



Article

# The Impact of Jumps and Leverage in Forecasting the Co-Volatility of Oil and Gold Futures

Manabu Asai <sup>1,†</sup> , Rangan Gupta <sup>2,†</sup> and Michael McAleer <sup>3,4,5,6,7,\*</sup> <sup>1</sup> Faculty of Economics, Soka University, Tokyo 192-8577, Japan<sup>2</sup> Department of Economics, University of Pretoria, Pretoria 0002, South Africa<sup>3</sup> Department of Finance, Asia University, Taichung 41354, Taiwan<sup>4</sup> Discipline of Business Analytics, University of Sydney Business School, Darlington, NSW 2006, Australia<sup>5</sup> Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA Rotterdam, The Netherlands<sup>6</sup> Department of Economic Analysis and ICAE, Complutense University of Madrid, 28040 Madrid, Spain<sup>7</sup> Institute of Advanced Science, Yokohama National University, Yokohama, Kanagawa 240-8501, Japan

\* Correspondence: michael.mcaleer@gmail.com

† These authors contributed equally to this work.

Received: 12 June 2019; Accepted: 27 August 2019; Published: 2 September 2019



**Abstract:** This paper investigates the impact of jumps in forecasting co-volatility in the presence of leverage effects for daily crude oil and gold futures. We use a modified version of the jump-robust covariance estimator of Koike (2016), such that the estimated matrix is positive definite. Using this approach, we can disentangle the estimates of the integrated co-volatility matrix and jump variations from the quadratic covariation matrix. Empirical results show that more than 80% of the co-volatility of the two futures contains jump variations and that they have significant impacts on future co-volatility but that the impact is negligible in forecasting weekly and monthly horizons.

**Keywords:** commodity markets; co-volatility; forecasting; jump; leverage effects; realized covariance; threshold estimation

## 1. Introduction

The severity and global nature of the recent financial crisis highlighted the risks associated with portfolios containing only conventional financial market assets [1]. Such a realization triggered an interest in considering investment opportunities in the energy (specifically oil) market [2]. In fact, the recent financialization of the commodity market [3,4] and, in particular, oil has resulted in an increased participation of hedge funds, pension funds, and insurance companies in the market, with investment in oil now being considered as a profitable alternative instrument in the portfolio decisions of financial institutions [5,6].

With gold traditionally considered as the most popular ‘safe haven’ [7–9], recent studies have analyzed volatility spillovers across the gold and oil markets [10–12], where volatility spillovers are defined as the delayed effect of a returns shock in one asset on the subsequent volatility or co-volatility in another asset [13]. For corresponding studies on co-movements in gold and oil returns, see References [14–17], and references cited therein. A literature review on return and volatility spillovers across asset classes can be found in Reference [18].

In this regard, it must be realized that modelling and forecasting the co-volatility of gold and oil markets is of paramount importance to international investors and portfolio managers in devising optimal portfolio and dynamic hedging strategies [19]. By definition, (partial) co-volatility spillovers occur when the returns shock from financial asset  $k$  affects the co-volatility between two financial assets,  $i$  and  $j$ , one of which can be asset  $k$  [13].

Against this backdrop, the objective of this paper is to forecast the daily co-volatility of gold and oil futures derived from 1-min intraday data over the period 27 September 2009 to 25 May 2017 (Although the variability of daily gold and oil price returns have traditionally been forecasted based on Generalized Autoregressive Conditional Heteroskedasticity (GARCH)-type models of volatility, recent empirical evidence suggests that the rich information contained in intraday data can produce more accurate estimates and forecasts of daily volatility (see Reference [20] for a detailed discussion)). In particular, realizing the importance of jumps, that is, discontinuities, in governing the volatility of asset prices [21–23], we investigate the impact of jumps by simultaneously accommodating leverage effects in forecasting the co-volatility of gold and oil markets, following the econometric approach of [24] (applied to three stocks traded on the New York Stock Exchange (NYSE)).

Although studies dealing with forecasting gold and oil market volatility has emphasized the role of jumps in forecasting realized volatility. (For a given fixed interval, realized volatility is defined as the sum of non-overlapping squared returns of high frequency within a day [25] which, in turn, presents volatility as an observed rather than a latent process.) (see, for example, References [26–28]), this paper would be seen to be the first attempt to incorporate their role in predicting the future co-volatility path of these two important commodities.

The remainder of the paper is structured as follows—Section 2 lays out the theoretical details of the econometric framework, while Section 3 presents the data, empirical results and analysis. Section 4 gives some concluding remarks.

## 2. Model Specification

Let  $X_t^*$  and  $Y_t^*$  denote latent log-prices at time  $s$  for two assets  $X$  and  $Y$ . Define  $p^*(\tau) = (X_t^*, Y_t^*)'$ , and let  $W(\tau)$  and  $Q(\tau)$  denote bivariate vectors of independent Brownian motions and counting processes, respectively. Let  $K(\tau)$  be the  $2 \times 2$  process controlling the magnitude and transmission of jumps, such that  $K(\tau)dQ(\tau)$  is the contribution of the jump process to the price diffusion. Under the assumption of a Brownian semimartingale with finite-activity jumps (BSMFJ),  $p^*(\tau)$  follows:

$$dp^*(\tau) = \mu(\tau)ds + \sigma(\tau)dW(\tau) + K(\tau)dQ(\tau), \quad 0 \leq \tau \leq T \quad (1)$$

where  $\mu(\tau)$  is a  $2 \times 1$  vector of continuous and locally-bounded variation processes, and  $\sigma(\tau)$  is the  $2 \times 2$  càdlàg matrix, such that  $\Sigma(\tau) = \sigma(\tau)\sigma'(\tau)$  is positive definite. Note that although we explain the framework with finite-activity jumps, the estimators used in this paper are applicable under infinite-activity jumps, as shown by Reference [29].

Assume that the observable log-price process is the sum of the latent log-price process in Equation (1) and the microstructure noise process. Denote the log-price process as  $p(\tau) = (X_\tau, Y_\tau)'$ . Consider non-synchronized trading times of the two assets and let  $\mathcal{Z}$  and  $\mathcal{E}$  be the set of transaction times of  $X$  and  $Y$ , respectively. Denote the counting process governing the number of observations traded in assets  $X$  and  $Y$  up to time  $T$  as  $n_T$  and  $m_T$ , respectively. By definition, the trades in  $X$  and  $Y$  occur at times  $\mathcal{Z} = \{\zeta_1, \zeta_2, \dots, \zeta_{n_T}\}$  and  $\mathcal{E} = \{\xi_1, \xi_2, \dots, \xi_{m_T}\}$ , respectively. For convenience, we set the opening and closing times as  $\zeta_1 = \xi_1 = 0$  and  $\zeta_{n_T} = \xi_{m_T} = T$ , respectively.

The observable log-price process is given by:

$$X_{\zeta_i} = X_{\zeta_i}^* + \varepsilon_{\zeta_i}^X \quad \text{and} \quad Y_{\xi_j} = Y_{\xi_j}^* + \varepsilon_{\xi_j}^Y, \quad (2)$$

where  $\varepsilon^X \sim \text{iid}(0, \sigma_{\varepsilon^X}^2)$ ,  $\varepsilon^Y \sim \text{iid}(0, \sigma_{\varepsilon^Y}^2)$ , and  $(\varepsilon^X, \varepsilon^Y)$  are independent of  $(X, Y)$ .

Define the quadratic covariation (QCov) of the log-price process over  $[0, T]$  as:

$$\text{QCov} = \text{plim}_{\Delta \rightarrow \infty} \sum_{i=1}^{\lfloor T/\Delta \rfloor} [p(i\Delta) - p((i-1)\Delta)] [p(i\Delta) - p((i-1)\Delta)]'. \quad (3)$$

By Proposition 2 of Reference [26], we obtain:

$$\text{QCov} = \text{ICov} + \sum_{0 < \tau \leq T} K(\tau)K'(\tau), \quad (4)$$

where

$$\text{ICov} = \int_0^T \Sigma(\tau) d\tau.$$

The first term on the right-hand side of (4) is the integrated co-volatility matrix over  $[0, T]$ , while the second term is the matrix of jump variability. We are interested in disentangling these two components from the estimates of QCov for the purpose of forecasting QCov.

There are several estimators for QCov and ICov (see the survey in Reference [24]). Among them, we use the estimators of [30] for QCov and [29] for ICov, respectively. Especially, the estimator of Reference [29] is consistent under non-synchronized trading times, jumps and microstructure noise for the bivariate process in (2) (see Appendix A for the detailed explanation of the calculation of these estimators). Note that the realized kernel (RK) estimator of Reference [31] is positive (semi-)definite and robust to microstructure noise under non-synchronized trading times. However, the robustness to jumps is still an open and unresolved issue for the multivariate RK estimator. Denote the estimators of QCov, ICov and jump component at day  $t$  as  $\hat{\Sigma}_t$ ,  $\hat{C}_t$  and  $\hat{J}_t$ , respectively, where  $\hat{J}_t = \hat{\Sigma}_t - \hat{C}_t$  (Jump variations can be defined strictly, as discussed in Reference [32]). By the definitions in (1)–(4), the estimators should be positive (semi-) definite, but there is no guarantee for it. For this purpose, Reference [24] suggest regularizing the estimated covariance matrix by the use of thresholding.

As shown by References [33–35], the regularized estimator has consistency, assuming a sparsity structure. Define the thresholding operator for a square matrix  $A$  as:

$$\mathcal{T}_h(A) = [a_{ij}\mathbf{1}(|a_{ij}| \geq h)], \quad (5)$$

which can be regarded as  $A$  thresholded at  $h$ . Define the Frobenius norm by  $\|A\|_F^2 = \text{tr}(AA')$ . For the selection of  $h$ , we follow Reference [34]. In order to obtain  $\tilde{A} = \mathcal{T}_h(\hat{A})$ , we minimize the distance by the Frobenius norm  $\|\mathcal{T}_h(\hat{A}) - \hat{A}\|_F^2$ , with the restriction that  $\tilde{A}$  is positive semi-definite. As in Reference [24], we obtain  $\tilde{C}_t = \mathcal{T}_h(\hat{C}_t)$  and  $\tilde{J}_t = \mathcal{T}_h(\hat{J}_t)$ , which are consistent and positive semi-definite. Note that  $\hat{\Sigma}_t$  is positive semi-definite as it is the sample analogue of QCov. In addition, we also disentangling observed return series into continuous and jump components, by applying the technique of Reference [36] (see Appendix A.3).

In order to examine the effects of jump and leverage in forecasting co-volatility, we consider four kinds of specifications, including the three models introduced by Reference [24]. Let  $\tilde{\Sigma}_{t(l)}$  denote the  $l$ -horizon average, defined by:

$$\tilde{\Sigma}_{t(l)} = \frac{1}{l} (\tilde{\Sigma}_t + \dots + \tilde{\Sigma}_{t-l+1}).$$

In order to examine the impact of jumps and leverage for forecasting volatility and co-volatility, we use four kinds of heterogeneous autoregressive (HAR) models for forecasting the  $(i, j)$ -elements of  $\tilde{\Sigma}_{t(l)}$  ( $l = 1, 5, 22$ ), as follows:

$$\tilde{\Sigma}_{ij,t(l)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + e_{ij,t} \quad (6)$$

$$\tilde{\Sigma}_{ij,t(l)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t} \quad (7)$$

$$\tilde{\Sigma}_{ij,t(l)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + e_{ij,t} \quad (8)$$

$$\tilde{\Sigma}_{ij,t(l)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t}, \quad (9)$$

where  $r_{i,t}^- = r_{i,t} I(r_{i,t} < 0)$  and  $rc_{i,t}^- = rc_{i,t} I(rc_{i,t} < 0)$ , which are the negative parts of the observed return and its continuous part for the  $i$ -th asset. The second model accommodates the asymmetric

effects, as in the specification of the asymmetric BEKK model of Reference [37]. For  $i \neq j$ ,  $\gamma_a r_{i,t-1}^- r_{j,t-1}^-$  represents the ‘co-leverage’ effect, which is caused by simultaneous negative returns in two assets. In the third model, we use the previous values of the estimated continuous sample path component variation,  $\tilde{C}_t$ , rather than those of the estimated quadratic variation,  $\tilde{\Sigma}_t$ , following the volatility forecasting models of References [21,23]. We exclude weekly and monthly effects of the jump component,  $\tilde{J}_t$ , in order to evaluate the impact of a single jump on future volatility and co-volatility. Note that  $\tilde{C}_t$  and  $\tilde{J}_t$  are positive (semi-) definite by the thresholding in (5). In addition to jump variability, the fourth model includes the asymmetric effect. Note that we use the continuous components of returns rather than the observed returns for the fourth model. We refer to Equations (6)–(9) as the HAR, HAR-A, HAR-TCJ and HAR-TCJA models, respectively. We estimate these models by ordinary least squares (OLS) and use the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator, with bandwidth 25 (see Reference [38]). We will examine the four models (6)–(9) in the next section.

### 3. Empirical Analysis

We examine the effects on jumps and leverage in forecasting co-volatility, using the estimates of QCov, ICov and jump variation, for two futures contracts traded on the New York Mercantile Exchange (NYMEX), namely West Texas Intermediate (WTI) Crude Oil and Gold. With the CME Globex system, the trades at NYMESX cover 24 h (The futures price data, in continuous format, are obtained from <http://www.kibot.com/>). Based on the vector of returns for the two futures for a 1-min interval of trading day at  $t$ , we calculated the daily values of  $\tilde{\Sigma}_t$ ,  $\tilde{C}_t$  and  $\tilde{J}_t$ , as explained in the previous section and also the corresponding open-close returns and their continuous components,  $r_t$  and  $r_{ct}$ , respectively, for the two futures. The sample period starts on 27 September 2009 and ends on 25 May 2017, giving 1978 observations. The sample is divided into two periods—the first 1000 observations are used for in-sample estimation, while the last 978 observations are used for evaluating the out-of-sample forecasts.

Table 1 presents the descriptive statistics of the returns,  $r_t$ , and estimated QCov,  $\tilde{\Sigma}_t$ . The empirical distribution of the returns is highly leptokurtic. Regarding volatility, their distributions are skewed to the right, with evidence of heavy tails in the two series. More than 90% of the sample period contains significant jumps in volatility. For co-volatility, the empirical distribution is highly leptokurtic and co-jump variations were found for 80% of the period.

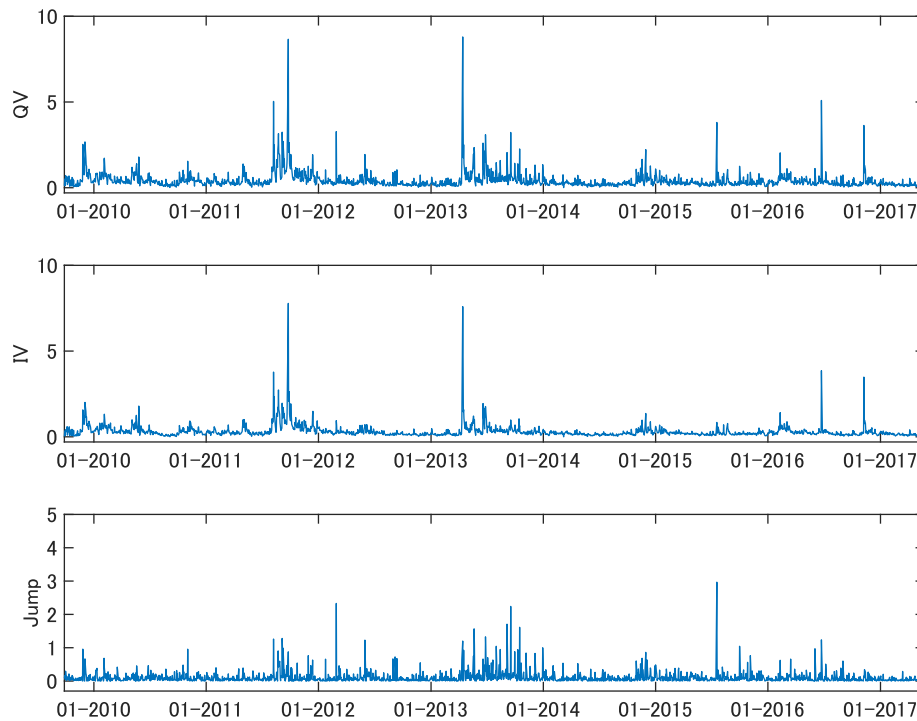
**Table 1.** Descriptive statistics of returns, volatility and co-volatility.

Stock	Mean	Std. Dev.	Skew.	Kurt.	Jump
Return					
Crude Oil	0.0256	0.6824	−0.4784	8.0378	0.9267
Gold	0.0185	1.4887	−0.0573	5.0225	0.9459
Volatility					
Crude Oil	0.4238	0.5070	7.2211	90.688	0.9267
Gold	1.5152	1.6510	3.3689	19.071	0.9459
Co-Volatility (Crude Oil, Gold)	0.1368	0.2844	−0.7851	37.359	0.8049

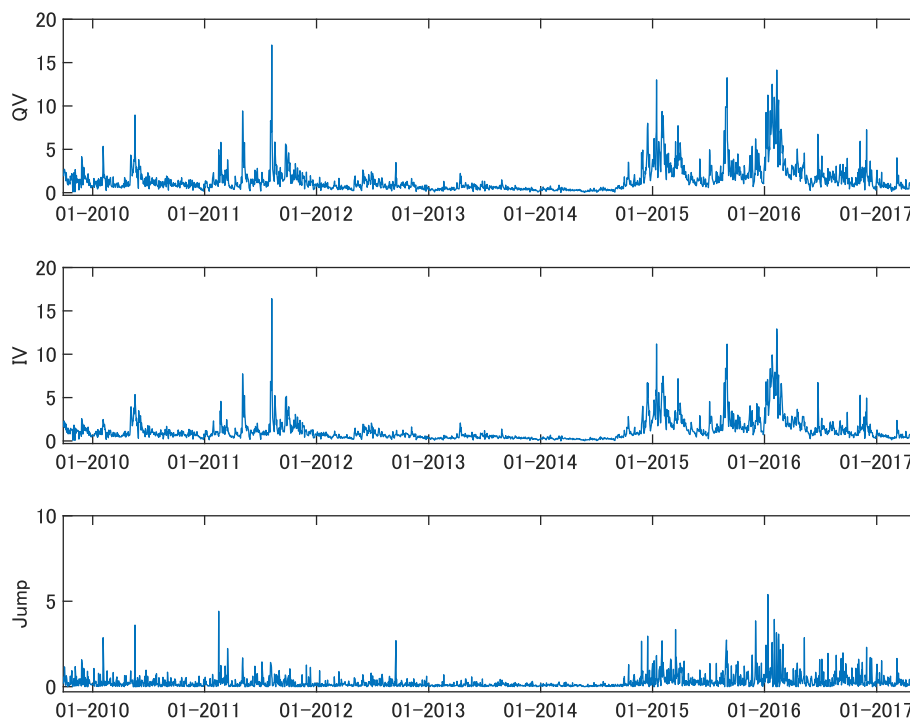
Note: The sample period is from 27 September 2009 to 25 May 2017. ‘Jump’ denotes the percentage of occurrence of significant jumps.

Figures 1 and 2 show the estimates of quadratic variation, integrated volatility, and jump variability, namely the diagonal elements of  $\tilde{\Sigma}_t$ ,  $\tilde{C}_t$ , and  $\tilde{J}_t$ , respectively. It is known that spot and future prices of crude oil are effected by a variety of geopolitical and economic events. For instance, the estimates of volatility in Figure 1 are relatively high for the period following the Arab Spring of 2011, and the extreme jump in 2015 is caused by the oversupply and the technological advancements of US shale oil production. On the other hand, the spot and futures prices of Gold reflect news and

recessions, as investigated by Reference [39]. The estimates of volatility are high in 2011 during the European debt crisis, and the last one-third in Figure 2. For the latter, it corresponds to China’s economy growing at its slowest pace for 24 years in 2014.

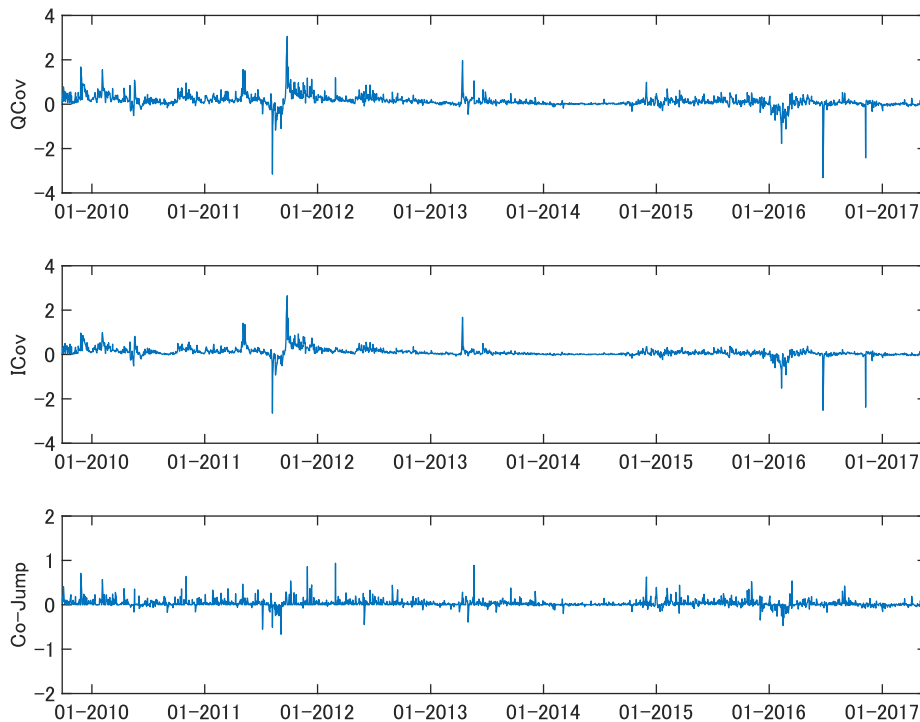


**Figure 1.** Estimates of Quadratic Variation, Integrated Volatility, and Jump Variability for Crude Oil Futures. Figure 1 shows the (1,1)-elements of  $\hat{\Sigma}_t$ ,  $\hat{C}_t$ , and  $\hat{J}_t$ .



**Figure 2.** Estimates of Quadratic Variation, Integrated Volatility, and Jump Variability for Gold Futures. Figure 2 shows the (2,2)-elements of  $\hat{\Sigma}_t$ ,  $\hat{C}_t$ , and  $\hat{J}_t$ .

Figure 3 illustrates the estimates of quadratic covariation, integrated co-volatility and jump co-variability, namely the (2, 1)-element of  $\hat{\Sigma}_t$ ,  $\hat{C}_t$ , and  $\hat{J}_t$ . Figure 3 indicates that crude oil and gold futures are negatively correlated for the first half in 2011, but the sign changes for the latter half. A large and positive co-jump variability is found in 2013, which reflects the political unrest in Egypt and the values of Dow Jones Industrial Average kept increasing with a rapid trend.



**Figure 3.** Estimates of Quadratic Covariation, Integrated Covolatility, and Jump Co-variability for Crude Oil and Gold Futures. Figure 3 shows the (2, 1)-elements of  $\hat{\Sigma}_t$ ,  $\hat{C}_t$ , and  $\hat{J}_t$ .

In the following empirical analysis based on the four models (6)–(9), we:

- examine the robustness of the positive effects of jump components under microstructure noise for the volatility equation ( $i = j$ );
- test the robustness of the leverage effects under jump and microstructure noise for the volatility equation;
- investigate the effects of co-jumps and co-leverage for the co-volatility equation;
- compare the out-of-sample forecasts of the above models.

For the above models for the volatility equation, the estimates of  $\gamma_j$  are expected to be positive. However, the empirical results of Reference [21] indicate that the estimates of  $\gamma_j$  are generally insignificant. Reference [23] noted that the puzzle is due to the small sample bias of the integrated volatility, and found that the estimates of  $\gamma_j$  are positive and significant. Since the estimators of quadratic variation and integrated volatility used in Reference [23] are biased in the presence of microstructure noise, we re-examine the robustness of the result and this is the motivation of (a). With respect to (b), we also test  $\gamma_a > 0$  under microstructure noise in order to check the robustness of the result of Reference [40]. Among (a)–(d), (c) is the main purpose of the current empirical analysis.

For volatility equation ( $i = j$ ), the estimates of  $\gamma_j$  and  $\gamma_a$  are expected to be positive and significant. On the other hand, their signs are not determined for the co-volatility equation ( $i \neq j$ ). As noted in (d), we compare the forecasting performances of the HAR, HAR-TCJ and HAR-TCJA models for daily, weekly and monthly horizons. For crude oil futures, Reference [41] found that HAR-TCJA is the best forecasting one-day-ahead volatility, while the HAR model is preferred for weekly and monthly

forecasts, after removing the effects of structural breaks. In addition to crude oil, we examine the forecasting performances of the volatility of gold futures and the co-volatility of the two futures.

As stated in the previous section, we use the HAC estimator for the covariance matrix of the OLS estimators. To ensure the statistical adequacy of the approach, we conduct three kinds of conventional diagnostic tests. The first is the White test for heteroskedasticity. When the number of explanatory variables is  $k$ , excluding the constant, the Wald statistic for the White test has the asymptotic  $\chi^2$  distribution with the degree-of-freedom parameter  $(k + 1)(k + 2)/2$  under the null of homoskedasticity. The second test is the ARCH(2) test with the auxiliary regression:

$$\hat{\varepsilon}_{ij,t}^2 = \delta_0 + \delta_1 \hat{\varepsilon}_{ij,t-1}^2 + \delta_2 \hat{\varepsilon}_{ij,t-1}^2 + \text{error}_{ij,t},$$

where  $\hat{\varepsilon}_{ij,t}$  is the residuals. Under the null hypothesis of homoskedasticity, that is  $\delta_1 = \delta_2 = 0$ , the Wald statistic has the asymptotic  $\chi^2$  distribution with the degree-of-freedom parameter 2. The third one is the test for autocorrelation with the auxiliary regression:

$$\hat{\varepsilon}_{ij,t} = \rho_0 + \rho_1 \hat{\varepsilon}_{ij,t-1} + \dots + \rho_{10} \hat{\varepsilon}_{ij,t-1} + \text{error}_{ij,t}.$$

Under the null hypothesis of no autocorrelation, that is  $\rho_1 = \dots = \rho_{10} = 0$ , the Wald statistic has the asymptotic  $\chi^2$  distribution with the degree-of-freedom parameter 10.

We estimate each model using the first 1000 observations, and obtain a forecast,  $\tilde{\Sigma}_{1001}^f$ . We re-estimate each model fixing the sample size at 1000, and obtain new forecasts based on updated parameter estimates. For evaluating the forecasting performance of the different models, we consider the Mincer and Zarnowitz (MZ) [42] regression, namely:

$$\tilde{\Sigma}_{ij,t} = \alpha_0 + \alpha_1 \tilde{\Sigma}_{ij,t}^f + \text{error}_{ij,t}, \quad t = 1001, \dots, 1978.$$

We estimate the MZ regression via the OLS and report  $R^2$  for comparing forecasting performance. Intuitively, a model with the highest  $R^2$  indicates the accuracy of forecasts. We also use the heteroskedasticity-adjusted root mean square error suggested in Reference [43], namely:

$$\text{HRMSE} = \sqrt{\frac{1}{978} \sum_{t=1001}^{1978} \left( \frac{\tilde{\Sigma}_{ij,t} - \tilde{\Sigma}_{ij,t}^f}{\tilde{\Sigma}_{ij,t}} \right)^2}.$$

For the latter, we examine equal forecast accuracy using the Diebold-Mariano (DM) [44] test at the 5% significance level, and use the HAC covariance matrix estimator, with bandwidth 25. We also examine the forecasts of  $\tilde{\Sigma}_{ij,t(5)}$  and  $\tilde{\Sigma}_{ij,t(22)}$  in the same manner.

As an application of forecasts of covariance matrix, we examine performances of portfolio weights determined by the forecasts. We consider a portfolio of two options,  $r_{pt} = w_t' \{p(t) - p(t-1)\}$ , where  $p(s)$  is the vector of log-prices stated in Section 2 and  $w_t$  is the vector of portfolio weights. With the forecasts of covariance matrix,  $\tilde{\Sigma}_t^f$ , the weights of the minimum variance portfolio are given by:

$$w_t^* = (\iota' \tilde{\Sigma}_t^{f-1} \iota)^{-1} \tilde{\Sigma}_t^{f-1} \iota, \quad (10)$$

where  $\iota = (1, 1)'$ . Given the realized value  $\tilde{\Sigma}_t$ , the corresponding portfolio variance is evaluated as:

$$\sigma_{pt}^2 = w_t^{*'} \tilde{\Sigma}_t w_t^*. \quad (11)$$

We calculate the HRMSE of  $\sigma_{pt}^2$  for the four models, and compare the values by the DM test. In the same manner, we compare the portfolio weights of the minimum variance portfolio based on the forecasts of  $\tilde{\Sigma}_{ij,t(5)}$  and  $\tilde{\Sigma}_{ij,t(22)}$ .

Table 2 shows the estimates of the daily regressions for the first 1000 observations. For the jump parameter,  $\gamma_j$ , the estimates are positive and significant at five percent level in all cases. The results for the volatilities support the empirical evidence of Reference [23]. For the asymmetric effect, the estimates of  $\gamma_a$  are positive and significant in all cases, supporting the negative relationship between return and one-step-ahead volatility, as in Reference [40]. The results for the co-volatility equation indicate that a pair of negative returns and/or co-jumps of two assets increases future co-volatility. Either of HAR-A and HAR-TCJA model gives the highest  $\bar{R}^2$  in all cases.

Table 3 presents the diagnostic statistics for the daily regressions. White and ARCH(2) reject the null hypothesis of homoskedasticity at the five percent level in all cases. The test for 10th-order autocorrelation rejected the null hypothesis of no autocorrelation for the volatility of Gold futures. Hence, it is justified to use the HAC covariance matrix estimates for the OLS estimates in Table 2. The corresponding results for the weekly and monthly regressions are omitted to save space.

Table 4 presents  $R^2$  of the MZ regressions and HRMSE for the daily regressions. The HAR-A has the highest  $R^2$  for four-ninth of the cases, while the results of the DM tests regarding the HRMSE indicate that there are no significant differences for the four models in all cases.

Table 5 reports the estimates of the weekly regressions. The estimates of the jump parameter,  $\gamma_j$ , and the parameter of the asymmetric effect,  $\gamma_a$ , are positive and significant. For the weekly regressions, the HAR-A model gives the highest  $\bar{R}^2$  values in all cases. Table 6 gives the  $R^2$  values of the MZ regressions, and HRMSE for the out-of-sample forecasts for the weekly regressions. Tables 5 and 6 indicate that the values of  $R^2$  (and  $\bar{R}^2$ ) are higher than those for the daily regressions in Tables 2 and 4, respectively. Table 6 shows that the HAR-A model gives the highest  $\bar{R}^2$  for volatility. For co-volatility,  $R^2$  for MZ tends to select the HAR model, while HRMSE chooses the HAR-A model. However, the DM tests show that there are no significant differences for the four models for co-volatility.

Table 7 shows the in-sample estimates of the monthly regressions, while Table 8 reports the results of the corresponding out-of-sample forecasts. As in the weekly regressions, the HAR-A model gives the highest  $\bar{R}^2$  values in all cases. Tables 7 and 8 indicate that the values of  $R^2$  are higher than those for the daily regressions in Tables 5 and 6, respectively. Table 8 shows that the HAR-A model is the best model in all cases and, moreover, there are significant differences between HAR-A model and (HAR-TCJ, HAR-TCJA) models in forecasting volatility.

Table 9 presents the results for forecasts regarding the portfolio weights of the minimum variance portfolio. Using the portfolio weights defined in (10), we evaluated the portfolio variance applying the weights to the realized covariance as in Equation (11). For the forecasts via the daily regression, the HAR-TCJA model has the highest  $R^2$  value for the MZ regression, while the HAR-A gives the smallest HRMSE value. The DM test indicates there is no significant difference in the four models for the forecasts based on the daily regressions. On the other hand, the HAR-A model has the highest  $R^2$  value and smallest HRMSE value for the weekly and monthly regressions.

The empirical results for the volatility models support the findings of References [23,40,41]. Regarding co-volatility, the impacts of co-jumps of two assets are positive and significant for the daily, weekly, and monthly regressions. Although the four models show similar forecasting performance for the daily regressions, the weekly and monthly regressions prefer the HAR-A model. We may improve the HAR-TCJ and HAR-TCJA models by accommodating the positive and negative jumps of Reference [45], in addition to the weekly and monthly averages of jumps and leverage effects.



**Table 2.** In-sample estimates for daily regressions.

$$\begin{aligned} \text{HAR} & \tilde{\Sigma}_{ij,t} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + e_{ij,t} \\ \text{HAR-A} & \tilde{\Sigma}_{ij,t} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t} \\ \text{HAR-TCJ} & \tilde{\Sigma}_{ij,t} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + e_{ij,t} \\ \text{HAR-TCJA} & \tilde{\Sigma}_{ij,t} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t} \end{aligned}$$

Model	$\gamma_0$	$\gamma_d$	$\gamma_w$	$\gamma_m$	$\gamma_j$	$\gamma_a$	$R^2$	$\bar{R}^2$
Volatility: Crude Oil								
HAR	0.0958 (0.0021)	0.4428 (0.0054)	0.1047 (0.0050)	0.2665 (0.0056)			0.3514	0.3494
HAR-A	0.1067 (0.0018)	0.2258 (0.0039)	0.2000 (0.0040)	0.2905 (0.0056)		0.1301 (0.0032)	0.4153	0.4129 <sup>†</sup>
HAR-TCJ	0.1239 (0.0017)	0.5182 (0.0083)	0.1012 (0.0064)	0.2879 (0.0064)	0.2410 (0.0112)		0.3546	0.3520
HAR-TCJA	0.1301 (0.0015)	0.2285 (0.0067)	0.2431 (0.0054)	0.3176 (0.0064)	0.2575 (0.0097)	0.1292 (0.0032)	0.4140	0.4110
Volatility: Gold								
HAR	0.2039 (0.0028)	0.4730 (0.0060)	0.2067 (0.0053)	0.1462 (0.0043)			0.4396	0.4378
HAR-A	0.0356 (0.0009)	0.3161 (0.0045)	0.3324 (0.0046)	0.1337 (0.0062)		0.0574 (0.0010)	0.4320	0.4296
HAR-TCJ	0.2520 (0.0027)	0.5629 (0.0082)	0.1512 (0.0057)	0.1767 (0.0048)	0.2913 (0.0094)		0.4414	0.4391
HAR-TCJA	0.2193 (0.0024)	0.4612 (0.0066)	0.2066 (0.0051)	0.1621 (0.0041)	0.2427 (0.0069)	0.1189 (0.0014)	0.5266	0.5242 <sup>†</sup>
Co-Volatility: Crude Oil & Gold								
HAR	0.0431 (0.0009)	0.3938 (0.0043)	0.3199 (0.0049)	0.0923 (0.0064)			0.4034	0.4016
HAR-A	0.1704 (0.0024)	0.3793 (0.0052)	0.2547 (0.0048)	0.1328 (0.0036)		0.1205 (0.0014)	0.5271	0.5252 <sup>†</sup>
HAR-TCJ	0.0515 (0.0008)	0.5028 (0.0065)	0.3268 (0.0060)	0.1447 (0.0071)	0.2047 (0.0057)		0.4116	0.4092
HAR-TCJA	0.0444 (0.0008)	0.4086 (0.0069)	0.3486 (0.0057)	0.1933 (0.0068)	0.1664 (0.0049)	0.0564 (0.0010)	0.4391	0.4362

Note: Standard errors are given in parentheses. ‘†’ denotes the model which has the highest  $\bar{R}^2$  value of the four models.

**Table 3.** Diagnostic tests for daily regressions.

Model	White	d.f.	p-Value	ARCH(2)	p-Value	AutoCor (10)	p-Value
Volatility: Crude Oil							
HAR	68.955	10	[0.0000]	7.3212	[0.0257]	12.849	[0.2322]
HAR-TCJ	137.08	15	[0.0000]	8.6057	[0.0135]	12.510	[0.2524]
HAR-TCJA	565.48	21	[0.0000]	15.150	[0.0005]	18.158	[0.0524]
HAR-A	280.45	15	[0.0000]	14.439	[0.0007]	17.012	[0.0741]
Volatility: Gold							
HAR	127.72	10	[0.0000]	58.588	[0.0007]	28.195	[0.0017]
HAR-TCJ	162.06	15	[0.0000]	66.637	[0.0000]	27.113	[0.0025]
HAR-TCJA	369.97	21	[0.0000]	27.265	[0.0000]	29.001	[0.0013]
HAR-A	284.90	15	[0.0000]	22.363	[0.0000]	30.453	[0.0007]
Co-Volatility: Crude Oil & Gold							
HAR	62.977	10	[0.0000]	25.708	[0.0000]	9.5972	[0.4765]
HAR-TCJ	74.695	15	[0.0000]	30.078	[0.0000]	9.6906	[0.4680]
HAR-TCJA	87.284	21	[0.0000]	17.846	[0.0000]	12.637	[0.2447]
HAR-A	67.424	15	[0.0000]	14.547	[0.0000]	15.075	[0.1294]

Note: The table presents the values of the diagnostic statistics for the OLS residuals. ‘White’ denotes the White test for heteroskedasticity, ‘ARCH(2)’ is the ARCH(2) test, and AutoCorr(10)’ is the test for 10th-order autocorrelation. ‘df’ denotes the degree-of-freedom. While the ARCH(2) test has the asymptotic  $\chi^2(2)$  under the null hypothesis, the asymptotic distribution of AutoCorr(10) test is  $\chi^2(10)$ .

**Table 4.** Out-of-sample forecast evaluation for daily regressions.

Model	R <sup>2</sup>	HRMSE	R <sup>2</sup> -J	HRMSE-J	R <sup>2</sup> -C	HRMSE-C
Volatility: Crude Oil (900 Times Jump)						
HAR	0.0796	2.5591	0.0818	2.6478	0.0590	1.1057
HAR-A	0.0987	2.0655	0.1023	2.1329	0.0702	1.0018
HAR-TCJ	0.0928	2.0476	0.0944	2.1156	0.0775 <sup>†</sup>	0.9635
HAR-TCJA	0.0995 <sup>†</sup>	2.0205	0.1037 <sup>†</sup>	2.0869	0.0641	0.9672
Volatility: Gold (950 Times Jump)						
HAR	0.6900	0.9009	0.6918	0.9066	0.6017	0.7857
HAR-A	0.6933 <sup>†</sup>	0.8376	0.6944 <sup>†</sup>	0.8431	0.6292 <sup>†</sup>	0.7249
HAR-TCJ	0.6907	0.9103	0.6929	0.9132	0.5891	0.8534
HAR-TCJA	0.6925	0.8493	0.6938	0.8521	0.6211	0.7945
Co-Volatility: Crude Oil & Gold (616 Times Co-Jump)						
HAR	0.2199 <sup>†</sup>	40.906	0.1897	45.317	0.3247 <sup>†</sup>	32.033
HAR-A	0.2197	41.743	0.2007 <sup>†</sup>	45.906	0.2927	33.489
HAR-TCJ	0.2103	36.572	0.1832	40.316	0.3127	29.114
HAR-TCJA	0.2119	37.543	0.1971	41.018	0.2806	30.739

Note: The table presents R<sup>2</sup> for the MZ regression and heteroskedasticity-adjusted root mean squared error (HRMSE). R<sup>2</sup>-J (HRMSE-J) is R<sup>2</sup> (HRMSE) conditionally on having a jump at time t - 1, respectively, while R<sup>2</sup>-C (HRMSE-C) is conditional on no jump at time t - 1. ‘†’ indicates the model which has the highest R<sup>2</sup> value of the four models. For the DM test of equal forecast accuracy, there were no significant differences at the five percent level.

**Table 5.** In-sample estimates for weekly regressions.

HAR  $\tilde{\Sigma}_{ij,t(5)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + e_{ij,t}$   
HAR-A  $\tilde{\Sigma}_{ij,t(5)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t}$   
HAR-TCJ  $\tilde{\Sigma}_{ij,t(5)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + e_{ij,t}$   
HAR-TCJA  $\tilde{\Sigma}_{ij,t(5)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t}$

Model	$\gamma_0$	$\gamma_d$	$\gamma_w$	$\gamma_m$	$\gamma_j$	$\gamma_a$	R <sup>2</sup>	$\bar{R}^2$
Volatility: Crude Oil								
HAR	0.0203 (0.0006)	0.1731 (0.0015)	0.7803 (0.0028)	0.0070 (0.0017)			0.9208	0.9206
HAR-A	0.0221 (0.0005)	0.1387 (0.0015)	0.7954 (0.0029)	0.0108 (0.0017)		0.0206 (0.0008)	0.9237	0.9234 <sup>†</sup>
HAR-TCJ	0.0562 (0.0006)	0.1843 (0.0024)	0.8853 (0.0039)	0.0002 (0.0022)	0.2655 (0.0028)		0.9060	0.9057
HAR-TCJA	0.0571 (0.0006)	0.1441 (0.0026)	0.9050 (0.0039)	0.0043 (0.0023)	0.2678 (0.0026)	0.0179 (0.0008)	0.9081	0.9076
Volatility: Gold								
HAR	0.0446 (0.0008)	0.1793 (0.0016)	0.7970 (0.0023)	-0.0142 (0.0012)			0.9403	0.9402
HAR-A	0.0380 (0.0007)	0.1609 (0.0015)	0.8064 (0.0023)	-0.0168 (0.0012)		0.0236 (0.0003)	0.9456	0.9454 <sup>†</sup>
HAR-TCJ	0.1086 (0.0009)	0.2244 (0.0023)	0.8310 (0.0033)	-0.0058 (0.0016)	0.2440 (0.0019)		0.9240	0.9237
HAR-TCJA	0.1027 (0.0009)	0.2059 (0.0021)	0.8411 (0.0033)	-0.0084 (0.0016)	0.2351 (0.0015)	0.0217 (0.0003)	0.9285	0.9281
Co-Volatility: Crude Oil & Gold								
HAR	0.0099 (0.0002)	0.1441 (0.0012)	0.8321 (0.0020)	-0.0207 (0.0017)			0.9346	0.9344
HAR-A	0.0084 (0.0002)	0.1277 (0.0012)	0.8347 (0.0020)	-0.0120 (0.0016)		0.0121 (0.0002)	0.9367	0.9364 <sup>†</sup>
HAR-TCJ	0.0237 (0.0003)	0.1790 (0.0021)	0.9493 (0.0031)	0.0140 (0.0024)	0.1996 (0.0013)		0.9160	0.9157
HAR-TCJA	0.0222 (0.0003)	0.1594 (0.0022)	0.9538 (0.0031)	0.0241 (0.0023)	0.1916 (0.0011)	0.0117 (0.0003)	0.9180	0.9176

Note: Standard errors are given in parentheses. ‘†’ denotes the model which has the highest R<sup>2</sup> value of the four models.

**Table 6.** Out-of-sample forecast evaluation for weekly regressions.

Model	MZ $R^2$	HRMSE	$R^2$ -J	HRMSE-J	$R^2$ -C	HRMSE-C
Volatility: Crude Oil (900 Times Jump)						
HAR	0.8060	0.1980 <sup>c,d</sup>	0.7964	0.2011 <sup>c,d</sup>	0.8792	0.1578 <sup>c,d</sup>
HAR-A	0.8138 <sup>†</sup>	0.1996 <sup>c,d</sup>	0.8035 <sup>†</sup>	0.2032 <sup>c,d</sup>	0.8874 <sup>†</sup>	0.1523 <sup>c,d</sup>
HAR-TCJ	0.7321	0.2444	0.7286	0.2469	0.8064	0.2141
HAR-TCJA	0.7381	0.2465	0.7340	0.2492	0.8120	0.2128
Volatility: Gold (929 Times Jump)						
HAR	0.9748	0.1258 <sup>c,d</sup>	0.9750	0.1262 <sup>c,d</sup>	0.9639	0.1168
HAR-A	0.9755 <sup>†</sup>	0.1224 <sup>a,c,d</sup>	0.9756 <sup>†</sup>	0.1228 <sup>a,c,d</sup>	0.9679 <sup>†</sup>	0.1158
HAR-TCJ	0.9692	0.1592	0.9693	0.1595	0.9601	0.1546
HAR-TCJA	0.9697	0.1557 <sup>c</sup>	0.9698	0.1559 <sup>c</sup>	0.9647	0.1510
Co-Volatility: Crude Oil & Gold (616 Times Co-Jump)						
HAR	0.8884 <sup>†</sup>	9.2664	0.8695	11.509	0.9138 <sup>†</sup>	2.5693
HAR-A	0.8876	8.2840	0.8709 <sup>†</sup>	10.254	0.9102	2.5458
HAR-TCJ	0.8567	4.7787	0.8391	5.0081	0.8834	4.3606
HAR-TCJA	0.8564	4.0257	0.8409	3.9249	0.8799	4.1917

Note: The table presents  $R^2$  for the MZ regression and heteroskedasticity-adjusted root mean squared error (HRMSE).  $R^2$ -J (HRMSE-J) is  $R^2$  (HRMSE) conditionally on having a jump at time  $t - 1$ , respectively, while  $R^2$ -C (HRMSE-C) is conditional on no jump at time  $t - 1$ . '†' indicates the model which has the highest  $R^2$  value of the four models. For the DM test of equal forecast accuracy, 'a', 'b', 'c', and 'd' denote significant improvements in forecasting performance with respect to the HAR, HAR-A, HAR-TCJ and HAR-TCJA models, respectively.

**Table 7.** In-sample estimates for monthly regressions.

$$\begin{aligned}
 \text{HAR} & \quad \tilde{\Sigma}_{ij,t(22)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + e_{ij,t} \\
 \text{HAR-A} & \quad \tilde{\Sigma}_{ij,t(22)} = \gamma_0 + \gamma_d \tilde{\Sigma}_{ij,t-1} + \gamma_w \tilde{\Sigma}_{ij,t-1(5)} + \gamma_m \tilde{\Sigma}_{ij,t-1(22)} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t} \\
 \text{HAR-TCJ} & \quad \tilde{\Sigma}_{ij,t(22)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + e_{ij,t} \\
 \text{HAR-TCJA} & \quad \tilde{\Sigma}_{ij,t(22)} = \gamma_0 + \gamma_d \tilde{C}_{ij,t-1} + \gamma_w \tilde{C}_{ij,t-1(5)} + \gamma_m \tilde{C}_{ij,t-1(22)} + \gamma_j \tilde{J}_{ij,t-1} + \gamma_a r_{i,t-1}^- r_{j,t-1}^- + e_{ij,t}
 \end{aligned}$$

Model	$\gamma_0$	$\gamma_d$	$\gamma_w$	$\gamma_m$	$\gamma_j$	$\gamma_a$	$R^2$	$\bar{R}^2$
Volatility: Crude Oil								
HAR	0.0010 (0.0001)	0.0148 (0.0003)	0.0261 (0.0003)	0.9573 (0.0005)			0.9913	0.9913
HAR-A	0.0015 (0.0001)	0.0049 (0.0002)	0.0305 (0.0003)	0.9584 (0.0005)		0.0059 (0.0001)	0.9917	0.9917 <sup>†</sup>
HAR-TCJ	0.0441 (0.0003)	0.0117 (0.0004)	0.0430 (0.0008)	1.0977 (0.0016)	0.0752 (0.0009)		0.9741	0.9740
HAR-TCJA	0.0444 (0.0003)	-0.0031 (0.0003)	0.0503 (0.0008)	1.0992 (0.0016)	0.0760 (0.0008)	0.0066 (0.0002)	0.9746	0.9745
Volatility: Gold								
HAR	-0.0017 (0.0003)	0.0201 (0.0003)	0.0288 (0.0004)	0.9518 (0.0005)			0.9931	0.9931
HAR-A	-0.0033 (0.0003)	0.0157 (0.0002)	0.0311 (0.0004)	0.9511 (0.0005)		0.0057 (0.0001)	0.9936	0.9936 <sup>†</sup>
HAR-TCJ	0.0819 (0.0004)	0.0198 (0.0005)	0.0335 (0.0007)	1.0575 (0.0008)	0.0850 (0.0008)		0.9820	0.9819
HAR-TCJA	0.0805 (0.0004)	0.0154 (0.0004)	0.0360 (0.0007)	1.0569 (0.0008)	0.0829 (0.0007)	0.0052 (0.0001)	0.9824	0.9823
Co-Volatility: Crude Oil & Gold								
HAR	-0.0004 (0.0001)	0.0166 (0.0002)	0.0363 (0.0003)	0.9482 (0.0005)			0.9922	0.9922
HAR-A	-0.0007 (0.0001)	0.0133 (0.0002)	0.0369 (0.0003)	0.9500 (0.0005)		0.0025 (0.0001)	0.9924	0.9924 <sup>†</sup>
HAR-TCJ	0.0184 (0.0001)	0.0131 (0.0004)	0.0279 (0.0006)	1.1670 (0.0011)	0.0551 (0.0006)		0.9790	0.9790
HAR-TCJA	0.0181 (0.0001)	0.0093 (0.0004)	0.0288 (0.0006)	1.1690 (0.0011)	0.0536 (0.0006)	0.0023 (0.0001)	0.9792	0.9791

Note: Standard errors are given in parentheses. '†' denotes the model which has the highest  $\bar{R}^2$  value of the four models.

**Table 8.** Out-of-sample forecast evaluation for monthly regressions.

Model	MZ $R^2$	HRMSE	$R^2$ -J	HRMSE-J	$R^2$ -C	HRMSE-C
Volatility: Crude Oil (900 Times Jump)						
HAR	0.9794	0.0542 <sup>c,d</sup>	0.9786	0.0552 <sup>c,d</sup>	0.9856	0.0424 <sup>c,d</sup>
HAR-A	0.9798 <sup>†</sup>	0.0542 <sup>c,d</sup>	0.9790 <sup>†</sup>	0.0551 <sup>c,d</sup>	0.9856 <sup>†</sup>	0.0420 <sup>c,d</sup>
HAR-TCJ	0.8563	0.1386	0.8571	0.1384	0.8625	0.1404
HAR-TCJA	0.8581	0.1381	0.8589	0.1379	0.8640	0.1403
Volatility: Gold (929 Times Jump)						
HAR	0.9978	0.0340 <sup>c,d</sup>	0.9978	0.0342 <sup>c,d</sup>	0.9975	0.0299 <sup>c,d</sup>
HAR-A	0.9978 <sup>†</sup>	0.0339 <sup>c,d</sup>	0.9978 <sup>†</sup>	0.0341 <sup>c,d</sup>	0.9976 <sup>†</sup>	0.0299 <sup>c,d</sup>
HAR-TCJ	0.9944	0.0864	0.9945	0.0865	0.9933	0.0834
HAR-TCJA	0.9945	0.0856	0.9945	0.0858	0.9939	0.0827
Co-Volatility: Crude Oil & Gold (616 Times Co-Jump)						
HAR	0.9888	3.0184	0.9855	1.8945	0.9941	4.3019
HAR-A	0.9888 <sup>†</sup>	2.9907	0.9856 <sup>†</sup>	1.9492	0.9941 <sup>†</sup>	4.2070
HAR-TCJ	0.9668	8.7480	0.9652	3.8293	0.9698	13.483
HAR-TCJA	0.9670	8.6250	0.9657	3.7655	0.9696	13.299

Note: The table presents  $R^2$  for the MZ regression and heteroskedasticity-adjusted root mean squared error (HRMSE).  $R^2$ -J (HRMSE-J) is  $R^2$  (HRMSE) conditionally on having a jump at time  $t - 1$ , respectively, while  $R^2$ -C (HRMSE-C) is conditional on no jump at time  $t - 1$ . '†' indicates the model which has the highest  $R^2$  value of the four models. For the Diebold-Mariano test of equal forecast accuracy, 'a', 'b', 'c', and 'd' denote significant improvements in forecasting performance with respect to the HAR, HAR-A, HAR-TCJ and HAR-TCJA respectively.

**Table 9.** Out-of-sample forecast evaluation for minimum variance portfolio weights.

Model	Daily Regression		Weekly Regression		Monthly Regression	
	MZ $R^2$	HRMSE	MZ $R^2$	HRMSE	MZ $R^2$	HRMSE
HAR	0.8202	0.1801	0.9938	0.0231 <sup>c,d</sup>	0.9998	0.0040 <sup>c,d</sup>
HAR-A	0.8315	0.1709	0.9940 <sup>†</sup>	0.0228 <sup>c,d</sup>	0.9999 <sup>†</sup>	0.0040 <sup>c,d</sup>
HAR-TCJ	0.8298	0.1790	0.9919	0.0294	0.9994	0.0101
HAR-TCJA	0.8390 <sup>†</sup>	0.1775	0.9921	0.0291	0.9994	0.0100

The table reports  $R^2$  of the MZ regression and heteroskedasticity-adjusted root mean squared error (HRMSE). '†' denotes the model which has the highest  $R^2$  value of the four models. For the DM test of equal forecast accuracy, 'a', 'b', 'c', and 'd' denote significant improvements in forecasting performance with respect to the HAR, HAR-A, HAR-TCJ and HAR-TCJA respectively.

#### 4. Concluding Remarks

The paper investigated the impacts of co-jumps and leverage of crude oil and gold futures in forecasting co-volatility. We use the approach of Reference [24] for disentangling the estimates of the integrated co-volatility matrix and jump variations so that they are positive (semi-) definite for coherence of the estimator. The empirical results showed that more than 80% of the co-volatility of the two futures has non-negligible jump co-variations, and that the co-jumps of the two futures have significant impacts on future co-volatility, but that the impacts are minor in forecasting weekly and monthly horizons.

The empirical results also showed that the impacts of the co-leverage effects caused by the negative returns of two assets are significant, but the impact decreases in forecasting longer horizons. The results of in-sample and out-of-sample forecasts showed that the datasets generally prefer the HAR-A model, which was not used in Reference [24].

Overall, the analysis given in the paper should be useful for investment analysis, and both public and private policy prescription, in terms of accommodating jumps and leverage, as well as choice of models and time frequency horizons, in forecasting the co-volatility of oil and gold futures. Extensions to other financial, precious and semi-precious metals, and alternative sources of renewable and non-renewable energy, are the subject of ongoing research.

**Author Contributions:** These authors contributed equally to this work.

**Funding:** The first author acknowledges the financial support of the Japan Ministry of Education, Culture, Sports, Science and Technology, Japan Society for the Promotion of Science (JSPS KAKENHI Grant Number 19K01594), and Australian Academy of Science. The third author is most grateful for the financial support of the Australian Research Council, Ministry of Science and Technology (MOST), Taiwan, and the Japan Society for the Promotion of Science.

**Acknowledgments:** The authors are most grateful to Yoshi Baba and four anonymous reviewers for very helpful comments and suggestions.

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

## Abbreviations

The following abbreviations are used in this manuscript:

DM	Diebold-Mariano
HAR	Heterogeneous Autoregression
HRMSE	Heteroskedasticity-adjusted Root Mean Squared Error
ICov	Integrated Co-volatility
MZ	Mincer-Zarnowitz
OLS	Ordinary Least Squares
QCov	Quadratic Covariation
RK	Realized Kernel

## Appendix A

We explain the calculation of the estimators of QCov and ICov in (4), and jump components in returns. Parts of the following technical section follow [24] closely.

### Appendix A.1. Estimator of Quadratic Covariation

We introduce below the pre-averaged Hayashi–Yoshida (PHY) estimator, proposed by [30] for improving the estimator of [46] for non-synchronized trading times. The PHY estimator is a consistent estimator of QCov under non-synchronized trading times, jumps and microstructure noise, as shown by [29].

Consider a sequence,  $c_n$ , of integers and a number,  $\kappa_0 \in (0, \infty)$ , satisfying  $c_n = \kappa_0 \sqrt{n} + o(n^{1/4})$ , where  $n$  is the observation frequency. Assume that a continuous function  $g : [0, 1] \rightarrow \mathfrak{R}$  is piecewise  $C^1$  with piecewise Lipschitz derivative  $g'$  that satisfies  $g(0) = g(1) = 0$  and  $\kappa_{HY} = \int_0^1 g(x) dx \neq 0$ . We also consider pre-averaged observation data of  $X$  and  $Y$  based on the sampling designs  $\mathcal{T}$  and  $\mathfrak{E}$  as:

$$\begin{aligned}\bar{X}^i &= \sum_{p=1}^{c_n-1} g\left(\frac{p}{c_n}\right) \left(X_{\zeta_{i+p}} - X_{\zeta_{i+p-1}}\right) \\ \bar{Y}^j &= \sum_{q=1}^{c_n-1} g\left(\frac{q}{c_n}\right) \left(Y_{\zeta_{j+q}} - X_{\zeta_{j+q-1}}\right), \quad i, j = 0, 1, \dots\end{aligned}$$

Then the PHY estimator is defined by:

$$\mathbb{P}HY(X, Y) = \frac{1}{(c_n \kappa_{HY})^2} \sum_{\substack{i, j=0 \\ \max(\zeta_{i+c_n}, \zeta_{j+c_n}) \leq T}}^{\infty} \bar{X}^i \bar{Y}^j \bar{K}^{ij}, \quad (\text{A1})$$

where  $\bar{K}^{ij} = \mathbf{1}([\zeta_i, \zeta_{i+c_n}] \cap [\zeta_j, \zeta_{j+c_n}] \neq \emptyset)$ . Under non-synchronized trading times and microstructure noise without jumps, the PHY estimator has the consistency and asymptotic mixed normality, as shown by [30]. For the quadratic covariation of  $(X, Y)$ , ref. [29] shows that the PHY estimator

is consistent. Following [29], we use the specifications  $g(x) = \min(x, 1 - x)$ ,  $c_n = \lceil \kappa_0 \sqrt{n} \rceil$  and  $\kappa_0 = 0.15$ .

We can obtain the estimator of quadratic variation of  $X$ , by using  $\mathbb{P}\text{HY}(X, X)$ . Hence, we can construct a consistent estimator of QCov for the bivariate case. As noted above, the estimator of QCov is generally positive definite, as it is the sample analogue of preaveraged data.

#### Appendix A.2. Estimator of Integrated Co-Volatility

This section explains the estimator of [29] for ICov under non-synchronized trading times, jumps and microstructure noise. A truncation technique is used by [29] for removing the jump components in order to obtain the pre-averaged truncated Hayashi–Yoshida (PTHY) estimator:

$$\mathbb{P}\text{THY}(X, Y) = \frac{1}{(c_n \kappa_{HY})^2} \sum_{i,j=0}^{\infty} \bar{X}^i \bar{Y}^j \bar{K}^{ij} \bar{\Psi}^{ij}, \quad (\text{A2})$$

$$\max(\zeta_{i+c_n}, \zeta_{j+c_n}) \leq T$$

where  $\bar{\Psi}^{ij} = \mathbf{1}(|\bar{X}^i|^2 \leq \omega^X(\zeta_i), |\bar{Y}^j|^2 \leq \omega^Y(\zeta_j))$ , and  $\omega^X(t)$  and  $\omega^Y(t)$  are sequences of positive-valued stochastic process. Note that a similar idea is proposed by [47] for the truncation in the univariate case. The consistency and asymptotic mixed normality of the PTHY estimator were shown by [29]. Furthermore, the consistency of the quadratic co-variation of the jump component was shown by [29], using the difference between the PHY and PTHY estimators.

For the process of the threshold value of  $X$ :

$$\omega^X(\zeta_i) = 2 \log(N)^{1+\epsilon} \hat{\sigma}_{\zeta_i}^2,$$

was used by [36] with  $\epsilon = 0.2$ , where:

$$\hat{\sigma}_{\zeta_i}^2 = \frac{\mu_1^{-2}}{M - 2c_n + 1} \sum_{p=i-M}^{i-2c_n} |\bar{X}^p| |\bar{X}^{p+c_n}|, \quad i = M, M+1, \dots, N,$$

and  $\hat{\sigma}_{\zeta_i}^2 = \hat{\sigma}_{\zeta_M}^2$  if  $i < M$ . Here,  $\mu_1 = \sqrt{2/\pi}$ ,  $M = \lceil N^{3/4} \rceil$ , and  $N$  is the number of the available pre-averaged data  $\bar{X}^i$ . We can obtain  $\omega^Y(\zeta_j)$  in the same manner.

We construct a consistent estimator of ICov for the bivariate case,  $\hat{C}_t$ , using the estimators based on  $\mathbb{P}\text{THY}(X, X)$  and  $\mathbb{P}\text{THY}(X, Y)$ . In order to guarantee the positive semi-definiteness of the estimators of ICov and the jump component, we use the approach of [22], as explained above.

Using estimation techniques for  $\mathbb{P}\text{THY}(X, X)$  and  $\mathbb{P}\text{THY}(X, Y)$  (Equation (A2)), we can construct a consistent estimator of the  $2 \times 2$  integrated co-volatility matrix at day  $t$ ,  $\hat{C}_t$ , under jumps and microstructure noise. We can also obtain the estimator of QCov, which we denote as  $\hat{\Sigma}_t$ , by using  $\mathbb{P}\text{HY}(X, X)$  and  $\mathbb{P}\text{HY}(X, Y)$  (Equation (A1)), which yields the jump estimator,  $\hat{J}_t = \hat{\Sigma}_t - \hat{C}_t$ . We obtain the final estimates as  $\tilde{\Sigma}_t = \mathcal{T}_h(\hat{\Sigma}_t)$ ,  $\tilde{C}_t = \mathcal{T}_h(\hat{C}_t)$  and  $\tilde{J}_t = \mathcal{T}_h(\hat{J}_t)$ , applying the threshold operator defined by (5).

#### Appendix A.3. Estimation of Jump Component in Returns

A simple methodology was suggested by [36] to decompose asset returns sampled at a high frequency into their base components (continuous and jumps). For the process of log-prices,  $X$ , the return is defined by  $r^X = \sum_{i=2}^{mT} (X_{\zeta_i} - X_{\zeta_{i-1}}) = X_m - X_1$ , where  $m = \max\{i : \zeta_i \leq T\}$ . Following the idea of [36], we define continuous, jump, and noise components of the return by:

$$rc^X = \frac{1}{c_n \kappa_{HY}} \sum_{i=0}^{m-c_n} \bar{X}^i \mathbf{1}(|\bar{X}^i|^2 \leq \omega^X(\zeta_i)),$$

$$rj^X = \frac{1}{c_n \kappa_{HY}} \sum_{i=0}^{m-c_n} \bar{X}^i \mathbf{1}(|\bar{X}^i|^2 \geq \omega^X(\zeta_i)),$$

$$rn^X = r^X - rc^X - rj^X.$$

The decomposition of  $R$  into  $RC$ ,  $RJ$ , and  $RN$  corresponds to that of  $\tilde{\Sigma}_t$  into  $\tilde{C}_t$ ,  $\tilde{J}_t$ , and  $\tilde{\Sigma}_t - \tilde{C}_t - \tilde{J}_t$ . In the empirical analysis, we use the continuous component of the return for the HAR-TCJA model, and we use the observed return for the HAR-A model.

## References

- Muteba Mwamba, J.W.; Hammoudeh, S.; Gupta, R. Financial Tail Risks in Conventional and Islamic Stock Markets: A Comparative Analysis. *Pac. Basin Financ. J.* **2017**, *42*, 60–82. [[CrossRef](#)]
- Bahloul, W.; Balcilar, M.; Cunado, J.; Gupta, R. The Role of Economic and Financial Uncertainties in Predicting Commodity Futures Returns and Volatility: Evidence from a Nonparametric Causality-in-Quantiles Test. *J. Multinatl. Financ. Manag.* **2018**, *45*, 52–71. [[CrossRef](#)]
- Silvennoinen, A.; Thorp, S. Financialization, Crisis and Commodity Correlation Dynamics. *J. Int. Financ. Mark. Inst. Money* **2013**, *24*, 42–65. [[CrossRef](#)]
- Tang, K.; Xiong, W. Index Investment and the Financialization of Commodities. *Financ. Anal. J.* **2012**, *68*, 54–74. [[CrossRef](#)]
- Büyüksahin, B.; Robe, M.A. Speculators, Commodities and Crossmarket Linkages. *J. Int. Money Financ.* **2014**, *42*, 38–70. [[CrossRef](#)]
- Fattouh, B.; Kilian, L.; Mahadeva, L. The Role of Speculation in Oil Markets: What Have We Learned So Far? *Energy J.* **2013**, *34*, 7–33. [[CrossRef](#)]
- Baur, D.G.; Lucey, B.M. Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold. *Financ. Rev.* **2010**, *45*, 217–229. [[CrossRef](#)]
- Baur, D.G.; McDermott, T.K. Is Gold a Safe Haven? International Evidence. *J. Bank. Financ.* **2010**, *34*, 1886–1898. [[CrossRef](#)]
- Reboredo, J.C. Is Gold a Safe Haven or a Hedge for the US Dollar? Implications for Risk Management. *J. Bank. Financ.* **2013**, *37*, 2665–2676. [[CrossRef](#)]
- Ewing, B.T.; Malik, F. Volatility Transmission between Gold and Oil Futures under Structural Breaks. *Int. Rev. Econ. Financ.* **2013**, *25*, 113–121. [[CrossRef](#)]
- Mensi, W.; Beljid, M.; Boubaker, A.; Managi, S. Correlations and Volatility Spillovers across Commodity and Stock Markets: Linking Energies, Food, and Gold. *Econ. Model.* **2013**, *32*, 15–22. [[CrossRef](#)]
- Yaya, O.S.; Tumala, M.M.; Udomboso, C.G. Volatility Persistence and Returns Spillovers between Oil and Gold Prices: Analysis Before and After the Global Financial Crisis. *Resour. Policy* **2016**, *49*, 273–281. [[CrossRef](#)]
- Chang, C.-L.; Li, Y.-Y.; McAleer, M. Volatility Spillovers between Energy and Agricultural Markets: A Critical Appraisal of Theory and Practice. *Energies* **2018**, *11*, 1595. [[CrossRef](#)]
- Reboredo, J.C. Is Gold a Hedge or Safe Haven against Oil Price Movements? *Resour. Policy* **2013**, *38*, 130–137. [[CrossRef](#)]
- Bampinas, G.; Panagiotidis, T. On the Relationship between Oil and Gold before and after Financial Crisis: Linear, Nonlinear and Time-Varying Causality Testing. *Stud. Nonlinear Dyn. Econom.* **2015**, *19*, 657–668. [[CrossRef](#)]
- Balcilar, M.; Ozdemir, Z.A.; Shahbaz, M. On the Time-Varying Links between Oil and Gold: New Insights from the Rolling and Recursive Rolling Approaches. *Int. J. Financ. Econ.* **2019**, *24*, 1047–1065. [[CrossRef](#)]
- Coronado, S.; Jiménez-Rodríguez, R.; Rojas, O. An Empirical Analysis of the Relationships between Crude Oil, Gold and Stock Markets. *Energy J.* **2018**, *39*, 193–207. [[CrossRef](#)]
- Tiwari, A.K.; Cunado, J.; Gupta, R.; Wohar, M.E. Volatility Spillovers across Global Asset Classes: Evidence from Time and Frequency Domains. *Q. Rev. Econ. Financ.* **2018**, *70*, 194–202. [[CrossRef](#)]

19. Chang, C.-L.; McAleer, M.; Wang, Y. Testing Co-volatility Spillovers for Natural Gas Spot, Futures and ETF Spot using Dynamic Conditional Covariances. *Energy* **2018**, *151*, 984–997. [[CrossRef](#)]
20. Degiannakis, S.; Filis, G. Forecasting Oil Price Realized Volatility using Information Channels from Other Asset Classes. *J. Int. Money Financ.* **2017**, *76*, 28–49. [[CrossRef](#)]
21. Andersen, T.G.; Bollerslev, T.; Diebold, F.X. Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility. *Rev. Econ. Stat.* **2007**, *89*, 701–720. [[CrossRef](#)]
22. Bollerslev, T.; Kretschmer, U.; Pigorsch, C.; Tauchen, G. A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects. *J. Econom.* **2009**, *150*, 151–166.
23. Corsi, F.; Pirino, D.; Renò, R. Threshold Bipower Variation and The Impact of Jumps on Volatility Forecasting. *J. Econom.* **2010**, *159*, 276–288. [[CrossRef](#)]
24. Asai, M.; McAleer, M. The Impact of Jumps and Leverage in Forecasting Co-Volatility. *Econ. Rev.* **2017**, *36*, 638–650. [[CrossRef](#)]
25. Andersen, T.; Bollerslev, T. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *Int. Econ. Rev.* **1998**, *39*, 657–668. [[CrossRef](#)]
26. Demirer, R.; Gkillas, K.; Gupta, R.; Pierdzioch, C. Time-varying Risk Aversion and Realized Gold Volatility. *N. Am. J. Econ. Financ.* **2019**, *50*, 101048. [[CrossRef](#)]
27. Prokopczuk, M.; Symeonidis, L.; Wese Simen, C. Do Jumps Matter for Volatility Forecasting? Evidence from Energy Markets. *J. Futures Mark.* **2015**, *36*, 758–792. [[CrossRef](#)]
28. Sévi, B. Forecasting the Volatility of Crude Oil Futures using Intraday Data. *Eur. J. Oper. Res.* **2014**, *235*, 643–659. [[CrossRef](#)]
29. Koike, Y. Estimation of Integrated Covariances in the Simultaneous Presence of Non-synchronicity, Microstructure Noise and Jumps. *Econom. Theory* **2016**, *32*, 533–611. [[CrossRef](#)]
30. Christensen, K.; Kinnebrock, S.; Podolskij, M. Pre-Averaging Estimators of the Ex-Post Covariance Matrix in Noisy Diffusion Models with Non-Synchronous Data. *J. Econom.* **2012**, *159*, 116–133. [[CrossRef](#)]
31. Barndorff-Nielsen, O.E.; Hansen, P.R.; Lunde, A.; Shephard, N. Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of The Covariation of Equity Prices with Noise and Non-Synchronous Trading. *J. Econom.* **2011**, *162*, 149–169. [[CrossRef](#)]
32. Christensen, K.; Oomen, R.; Podolskij, M. Fact or Friction, Jumps at Ultra High Frequency. *J. Financ. Econ.* **2018**, *114*, 576–599. [[CrossRef](#)]
33. Bickel, P.J.; Levina, E. Regularized Estimation of Large Covariance Matrices. *Ann. Stat.* **2008**, *36*, 199–277. [[CrossRef](#)]
34. Bickel, P.J.; Levina, E. Covariance Regularization by Thresholding. *Ann. Stat.* **2008**, *36*, 2577–2604. [[CrossRef](#)]
35. Tao, M.; Wang, Y.; Yao, Q.; Zou, J. Large Volatility Matrix Inference via Combining Low-Frequency and High-Frequency Approaches. *J. Am. Stat. Assoc.* **2011**, *106*, 1025–1040. [[CrossRef](#)]
36. Ait-Sahalia, Y.; Jacod, J. Analyzing the Spectrum of Asset Returns: Jump and Volatility Components in High Frequency Data. *J. Econ. Lit.* **2012**, *50*, 1007–1050. [[CrossRef](#)]
37. Kroner, K.F.; Ng, V.K. Modeling Asymmetric Co-movements of Assets Returns. *Rev. Financ. Stud.* **1998**, *11*, 817–844. [[CrossRef](#)]
38. Newey, W.K.; West, V.K. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* **1987**, *55*, 703–708. [[CrossRef](#)]
39. Smales, L.A. News Sentiment in the Gold Futures Market. *J. Bank. Financ.* **2014**, *49*, 275–286. [[CrossRef](#)]
40. Corsi, F.; Renò, R. Discrete-time Volatility Forecasting with Persistent Leverage Effect and the Link with Continuous-time Volatility Modeling. *J. Bus. Econ. Stat.* **2012**, *30*, 368–380. [[CrossRef](#)]
41. Wen, F.; Gong, X.; Cai, S. Forecasting the Volatility of Crude Oil Futures using HAR-Type Models with Structural Breaks. *Energy Econ.* **2016**, *59*, 400–413. [[CrossRef](#)]
42. Mincer, J.A.; Zarnowitz, V. The Evaluation of Economic Forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*; Mincer, J.A., Ed.; National Bureau of Economic Research: New York, NY, USA, 1969; pp. 3–46.
43. Bollerslev, T.; Ghysels, E. Periodic Autoregressive Conditional Heteroscedasticity. *J. Bus. Econ. Stat.* **1996**, *14*, 139–151.
44. Diebold, F.; Mariano, R. Comparing Predictive Accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.



45. Patton, A.J.; Sheppard, K. Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility. *Rev. Econ. Stat.* **2015**, *97*, 683–697. [[CrossRef](#)]
46. Hayashi, T.; Yoshida, N. On Covariance Estimation of Nonsynchronously Observed Diffusion Processes. *Bernoulli* **2005**, *11*, 359–379. [[CrossRef](#)]
47. Aït-Sahalia, Y.; Jacod, J.; Li, J. Testing for Jumps in Noisy High Frequency Data. *J. Econom.* **2012**, *168*, 207–322. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).