-jump tests in our empirical analysis, we additionally carried out a series of Monte Carlo experiments to examine the relative performances of a number of widely used co-jump tests, including: (i) the BLT co-jump test in Bollerslev et al. (2008); (ii) the JT co-jump test in Jacod and Todorov (2009); (iii) the co-exceedance rule in Gilder et al. (2014) using the LM test in Lee and Mykland (2007); and (iv) the co-exceedance rule using the intersection of the LM test and BNS test in Barndorff-Nielsen and Shephard (2006). While univariate jump tests have been researched extensively[1], the co-jump testing literature used in our empirical analysis is relatively nascent. One strand of the literature identifies co-jumps based on the identification of jumps in a portfolio (e.g., Bollerslev et al. 2008). Another strand uses univariate jump tests to identify co-jumps in multivariate processes (e.g., Gilder et al. 2014). A third strand develops co-jump tests by direct examination of multiple price processes (see Jacod and Todorov 2009; Mancini and Gobbi 2012; Bandi and Reno 2016; Bibinger and Winkelmann 2015; Caporin et al. 2017).

Therefore, based on our Monte Carlo experiments, our key finding is that the BLT test and the co-exceedance rule of the LM test outperform other tests, while the co-exceedance rule of the LM test has larger power and a smaller empirical size than that of the BLT test. The JT test, the BLT test, and the co-exceedance rule of the LM test all have good powers, except for the co-exceedance rule of the intersection of the LM and BNS tests, and this result is robust across all the data-generating processes (DGPs) considered in our experiment. Additionally, the BLT test and the co-exceedance rule of the LM test both have empirical sizes close to the nominal size in all DGPs. The empirical size of the JT test is mixed, under- or over-sized in different DGPs, and the co-exceedance rule of the LM and BND test is significantly undersized. Interestingly, adding idiosyncratic jumps into the DGPs increases the size of the JT test and lowers the size of the other three tests.

The rest of this paper is organized as follows. Section 2 outlines the setup and summarizes the co-jump tests examined in the paper. Section 3 discusses the Monte Carlo experiments. Section 4 outlines all volatility forecasting models used in the sequel. Section 5 summarizes the setup of our forecasting experiments. Section 6 briefly discusses the data used in our analysis, and Section 7 summarizes our key empirical findings. Concluding remarks are contained in Section 9.

2. Co-Jump Identification

2.1. Setup

Let P_t be the log price of an asset at time t. We assume P_t follows an Itô semimartingale:

$$P_t = P_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t \tag{1}$$

In the above equation, both b_s and σ_s are adapted, càdlàg, and locally bounded. W_s is a standard Brownian motion and J_t is a compound Poisson process (CPP):

$$J_t = \sum_{i=1}^{N_t} Y_i \tag{2}$$

where N_t is a Poisson process, representing the number of jumps in the interval $[0, t]$. The jump magnitudes Y_i's are *iid* random variables.

Consider a finite time horizon $[0, t]$ that contains n high-frequency observations of the log price process. A typical time horizon is one day. Let $\Delta_n = t/n$ be the sampling frequency. The intraday return at the equidistant interval is $[(i-1)\Delta_n, i\Delta_n]$ is $r_i = P_{i\Delta_n} - P_{(i-1)\Delta_n}$.

In the high-frequency literature, it is widely recognized that realized volatility (RV)[2] is a consistent estimator of quadratic variation, namely:

$$RV_t = \sum_{i=1}^{n} r_i^2 \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^2 ds + \sum_{s \le T} (\Delta P_s)^2 = QV_t = IV_t + JV_t, \tag{3}$$

where u.c.p. denotes convergence in probability, uniformly in time. The integrated volatility ($IV_t = \int_0^t \sigma_s^2 ds$) measures the contribution of the continuous component to the total quadratic variation. The jump variation $JV_t = \sum_{s \leq T} (\Delta P_s)^2$ measures the contribution from the discontinuous (jump) component.

There are many estimators of IV. For example, Barndorff-Nielsen and Shephard (2004) propose a bipower variation estimation BPV_t to estimate integrated volatility. Namely, define the standardized realized bipower variation as:

$$BPV_t = (\mu_1)^{-2} \sum_{i=2}^{n} |r_i||r_{i-1}|, \tag{4}$$

where $\mu_1 = E(|Z|) = \sqrt{2}\Gamma(1)/\Gamma(1/2) = \sqrt{2/\pi}$, where Z is a standard normal random variable and $\Gamma(\cdot)$ denotes the gamma function.

2.2. Co-Jump Tests

In this section, we briefly summarize the co-jump testing methods used in the sequel.

2.2.1. BLT Co-Jump Test

Bollerslev et al. (2008) proposed a BLT test to detect co-jumps in a large ensemble of stocks. They developed a theoretical foundation that showed how only co-jumps (not idiosyncratic jumps) could be detected in a large equal-weighted index. Let M denote the total number of assets under consideration. The BLT mean cross-product test statistic is defined as:

$$mcp_{t,i} = \frac{2}{M(M-1)} \sum_{j=1}^{M-1} \sum_{l=j+1}^{M} r_i^j r_i^l, \quad i = 1, \cdots, n \text{ and } t = 1, \cdots, T, \tag{5}$$

where $r_i^j = P^j_{i\Delta_n} - P^j_{(i-1)\Delta_n}$ for $j = 1, \cdots, M$.

The mcp statistic has a nonzero means; the corresponding studentized test statistic is:

$$z_{mcp,t,i} = \frac{mcp_{t,i} - \overline{mcp_t}}{s_{mcp,t}}, \quad \text{for} \quad i = 1, \cdots, n \quad and \quad t = 1, \cdots, T. \tag{6}$$

where

$$\overline{mcp_t} = \frac{1}{n} mcp_t = \frac{1}{n} \sum_{i=1}^{n} mcp_{t,i}, \tag{7}$$

$$s_{mcp,t} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (mcp_{t,i} - \overline{mcp_t})^2}. \tag{8}$$

The critical value under the null is calculated by bootstrapping the test statistics $z_{mcp,t,i}$ using Monte Carlo simulations.

2.2.2. JT Co-Jump Test

Jacod and Todorov (2009) constructed two test statistics to identify co-jumps under two different null hypotheses of joint and disjoint jumps. The test statistics were proposed for detecting co-jumps in a bivariate process. Co-jumps among multivariate processes can be detected from the combination of bivariate processes. The test statistics of the common jump $\Phi_n^{(j)}$ and disjoint jump $\Phi_n^{(d)}$ are defined as:

$$\Phi_n^{(j)} = \frac{V(f, k\Delta_n)_t}{V(f, \Delta_n)_t} \quad \text{and} \quad \Phi_n^{(d)} = \frac{V(f, \Delta_n)_t}{\sqrt{V(g_1, \Delta_n)_t V(g_2, \Delta_n)_t}}, \tag{9}$$

where k is an integer greater than 1 and $V(f, k\Delta_n)_t$ is defined as:

$$V(f, k\Delta_n)_t = \sum_{i=1}^{[t/k\Delta_n]} f(P_{i\Delta_n} - P_{(i-1)\Delta_n}). \tag{10}$$

The functions $f(x)$, $g_1(x)$, and $g_2(x)$ used in the cited paper are given as follows:

$$f(x) = (x_1 x_2)^2, \; g_1(x) = (x_1)^4, \; g_2(x) = (x_2)^4. \tag{11}$$

They show that the test statistics for the null hypothesis of disjoint jumps $\Phi_n^{(d)}$ converge in probability to 0 on $\Omega_T^{(d)}$ and the null hypothesis with common jumps $\Phi_n^{(j)}$ converges in probability to 1 on $\Omega_T^{(j)}$ (refer to Theorem 3.1 therein). Here, $\Omega_T^{(j)}$ and $\Omega_T^{(d)}$ are defined as:

$$\Omega_T^{(j)} = \{\omega\text{: on } [0, t] \text{ the process } r_i^1 r_i^2 \text{ is not identically } 0\}, \tag{12}$$

$$\begin{aligned} \Omega_T^{(d)} = \{\omega\text{: on } [0, t] \text{ the processes } r_i^1 \text{ and } r_i^2 \text{ are} \\ \text{not identically } 0\text{, but the process } r_i^1 r_i^2 \text{ is}\}. \end{aligned} \tag{13}$$

The authors constructed critical regions of the two statistics as:

$$C_n^{(j)} = \{|\Phi_n^{(j)} - 1| \geq c_n^{(j)}\} \quad \text{and} \quad C_n^{(d)} = \{\Phi_n^{(d)} \geq c_n^{(d)}\}, \tag{14}$$

where $c_n^{(j)} = \hat{V}_n^{(j)}/\sqrt{\alpha}$ and $c_n^{(d)} = \hat{V}_n^{\prime(d)}/\alpha$ with significance level α. The term $\hat{V}_n^{(j)}$ is defined as:

$$\hat{V}_n^{(j)} = \frac{\sqrt{\Delta_n (k-1) \hat{F}_t^{\prime n}}}{V(f, \Delta_n)_t} \tag{15}$$

with $\hat{F}_t^{\prime n}$ given by:

$$\hat{F}_t^{\prime n} = \frac{2}{k_n \Delta_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n-1} \sum_{m \in I_n(i)} (r_i^1)^2 (r_i^2)^2 \times (r_i^1 r_m^2 + r_m^1 r_i^2)^2 1_{\{|r_i| > \alpha\Delta_n^{\omega}, |r_m| \leq \alpha\Delta_n^{\omega}\}}, \tag{16}$$

where $I_n(i) = I_{n,-}(i) \cup I_{n,+}(i)$ and $I_{n,-}(i) = \{i - k_n, i - k_n + 1, \cdots, i - 1\}$ if $i > k_n$ and $I_{n,+}(i) = \{i + 2, i + 3, \cdots, i + k_n + 1\}$. The term $\hat{V}_n^{('d)}$ is defined as:

$$\hat{V}_n^{('d)} = \frac{\Delta_n \hat{F}_t^n + \hat{A}_T^{\prime n}}{V(g_1, \Delta_n)_t V(g_2, \Delta_n)_t} \tag{17}$$

where

$$\hat{F}_t^n = \frac{2}{k_n \Delta_n} \sum_{i=1+k_n}^{[t/\Delta_n]-k_n-1} \sum_{m \in I_n(i)} ((r_i^1)^2 (r_m^2)^2 + (r_m^1)^2 (r_i^2)^2) \times 1_{\{|r_i| > \alpha\Delta_n^{\omega}, |r_m| \leq \alpha\Delta_n^{\omega}\}} \tag{18}$$

$$\hat{A}_t^{\prime n} = \frac{1}{\Delta_n} \sum_{i=1}^{t/\Delta_n} f(r_i) 1_{\{|r_i| \leq \alpha\Delta_n^{\omega}\}}. \tag{19}$$

The local window $k_n = 1/\sqrt{\Delta_n}$ and the truncation level of $\alpha\Delta_n^{\omega} = 0.03 \times \Delta_n^{0.49}$[3].

2.2.3. Univariate Jump Co-Exceedance Rule

Gilder et al. (2014) proposed a co-exceedance-based co-jump detection method by applying univariate jump tests to individual stocks to identify co-jumps. They selected three univariate jump tests developed by Barndorff-Nielsen and Shephard (2006), Lee and Mykland (2007) and Andersen et al. (2010), respectively. The idea was to find the intersection between ABD jump test results and BNS jump test results (ABD∩BNS), the intersection between ABD jump test results and LM jump test results (ABD∩LM), the intersection between BNS jump test results and LM jump test results (BNS∩LM), and the intersection among three jump tests results (ABD∩LM∩BNS), and retain only the intersection of the identified jumps to avoid spurious test results. Co-jumps are detected if more than one univariate process has jumps at the same time interval.

In the Lee and Mykland (2007) jump test, they used the ratio of realized returns to the estimated instantaneous volatility and constructed a nonparametric jump test to identify the exact timings of the jumps at the intraday level. The test statistics to identify the jump during interval $(t + l/n, t + (l+1)/n)$ is:

$$L_{(t+(l+1)/n)} = \frac{X_{t+(l+1)/n} - X_{t+l/n}}{\hat{\sigma}_{t+(l+1)/n}}, \tag{20}$$

where

$$\hat{\sigma}^2_{t+(l+1)/n} \equiv \frac{1}{K-2} \sum_{i=l-K+1}^{l-2} |X_{t+(i+1)/n} - X_{t+i/n}\, ||\, X_{t+i/n} - X_{t+(i-1)/n}|. \tag{21}$$

Here, K is the window size of a local movement of the process. It is also shown that:

$$\frac{\max_{l \in \bar{A}_n} |L_{(t+(l+1)/n)}| - C_n}{S_n} \to \varepsilon, \quad as \;\; \Delta_n \to 0, \tag{22}$$

where ε has a cumulative distribution function $P(\varepsilon \leq x) = \exp(-e^{-x})$,

$$C_n = \frac{(2\log n)^{1/2}}{c} - \frac{\log \pi + \log(\log n)}{2c(2\log n)^{1/2}} \quad \text{and} \quad S_n = \frac{1}{c(2\log n)^{1/2}} \tag{23}$$

c $\approx$ 0.7979 and $\bar{A}_n$ is the set of $l \in \{0, 1, \cdots, n\}$, so that there are no jumps in $(t + l/n, t + (l+1)/n]$.

Barndorff-Nielsen and Shephard (2006) developed a non-parametric (Hausman 1978) type test using the difference between the realized quadratic variation and realized bipower variation. Realized quadratic variation was estimated using realized volatility. The adjusted jump ratio test statistics can be defined as,

$$BNS = \frac{M^{1/2}}{\sqrt{\vartheta \max(1, \frac{QPV}{(\mu_1^2 BPV)^2})}}(1 - \frac{BPV}{RV}) \xrightarrow{d} N(0,1) \tag{24}$$

where RV and BPV are the same as in (3) and (4), $\vartheta = \pi^2/4 + \pi - 5 \approx 0.6090$. The realized quad-power variation QPV is used to estimate integrated quarticity ($\int_0^t \sigma_s^4 ds$) and can be given as:

$$QPV = \frac{1}{\Delta_n} \sum_{j=4}^{n} |\Delta_j X||\Delta_{j-1} X||\Delta_{j-2} X||\Delta_{j-3} X| \xrightarrow{P} \mu_1^4 \int_0^t \sigma_s^4 ds. \tag{25}$$

The authors show that the null hypothesis of no jumps is rejected if the test statistic BNS is in excess of the critical value Φ_α, where α is the test significance level, leading to a conclusion that there are jumps during the tested time interval.

3. Monte Carlo Experiments

The co-jump tests discussed above are widely used in the empirical literature to identify co-jumps among asset prices. In this section, we carry out a series of Monte Carlo experiments to evaluate the finite sample properties of the above co-jump tests: (i) BLT co-jump test; (ii) JT co-jump test; (iii) co-exceedance rule using the LM jump test; and (iv) the co-exceedance rule using the intersection of the LM test and BNS test. It is possible that both co-jumps and idiosyncratic jumps are present in the same period. Hence, it is important to check which test has the best performance in identifying co-jumps from idiosyncratic jumps.

We generate log prices for $M = 10$ assets using a Milstein discretion scheme from the following data-generating process (DGP):

$$dP_t^{(i)} = \mu dt + \sqrt{V_t^{(i)}} dW_{1,t}^{(i)} + dCJ_t^{(i)} + dIJ_t^{(i)}, \tag{26}$$

$$dV_t^{(i)} = \kappa_v(\theta_v - V_t^{(i)})dt + \zeta\sqrt{V_t^{(i)}} dW_{2,t}^{(i)}, \tag{27}$$

where $W_{1,t}^{(i)}$ and $W_{2,t}^{(i)}$ are two standard Brownian motions with correlation ρ, where $\rho = \{0, -0.5\}$. The co-jump and idiosyncratic jump are denoted by $CJ_t^{(i)}$ and $IJ_t^{(i)}$, respectively. The jump component $J_t^{(i)}$ (the sum of $CJ_t^{(i)}$ and $IJ_t^{(i)}$) are simulated as compound Poisson processes N_t with intensity λ_i:

$$J_t^{(i)} = \sum_{j=1}^{N_t} Y_j, \tag{28}$$

where Y_j's are *iid* normal, which characterizes the jump size. Under the null hypothesis, we first simulate DGPs without jumps as the first scenario. Then, we add idiosyncratic jumps $IJ_t^{(i)}$ onto the sample paths as the second scenario of the null hypothesis. Under the alternative, we add co-jumps $CJ_t^{(i)}$ to the null DGPs. When adding co-jumps to the null DGPs, we set the number of co-jumps $CJ_t^{(i)}$ among 10 assets equal to $m = 3$, which is in line with what we find in real data[4]. Table 1 shows the parameter settings in the Monte Carlo experiments. We follow the parameter settings from (Huang and Tauchen 2005; Aït-Sahalia et al. 2009; Corradi et al. 2018). The jump intensity λ is set to be either 0.1 or 0.2. The sampling frequency is $\Delta_n = \frac{1}{78}$, which is consistent with 5-min sampling frequencies. Table 2 shows details of the data generating processes. In all experiments, we performed 1000 Monte Carlo replications.

Table 1. Monte Carlo experiments—parameter settings *.

Jump Intensity	$\lambda = \{0.1, 0.2\}$
Leverage Effect	$\rho = \{0, -0.5\}$
Jump Distribution	$J_t \overset{i.i.d}{\sim} N(0, \sigma)$, where $\sigma = \{0.3, 0.4\}$
Sampling Frequency	$\Delta_n = \{\frac{1}{78}\}$
Other Parameters	$\mu, \kappa_v, \theta_v, \zeta, N = \{0.05, 5, 0.0144, 0.5, 10\}$

* Notes: Entries in this table show the parameter settings in Equations (26) and (27) in the Monte Carlo experiments. N is the number of simulated log prices of assets in the data generating process(es) (DGP). Moreover, $\frac{1}{78}$ indicates the 5-min sampling frequency. See Section 3 for complete details.

Table 2. Monte Carlo experiments—data generating process *.

	λ_{IJ}	λ_{CJ}	σ_{IJ}	σ_{CJ}	ρ	Δ_n	μ	κ_v	θ_v	ζ	N
DGP1					0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP2					−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP3	0.1		0.3		0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP4	0.2		0.3		0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP5	0.1		0.4		0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP6	0.2		0.4		0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP7	0.1		0.3		−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP8	0.2		0.3		−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP9	0.1		0.4		−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP10	0.2		0.4		−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP11		0.1		0.3	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP12		0.2		0.3	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP13		0.1		0.4	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP14		0.2		0.4	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP15		0.1		0.3	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP16		0.2		0.3	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP17		0.1		0.4	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP18		0.2		0.4	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP19	0.1	0.1	0.3	0.3	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP20	0.2	0.2	0.3	0.3	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP21	0.1	0.1	0.4	0.4	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP22	0.2	0.2	0.4	0.4	0	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP23	0.1	0.1	0.3	0.3	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP24	0.2	0.2	0.3	0.3	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP25	0.1	0.1	0.4	0.4	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10
DGP26	0.2	0.2	0.4	0.4	−0.5	$\frac{1}{78}$	0.05	5	0.0144	0.5	10

* Notes: DGPs 1–2 are continuous processes. DGPs 3–10 are continuous processes plus idiosyncratic jumps characterized by various jump size densities. DGPs 11–18 are continuous processes plus co-jumps characterized by various jump size densities. DGPs 19–26 are continuous processes plus co-jumps and idiosyncratic jumps characterized by various jump size densities. λ_{CJ} is the co-jump intensity and λ_{IJ} is the idiosyncratic jump intensity. ρ is the leverage effect. See Section 3 for complete details.

The co-exceedance rule using the LM jump test and the intersection of the LM test and BNS test are implemented using the following approach[5],

$$\sum_{i=1}^{M} 1\{Jump_s^{(i)} > 0\} \begin{cases} \geq 2, \text{co-jump} \\ \leq 1, \text{no co-jump} \end{cases} \tag{29}$$

where $1\{Jump_s^{(i)} > 0\}$ is an indicator function equals 1 when a jump is detected in asset i during the intraday interval that contains time s, and takes value 0 when no jump is detected.

Since the JT test is a bivariate co-jump test, we use the following rule to detect co-jumps:

$$\sum_{1 \leq i < j \leq M} 1\{Cojump_s^{(i,j)} > 0\} \begin{cases} \geq 1, \text{co-jump} \\ = 0, \text{no co-jump} \end{cases} \tag{30}$$

where $1\{Cojump_s^{(i,j)} > 0\}$ is an indicator function equals 1 when a co-jump is identified between asset i and asset j at time s. We use the test statistics $\Phi_n^{(d)}$ in the JT co-jump test to identify co-jumps.

The results are reported in Tables 3–6, where the significance level is 5%. Table 3 contains the empirical sizes of the co-jump tests when the underlying process is continuous (i.e., DGPs 1–2). At the 5% significance level, the range of the empirical size for the BLT test is (0.0687, 0.0718); for the JT test, it is (0.0143, 0.0220); for the co-exceedance rule of the LM test, it is (0.0426, 0.0435); and for the co-exceedance rule of the intersection of the LM and BNS tests, it is (0.0014,0.0010). Hence, the BLT test and co-exceedance rule of the JT test are close to the nominal sizes of all DGPs. The JT test is undersized and the co-exceedance rule of the LM–BNS test is far below the nominal size.

Table 3. Empirical sizes of co-jump tests under continuous processes *.

Case	BLT	JT	LM	LM–BNS
DGP1	0.0687	0.0220	0.0426	0.0014
DGP2	0.0718	0.0143	0.0435	0.0010

* Notes: Entries are rejection frequencies based on applications of the BLT co-jump test, JT co-jump tests, the co-exceedance rule of the LM test, and the co-exceedance rule of the LM–BNS test. Results for the 0.05 nominal size are reported. There are 1000 Monte Carlo replications. See Sections 2.2 and 3 for complete details.

Table 4 summarizes the empirical sizes of the co-jump tests when there are only idiosyncratic jumps (i.e., DGPs 3–10). Adding idiosyncratic jumps significantly increases the size of the JT test. For example, the size increases from 0.0143 to 0.1771 at a 5% significance level in DGPs 2 and 7, and these results with other DGPs are quite similar. For the BLT test, the co-exceedance rule of the LM test and the co-exceedance rule of the LM–BNS test, adding idiosyncratic jumps, slightly lower its size. Similarly, the BLT test and co-exceedance rule of the JT test are close to the nominal size of all DGPs. The JT test is significantly oversized and the co-exceedance rule of the LM–BNS test is far below the nominal size.

Table 4. Empirical sizes of co-jump tests under Monte Carlo simulations with idiosyncratic jumps *.

Case	BLT	JT	LM	LM–BNS
DGP3	0.0565	0.1892	0.0478	0.0015
DGP4	0.0506	0.3751	0.0520	0.0022
DGP5	0.0541	0.2486	0.0480	0.0017
DGP6	0.0476	0.4510	0.0523	0.0028
DGP7	0.0554	0.1771	0.0466	0.0019
DGP8	0.0480	0.3982	0.0512	0.0027
DGP9	0.0526	0.2178	0.0468	0.0022
DGP10	0.0452	0.4543	0.0515	0.0034

* Notes: see notes in Table 3.

Table 5 shows the empirical power of the co-jump tests when there are no idiosyncratic jumps. The JT test, the BLT test, and the co-exceedance rule of the LM test have significantly larger powers compared to the co-exceedance rule of the intersection of the LM–BNS test. As noted in the table, at the 5% significance level, the range of the empirical power of the BLT test is (0.9167, 0.9362), and the JT test is (0.9451, 0.9891), while the co-exceedance rule of the LM test is (0.9362, 0.9722), and the intersection of the LM–BNS test is (0.1702, 0.3409). Jump intensity and jump size have positive effects on power in most cases. For example, in DGPs 11 and 13, as the co-jump size σ_{CJ} increases from 0.3 to 0.4, the empirical power of the BLT co-jump test increases from 0.9255 to 0.9362, the co-exceedance rule of the LM test increases from 0.9362 to 0.9521, the co-exceedance rule of the intersection of the LM–BNS test increases from 0.1702 to 0.2606, and the JT test increases from 0.9451 to 0.9670.

Table 5. Empirical power of co-jump tests under Monte Carlo simulations with co-jumps *.

Case	BLT	JT	LM	LM–BNS
DGP11	0.9255	0.9451	0.9362	0.1702
DGP12	0.9167	0.9836	0.9495	0.2121
DGP13	0.9362	0.9670	0.9521	0.2606
DGP14	0.9242	0.9891	0.9697	0.3409
DGP15	0.9202	0.9451	0.9415	0.2340
DGP16	0.9242	0.9836	0.9672	0.2020
DGP17	0.9309	0.9560	0.9628	0.3085
DGP18	0.9293	0.9891	0.9722	0.3409

* Notes: see notes in Table 3.

Table 6 summarizes the empirical power of the co-jump tests when there are also idiosyncratic jumps. Similarly, the JT co-jump test, the BLT test, and the co-exceedance rule of the LM test have significantly larger powers than the co-exceedance rule of the intersection of the LM and BNS tests. It is also worth noting that adding idiosyncratic jumps lowers the empirical power of the co-jump tests, particularly for the BLT test and the co-exceedance rule of the LM test. The jump size impacts power positively for all co-jump tests. The jump intensity does not impact the power unanimously. For example, the jump intensity has negative effects on power in the BLT test, LM test, and LM–BNS, but it has positive effects on the power of the JT test. We have different results on the effects of jump intensity because when jump intensity increases, it increases both the co-jump intensity and idiosyncratic jump intensity.

Table 6. Empirical power of co-jump tests under Monte Carlo simulations with co-jumps and idiosyncratic jumps *.

Case	BLT	JT	LM	LM–BNS
DGP19	0.9126	0.9394	0.9029	0.2087
DGP20	0.8456	0.9798	0.9332	0.2166
DGP21	0.9175	0.9697	0.9126	0.3155
DGP22	0.8687	0.9798	0.9516	0.3433
DGP23	0.8932	0.9596	0.9223	0.2524
DGP24	0.8687	0.9747	0.9424	0.2097
DGP25	0.9126	0.9697	0.9515	0.3350
DGP26	0.8825	0.9798	0.9516	0.3295

* Notes: See notes in Table 3.

Overall, the BLT test, the JT test, and the co-exceedance rule of the LM test all have good testing power, except for the co-exceedance rule of the LM–BNS test. Additionally, the BLT test and the co-exceedance rule of the LM test have empirical sizes close to the nominal sizes in all DGPs, while the JT test is mixed, under- or over-sized in different DGPs, and the co-exceedance rule of the LM–BNS test is significantly undersized in all DGPs. The co-exceedance rule of the LM test has good testing power and is more conservative under the null. Hence, this test is very unlikely to yield the false discovery of co-jumps. Therefore, we used this test in the following.

4. Forecasting Methods

4.1. Forecasting Models

In this section, we summarize all models we used in predicting sector-level equity volatility. With the recent availability of high-frequency financial data, the HAR-type models are now widely used in the literature for forecasting the time-varying return variance. The HAR-RV model, first proposed by Corsi (2009), is based on the idea of the HARCH model and the heterogeneous market hypothesis developed by Müller et al. (1997), in which the authors believe that different types of traders react to the market on different horizons and, consequently, affect the market differently. So the HAR-RV model incorporates different kinds of market volatilities over different time horizons to forecast volatility. One simple version of the HAR-RV model is:

$$RV_{t+h} = \beta_0 + \beta_d RV_t + \beta_w RV_{t-5,t} + \beta_m RV_{t-22,t} + \epsilon_{t+h}, \tag{31}$$

where RV_{t+h} is the realized volatility defined in Equation (3) at day $t+h$. The multi-period normalized realized measures for realized volatility $RV_{t-h,t}$ are defined as the average of the corresponding one-period measures. Namely,

$$RV_{t-h,t} = h^{-1}[RV_{t-h+1} + RV_{t-h+2} + \cdots + RV_t]. \tag{32}$$

Built on the HAR-RV model, Andersen et al. (2007) further separated jumps from the continuous part and developed a new HAR-RV-CJ prediction model that contained the jump component, the continuous component, and the lag terms over different time horizons. More specifically, the HAR-RV-CJ model is given by:

$$RV_{t,t+h} = \beta_0 + \beta_{cd} C_t + \beta_{cw} C_{t-5,t} + \beta_{cm} C_{t-22,t} + \beta_{jd} J_t + \beta_{jw} J_{t-5,t} + \beta_{jm} J_{t-22,t} + \epsilon_{t+h}, \tag{33}$$

where C_t is the variation of the continuous component on day t estimated by BPV defined in Equation (4) and J_t is the corresponding jump variation estimated by

$$J_t = RV_t - BPV_t \tag{34}$$

Since the sampling interval in the empirical analysis is finite ($\Delta_n > 0$), the above-defined J_t can be negative in some cases. Andersen et al. (2007) suggested excluding the negative values of the estimates and truncating the jump measurements at 0. The jump component variation is calculated as:

$$J_t = \max[RV_t - BPV_t, 0] \tag{35}$$

The regressors $J_{t-h,t}$ and $C_{t-h,t}$ are defined in the same way as $RV_{t-h,t}$ above.

In the sequel, we estimate various HAR-RV-J models that incorporate co-jumps (CJ_t) and idiosyncratic jumps (IJ_t) and analyze the predictive content for co-jumps and idiosyncratic jumps in the forecasting volatility.

Specification 1: HAR-RV-C model (benchmark model):

$$\log RV_{t+1} = \beta_0 + \beta_{cd} \log C_t + \beta_{cw} \log C_{t-5,t} + \beta_{cm} \log C_{t-22,t} + \epsilon_{t+1}. \tag{36}$$

Specification 2: HAR-RV-C-CJ model with co-jumps (CJ_t):

$$\begin{aligned} \log RV_{t+1} = \beta_0 &+ \beta_{cd} \log C_t + \beta_{cw} \log C_{t-5,t} + \beta_{cm} \log C_{t-22,t} + \beta_{cjd} \log CJ_t \\ &+ \beta_{cjw} \log CJ_{t-5,t} + \beta_{cjm} \log CJ_{t-22,t} + \epsilon_{t+1}. \end{aligned} \tag{37}$$

Specification 3: HAR-RV-C-IJ model with idiosyncratic jumps (IJ_t):

$$\begin{aligned} \log RV_{t+1} = \beta_0 &+ \beta_{cd} \log C_t + \beta_{cw} \log C_{t-5,t} + \beta_{cm} \log C_{t-22,t} + \beta_{ijd} \log IJ_t \\ &+ \beta_{ijw} \log IJ_{t-5,t} + \beta_{ijm} \log IJ_{t-22,t} + \epsilon_{t+1}. \end{aligned} \tag{38}$$

Specification 4: HAR-RV-C-CJ-IJ model with co-jumps (CJ_t) and idiosyncratic jumps (IJ_t):

$$\begin{aligned} \log RV_{t+1} = \beta_0 &+ \beta_{cd} \log C_t + \beta_{cw} \log C_{t-5,t} + \beta_{cm} \log C_{t-22,t} + \beta_{cjd} \log CJ_t \\ &+ \beta_{cjw} \log CJ_{t-5,t} + \beta_{cjm} \log CJ_{t-22,t} + \beta_{ijd} \log IJ_t \\ &+ \beta_{ijw} \log IJ_{t-5,t} + \beta_{ijm} \log IJ_{t-22,t} + \epsilon_{t+1} \end{aligned} \tag{39}$$

The regressors include the daily, weekly, and monthly lag terms. All of the model specifications are summarized in Table 7.

Table 7. Brief descriptions of forecasting models *.

Model	Description
HAR-RV-C	Future RV depends on the lags of the continuous component of the RV.
HAR-RV-C-CJ	Future RV depends on the lags of the continuous component of RV and the co-jump variation.
HAR-RV-C-IJ	Future RV depends on the lags of the continuous component of the RV and the idiosyncratic jump variation.
HAR-RV-C-CJ-IJ	Future RV depends on the lags of the continuous component of RV, co-jump variation, and idiosyncratic jump variation.

* Notes: This table reports the models in the forecasting experiments. The benchmark model is the HAR-RV-C model. For complete details, refer to Section 4.

4.2. Co-Jump and Idiosyncratic Jump Variations

The co-exceedance rule of the LM test is utilized to identify co-jumps and idiosyncratic jumps among M assets, as defined in (29). The co-jump variation is measured using the power transformation of the instantaneous return:[6]

$$CJ_t = \sum_{s=1}^{n} r_s^2 * I_s^{CJ} \tag{40}$$

where I_s^{CJ} is a co-jump indicator for the target asset at intraday interval s on day t and r_s is the return over the same interval. The number of observations per day is $n = 78$. The idiosyncratic jump variation is

$$IJ_t = \sum_{s=1}^{n} r_s^2 * I_s^{IJ} \tag{41}$$

where I_s^{IJ} is an idiosyncratic jump indicator for the target asset at the intraday interval s on day t.

The ratios of co-jump and idiosyncratic jump contributions to the total variations are calculated as:

$$ratio^{CJ} = \frac{CJ_t}{RV_t} \quad \text{and} \quad ratio^{IJ} = \frac{IJ_t}{RV_t}. \tag{42}$$

The proportions of co-jumps and idiosyncratic jumps to the total number of jumps are defined as:

$$Porp^{CJ} = \frac{N^{CJ}}{N^{CJ} + N^{IJ}} \quad \text{and} \quad Porp^{IJ} = \frac{N^{IJ}}{N^{CJ} + N^{IJ}}, \tag{43}$$

where N^{CJ} is the number of co-jumps and N^{IJ} is the number of idiosyncratic jumps[7]. The frequencies of co-jumps and idiosyncratic jumps are defined as the percentage of days identified as having co-jumps and idiosyncratic jumps, respectively:

$$Frequency^{CJ} = \frac{1}{T}\sum_{i=1}^{T} 1_{\{\sum_{s=1}^{M} I_s^{CJ}\}>0} \quad \text{and} \quad Frequency^{IJ} = \frac{1}{T}\sum_{i=1}^{T} 1_{\{\sum_{s=1}^{M} I_s^{IJ}\}>0}, \tag{44}$$

where T is the total number of days in the sample.

5. Setup of the Forecasting Experiment

For the forecasting experiment, we carried out both an in-sample regression analysis and an out-of-sample forecast analysis. We used a rolling window estimation scheme, with window size $T = 450$ (i.e., approximately two years). We obtained sequences of daily out-of-sample volatility forecasts for the sample period from 13 November 2008 to 31 December 2019, which contained 2801 trading days. The in-sample regression analysis was also conducted using the same sample periods.

We used in-sample adjusted R^2, the mean absolute forecasting error (MAFE), and out-of-sample R^2 to evaluate forecasting results among candidate models.

The out-of-sample R^2 is defined as:

$$\text{out-of-sample } R^2 = 1 - \frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{T}(y_t - \bar{y}_t)^2} \tag{45}$$

where $\bar{y}_t$ is the historical average of volatility.

Moreover, we implemented the model confidence set (MCS) test of Hansen et al. (2011) to compare the forecasting performance amongst candidate models, in which a large p-value indicates a superior performance. The MCS test has a sequence of statistic tests, which are utilized to determine the "Superior Set Models" (SSM)[8] from a collection of models, where the null hypothesis of equal predictive ability among candidate models is not rejected at a certain confidence level. Let $d_{jh,i}$ denote the loss difference between the forecasting model j and h:

$$d_{jh,i} = l(\epsilon_i^j) - l(\epsilon_i^h), \quad j,h = 1,\cdots,Q,\ i = 1,\cdots,n, \tag{46}$$

where ϵ_i^j is the prediction error in model j at time span i, ϵ_i^h is the prediction error in the other model h, and $l()$ represents the quadratic loss function. The null hypothesis of the equal predictive ability for a given set of models Q can be given as:

$$H_{0,Q} : c_{j,h} = 0, \quad \text{for all} \quad j,h = 1,2,\cdots,Q, \tag{47}$$

$$H_{1,Q} : c_{j,h} \neq 0, \quad \text{for all} \quad j,h = 1,2,\cdots,Q, \tag{48}$$

where $c_{j,h} = E(d_{j,h})$ is finite- and time-independent. The first test statistic is constructed as:

$$T_{j,h} = \frac{\bar{d}_{j,h}}{\sqrt{\widehat{var}(\bar{d}_{j,h})}} \tag{49}$$

where $\bar{d}_{j,h} = n^{-1}\sum_{i=1}^{n} d_{jh,i}$ and $\widehat{var}(\bar{d}_{j,h})$ is the bootstrapped estimate of $var(\bar{d}_{j,h})$. We performed a stationary bootstrap procedure of 2000 resamples. In Hansen et al. (2011), the first test statistics maps naturally into the following test statistic:

$$T_{R,Q} = \max_{j,h \in Q} \mid T_{j,h} \mid \tag{50}$$

where $T_{j,h}$ is calculated using Equation (49). Since the asymptotic distribution of the test statistic is nonstandard, the relevant distribution can be estimated with bootstrap methods. We calculate the p-values for the candidate forecasting models. The MCS test used to obtain the SSM can be summarized as the following step:

1. Set $Q = Q_0$.
2. Test for the null hypothesis of equal predictive ability among candidate models in Q with significance level α. If the null hypothesis cannot be rejected, then the set SSM is the same as the set of all models. Otherwise, determine the worst model in the set Q.
3. Remove the worst model from Q, and proceed to step 2.

The choice of the worst model to be eliminated in step 2 is determined using an elimination rule that is consistent with the test statistics defined in Equation (49), which are

$$e_{R,Q} = \arg\max_{j} \left\{ \sup_{j \in Q} \frac{\bar{d}_{j,h}}{\sqrt{\widehat{var}(\bar{d}_{j,h})}} \right\} \tag{51}$$

6. Data Description

We collected data for nine SPDR exchange-traded funds (ETFs), which were designed to track nine sectors of the S&P 500 index, including the material sector (XLB), energy sector (XLE), financial sector (XLF), industrial sector (XLI), technology sector (XLK), consumer staples sector (XLP), utility sector (XLU), healthcare sector (XLV), and consumer discretionary sector (XLY). The sources of nine sector ETF intraday observations were from the trade and quote database (TAQ) through the Wharton research data service (WRDS). Table 8 shows the details of the forecasting targets. The full sample period was from 3 January 2006 to 31 December 2019 for a total of 3523 trading days. We excluded overnight returns from our sample. The previous tick method derived from Gençay et al. (2001) was used to filter out price data. Raw data were cleaned and outliers were removed based on the procedures in Brownlees and Gallo (2006). We used the 5-min sampling frequency to reduce the effect of microstructure noise, yielding 78 observations per day in most cases.

Table 8. Target variables in the forecasting experiment *.

Target Name	Description	Transformation	Frequency
XLF	Financial Sector SPDR Fund	$ln(X_t) - ln(X_{t-1})$	Daily
XLK	Technology Sector SPDR Fund	$ln(X_t) - ln(X_{t-1})$	Daily
XLP	Consumer Staples Select Sector SPDR Fund	$ln(X_t) - ln(X_{t-1})$	Daily
XLU	Utilities Select Sector SPDR Fund	$ln(X_t) - ln(X_{t-1})$	Daily
XLY	Consumer Discretionary Select Sector SPDR Fund	$ln(X_t) - ln(X_{t-1})$	Daily

* Notes: This table reports the prediction targets. Data transformations used in the forecasting experiments are presented in the third column of the table. See Section 6 for further details.

7. Empirical Findings

In this section, we discuss the main empirical findings. We begin by discussing the results from identifying co-jumps and idiosyncratic jumps among nine sector ETFs in the S&P 500 market. The first sense of the prevalence of co-jumps and idiosyncratic jumps can be formulated by inspecting Table 9 and Figures 1–3. Then we turn to the one-day-ahead predictive performance of the forecasting models outlined in Section 4. We report the model confidence set results (Table 10), in-sample adjusted R^2 results (Table 11), out-of-sample R^2 results (Table 12), and additionally, the mean absolute forecasting error (MAFE) results (Table 13). Our main empirical findings are summarized as follows.

First, both co-jumps and idiosyncratic jumps occurred frequently from January 2007 to December 2019, especially during significant market turmoil, such as the 2008 financial crisis and the 2011 European debt crisis. Table 9 shows the summary statistics for co-jumps and idiosyncratic jumps. While both types of jumps occurred frequently from January 2008 to December 2019, the frequency of co-jumps was much higher. Moreover, the descriptive statistics also suggested that the magnitude (including both the upside jump and downside jump) of the co-jumps was typically larger than that of the idiosyncratic jumps. All these results are in line with the fact that there were two major crises in the sample period, during which the systematic jump risk would be more prevalent than the idiosyncratic jump risk.

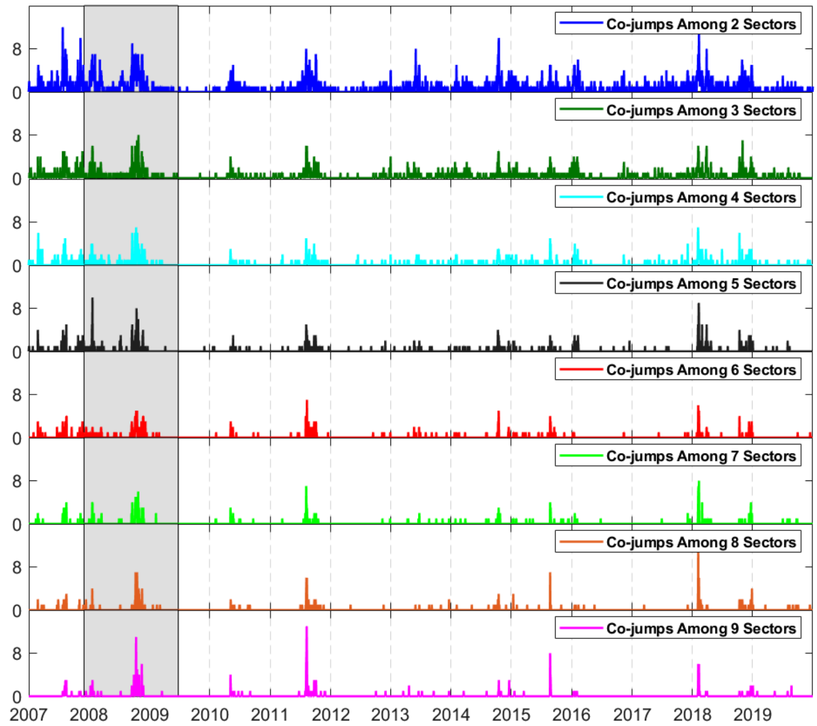

Figure 1. Number of co-jumps under different scenarios: Figure 1 shows the number of co-jumps identified under different scenarios, i.e., across nine sectors in the S&P 500 market from January 2007 to December 2019. The nine sectors are materials, energy, financial, industrial, technology, consumer staples, utility, healthcare, and consumer discretionary. The co-exceedance rule of the LM test is utilized to detect co-jumps among sectors. The shaded area in the figure denotes the financial crisis period from December 2007 to June 2009. See Sections 4 and 7 for further details.

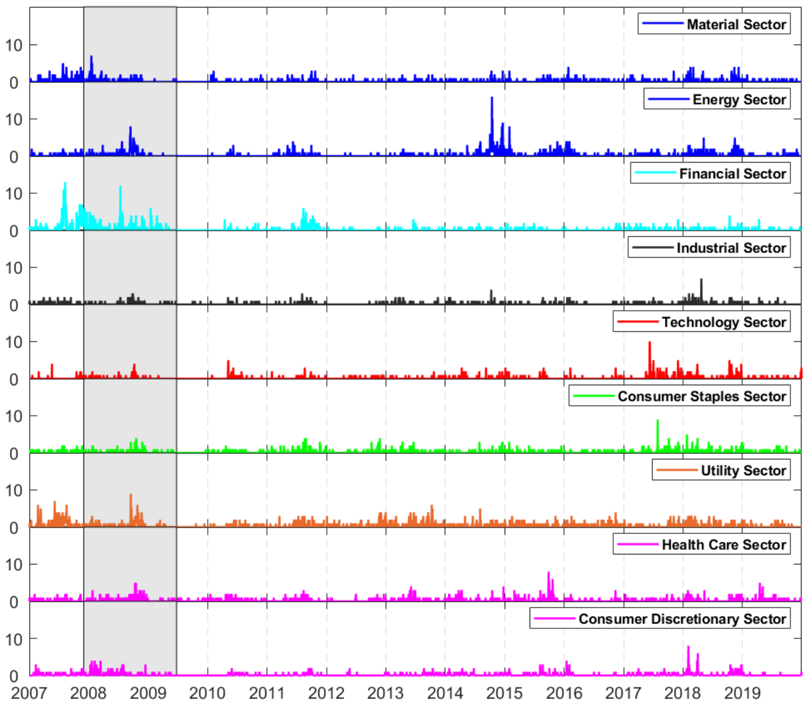

Figure 2. Number of sector-level idiosyncratic jumps: see notes in Figure 1. Figure 2 shows the number of idiosyncratic jumps in nine sectors in the S&P 500 market from January 2007 to December 2019. See Sections 4 and 7 for further details.

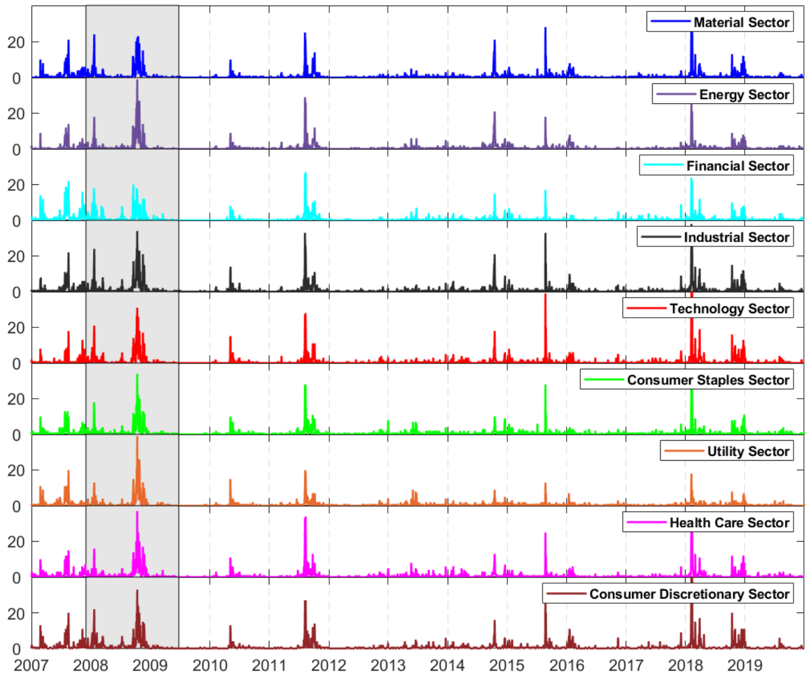

Figure 3. Number of sector-level co-jumps: see notes for Figure 1. Figure 3 shows the number of co-jumps in nine sectors in the S&P 500 market from January 2007 to December 2019. See Sections 4 and 7 for further details.

Table 9. Descriptive statistics for co-jumps and idiosyncratic jumps (2007:1–2019:12) *.

	Sector	XLB	XLE	XLF	XLI	XLK	XLP	XLU	XLV	XLY
Co-jump	Frequency (%)	30.10%	24.45%	25.31%	30.78%	24.48%	32.73%	29.55%	29.49%	29.03%
	Prop. (%)	79.36%	70.08%	72.79%	87.46%	83.52%	79.76%	64.37%	78.23%	86.09%
	upside jump mean (%)	0.63%	0.78%	0.72%	0.56%	0.57%	0.43%	0.61%	0.49%	0.55%
	downside jump mean (%)	−0.61%	−0.76%	−0.70%	−0.54%	−0.53%	−0.43%	−0.60%	−0.47%	−0.54%
	St. dev.	0.0074	0.0090	0.0090	0.0064	0.0064	0.0052	0.0072	0.0055	0.0067
	Ratio of JV to RV (%)	27.96%	22.71%	23.50%	28.58%	22.74%	30.40%	27.45%	27.39%	26.97%
Idiosyncratic jump	Frequency (%)	15.01%	17.42%	14.82%	10.39%	10.36%	15.19%	24.05%	16.56%	10.36%
	Prop. (%)	20.64%	29.92%	27.21%	12.54%	16.48%	20.24%	35.63%	21.77%	13.91%
	upside jump mean (%)	0.50%	0.59%	0.67%	0.42%	0.38%	0.31%	0.43%	0.39%	0.40%
	downside jump mean (%)	−0.51%	−0.59%	−0.62%	−0.43%	−0.39%	−0.31%	−0.45%	−0.40%	−0.38%
	St. dev.	0.0056	0.0064	0.0077	0.0048	0.0042	0.0034	0.0047	0.0043	0.0043
	Ratio of JV to RV (%)	13.94%	16.18%	13.77%	9.65%	9.62%	14.11%	22.34%	15.38%	9.62%

* Notes: This table shows the descriptive statistics for co-jumps and idiosyncratic jumps identified from nine sector ETFs. The frequency denotes the percentage of days having co-jumps or idiosyncratic jumps. JV is the jump variation and RV is the total variation calculated by realized volatility. XLB is the material select sector ETF; XLE is the energy select sector ETF; XLF is the financial select sector EFT; XLI is the industrial select sector EFT; XLK is the technology select sector EFT; XLP is the consumer staples select sector EFT; XLU is the utility sector ETF; XLV is the healthcare sector ETF, and XLY is the consumer discretionary sector ETF. See Sections 5 and 7 for further details.

Table 10. Model confidence set test *.

Sector	HAR-RV-C	HAR-RV-C-CJ	HAR-RV-C-IJ	HAR-RV-C-CJ-IJ
	Forecasting Period: 2008:11–2019:12			
XLF	0.6560 *	1.0000 *	0.6560 *	0.6560 *
XLK	0.2165 *	0.4195 *	0.1415 *	1.0000 *
XLP	0.1220 *	0.1220 *	0.0880	1.0000 *
XLU	0.3860 *	0.3860 *	0.2975 *	1.0000 *
XLY	0.7780 *	0.9385 *	0.7780 *	1.0000 *

* Notes: See notes in Table 9. This table reports the model confidence test (MCS) p-values associated with the MCS test discussed in Section 5. Moreover, the 90% model confidence set members are identified by models with asterisks.

Table 11. In-sample adjusted R^2 *.

Sector	HAR-RV-C	HAR-RV-C-CJ	HAR-RV-C-IJ	HAR-RV-C-CJ-IJ
	Forecasting Period: 2008:11–2019:12			
XLF	0.7761	0.7765	0.7761	0.7768
XLK	0.6748	0.6758	0.6758	0.6763
XLP	0.6430	0.6441	0.6439	0.6443
XLU	0.6157	0.6189	0.6175	0.6200
XLY	0.7458	0.7470	0.7462	0.7470

* Notes: See notes in Table 10. Entries in this table include in-sample adjusted R^2. All results are from forecasting experiments that utilized 5-min frequency data. See Sections 5 and 7 for further details.

Table 12. Out-of-sample R^2 *.

Sector	HAR-RV-C	HAR-RV-C-CJ	HAR-RV-C-IJ	HAR-RV-C-CJ-IJ
	Forecasting Period: 2008:11–2019:12			
XLF	0.6765	0.6866	0.6823	0.6823
XLK	0.5680	0.5772	0.5649	0.5817
XLP	0.5567	0.5602	0.5556	0.5696
XLU	0.4927	0.4970	0.4914	0.5047
XLY	0.6804	0.6838	0.6802	0.6841

* Notes: See notes in Table 10. This table reports the out-of-sample R^2 for all forecasting experiments.

Table 13. Mean absolute forecasting error *.

Sector	HAR-RV-C	HAR-RV-C-CJ	HAR-RV-C-IJ	HAR-RV-C-CJ-IJ
	Forecasting Period: 2008:11–2019:12			
XLF	0.2293	0.2251	0.2276	0.2267
XLK	0.2371	0.2348	0.2360	0.2320
XLP	0.2062	0.2056	0.2063	0.2031
XLU	0.2070	0.2061	0.2069	0.2042
XLY	0.2225	0.2211	0.2214	0.2205

* Notes: See notes in Table 10. This table reports the mean absolute forecasting errors (MAFEs) for all forecasting experiments.

Second, co-jumps occurred more frequently than idiosyncratic jumps. Figure 1 shows the number of co-jumps under different scenarios, i.e., across a different number of sectors from January 2007 to December 2019. Not surprisingly, co-jumps detected under scenarios with fewer sectors occurred more frequently than co-jumps identified under scenarios with a larger number of sectors. Figures 2 and 3 depict the numbers of co-jumps and idiosyncratic jumps, respectively. The results in Figures 2 and 3 indicate that co-jumps and idiosyncratic jumps were more densely populated during significant market turmoils, including the 2008 financial crisis, the 2011 European debt crisis, the 2015–2016 stock market selloff, and the stock market dip in December 2018–January 2019. This suggested higher levels of systematic jump risk and idiosyncratic jump risk during those market downturns.

Moreover, as expected, the financial sector had more idiosyncratic jumps during the 2008 financial crisis compared with other sectors, which is consistent with the fact that this crisis was mainly triggered by the collapse of the financial sector.

Third, co-jumps and idiosyncratic jumps demonstrate marginal predictive content for volatility prediction, and this result was robust across different forecasting targets. The results in the model confidence set test (Table 10), in-sample adjusted R^2 (Table 11), out-of-sample R^2 (Table 12), and mean absolute forecasting error (Table 13) indicate that the HAR-RV-C-CJ-IJ forecasting model, which contains co-jumps and idiosyncratic jumps variations, outperforms other forecasting models (HAR-RV-C, HAR-RV-C-CJ, and HAR-RV-C-IJ) in most scenarios. Additionally, when comparing the HAR-RV-C benchmark model with HAR-RV-C-CJ, it is clear that adding co-jump variations into the model reduces the forecasting error and improves in-sample adjusted R^2 and out-of-sample R^2 in most cases. This suggests both systematic risk and idiosyncratic risk are useful in predicting volatility.

Fourth, the predictive content of co-jumps is higher than that of idiosyncratic jumps. Evidently, the model with co-jump variations (HAR-RV-C-CJ) represents the majority of entries with smaller MAFEs and larger out-of-sample R^2 and in-sample R^2 than the model with idiosyncratic jump variations (HAR-RV-C-IJ). For example, in Table 13, the HAR-RV-C-CJ model yields the lowest MAFEs for all sectors, and these MAFEs are comparatively lower than models using idiosyncratic jump information[9]. The same finding holds with the model confidence set testing results in Table 10. Comparing the HAR-RV-C-IJ model with idiosyncratic jump variations, the HAR-RV-C-CJ model with co-jump variations has larger p-values in all sectors. In Tables 11 and 12, the HAR-RV-C-CJ model has higher in-sample R^2 and out-of-sample R^2 compared with the HAR-RV-C-IJ model. All these results suggest that systematic risk is more useful than idiosyncratic risk in volatility forecasting.

8. Policy Implications

In this paper, we distinguished jumps in sector-level ETF prices into idiosyncratic jumps, which represent large-scale price movements that occur mostly within one sector, and co-jumps, which are defined as large-scale price movements that take place across more than one sector around the same time. Specifically, idiosyncratic jumps reflect the influence of events that are mostly confined within one sector, while co-jumps, by the nature of price co-movements among more than one sector at the same time, capture the impact of events that can spill over to other sectors, leading to financial risk contagion.

In our empirical experiment, the augmentation of co-jumps leads to improvements in the forecasting of realized volatility for each of the five S&P 500 sectors. Co-jumps also demonstrate higher predictive power compared to idiosyncratic jumps and achieve even better predictive accuracy when combined with idiosyncratic jumps. These results put co-jumps as a key source of volatility for S&P 500 sectors, aside from within-sector risks or events that are captured by idiosyncratic jumps. Spillovers from other sectors account for a significant portion of volatility for each individual market sector, which should be adequately noticed by researchers and policymakers. Moreover, we also found that co-jumps occurred two to three times more frequently than idiosyncratic jumps in our sample period from 2007 to 2019, suggesting within-sector risks or events are more likely to spill over to other sectors.

For policymakers, it is crucial to be vigilant to within-sector risks or events to prevent systemic risks to the financial market, given the outsized effects on other sectors and the frequency of such scenarios. Significant risks or events of one sector could spill over to other sectors and cause increased volatility, resulting in contagion among different sectors.

9. Concluding Remarks

In this paper, we examined the usefulness of co-jump and idiosyncratic jump variation measures in forecasting the volatility of sector ETFs. The co-jumps are meant to represent the systematic risks across multiple assets, sectors, and markets, while the idiosyncratic jumps capture the idiosyncratic risks in an individual sector. In particular, using 5-min return data, we utilized extant jump and co-jump tests to disentangle individual sector idiosyncratic jumps and multiple sector co-jumps from January 2007 to December 2019. Perhaps not surprisingly, we found that co-jumps and idiosyncratic jumps occurred more frequently during the financial market turmoils, such as the 2008 financial crisis and the 2011 European debt crisis, affirming the severity of financial stress during the market turmoil. Moreover, we found that co-jumps have larger magnitudes and variations compared with idiosyncratic jumps, which suggests that systematic risk has a larger impact than idiosyncratic risk on the financial market. Additionally, we showed (via a series of forecasting experiments that utilized heterogeneous autoregressive (HAR) models) that volatility forecasting accuracy can be improved when the original HAR forecasting models are augmented by including co-jump and idiosyncratic jump variation measures. We also find that the predictive content of co-jumps is higher than that of idiosyncratic jumps, suggesting that systematic risk is more useful than idiosyncratic risk in volatility forecasting.

In addition, we carried out Monte Carlo simulations to examine the relative performances of four existing co-jump tests, including the BLT co-jump test by Bollerslev et al. (2008), the JT co-jump test by Jacod and Todorov (2009), the co-exceedance rule by Gilder et al. (2014) using the LM test by Lee and Mykland (2007), and the co-exceedance rule using the intersection of the LM test and BNS tests by Barndorff-Nielsen and Shephard (2006). We find that the BLT test and the co-exceedance rule of the LM test outperform the JT co-jump test and co-exceedance rule of the intersection of the LM and BNS tests. The BLT test, the JT test, and the co-exceedance rule of the LM test all have good testing power, except for the co-exceedance rule of the LM–BNS test. The BLT test and the co-exceedance rule of the LM test have empirical sizes close to the nominal sizes in all DGPs, while the JT test is mixed, under- or over-sized in different DGPs, and the co-exceedance rule of the LM–BNS test is significantly undersized.

Author Contributions: Conceptualization, W.P. and C.Y.; methodology, W.P. and C.Y.; software, W.P. and C.Y.; validation, W.P. and C.Y.; formal analysis, W.P. and C.Y.; investigation, W.P. and C.Y.; resources, W.P. and C.Y.; data curation, W.P. and C.Y.; writing—original draft preparation, W.P. and C.Y.; writing—review and editing, W.P. and C.Y.; visualization, W.P. and C.Y.; supervision, W.P. and C.Y.; project administration, W.P. and C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors are grateful to Rodger Klein, Yuan Liao, Norman Swanson, and Xiye Yang for the useful comments and suggestions on the topics explored in this paper. We also thank the participants of the Southwestern Finance Association Annual Meeting, the Midwest Economic Association Annual Meeting, and workshops at Rutgers University for the helpful comments.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. For literature in this area, see (Barndorff-Nielsen and Shephard 2006; Lee and Mykland 2007; Jiang and Oomen 2008; Aït-Sahalia et al. 2009; Huang and Tauchen 2005; Mancini 2009; Podolskij and Ziggel 2010; Corradi et al. 2018; Boswijk et al. 2018; Mukherjee et al. 2020, and the references cited therein).
2. Bandi and Russell (2008) show in the presence of market microstructure noise, realized variance does not identify daily integrated variance of the frictionless equilibrium. They propose a more general treatment of the effect of market microstructure noise on realized variance estimates.
3. We follow the parameter settings from Jacod and Todorov (2009).
4. Among 10 assets, co-jumps occur more frequently among a small portion of all assets (e.g., $m = 2, 3, 4$) than a large portion of all assets (e.g., $m = 5, 6, 7, 8, 9, 10$) at each time.
5. Gilder et al. (2014) also implement the same approach.

[6] One approach to measure jump variation is to use the difference between realized volatility and bipower variation, as shown in Equations (34) and (35). Another approach to examine the jump power variation is formed using power transformation of the instantaneous return, i.e., $|\ r_s\ |^q$. Key papers in this area include (Ding et al. 1993; Ding and Granger 1996; Todorov and Tauchen 2010; Barndorff-Nielsen et al. 2008, and the references cited there in).

[7] The co-exceedance rule of the LM test is utilized to identify co-jumps and idiosyncratic jumps among assets.

[8] The set of models that consists of the best model(s).

[9] Overall, the mean absolute forecasting errors in Table 13 are comparatively smaller with respect to results in the relevant literature (e.g., see Qiu et al. 2019).

References

Aït-Sahalia, Yacine, and Jean Jacod. 2009. Testing for jumps in a discretely observed process. *The Annals of Statistics* 37: 184–222. [CrossRef]

Andersen, Torben G., and Jesper Lund. 1997. Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics* 77: 343–77. [CrossRef]

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold. 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics* 89: 701–20. [CrossRef]

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2001. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96: 42–55. [CrossRef]

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2003. Modeling and forecasting realized volatility. *Econometrica* 71: 579–625. [CrossRef]

Andersen, Torben G., Tim Bollerslev, Per Frederiksen, and Morten Ørregaard Nielsen. 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. *Journal of Applied Econometrics* 25: 233–61. [CrossRef]

Bandi, Federico M., and Jeffrey R. Russell. 2008. Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies* 75: 339–69. [CrossRef]

Bandi, Federico M., and Roberto Reno. 2016. Price and volatility co-jumps. *Journal of Financial Economics* 119: 107–46. [CrossRef]

Barndorff-Nielsen, Ole E., and Neil Shephard. 2004. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2: 1–37. [CrossRef]

Barndorff-Nielsen, Ole E., and Neil Shephard. 2006. Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4: 1–30. [CrossRef]

Barndorff-Nielsen, Ole E., Silja Kinnebrock, and Neil Shephard. 2008. Measuring Downside Risk-Realised Semivariance. Available online: https://ssrn.com/abstract=1262194 (accessed on 29 June 2022).

Bibinger, Markus, and Lars Winkelmann. 2015. Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics* 184: 361–78. [CrossRef]

Bollerslev, Tim, Tzuo Hann Law, and George Tauchen. 2008. Risk, jumps, and diversification. *Journal of Econometrics* 144: 234–56. [CrossRef]

Boswijk, H. Peter, Roger J. A. Laeven, and Xiye Yang. 2018. Testing for self-excitation in jumps. *Journal of Econometrics* 203: 256–66. [CrossRef]

Brandt, Michael W., and Christopher S. Jones. 2006. Volatility forecasting with range-based EGARCH models. *Journal of Business & Economic Statistics* 24: 470–86.

Brownlees, Christian T., and Giampiero M. Gallo. 2006. Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis* 51: 2232–45.

Caporin, Massimiliano, Aleksey Kolokolov, and Roberto Renò. 2017. Systemic co-jumps. *Journal of Financial Economics* 126: 563–91. [CrossRef]

Corsi, Fulvio. 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7: 563–91. [CrossRef]

Corsi, Fulvio, Davide Pirino, and Roberto Reno. 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics* 159: 276–88. [CrossRef]

Corradi, Valentina, Mervyn J. Silvapulle, and Norman R. Swanson. 2018. Testing for jumps and jump intensity path dependence. *Journal of Econometrics* 204: 248–67. [CrossRef]

Ding, Zhuanxin, and Clive W. J. Granger. 1996. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics* 73: 185–215. [CrossRef]

Ding, Zhuanxin, Clive W. J. Granger, and Robert F. Engle. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1: 83–106. [CrossRef]

Gençay, Ramazan, Michel Dacorogna, Ulrich A. Muller, Olivier Pictet, and Richard Olsen. 2001. *An Introduction to High-Frequency Finance*. Cambridge: Academic Press.

Ghysels, Eric, and Arthur Sinko. 2011. Volatility forecasting and microstructure noise. *Journal of Econometrics* 160: 257–71. [CrossRef]

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov. 2006. Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131: 59–95. [CrossRef]

Gilder, Dudley, Mark B. Shackleton, and Stephen J. Taylor. 2014. Cojumps in stock prices: Empirical evidence. *Journal of Banking & Finance* 40: 443–59.

Hansen, Peter R., Asger Lunde, and James M. Nason. 2011. The model confidence set. *Econometrica* 79: 453–97. [CrossRef]

Hansen, Peter R., and Asger Lunde. 2005. A forecast comparison of volatility models: Does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics* 20: 873–89. [CrossRef]

Harvey, Andrew, Esther Ruiz, and Neil Shephard. 1994. Multivariate stochastic variance models. *The Review of Economic Studies* 61: 247–64. [CrossRef]

Hausman, Jerry A. 1978. Specification tests in econometrics. *Econometrica: Journal of the Econometric Society* 46: 1251–71. [CrossRef]

Huang, Xin, and George Tauchen. 2005. The relative contribution of jumps to total price variance. *Journal of Financial Econometrics* 3: 456–99. [CrossRef]

Jacod, Jean, and Viktor Todorov. 2009. Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics* 37: 1792–838. [CrossRef]

Jiang, George J., and Roel CA Oomen. 2008. Testing for jumps when asset prices are observed with noise—A "swap variance" approach. *Journal of Econometrics* 144: 352–70. [CrossRef]

Koopman, Siem Jan, Borus Jungbacker, and Eugenie Hol. 2005. Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance* 12: 445–75.

Lahaye, Jérôme, and Christopher Neely. 2020. The role of jumps in volatility spillovers in foreign exchange markets: Meteor shower and heat waves revisited. *Journal of Business & Economic Statistics* 38: 410–27.

Lee, Suzanne S., and Per A. Mykland. 2007. Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies* 21: 2535–63. [CrossRef]

Mancini, Cecilia. 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36: 270–96. [CrossRef]

Mancini, Cecilia, and Fabio Gobbi. 2012. Identifying the brownian covariation from the co-jumps given discrete observations. *Econometric Theory* 28: 249–73. [CrossRef]

Mukherjee, Arpita, Weijia Peng, Norman R. Swanson, and Xiye Yang. 2020. Financial econometrics and big data: A survey of volatility estimators and tests for the presence of jumps and co-jumps. In *Handbook of Statistics*. Amsterdam: Elsevier, vol. 42, pp. 3–59.

Müller, Ulrich A., Michel M. Dacorogna, Rakhal D. Davé, Richard B. Olsen, Olivier V. Pictet, and Jacob E. Von Weizsäcker. 1997. Volatilities of different time resolutions—Analyzing the dynamics of market components. *Journal of Empirical Finance* 4: 213–39. [CrossRef]

Patton, Andrew J., and Kevin Sheppard. 2015. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics* 97: 683–97. [CrossRef]

Podolskij, Mark, and Daniel Ziggel. 2010. New tests for jumps in semimartingale models. *Statistical Inference for Stochastic Processes* 13: 15–41. [CrossRef]

Qiu, Yue, Xinyu Zhang, Tian Xie, and Shangwei Zhao. 2019. Versatile HAR model for realized volatility: A least square model averaging perspective. *Journal of Management Science and Engineering* 4: 55–73. [CrossRef]

Todorov, Viktor, and George Tauchen. 2010. Activity signature functions for high-frequency data analysis. *Journal of Econometrics* 154: 125–38. [CrossRef]