ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

https://www.econometricsociety.org/

Econometrica, Vol. 88, No. 4 (July, 2020), 1515–1551

REALIZED SEMICOVARIANCES

TIM BOLLERSLEV
Department of Economics, Duke University, NBER, and CREATES

JIA LI
Department of Economics, Duke University

ANDREW J. PATTON
Department of Economics, Duke University

ROGIER QUAEDVLIEG
Department of Business Economics, Erasmus University Rotterdam

Econometrica, Vol. 88, No. 4 (July, 2020), 1515–1551

REALIZED SEMICOVARIANCES

TIM BOLLERSLEV
Department of Economics, Duke University, NBER, and CREATES

JIA LI
Department of Economics, Duke University

ANDREW J. PATTON
Department of Economics, Duke University

ROGIER QUAEDVLIEG
Department of Business Economics, Erasmus University Rotterdam

We propose a decomposition of the realized covariance matrix into components based on the signs of the underlying high-frequency returns, and we derive the asymptotic properties of the resulting *realized semicovariance* measures as the sampling interval goes to zero. The first-order asymptotic results highlight how the same-sign and mixed-sign components load differently on economic information related to stochastic correlation and jumps. The second-order asymptotic results reveal the structure underlying the same-sign semicovariances, as manifested in the form of co-drifting and dynamic "leverage" effects. In line with this anatomy, we use data on a large cross-section of individual stocks to empirically document distinct dynamic dependencies in the different realized semicovariance components. We show that the accuracy of portfolio return variance forecasts may be significantly improved by exploiting the information in realized semicovariances.

KEYWORDS: High-frequency data, realized variances, semicovariances, co-jumps, volatility forecasting.

1. INTRODUCTION

THE COVARIANCE MATRIX OF ASSET RETURNS arguably constitutes the most crucial input for asset pricing, portfolio, and risk management decisions. Correspondingly, there is a substantial literature devoted to the estimation, modeling, and prediction of covariance matrices dating back more than half a century (e.g., Kendall (1953), Elton and Gruber (1973), and Bauwens, Laurent, and Rombouts (2006)). Meanwhile, a rapidly-growing recent literature has forcefully advocated the use of high-frequency intraday data to more reliably estimate lower-frequency return covariance matrices (e.g., Andersen, Bollerlsev,

Tim Bollerslev: boller@duke.edu
Jia Li: jl410@duke.edu
Andrew J. Patton: andrew.patton@duke.edu
Rogier Quaedvlieg: quaedvlieg@ese.eur.nl
We would like to thank the four anonymous referees for their helpful comments, which greatly improved the paper. We would also like to thank conference and seminar participants at Aarhus, Boston, Ca'Foscari Venice, Chicago, Cologne, Gerzensee, ITAM, Konstanz, Lancaster, Padova, Pennsylvania, PUC Rio, QUT, Rutgers, and Toulouse for helpful comments and suggestions. Bingzhi Zhao kindly provided us with the cleaned high-frequency data underlying our empirical investigations. Patton and Quaedvlieg gratefully acknowledge support from, respectively, Australian Research Council Discovery Project 180104120 and Netherlands Organisation for Scientific Research Grant 451-17-009. This paper subsumes the 2017 paper "Realized Semicovariances: Looking for Signs of Direction Inside the Covariance Matrix" by Bollerslev, Patton, and Quaedvlieg.

 https://doi.org/10.3982/ECTA17056

Diebold, and Labys (2003), Barndorff-Nielsen and Shephard (2004b), and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2011)).

Set against this background, we propose a decomposition of the realized covariance matrix into three *realized semicovariance matrix* components dictated by the signs of the underlying high-frequency returns. The realized semicovariance matrices may be seen as a high-frequency multivariate extension of the semivariances originally proposed in the finance literature (e.g., Markowitz (1959), Mao (1970), Hogan and Warren (1972, 1974), and Fishburn (1977)). Our high-frequency theoretical analysis is inspired by and extends the pioneering work of Barndorff-Nielsen, Kinnebrock, and Shephard (2010).

To fix ideas, let $X_t = (X_{1,t}, \dots, X_{d,t})^\top$ denote a d-dimensional log-price process, sampled on a regular time grid $\{i\Delta_n : 0 \le i \le [T/\Delta_n]\}$ over some fixed time span $T > 0$. Let the ith return of X be denoted by $\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}$. The *realized covariance matrix* (Barndorff-Nielsen and Shephard (2004b)) is then defined as

$$\widehat{C} \equiv \sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n X)(\Delta_i^n X)^\top.$$

If we let $p(x) \equiv \max\{x, 0\}$ and $n(x) \equiv \min\{x, 0\}$ denote the component-wise positive and negative elements of the real vector x, the corresponding "positive," "negative," and "mixed" *realized semicovariance matrices* are then simply defined as

$$\begin{aligned} \widehat{P} &\equiv \sum_{i=1}^{[T/\Delta_n]} p(\Delta_i^n X) p(\Delta_i^n X)^\top, \qquad \widehat{N} \equiv \sum_{i=1}^{[T/\Delta_n]} n(\Delta_i^n X) n(\Delta_i^n X)^\top, \\ \widehat{M} &\equiv \sum_{i=1}^{[T/\Delta_n]} \big(p(\Delta_i^n X) n(\Delta_i^n X)^\top + n(\Delta_i^n X) p(\Delta_i^n X)^\top \big). \end{aligned} \tag{1}$$

Note that $\widehat{C} = \widehat{P} + \widehat{N} + \widehat{M}$ for any sampling frequency Δ_n. The concordant realized semicovariance matrices, $\widehat{P}$ and $\widehat{N}$, are defined as sums of vector outer-products and thus are positive semidefinite. By contrast, the mixed semicovariance matrix, $\widehat{M}$, has diagonal elements that are identically zero, and thus is necessarily indefinite.

As an illustration of the different dynamic dependencies and information conveyed by the realized semicovariances, Figure 1 plots the daily realized covariance averaged across 500 randomly-selected pairs of S&P 500 stocks, together with its concordant ($\widehat{P} + \widehat{N}$) and mixed ($\widehat{M}$) semicovariance components.[1] The mixed component is, of course, always negative, while the concordant component is always positive. The two components are typically similar in magnitude during "normal" time periods, while in periods of high volatility the concordant component increases substantially more than the mixed component declines, in line with the widely-held belief that during periods of financial market stress, correlations between financial assets tend to increase. As such, the (total) realized covariance is largely determined by the concordant realized semicovariance components in these "crisis" periods.

[1]Each day, we randomly draw 500 pairs of assets and compute $\overline{C} \equiv (1/500) \sum_{j \neq k} \widehat{C}_{jk}$. $\overline{P}$, $\overline{N}$, and $\overline{M}$ are computed similarly. A detailed description of the data is provided in Section 4 below. To avoid cluttering the figure, we sum $\overline{P}$ and $\overline{N}$ into a single concordant component, and smooth the daily measures using a day $t - 25$ to day $t + 25$ moving average.

To help understand these empirical features, consider a simple setting in which the vector log-price process X_t is generated by a Brownian motion with constant drift b, unit volatility, and constant correlation ρ. Although stylized, this simple model captures the central force in the first-order asymptotic behavior of the semicovariance estimators in the no-jump setting; we provide a more general analysis below. Normalizing $T = 1$, by the law of large numbers the probability limits (as $\Delta_n \to 0$) of the (j, k) off-diagonal elements of the realized semicovariance matrices are then given by

$$\operatorname{plim} \widehat{P}_{jk} = \operatorname{plim} \widehat{N}_{jk} = \psi(\rho), \qquad \operatorname{plim} \widehat{M}_{jk} = -2\psi(-\rho),$$

where

$$\psi(\rho) = (2\pi)^{-1}\big(\rho \arccos(-\rho) + \sqrt{1-\rho^2}\big) \tag{2}$$

corresponds to $\mathbb{E}[Z_1 Z_2 1_{\{Z_1<0, Z_2<0\}}]$ for (Z_1, Z_2) bivariate standard normally distributed with correlation ρ. As these expressions illustrate, the limiting values of the semicovariance components depend crucially on the value of ρ. As ρ increases to 1, the limiting value of the concordant component $\widehat{P} + \widehat{N}$ also approaches 1, while the mixed component $\widehat{M}$ approaches zero, and vice versa when ρ decreases to -1. This is consistent with the empirical observation from Figure 1 that the concordant component accounts for most of the covariance in periods of market stress, which are generally believed to be accompanied by increased positive correlations.

This simple diffusive setting highlights the potentially different information conveyed by the concordant and the mixed semicovariance components. It does not, however, reveal any differences between the $\widehat{P}$ and $\widehat{N}$ components as they have the same limits in this stylized setting. This is at odds with the intuition that these measures carry distinct economic information about good and bad news over each day. To illustrate, Figure 2 presents high-frequency returns on each of the 30 Dow Jones Industrial Average (DJIA) stocks on two different days. On September 18, 2013 (left panel), the Federal Reserve announced that it would not taper its asset-purchasing program, in contrast to what the market had been anticipating, and individual stocks responded with positive jumps at

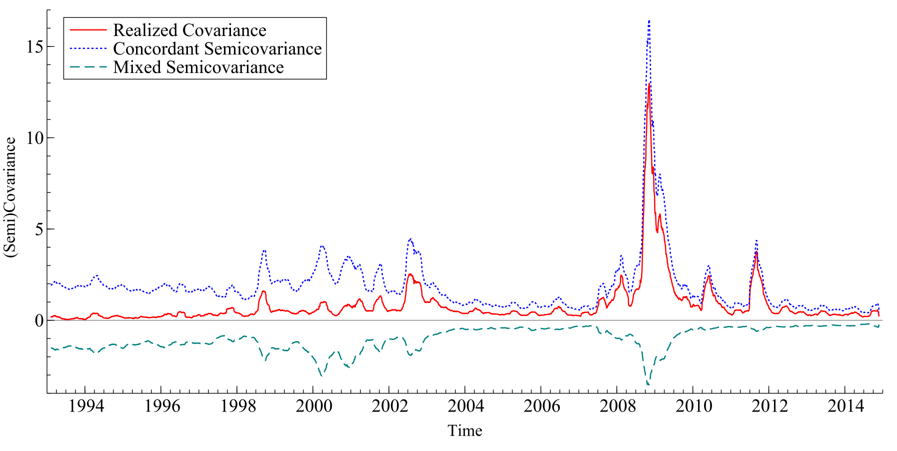

FIGURE 1.—Realized covariance decomposition. *Note*: The figure plots the time series of the concordant semicovariance ($\widehat{P} + \widehat{N}$), the mixed semicovariance ($\widehat{M}$), and the realized covariance ($\widehat{C}$). Each series is constructed as the moving average of the relevant daily realized measures averaged across 500 random pairs of S&P 500 stocks.

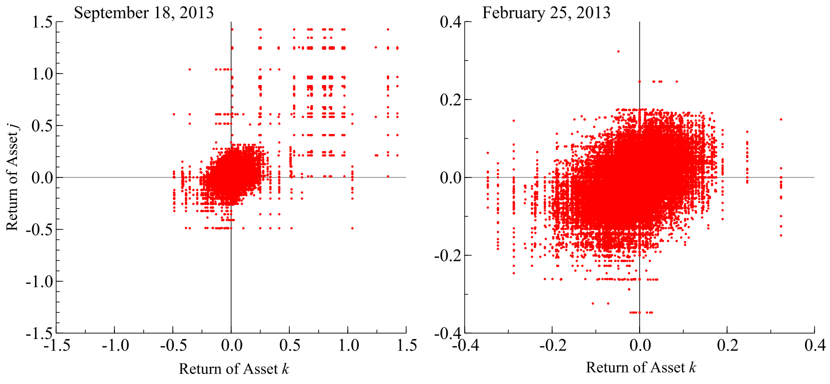

FIGURE 2.—Signed return-pairs for DJIA stocks. *Note*: The figure shows a scatter plot of the one-minute returns of each pair of the 30 Dow Jones Industrial Average stocks on two days in 2013. The left panel presents a day with an FOMC announcement that led to positive stock price jumps for many stocks. The right panel presents a day with steady downward price moves for many stocks.

the announcement time, resulting in much larger estimates of $\widehat{P}$ than $\widehat{N}$. By contrast, on February 25, 2013 (right panel), the DJIA drifted down by 1.5% over the course of the day amid concerns, according to market anecdotes,[2] about political uncertainty in Italy, and generated much larger estimates of $\widehat{N}$ than $\widehat{P}$. Hence, the empirical estimates of $\widehat{P}$ and $\widehat{N}$ can indeed differ depending on the directional content of the news, and whether it manifests in the form of price jumps or price drifts.

Motivated by these empirical observations, in Section 2 we derive first- and second-order asymptotics for the positive and the negative semicovariance estimators in a general Itô semimartingale setting, focusing particularly on a deeper understanding of the information they convey. Extending the work of Barndorff-Nielsen, Kinnebrock, and Shephard (2010), our limit theory identifies three channels through which $\widehat{P}$ and $\widehat{N}$ may differ: directional "co-jumps," a type of "co-drifting," and a specific form of "dynamic leverage effect." The co-jump channel manifests straightforwardly in the first-order asymptotics. In particular, assuming that the vector log-price process in the aforementioned example is subject to finite activity jumps, it follows readily that

$$\operatorname{plim} \widehat{P}_{jk} = \psi(\rho) + \sum_{0<s\leq 1} p(\Delta X_{j,s}) p(\Delta X_{k,s}),$$

$$\operatorname{plim} \widehat{N}_{jk} = \psi(\rho) + \sum_{0<s\leq 1} n(\Delta X_{j,s}) n(\Delta X_{k,s}),$$

where $\Delta X_{j,s}$ denotes the jump of the jth component of X at time s. By comparison, co-drifting and dynamic leverage effects manifest in second-order bias terms in a non-central limit theorem.

These terms are unique to our analysis of realized semicovariances, and, from a methodological perspective, they set our asymptotic analysis apart from standard high-frequency econometric analysis, in which central limit theorems are generally applied for the purpose of conducting statistical inference (see, e.g., Aït-Sahalia and Jacod (2014)). By contrast, the main purpose of our higher-order asymptotic results is to "dissect" the

[2]Source: https://money.cnn.com/2013/02/25/investing/stocks-markets/index.html.

semicovariance estimators, thereby allowing for additional theoretical and empirical insights by comparing and contrasting the relevant terms.

To gain more insights into the price behaviors illustrated in Figure 2, we employ a standard truncation technique (Mancini (2001)) to separately study the jump and diffusive information in positive and negative semicovariances. For the jump component, we establish a feasible central limit theorem that facilitates formal statistical inference. For the diffusive component, which is more complicated to study, we provide a standard error estimator that quantifies its sampling variability in a well-defined sense, and which, under more restrictive regularity conditions, also yields an asymptotically valid test.

Implementing the new inference procedures on high-frequency data for the 30 DJIA stocks over a nine-year period reveals strong evidence of significant differences in the $\widehat{P}$ and $\widehat{N}$ semicovariance components on many days. Consistent with economic intuition, large differences in the jump semicovariance components are typically associated with "sharp" public news announcements (e.g., FOMC announcements). Large differences in the diffusive semicovariance components are typically associated with more difficult-to-interpret news, which manifests in the form of common price drifts within the day.[3]

The above results naturally suggest that decomposing the realized covariance matrix into its semicovariance components may be useful for volatility forecasting. We investigate this idea by analyzing a large cross-section of stocks composed of all of the S&P 500 constituents. We show that out-of-sample forecasts of portfolio return volatility are significantly improved by using semicovariance measures. The gains from doing so increase with the number of stocks included in the portfolio, although in line with the gains from portfolio diversification, they appear to plateau at around 30–40 stocks in the portfolio. Further dissecting the forecasting gains, we find that the models that use realized semicovariances generally respond to new information faster than models that use only realized variances or semivariances (e.g., Corsi (2009) and Patton and Sheppard (2015)). Interestingly, during the financial crisis existing volatility models reduce the weight on recent information, whereas semicovariance-based models *increase* the weight, primarily due to an increase in the short-run importance of the negative semicovariance component.

The forecasting gains obtained through the use of the realized semicovariances are naturally linked to the early work on asymmetric volatility models (e.g., Kroner and Ng (1998) and Cappiello, Engle, and Sheppard (2006)). Our proposed tests for differences between positive and negative semicovariance components are also related to existing tests for asymmetric dependencies (e.g., Longin and Solnik (2001), Ang and Chen (2002), and Hong, Tu, and Zhou (2007)). Our work also connects to earlier empirical results on correlations between asset returns in "bear" versus "bull" markets, and notions of asymmetric tail dependencies (e.g., Patton (2004), Poon, Rockinger, and Tawn (2004), and Tjøstheim and Hufthammer (2013)), and to recent work on high-frequency based co-skewness and co-kurtosis measures (e.g., Neuberger (2012) and Amaya, Christoffersen, Jacobs, and Vasquez (2015)), and jumps and co-jumps (e.g., Das and Uppal (2004), Bollerslev, Law, and Tauchen (2008), Lee and Mykland (2008), Mancini and Gobbi (2012), Jacod and Todorov (2009), Aït-Sahalia and Xiu (2016), and Li, Todorov, and Tauchen (2017b)). In contrast to all of these studies, however, we retain the covariance matrix as the summary measure of dependence, and instead use information from signed high-frequency returns to "look inside" this matrix to obtain additional information about the inherent dependencies, both dynamically and cross-sectionally at a given point in time.

[3]The two days plotted in Figure 2 also correspond to these two scenarios, and are indeed detected by using the new inference method, as further discussed in Section 3 below.

The rest of the paper is organized as follows. Section 2 presents the first- and second-order asymptotic properties of the realized semicovariances. Readers primarily interested in empirical applications of the new measures may skip the more technical parts of Section 2. Section 3 discusses our empirical findings from the new semicovariance-based tests. Our results pertaining to the use of the realized semicovariances in the construction of improved volatility forecasts are discussed in Section 4. Section 5 concludes. Technical regularity conditions and proofs are deferred to the Appendix. Additional robustness checks and extensions are available in the Supplemental Material (Bollerslev, Li, Patton, and Quaedvlieg (2020)).

2. INFORMATION CONTENT OF REALIZED SEMICOVARIANCES: AN ASYMPTOTIC ANALYSIS

In this section, we demonstrate the differential information embedded in the realized semicovariance measures in equation (1) in an infill asymptotic framework. Sections 2.1 and 2.2 present the first- and the second-order asymptotics, respectively. Section 2.3 describes feasible inference methods. Below, for a matrix A, we denote its (j,k) element by A_{jk} and its transpose by $A^\top$. Convergence in probability and stable convergence in law are denoted by $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{\mathcal{L}\text{-}s}$, respectively. All limits are for the sample frequency $\Delta_n \to 0$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

2.1. *First-Order Asymptotic Properties*

Suppose that the log-price vector X_t is an Itô semimartingale of the form

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s + J_t, \tag{3}$$

where b is the $\mathbb{R}^d$-valued drift process, W is a d-dimensional standard Brownian motion, σ is the $d \times d$ dimensional stochastic volatility matrix, and J is a finitely active pure-jump process. We denote the spot covariance matrix of X by $c_t \equiv \sigma_t\sigma_t^\top$ and further set

$$v_{j,t} \equiv \sqrt{c_{jj,t}}, \qquad \rho_{jk,t} \equiv \frac{c_{jk,t}}{v_{j,t}v_{k,t}}. \tag{4}$$

That is, $v_{j,t}$ and $\rho_{jk,t}$ denote the *spot volatility* of asset j and the *spot correlation coefficient* between assets j and k, respectively. We explicitly allow for so-called "leverage effect" (i.e., dependence between changes in the price and changes in volatility), stochastic volatility of volatility, volatility jumps, and price-volatility co-jumps.

We begin by characterizing the first-order limiting behavior of the realized semicovariance estimators defined by equation (1) in the Introduction. Let ΔX_s denote the price jump occurring at time s, if a jump occurred, and set it to zero if no jump occurred at time s. Further define

$$\begin{aligned}
P^\dagger &\equiv \sum_{s\le T} p(\Delta X_s)p(\Delta X_s)^\top,\\
N^\dagger &\equiv \sum_{s\le T} n(\Delta X_s)n(\Delta X_s)^\top,\\
M^\dagger &\equiv \sum_{s\le T} \big(p(\Delta X_s)n(\Delta X_s)^\top + n(\Delta X_s)p(\Delta X_s)^\top\big).
\end{aligned}$$

These measures characterize the discontinuous parts of the semicovariance measures, as formally spelled out in the following theorem.

THEOREM 1: *Under Assumption 1 in the Appendix*, $(\widehat{P}, \widehat{N}, \widehat{M}) \xrightarrow{\mathbb{P}} (P, N, M)$, *where* P, N, *and* M *are* $d \times d$ *matrices with their* (j, k) *elements given by*

$$P_{jk} \equiv \int_0^T v_{j,s}v_{k,s}\psi(\rho_{jk,s})\,ds + P_{jk}^{\dagger},$$

$$N_{jk} \equiv \int_0^T v_{j,s}v_{k,s}\psi(\rho_{jk,s})\,ds + N_{jk}^{\dagger},$$

$$M_{jk} \equiv -2\int_0^T v_{j,s}v_{k,s}\psi(-\rho_{jk,s})\,ds + M_{jk}^{\dagger},$$

and $\psi(\cdot)$ *is defined in equation* (2).

It follows from Theorem 1 that each of the realized semicovariances contains both diffusive and jump covariation components. Importantly, the limiting variables P and N share exactly the same diffusive component, but their jump components differ. In particular,

$$\widehat{P} - \widehat{N} \xrightarrow{\mathbb{P}} P - N = P^{\dagger} - N^{\dagger}.$$

That is, the first-order asymptotic behavior of the concordant semicovariance differential (CSD) is fully characterized by the "directional co-jumps." Consequently, in line with the stylized model in equation (1) discussed in the Introduction, Theorem 1 cannot distinguish the information conveyed by $\widehat{P}$ and $\widehat{N}$ in periods when there are no jumps. Hence, in order to reveal the differential information inherent in the realized measures more generally, we turn next to a more refined second-order asymptotic analysis.

2.2. *Second-Order Asymptotic Properties*

Since the main theoretical lessons about the second-order asymptotic behavior of the concordant realized semicovariance components can be readily learned in a bivariate setting, we set $d = 2$ and focus on the analysis of $\widehat{P}_{12}$ and $\widehat{N}_{12}$ throughout this subsection. Correspondingly, we also write ρ_t in place of $\rho_{12,t}$ for simplicity. The joint analysis of all semicovariance components can be done in a similar manner. However, it is much more tedious to discuss, so for readability we defer this more general analysis to the Supplemental Material, Section S1.

We need to impose some additional structure on the volatility dynamics. In particular, we will assume that the stochastic volatility σ_t is also an Itô semimartingale of the form (see, e.g., equation (4.4.4) in Jacod and Protter (2012))

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s\,ds + \int_0^t \tilde{\sigma}_s\,dW_s + \widetilde{M}_t + \sum_{s\le t} \Delta\sigma_s 1_{\{\|\Delta\sigma_s\| > \underline{\sigma}\}}, \tag{5}$$

where $\tilde{b}$ is the drift, $\tilde{\sigma}$ is a $d \times d \times d$ tensor-valued process, and $\widetilde{M}$ is a local martingale that is orthogonal to the Brownian motion W.[4] A few remarks are in order. The

[4]By convention, the (j, k) element of the stochastic integral $\int_0^t \tilde{\sigma}_s\,dW_s$ equals $\sum_{l=1}^d \int_0^t \tilde{\sigma}_{jkl,s}\,dW_{l,s}$.

process $\tilde{\sigma}$ collects the loadings of the stochastic volatility matrix σ on the price Brownian shocks dW, and hence is naturally thought of as a multivariate quantification of a "leverage effect." We also allow σ to load on Brownian shocks that are independent of dW through the local martingale $\widetilde{M}$. The $\widetilde{M}$ process may also contain compensated "small" volatility jumps in the form of a purely discontinuous local martingale.[5] Meanwhile, the term $\sum_{s\le t}\Delta\sigma_s 1_{\{\|\Delta\sigma_s\|>\underline{\sigma}\}}$ collects the "large" volatility jumps (with an arbitrary but fixed threshold $\underline{\sigma}>0$), which often occur in response to major news announcements (see, e.g., Bollerslev, Li, and Xue (2018)). The Itô semimartingale setting also readily accommodates the well-established intraday periodicity in the volatility dynamics (see, e.g., Andersen and Bollerslev (1997)). Further regularity conditions regarding the σ process are collected in the Appendix.

Theorem 2, below, describes the $\mathcal{F}$-stable convergence in law of the normalized statistic $\Delta_n^{-1/2}(\widehat{P}_{12}-P_{12},\widehat{N}_{12}-N_{12})$. The limit variable turns out to be fairly complicated, but it may be succinctly expressed as

$$\begin{pmatrix} B \\ -B \end{pmatrix}+\begin{pmatrix} L \\ -L \end{pmatrix}+\begin{pmatrix} \zeta \\ -\zeta \end{pmatrix}+\begin{pmatrix} \tilde{\zeta}_P \\ \tilde{\zeta}_N \end{pmatrix}+\begin{pmatrix} \tilde{\xi}_P \\ \tilde{\xi}_N \end{pmatrix},$$

where, as we will detail below, B and L are bias terms, and $(\zeta,\tilde{\zeta},\xi)$ capture sampling variabilities that arise from various sources. In particular, we note that $\widehat{P}_{12}$ and $\widehat{N}_{12}$ load on the B and L bias terms in exactly opposite ways, which would cancel with each other in the (aggregated) realized covariance $\widehat{C}_{12}$. Before presenting the actual limit theorem, we begin by briefly describing each of these separate terms. Recall that the processes b, σ, v, ρ, and $\tilde{\sigma}$ have previously been introduced in equations (3), (4), and (5).

Bias components due to price drift, B. The first type of bias is related to the price drift, which is defined for the semicovariance estimator $\widehat{P}_{12}$ as

$$B=\frac{1}{2\sqrt{2\pi}}\int_0^T\left(\frac{b_{1,s}}{v_{1,s}}+\frac{b_{2,s}}{v_{2,s}}\right)v_{1,s}v_{2,s}(1+\rho_s)\,ds.$$

As mentioned above, the analogous bias term for $\widehat{N}_{12}$ is $-B$. Other things equal, this bias term is proportional to the average spot "Sharpe ratio" of the two assets

$$\frac{1}{2}\left(\frac{b_{1,s}}{v_{1,s}}+\frac{b_{2,s}}{v_{2,s}}\right).$$

Therefore, the B term tends to be more pronounced when the two assets drift in the same direction, akin to a "*co-drift*" type phenomenon.

Bias components due to continuous price-volatility covariation, L. The second bias term stems from the fact that the volatility matrix process σ_t may be partially driven by the Brownian motion W (i.e., $\tilde{\sigma}\neq 0$), corresponding to a "dynamic leverage" type effect. To more precisely describe this bias term for $\widehat{P}_{12}$, define $f_1(x)\equiv 1_{\{x_1\ge 0\}}\max\{x_2,0\}$ and

[5] We remind the reader that two local martingales are called orthogonal if their product is a local martingale (or equivalently, their predictable covariation process is identically zero). A local martingale is called purely discontinuous if it is orthogonal to all continuous local martingales. See Definition I.4.11 and Proposition I.4.15 in Jacod and Shiryaev (2003) for additional details.

$f_2(x) \equiv \max\{x_1, 0\}1_{\{x_2 \geq 0\}}$ and then set, for any 2×2 matrix A,

$$F_j(A) \equiv \mathbb{E}\left[f_j(AW_1)\int_0^1 W_s\,dW_s^\top\right], \quad j = 1, 2.$$

The bias term in $\widehat{P}_{12}$ due to the common price-volatility Brownian dependence may then be expressed as

$$L \equiv \sum_{j=1}^{2}\int_0^T \mathrm{Trace}\big[\tilde{\sigma}_{j,s}F_j(\sigma_s)\big]\,ds, \tag{6}$$

where $\tilde{\sigma}_{j,s}$ denotes the 2×2 matrix $[\tilde{\sigma}_{jkl,s}]_{1\leq k,l\leq 2}$. The bias term for $\widehat{N}_{12}$ may be defined similarly, and it can be shown to equal $-L$.

Diffusive sampling error spanned by price risk, ζ. The third component in the limit of $\widehat{P}_{12}$ captures the sampling variability in $p(\Delta_i^n X_1)p(\Delta_i^n X_2)$ that is spanned by the Brownian price shock $\sigma_t\,dW_t$. Formally,

$$\zeta \equiv \int_0^T \big(c_s^{-1}\gamma_s\big)^\top(\sigma_s\,dW_s), \tag{7}$$

where the γ_t process is defined as

$$\gamma_t \equiv \frac{(1+\rho_t)^2 v_{1,t}v_{2,t}}{2\sqrt{2\pi}}\begin{pmatrix} v_{1,t} \\ v_{2,t}\end{pmatrix}.$$

The analogous component for $\widehat{N}_{12}$ equals $-\zeta$. Note that the quadratic covariation matrix of the local martingale $(\zeta, -\zeta)$ equals $\int_0^T \Gamma_s\,ds$, where

$$\Gamma_t \equiv \begin{pmatrix} \gamma_t^\top c_t^{-1}\gamma_t & -\gamma_t^\top c_t^{-1}\gamma_t \\ -\gamma_t^\top c_t^{-1}\gamma_t & \gamma_t^\top c_t^{-1}\gamma_t \end{pmatrix}. \tag{8}$$

Diffusive sampling error orthogonal to price risk, $\tilde{\zeta}$. While ζ defined above captures the diffusive risk in the semicovariance spanned by the Brownian shocks to the price process, $\tilde{\zeta}$ captures the diffusive risk component orthogonal to those shocks. This limit variable may be represented by its $\mathcal{F}$-conditional distribution as

$$\tilde{\zeta} = \begin{pmatrix} \tilde{\zeta}_P \\ \tilde{\zeta}_N \end{pmatrix} = \int_0^T \bar{\gamma}_s^{1/2}\,d\widetilde{W}_s, \tag{9}$$

where $\widetilde{W}$ is a two-dimensional standard Brownian motion that is independent of the σ-field $\mathcal{F}$, and the $\bar{\gamma}$ process is defined by $\bar{\gamma}_t \equiv \overline{\Gamma}_t - \Gamma_t$, where

$$\overline{\Gamma}_t \equiv v_{1,t}^2 v_{2,t}^2 \begin{pmatrix} \Psi(\rho_t) - \psi(\rho_t)^2 & -\psi(\rho_t)^2 \\ -\psi(\rho_t)^2 & \Psi(\rho_t) - \psi(\rho_t)^2 \end{pmatrix}, \tag{10}$$

and

$$\Psi(\rho) \equiv \frac{3\rho\sqrt{1-\rho^2} + \big(1+2\rho^2\big)\arccos(-\rho)}{2\pi}, \tag{11}$$

with $\Psi(\rho)$ corresponding to $\mathbb{E}[Z_1^2 Z_2^2 1_{\{Z_1<0,Z_2<0\}}]$ for (Z_1, Z_2) standard normally distributed with correlation ρ.

Jump-induced sampling error, ξ. The price jumps also induce sampling errors. Let $\mathcal{T}_j$ for $j \in \{1, 2\}$ denote the collection of jump times of $(X_{j,t})_{t\in[0,T]}$, with the corresponding "signed" subsets denoted by

$$\mathcal{T}_{j+} \equiv \{\tau \in \mathcal{T}_j : \Delta X_{j,\tau} > 0\}, \qquad \mathcal{T}_{j-} \equiv \{\tau \in \mathcal{T}_j : \Delta X_{j,\tau} < 0\}.$$

For each $\tau \in \mathcal{T}_1 \cup \mathcal{T}_2$, associate the variables $(\kappa_\tau, \tilde{\xi}_{\tau-}, \tilde{\xi}_{\tau+})$ that are, conditionally on $\mathcal{F}$, mutually independent with the following conditional distributions: $\kappa_\tau \sim \text{Uniform}[0, 1]$, $\tilde{\xi}_{\tau-} \sim \mathcal{MN}(0, c_{\tau-})$, and $\tilde{\xi}_{\tau+} \sim \mathcal{MN}(0, c_\tau)$. Further define $\tilde{\eta}_\tau = (\tilde{\eta}_{1,\tau}, \tilde{\eta}_{2,\tau})^\top \equiv \sqrt{\kappa_\tau}\tilde{\xi}_{\tau-} + \sqrt{1-\kappa_\tau}\tilde{\xi}_{\tau+}$. The limiting variable $\xi = (\xi_P, \xi_N)^\top$ may then be expressed as

$$\xi_P \equiv \sum_{\tau\in\mathcal{T}_{1+}\cap\mathcal{T}_{2+}} (\Delta X_{1,\tau}\tilde{\eta}_{2,\tau} + \Delta X_{2,\tau}\tilde{\eta}_{1,\tau}) + \sum_{\tau\in\mathcal{T}_{1+}\setminus\mathcal{T}_2} \Delta X_{1,\tau} p(\tilde{\eta}_{2,\tau}) + \sum_{\tau\in\mathcal{T}_{2+}\setminus\mathcal{T}_1} \Delta X_{2,\tau} p(\tilde{\eta}_{1,\tau}),$$

$$\xi_N \equiv \sum_{\tau\in\mathcal{T}_{1-}\cap\mathcal{T}_{2-}} (\Delta X_{1,\tau}\tilde{\eta}_{2,\tau} + \Delta X_{2,\tau}\tilde{\eta}_{1,\tau}) + \sum_{\tau\in\mathcal{T}_{1-}\setminus\mathcal{T}_2} \Delta X_{1,\tau} n(\tilde{\eta}_{2,\tau}) + \sum_{\tau\in\mathcal{T}_{2-}\setminus\mathcal{T}_1} \Delta X_{2,\tau} n(\tilde{\eta}_{1,\tau}).$$

Note that the first component in ξ_P (resp. ξ_N) concerns times when both assets have positive (resp. negative) jumps, while the other two terms are active when one asset jumps upwards (resp. downwards) and the other asset does not jump. The latter terms involve the half-truncated doubly mixed Gaussian variable $p(\tilde{\eta}_\tau)$ (resp. $n(\tilde{\eta}_\tau)$). To the best of our knowledge, this limiting distribution is new to the literature.

With the definitions above, we are now ready to state the stable convergence in law of the realized semicovariances.

THEOREM 2: *Under Assumption 2 in the Appendix*,

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{P}_{12} - P_{12} \\ \widehat{N}_{12} - N_{12} \end{pmatrix} \xrightarrow{\mathcal{L}\text{-}s} \begin{pmatrix} B \\ -B \end{pmatrix} + \begin{pmatrix} L \\ -L \end{pmatrix} + \begin{pmatrix} \zeta \\ -\zeta \end{pmatrix} + \begin{pmatrix} \tilde{\zeta}_P \\ \tilde{\zeta}_N \end{pmatrix} + \begin{pmatrix} \tilde{\xi}_P \\ \tilde{\xi}_N \end{pmatrix}.$$

Theorem 2 depicts a non-central limit theorem for the positive and negative realized semicovariances, where $(\pm B, \pm L)$ represent bias terms, while $(\pm\zeta, \tilde{\zeta}, \xi)$ stem from various sources of "sampling errors." The latter sampling error terms are all formed as (local) martingales and have zero mean under mild integrability conditions. We note that the bias terms arise because the test functions used to define the realized semicovariances (e.g., $(x_1, x_2) \mapsto p(x_1)p(x_2) = \max\{x_1, 0\}\max\{x_2, 0\}$) are not globally even.[6] This phenomenon also appears in earlier work by Kinnebrock and Podolskij (2008), Barndorff-Nielsen, Kinnebrock, and Shephard (2010), and Li, Mykland, Renault, Zhang, and Zheng (2014); also see Chapter 5 of Jacod and Protter (2012). In the Supplemental Material Section S1, we provide a more general result for the joint convergence of all the realized semicovariance components, including the realized semivariances of Barndorff-Nielsen, Kinnebrock, and Shephard (2010) as special "diagonal" cases. We also demonstrate there how to recover

[6]Following Jacod and Protter (2012), we say that a function $f(\cdot)$ defined on $\mathbb{R}^d$ is globally even if $f(x) = f(-x)$ for all $x \in \mathbb{R}^d$; see page 135 in that book. Kinnebrock and Podolskij (2008) simply referred to such functions as even functions; see page 1057 of that paper.

that prior result in the case without price jumps, and further characterize the effect of jumps on the sampling variability of the realized semivariances.

The presence of the bias terms means that Theorem 2 is not directly suitable for the construction of confidence intervals for the (P, N) estimand; however, that is *not* our goal. Instead, the main insight derived from Theorem 2 is to reveal the differential second-order behavior of the realized semicovariances $\widehat{P}$ and $\widehat{N}$, about which the first-order asymptotics in Theorem 1 remains entirely silent in the absence of jumps. Indeed, while Theorem 1 states that $\widehat{P}$ and $\widehat{N}$ have the same limit in the no-jump case, Theorem 2 clarifies that they actually load on higher-order "signals" $\pm B$ and $\pm L$ in the exact opposite way. From a theoretical perspective, this therefore explains why $\widehat{P}$ and $\widehat{N}$ may behave differently. From an empirical perspective, it helps guide our understanding of the actual $\widehat{P}$ and $\widehat{N}$ estimates discussed in Section 3, and the use of these measures in the construction of improved volatility forecasts discussed in Section 4.

Theorem 2 focuses on the concordant semicovariance terms, $\widehat{P}$ and $\widehat{N}$. The asymptotic behavior of the mixed semicovariance term, $\widehat{M}$, is of particular interest when studying negatively correlated assets, such as in analyzing hedge portfolios. This term is considered jointly with all other elements of the semicovariance matrices in the more general limit theorem presented in the Supplemental Material. Meanwhile, Theorem 2 can also be used to help understand the behavior of the mixed semicovariances by computing $\widehat{P}$ and $\widehat{N}$ on a rotation of the original returns: use the original returns on the first asset, and the negative of the returns on the second asset.

2.3. *Tests Based on Concordant Semicovariance Differentials*

Even though Theorem 2 does not allow for the construction of standard confidence intervals, it is still possible to develop feasible inference methods for the difference $\widehat{P}-\widehat{N}$, that is, the CSD. In particular, the asymptotic theory in the previous subsection reveals three types of signals underlying the CSD: a directional co-jump effect (i.e., $P^{\dagger} - N^{\dagger}$), a co-drifting effect (i.e., $2B$), and a dynamic leverage effect (i.e., $2L$). Empirically, it is of great interest to separate the variation due to jumps from that due to the diffusive price moves. Below, we use the standard truncation method (see, e.g., Mancini (2001, 2009)) to achieve such a separation. As in Section 2.2, we consider a bivariate setting, or $d = 2$, and focus on the inference for $\widehat{P}_{12} - \widehat{N}_{12}$.

The truncation method involves a sequence $u_n \in \mathbb{R}^2_+$ of truncation thresholds satisfying $u_{j,n} \asymp \Delta_n^{\varpi}$ for some $\varpi \in (0, 1/2)$ and $j \in \{1, 2\}$. Under our maintained assumption of finite activity jumps, it can be shown that the index set

$$\widehat{\mathcal{I}} \equiv \{i : -u_n \leq \Delta_i^n X \leq u_n \text{ does not hold}\}$$

consistently estimates the jump times of the vector log-price process X.[7] In practice, it is important to choose the truncation threshold u_n adaptively so as to account for time-varying volatility, particularly its well-known U-shape intraday pattern; we discuss this further in connection with our empirical analyses below. As a result, the diffusive and

[7]See Proposition 1 of Li, Todorov, and Tauchen (2017b), for which the inequality $-u_n \leq \Delta_i^n X \leq u_n$ is interpreted element-by-element. Importantly, although jumps can be separately recovered in asymptotic theory, it is difficult to do so in practice. The realized semicovariance measures only concern time-aggregated jump characteristics, instead of individual jumps. In our analysis, the jump recovery is only used as a theoretical auxiliary tool to prove asymptotic results for time-aggregated quantities.

jump returns may be separated, allowing for the separate estimation of the *diffusive* components of the semicovariances using the "small" (non-jump) returns:

$$\widehat{P}^{\star} \equiv \sum_{i \notin \widehat{\mathcal{I}}} p(\Delta_i^n X) p(\Delta_i^n X)^{\top}, \qquad \widehat{N}^{\star} \equiv \sum_{i \notin \widehat{\mathcal{I}}} n(\Delta_i^n X) n(\Delta_i^n X)^{\top},$$

and the *jump* components of the semicovariances using the "jump" returns:

$$\widehat{P}^{\dagger} \equiv \sum_{i \in \widehat{\mathcal{I}}} p(\Delta_i^n X) p(\Delta_i^n X)^{\top}, \qquad \widehat{N}^{\dagger} \equiv \sum_{i \in \widehat{\mathcal{I}}} n(\Delta_i^n X) n(\Delta_i^n X)^{\top}.$$

The aforementioned jump detection result permits the truncated estimators to be analyzed in the following straightforward extension of Theorem 2.

PROPOSITION 1: *Under Assumption 2 in the Appendix, the following convergences hold jointly*:

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{P}_{12}^{\dagger} - P_{12}^{\dagger} \\ \widehat{N}_{12}^{\dagger} - N_{12}^{\dagger} \end{pmatrix} \xrightarrow{\mathcal{L}\text{-}s} \begin{pmatrix} \xi_P \\ \xi_N \end{pmatrix},$$

$$\Delta_n^{-1/2} \begin{pmatrix} \widehat{P}_{12}^{\star} - P_{12}^{\star} \\ \widehat{N}_{12}^{\star} - N_{12}^{\star} \end{pmatrix} \xrightarrow{\mathcal{L}\text{-}s} \begin{pmatrix} B \\ -B \end{pmatrix} + \begin{pmatrix} L \\ -L \end{pmatrix} + \begin{pmatrix} \zeta \\ -\zeta \end{pmatrix} + \begin{pmatrix} \tilde{\zeta}_P \\ \tilde{\zeta}_N \end{pmatrix},$$

where $P_{12}^{\star} = N_{12}^{\star} \equiv \int_0^T v_{1,s} v_{2,s} \psi(\rho_s)\, ds$.

Proposition 1, and consistent estimation of the spot covariances (discussed below), allow for feasible inference on each of the two separate CSD components, $\widehat{P}_{12}^{\dagger} - \widehat{N}_{12}^{\dagger}$ and $\widehat{P}_{12}^{\star} - \widehat{N}_{12}^{\star}$. We will refer to these as the jump or diffusive concordant semicovariance differentials (JCSD or DCSD), respectively.

We start with a discussion of how to implement the JCSD test, which is the simpler of the two as it admits a central limit theorem. In particular, it follows immediately from Proposition 1 that

$$\Delta_n^{-1/2} \big(\widehat{P}_{12}^{\dagger} - \widehat{N}_{12}^{\dagger} - \big(P_{12}^{\dagger} - N_{12}^{\dagger} \big) \big) \xrightarrow{\mathcal{L}\text{-}s} \xi_P - \xi_N,$$

where, as discussed above, ξ_P and ξ_N are defined in terms of doubly-mixed Gaussian variables. Hence, to consistently estimate the distribution of the limiting variable, we first need to estimate the spot covariance matrix before and after each detected jump time. In order to do so, we choose an integer sequence k_n of local windows that satisfies $k_n \to \infty$ and $k_n \Delta_n \to 0$, and set, for each i,

$$\hat{c}_{i-} \equiv \frac{1}{k_n \Delta_n} \sum_{l=1}^{k_n} \big(\Delta_{i-l}^n X \big) \big(\Delta_{i-l}^n X \big)^{\top} 1_{\{-u_n \le \Delta_{i-l}^n X \le u_n\}},$$

$$\hat{c}_{i+} \equiv \frac{1}{k_n \Delta_n} \sum_{l=1}^{k_n} \big(\Delta_{i+l}^n X \big) \big(\Delta_{i+l}^n X \big)^{\top} 1_{\{-u_n \le \Delta_{i+l}^n X \le u_n\}}.$$

Algorithm 1 describes the requisite steps for implementing the resulting JCSD test for the null hypothesis $P_{12}^{\dagger} = N_{12}^{\dagger}$, that is, equal directional jump covariation.

ALGORITHM 1—JCSD Test: Step 1. Draw random variables $(\kappa_i^*, \tilde{\xi}_{i-}^*, \tilde{\xi}_{i+}^*)$ that are mutually independent such that $\kappa_i^* \sim \text{Uniform}[0,1]$ and $\tilde{\xi}_{i\pm}^* \sim \mathcal{MN}(0, \hat{c}_{i\pm})$. Set $\tilde{\eta}_i^* = (\tilde{\eta}_{i,1}^*, \tilde{\eta}_{i,2}^*) = \sqrt{\kappa_i^*}\tilde{\xi}_{i-}^* + \sqrt{1-\kappa_i^*}\tilde{\xi}_{i+}^*$.

Step 2. Let $\Delta_i^n X_j^* \equiv \Delta_i^n X_j 1_{\{|\Delta_i^n X_j| > u_{j,n}\}}$ for $j \in \{1, 2\}$ and set

$$\xi_P^* = \Delta_n^{-1/2} \sum_{i \in \widehat{\mathcal{I}}} \big(p\big(\Delta_i^n X_1^* + \Delta_n^{1/2} \tilde{\eta}_{i,1}^*\big) p\big(\Delta_i^n X_2^* + \Delta_n^{1/2} \tilde{\eta}_{i,2}^*\big) - p\big(\Delta_i^n X_1^*\big) p\big(\Delta_i^n X_2^*\big)\big),$$

$$\xi_N^* = \Delta_n^{-1/2} \sum_{i \in \widehat{\mathcal{I}}} \big(n\big(\Delta_i^n X_1^* + \Delta_n^{1/2} \tilde{\eta}_{i,1}^*\big) n\big(\Delta_i^n X_2^* + \Delta_n^{1/2} \tilde{\eta}_{i,2}^*\big) - n\big(\Delta_i^n X_1^*\big) n\big(\Delta_i^n X_2^*\big)\big).$$

Step 3. Repeat steps 1–2 many times. Compute the $1-\alpha$ (resp. α) quantile of $\xi_P^* - \xi_N^*$ as the critical value of $\Delta_n^{-1/2}(\widehat{P}_{12}^\dagger - \widehat{N}_{12}^\dagger)$ for the null hypothesis $P_{12}^\dagger = N_{12}^\dagger$ in favor of the one-sided alternative $P_{12}^\dagger > N_{12}^\dagger$ (resp. $P_{12}^\dagger < N_{12}^\dagger$) at significance level α.

Algorithm 1 may be seen as a parametric bootstrap that exploits the approximate (parametric) doubly-mixed Gaussian distribution of the detected jump returns given the estimated spot covariances, with ξ_P^* and ξ_N^* being the bootstrap analogue of the original normalized estimators. While this type of simulation-based inference is often used in the study of jumps, a non-standard feature of Algorithm 1 is its use of the truncated return $\Delta_i^n X_j^* = \Delta_i^n X_j 1_{\{|\Delta_i^n X_j| > u_{j,n}\}}$, which shrinks the detected diffusive returns to zero. This shrinkage is needed in situations where exactly one asset jumps at time τ, so that the sampling variability contributed by the other (no-jump) asset, say j, is given by a half-truncated doubly-mixed Gaussian variable like $p(\tilde{\eta}_{\tau,j})$. This distribution may in turn be mimicked by $\Delta_n^{-1/2} p(\Delta_i^n X_j^* + \Delta_n^{1/2}\tilde{\eta}_{i,j}^*) = p(\tilde{\eta}_{i,j}^*)$, which differs from the "un-shrunk" variable $\Delta_n^{-1/2} p(\Delta_i^n X_j + \Delta_n^{1/2}\tilde{\eta}_{i,j}^*)$.

PROPOSITION 2: *Under Assumption 2 in the Appendix, the conditional distribution of $\xi_P^* - \xi_N^*$ given the data converges in probability to the $\mathcal{F}$-conditional distribution of $\xi_P - \xi_N$ under the uniform metric. Consequently, the test described in Algorithm 1 has asymptotic level α under the null $\{P_{12}^\dagger = N_{12}^\dagger\}$ and asymptotic power of 1 under one-sided alternatives.*

We turn next to the conduct of feasible inference using the DCSD statistic $\widehat{P}_{12}^\star - \widehat{N}_{12}^\star$. This involves some additional non-standard theoretical subtlety. Proposition 1 implies that

$$\Delta_n^{-1/2}\big(\widehat{P}_{12}^\star - \widehat{N}_{12}^\star\big) - 2B - 2L \xrightarrow{\mathcal{L}\text{-}s} 2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N, \tag{12}$$

where we recall that B and L capture co-drift and dynamic leverage effects, respectively. The first-order limiting variables $P_{12}^\star$ and $N_{12}^\star$ exactly cancel with each other in the $\widehat{P}_{12}^\star - \widehat{N}_{12}^\star$ difference. Consequently, as revealed by the convergence in (12), the remaining "signal" carried by the DCSD is given by the higher-order term $2B + 2L$, which is comparable in magnitude with the statistical noise term $2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N$ (defined as a local martingale). Since the signal-to-noise ratio does not diverge to infinity even in large samples, the resulting test is generally not consistent.

A further non-standard complication related to (12) stems from the fact that the limiting variable $2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N$ is generally not mixed Gaussian. Specifically, while $\tilde{\zeta}_P - \tilde{\zeta}_N$ is $\mathcal{F}$-conditional Gaussian, the remaining part

$$2\zeta = 2\int_0^T \big(c_s^{-1}\gamma_s\big)^\top (\sigma_s \, dW_s)$$

is generally not mixed Gaussian unless, of course, the stochastic volatility σ is independent of the Brownian motion W that drives the diffusive price moves.

Although these non-standard features of the limit theory prevent us from conducting formal tests in the most general setting, it is nevertheless possible to assess the sampling variability of $\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12}$ in a well-defined way. Indeed, it follows from (7) and (9) that the quadratic variation of the continuous local martingale $2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N$ is given by

$$\Sigma^{\star} \equiv 2\int_0^T v_{1,s}^2 v_{2,s}^2 \Psi(\rho_s)\,ds. \tag{13}$$

Therefore, $\sqrt{\Sigma^{\star}}$ may be naturally used as the standard error for gauging the sampling variability of the centered variable $\Delta_n^{-1/2}(\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12}) - (2B + 2L)$. The $\Sigma^{\star}$ variable is defined as an integrated functional of the spot covariance matrix and it may be consistently estimated using a nonparametric "plug-in" estimator $\widehat{\Sigma}^{\star}$ (see, e.g., Li, Todorov, and Tauchen (2017a)).

Comparing $\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12}$ with its standard error provides an econometrically disciplined approach for detecting "large" differences in positive and negative concordant semicovariances, as summarized in the following algorithm. We refer to it cautiously as a detection rule instead of a test; it may be formally interpreted as a test under additional conditions, as detailed below.

ALGORITHM 2—DCSD Detection: Step 1. Define $\hat{v}_{1,i}$, $\hat{v}_{2,i}$, and $\hat{\rho}_i$ implicitly by decomposing the spot covariance matrix estimator $\hat{c}_{i+}$ as

$$\hat{c}_{i+} = \begin{pmatrix} \hat{v}_{1,i}^2 & \hat{\rho}_i \hat{v}_{1,i}\hat{v}_{2,i} \\ \hat{\rho}_i \hat{v}_{1,i}\hat{v}_{2,i} & \hat{v}_{2,i}^2 \end{pmatrix},$$

and set

$$\widehat{\Sigma}^{\star} \equiv \frac{2\Delta_n}{[T/\Delta_n] - k_n + 1} \sum_{i=0}^{[T/\Delta_n]-k_n} \hat{v}_{1,i}^2 \hat{v}_{2,i}^2 \Psi(\hat{\rho}_i).$$

Step 2. Use the $1-\alpha$ quantile of a standard normal distribution z_α as the critical value for the t-statistic $\Delta_n^{-1/2}(\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12})/\sqrt{\widehat{\Sigma}^{\star}}$ for one-sided detection of deviations from $B + L = 0$.

The DCSD detection rule described in Algorithm 2 should be interpreted carefully in empirical work. Due to the aforementioned lack of mixed Gaussianity for the limiting variable under the most general conditions, the t-statistic $\Delta_n^{-1/2}(\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12})/\sqrt{\widehat{\Sigma}^{\star}}$ is generally not asymptotically standard normally distributed. For this reason, the asymptotic level of the proposed one-sided test is not guaranteed to be α. Instead, the t-statistic may be interpreted as a well-defined "signal-to-noise" measure, as opposed to a formal statistical test. These asymptotic results are also corroborated by the simulation results pertaining to the finite-sample properties of the JCSD test and DCSD detection rule presented in Section S2 of the Supplemental Material.

In lieu of these results, mixed Gaussianity can be formally restored under the additional assumption that the volatility process σ is independent of the Brownian motion W, which in turn allows us to analyze the size and power properties of the DCSD detection rule as a formal test. Note that in this "no-leverage" case with σ being independent of W, without loss of generality we can fix $\tilde{\sigma} \equiv 0$ in the volatility model in (5), which further implies that $L \equiv 0$ (recall equation (6)). Hence, the statistic $\Delta_n^{-1/2}(\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12})$ is centered at $2B$, and the signal-to-noise ratio equals $2B/\sqrt{\Sigma^{\star}}$.

Formally, the null hypothesis for DCSD is composed of the following collection of sample paths:

$$\Omega_0 \equiv \{\omega \in \Omega : B(\omega) = 0\}.$$

We remind the reader that in non-ergodic high-frequency settings, it is standard to consider hypotheses as events consisting of sample paths of interest (see, e.g., Aït-Sahalia and Jacod (2014) and the many references therein). Correspondingly, to state a one-sided alternative hypothesis, we fix a constant $R > 0$, and consider:

$$\Omega_a \equiv \left\{\omega \in \Omega : \frac{2B(\omega)}{\sqrt{\Sigma^{\star}(\omega)}} \geq R\right\}.$$

Intuitively, R measures the "distance" between the null and the alternative by drawing a lower bound for the (random) signal-to-noise ratio $2B/\sqrt{\Sigma^{\star}}$, which, not surprisingly, determines the power of the test. Proposition 3, below, characterizes the size and power properties of the resulting DCSD test described in Algorithm 2, where we use $\Phi(\cdot)$ to denote the cumulative distribution function of the standard normal distribution.

PROPOSITION 3: *Suppose that* (i) *Assumption 2 in the Appendix holds*, *and* (ii) *the* (b_t) *and* (σ_t) *processes are independent of* W, *and* $\tilde{\sigma}_t \equiv 0$. *Then*, *the critical region* $C_n^+ \equiv \{\Delta_n^{-1/2}(\widehat{P}^{\star}_{12} - \widehat{N}^{\star}_{12})/\sqrt{\widehat{\Sigma}^{\star}} > z_\alpha\}$ *associated with the* (*positive*) *one-sided DCSD test has asymptotic level* α *in restriction to* Ω_0, *that is*,

$$\lim_{n\to\infty} \mathbb{P}(C_n^+ | \Omega_0) = \alpha.$$

Moreover, *for any* $R > 0$,

$$\liminf_{n\to\infty} \mathbb{P}(C_n^+ | \Omega_a) \geq 1 - \Phi(z_\alpha - R) > \alpha.$$

Proposition 3 shows that, under the additional "no-leverage" assumption, the one-sided DCSD test controls size under the null with "no co-drift" (i.e., $B = 0$), and is asymptotically unbiased with strictly non-trivial power under the alternative. The asymptotic power is higher when the signal-to-noise ratio $2B/\sqrt{\Sigma^{\star}}$ is higher, and the power is at least 50% when $R = z_\alpha$. Also, even though the proposition focuses on a positive one-sided test, these same results readily extend to a negative one-sided test, or two-sided tests.

2.4. *Discussion of Prior Related Studies*

Putting our results further into perspective, Theorem 2 naturally extends Barndorff-Nielsen, Kinnebrock, and Shephard's (2010) asymptotic theory (Proposition 2) from the univariate setting and realized semivariances to a multivariate setting and realized semicovariances. However, the results in Barndorff-Nielsen, Kinnebrock, and Shephard (2010) rely on the theory of Kinnebrock and Podolskij (2008), and hence rule out the presence of price or volatility jumps. Jacod and Protter (2012) provided a more general result allowing for volatility jumps (see Theorem 5.3.5). By comparison, we allow for both price and volatility jumps. Price jumps in turn result in an interesting limiting distribution that involves truncated doubly-mixed Gaussian variables (i.e., $p(\tilde{\eta}_\tau)$ and $n(\tilde{\eta}_\tau)$), which, to the best of our knowledge, are new to the literature on jump-related inference (see, e.g., Aït-Sahalia and Jacod (2009), Jacod and Todorov (2009), and Chapter 14 of Aït-Sahalia and Jacod (2014)).

The higher-order bias terms that appear in the second-order asymptotics arise from the fact that the realized semicovariances are not globally even transformations of the high-frequency returns. Results in Kinnebrock and Podolskij (2008), Barndorff-Nielsen, Kinnebrock, and Shephard (2010), and Chapter 5 of Jacod and Protter (2012) share that same feature. In the probability literature, a closely related limit theorem was also proved by Jacod (1997) and reported as Theorem IX.7.3 in Jacod and Shiryaev (2003). Another interesting example of this type of higher-order bias terms was provided by Li et al. (2014), in their analysis of realized tricity (see their Theorem 2) and a test for endogenous sampling times.[8]

The feasible inference developed in Section 2.3 also further sets our analysis apart from that of Barndorff-Nielsen, Kinnebrock, and Shephard (2010). In the absence of price jumps, in particular, our DCSD inference about the diffusive component involves a nonlinear transform of the spot covariance matrix, $\Sigma^\star \equiv 2\int_0^T v_{1,s}^2 v_{2,s}^2 \Psi(\rho_s)\,ds$. Correspondingly, our inference relies on the estimator for general integrated volatility functionals recently developed by Li and Xiu (2016) and Li, Todorov, and Tauchen (2017a); see also Renò (2008) and Kristensen (2010) for earlier results on spot volatility estimation. By contrast, the feasible inference of Barndorff-Nielsen, Kinnebrock, and Shephard (2010), available under more restrictive conditions, only requires the estimation of integrated quarticity.

In further contrast to Barndorff-Nielsen, Kinnebrock, and Shephard (2010), who did not consider jumps, our JCSD test is explicitly geared toward price jumps. By focusing on the differential between positive and negative directional co-jumps, our test also serves a different empirical purpose than other tests for the presence of jumps and co-jumps (see, e.g., Barndorff-Nielsen and Shephard (2006), Aït-Sahalia and Jacod (2009), Jacod and Todorov (2009), and Caporin, Kolokolov, and Renò (2017)). Our feasible inference described in Algorithm 1 may be seen as a parametric bootstrap, and as such is naturally related to the high-frequency bootstrap methods developed by Gonçalves and Meddahi (2009), Dovonon, Goncalves, Hounyo, and Meddahi (2019), among others. However, while the latter bootstrap methods are designed to mimic the sampling variability of aggregated Brownian shocks, we aim to recover that of the truncated jump returns.

3. EMPIRICAL SEMICOVARIANCE TESTS

We begin our empirical investigations by looking at the realized semicovariances for the 30 Dow Jones Industrial Average (DJIA) constituent stocks as of the last day of our sample period. We use one-minute returns, obtained from the Trades and Quotes (TAQ)

[8]The (realized) tricity estimator is defined as the sum of cubic powers of high-frequency returns. Theorem 2 of Li et al. (2014) establishes a non-central limit theorem for this estimator allowing for random sampling. In the special case with regular sampling (corresponding to $h_t = 1$ in the notation in that paper), the bias term has the form (using the notation in the present paper) $3\int_0^T \sigma_s^2 b_s\,ds + 3\int_0^T \sigma_s^3\,dW_s + (3/2)\langle\sigma^2, X\rangle_t$, where $\langle\sigma^2, X\rangle_t$ denotes the quadratic variation between σ^2 and X, corresponding to a leverage type effect. These three components play similar roles to the B, ζ, and L terms in Theorem 2, respectively. It may be interesting to further generalize the results in the present paper to a setting with random sampling using the technique of Li et al. (2014).

database, spanning the period from January 2006 to December 2014, a total of 2,265 trading days. Our choice of a one-minute sampling frequency and recent sample period is dictated by the need to reliably estimate the spot covariance matrix, which is used in the tests described in the previous section.[9] To more clearly pinpoint within-day market-wide price moves and economic events, we exclude the returns for the first half-hour of the trading day in the calculation of the statistics.

We calculate the JCSD and DCSD semicovariance-based tests for all of the 435 unique stock-pairs and 2,265 days in the sample, resulting in close to one million test statistics for each of the two tests.[10]

We use one-sided tests at the 5% significance level. The average rejection rates for the JCSD tests far exceed the nominal level, rejecting in favor of positive (negative) co-jumps for about 30% (32%) of the pairs. The DCSD algorithm detects significantly positive (negative) differences for 13% (10%) of all stock pairs.[11]

To shed additional light on these test results, Table I lists the days with the most rejections for each of the tests for each of the nine years in the sample, along with a short description of the most important economic events that occurred on these days. Panel A shows that all but one of the days with the most rejections for the JCSD test are associated with FOMC statements or changes in the federal funds rate. This finding is consistent with the prior literature that links jumps in asset prices with public news announcements (e.g., Andersen, Bollerslev, and Diebold (2007), Lee and Mykland (2008), Lee (2012), and Caporin, Kolokolov, and Renò (2017)). It is also in line with the literature on testing for co-jumps and the finding that those jumps are associated with economy-wide news (e.g., Bollerslev, Law, and Tauchen (2008) and Lahaye, Laurent, and Neely (2011)).[12]

In contrast to the "sharp" economic events associated with the JCSD test, Panel B shows that the days with the most rejections for the DCSD test are typically associated with "softer," more difficult-to-interpret information. Kyle (1985)-type models can be used to establish a more formal economic link between the "soft" information and price drift (which drives the DCSD test through the co-drift effect). In these models, informed agents trade strategically with liquidity traders to maximize their profit, and they do so patiently in order to manage the market marker's belief. As shown by Back (1992), informed traders' optimal order flow is smooth (i.e., differentiable) in time, which in turn determines the drift of the equilibrium price. In a setting with stochastic liquidity, Collin-Dufresne and Fos (2016) further showed that, in equilibrium, the price drift exhibits mean reversion toward the asset's true value, and mean reversion is strong (weak) when the short-term liquidity is high (low) relative to the long-term liquidity. This theory suggests

[9]The choice of a one-minute sampling frequency mirrors that of Li, Todorov, Tauchen, and Chen (2017) in their estimation of spot covariances. It is also supported by the "signature plots" in Section S8 of the Supplemental Material. For additional discussion of market microstructure effects, see Zhang, Mykland, and Aït-Sahalia (2005), Hansen and Lunde (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), and Jacod, Li, and Zheng (2017).

[10]We rely on the dynamic threshold advocated by Bollerslev and Todorov (2011a, 2011b) based on three times the trailing bipower variation, as originally defined by Barndorff-Nielsen and Shephard (2004a, 2006), adjusted for the intraday periodicity in the volatility. We set the local window $k_n = 45$.

[11]We do not intend to make a formal statistical statement jointly across all pairs. Instead, we view the rejection frequencies as simple summary statistics of the pairwise test results.

[12]A detailed comparison of the dates in Table I with those reported by all of the studies cited here is beyond the scope of this paper. However, we note that some of the dates overlap with those in Caporin, Kolokolov, and Renò (2017), for example, while others do not, as the different focus of the tests naturally leads to some variation in the days with the most significant jumps.

TABLE I

TOP REJECTION DAYS BY YEAR[a]

Year	Date	Direction	%	Headline Event
				Panel A. JCSD Test
2006	June 29	+	100	Fed raises short-term rate by a quarter-percentage point.
2007	September 18	+	99	Fed cuts short-term rate by a half-percentage point.
2008	December 16	+	100	Fed cuts short-term rate by a quarter-percentage point.
2009	March 18	+	100	Fed announces it will buy up to $300 billion in long-term Treasuries.
2010	August 10	+	98	Fed announces it will continue Quantitative Easing.
2011	September 22	+	100	Fed announces Operation Twist.
2012	September 13	+	100	Fed announces it will continue buying Mortgage Backed Securities.
2013	September 18	+	98	Fed announces it will sustain the asset buying program.
2014	August 5	−	97	Russian troops are reported lining on the borders of Ukraine.
				Panel B. DCSD Test
2006	July 19	+	57	Bernanke explains to the Senate Banking Committee how the Fed sees the economic slowdown.
2007	August 29	+	58	Bernanke writes letter to senator that Fed is monitoring and ready to step in if necessary.
2008	January 2	−	67	Markets react to poor manufacturing, housing and credit news.
2009	March 23	+	90	Obama administration announces its plan to buy $1 trillion in bad bank assets.
2010	July 7	+	76	EU reveals its first list of stress test banks.
2011	June 1	−	83	Moody's cuts Greece's bond rating by three notches.
2012	June 21	−	84	Rumors of Moody's downgrade for global banks.
2013	February 25	−	83	Political uncertainty surrounding Italian elections.
2014	February 3	−	85	Janet Yellen sworn in as the new Fed chair.

[a]The table reports the top rejection dates by year for the semicovariance-based tests across all 435 DJIA stock-pairs. The first column gives the date, the second gives the direction in which the rejections occurred, and the third provides the fraction of pairs of stocks for which the test rejects at the 5% level in that direction. The final column summarizes headline economic news events for the different days. Panel A reports the results for the jump CSD test, while Panel B is based on the diffusive CSD test.

that, ceteris paribus, price drift is greater in magnitude when there is a greater degree of mispricing, or when the revelation of the private information is imminent.

To illustrate the distinct price dynamics on the "sharp" and "soft" news event days identified by the JCSD and DCSD tests, Figure 3 plots cumulative returns for each of the 30 DJIA stocks for two days from Table I: September 18, 2013 and February 25, 2013.[13] (These are also the dates presented in Figure 2.) For the event detected by the JCSD test (left panel), all stocks experienced a large positive price change (of just over 1% on average) at 2 PM, following the release of the FOMC meeting statement which declared that the Fed would sustain its asset-purchasing program. By contrast, for the event detected by the DCSD test (right panel), we observe slow and steadily decreasing price paths throughout the day for all of the stocks. The total daily return is large, with the median daily return of around negative 2%, but no "extreme" one-minute returns occurred for any of the stocks during the course of that day.

The outcome of the DCSD or JCSD tests and the differences in the within-day price dynamics further translate into different dynamic dependencies in the semicovariance components across days.[14] The total realized covariance $\widehat{C}$, in particular, generally appears

[13]Section S3 in the Supplemental Material presents plots for all of the dates listed in Table I.

[14]A summary table with the correlations conditional on different DCSD and JCSD event days is available in Supplemental Material Section S4.

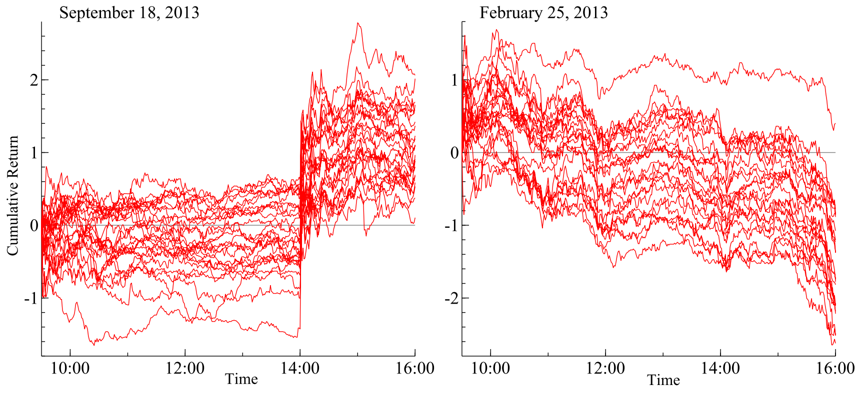

FIGURE 3.—DJIA Cumulative returns on representative event days. *Note*: The figure plots the cumulative return of the 30 Dow Jones Industrial Average stocks on two of the event dates associated with market-wide jump CSD and diffusive CSD events from Table I.

more persistent following days with diffusive events, as detected by the DCSD algorithm. On the other hand, the persistence of $\widehat{P}$ increases primarily following positive DCSD events, while the persistence of $\widehat{N}$ is mostly higher following negative DCSD events, again consistent with the idea that certain types of "soft" news are processed only slowly over multiple days. By comparison, only negative JCSD jump events appear to affect the average persistence of either component.

4. FORECASTING WITH REALIZED SEMICOVARIANCES

The results discussed in the previous section highlight the additional information and economic insights afforded by the realized semicovariances beyond those from standard realized covariances. The results also point to the existence of different dynamic dependencies conditional on different days. In this section, we further explore these empirical differences from the perspective of forecasting future variances and covariances.

To allow for the construction of larger-dimensional portfolios, we expand our previous sample of 30 DJIA stocks to include all of the S&P 500 constituent stocks, and we consider a longer sample period, from January 1993 to December 2014, a total of 5,541 trading days. In order to reliably estimate models for covariances and semicovariances, we include only stocks with at least 2,000 daily observations, resulting in a total of 749 stocks. Most of these stocks are not as actively traded as the DJIA stocks, especially during the earlier part of the sample. Correspondingly, since we only require consistent estimates for this part of our analysis, we rely on a coarser 15-minute sampling scheme to construct the realized measures. Finally, similarly to most existing work on volatility forecasting (e.g., Hansen, Huang, and Shek (2012), and Noureldin, Shephard, and Sheppard (2012)), we focus on the intra-daily period excluding the overnight returns.[15]

4.1. *Vector Autoregressions for Realized Semicovariances*

Figure 1 discussed in the Introduction already points to the existence of different dynamic dependencies in the average combined concordant and discordant semicovariance

[15]Empirical results that include the overnight returns are presented in Supplemental Material Section S6.3. All of our main empirical findings remain qualitatively unaltered.

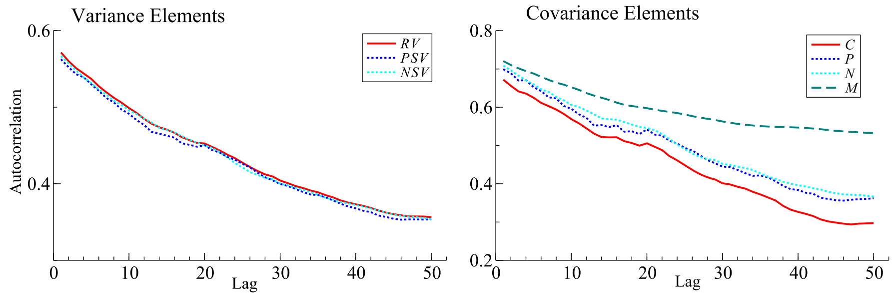

FIGURE 4.—Autocorrelations. *Note*: The graph plots the autocorrelation functions for the different realized semicovariance elements. All of the estimates are averaged across 1,000 randomly selected S&P 500 pairs of stocks, and bias-adjusted following the approach of Hansen and Lunde (2014).

components. To more directly highlight these differences, Figure 4 plots the lag 1 through 50 autocorrelations for each of the individual realized semicovariance components averaged across 1,000 randomly selected S&P 500 pairs of stocks.[16] While the autocorrelations for the realized variances (RV) and the positive and negative realized semivariances (PSV and NSV) shown in the left panel are almost indistinguishable, there is a clear ordering in the rate of decay of the autocorrelations for the realized semicovariance elements shown in the right panel. Most noticeably, the autocorrelations for the (total) covariances $\widehat{C}$ are systematically below those for the three realized semicovariance elements, with the mixed $\widehat{M}$ component exhibiting the highest overall persistence. The fact that almost all of our pairs of stocks are positively correlated means that few co-jumps will appear in $\widehat{M}$, and so by the asymptotic theory in Section 2, $\widehat{M}$ is mostly composed of diffusive covariation, while $\widehat{P}$ and $\widehat{N}$ contain both diffusive and co-jump components. Previous work (e.g., Maheu and McCurdy (2004), and Andersen, Bollerslev, and Diebold (2007)) has found that the diffusive component of volatility is generally more persistent than the jump component, and our finding that $\widehat{M}$ appears more persistent than $\widehat{P}$ and $\widehat{N}$ is consistent with those findings.

These differences also naturally suggest that more accurate volatility and covariance forecasts may be obtained by separately modeling the realized semicovariance components that make up the realized covariance. To more directly investigate this, we estimate a vector version of the popular HAR model of Corsi (2009), in which each of the elements in the realized semicovariance matrix is allowed to depend on its own daily, weekly, and monthly lags, as well as the lags of the other realized semicovariance components.[17] Specifically, for each pair of assets (j, k), we estimate the following three-dimensional

[16]We rely on the estimator of Hansen and Lunde (2014) to account for measurement errors in the realized measures. Section S5 of the Supplemental Material provides the corresponding unadjusted autocorrelation functions.

[17]Building on the new ideas and theoretical results first presented here, Bollerslev, Patton, and Quaedvlieg (2020) have recently shown how the realized semicovariances may similarly be used in the construction of improved multivariate realized GARCH-type models.

vector autoregression:

$$\begin{pmatrix} \widehat{P}_{jk,t} \\ \widehat{N}_{jk,t} \\ \widehat{M}_{jk,t} \end{pmatrix} = \begin{pmatrix} \phi_{jk,P} \\ \phi_{jk,N} \\ \phi_{jk,M} \end{pmatrix} + \Phi_{jk,\text{Day}} \begin{pmatrix} \widehat{P}_{jk,t-1} \\ \widehat{N}_{jk,t-1} \\ \widehat{M}_{jk,t-1} \end{pmatrix} + \Phi_{jk,\text{Week}} \begin{pmatrix} \widehat{P}_{jk,t-2:t-5} \\ \widehat{N}_{jk,t-2:t-5} \\ \widehat{M}_{jk,t-2:t-5} \end{pmatrix}$$

$$+ \Phi_{jk,\text{Month}} \begin{pmatrix} \widehat{P}_{jk,t-6:t-22} \\ \widehat{N}_{jk,t-6:t-22} \\ \widehat{M}_{jk,t-6:t-22} \end{pmatrix} + \begin{pmatrix} \epsilon^{P}_{jk,t} \\ \epsilon^{N}_{jk,t} \\ \epsilon^{M}_{jk,t} \end{pmatrix}, \tag{14}$$

where $\widehat{P}_{t-l:t-k} \equiv \frac{1}{k-l+1} \sum_{s=l}^{k} \widehat{P}_{t-s}$, with the other components defined analogously.

The first three columns of Table II report the resulting parameter estimates averaged across 500 randomly selected (j, k) pairs of stocks. The table reveals a clear block structure in the coefficients of this general specification. Most notably, the dynamic dependencies in $\widehat{P}$ and $\widehat{N}$ are almost exclusively driven by the lagged $\widehat{N}$ terms, while the dynamic behavior of the mixed $\widehat{M}$ elements is primarily determined by their own lags, with the monthly lag receiving the largest weight.

The last two columns of Table II report the parameter estimates from regressing the realized covariances $\widehat{C}$ on the lagged realized semicovariances and the lagged covariances.[18] The model with individual semicovariances clearly reveals the most important

TABLE II

SEMICOVARIANCE HAR ESTIMATES[a]

	$\widehat{P}_{jk,t}$	$\widehat{N}_{jk,t}$	$\widehat{M}_{jk,t}$	$\widehat{C}_{jk,t}$	
$\widehat{P}_{jk,t-1}$	0.038 (0.67)	0.050 (0.67)	−0.035 (0.63)	0.052 (0.60)	
$\widehat{P}_{jk,t-2:t-5}$	0.004 (0.44)	0.057 (0.47)	−0.002 (0.34)	0.059 (0.44)	
$\widehat{P}_{jk,t-6:t-22}$	−0.074 (0.42)	0.023 (0.31)	0.099 (0.42)	0.048 (0.43)	
$\widehat{N}_{jk,t-1}$	0.248 (0.99)	0.192 (0.99)	−0.096 (0.95)	0.344 (0.94)	
$\widehat{N}_{jk,t-2:t-5}$	0.312 (0.99)	0.250 (0.97)	−0.090 (0.71)	0.472 (0.94)	
$\widehat{N}_{jk,t-6:t-22}$	0.349 (0.90)	0.206 (0.70)	−0.021 (0.31)	0.534 (0.84)	
$\widehat{M}_{jk,t-1}$	−0.075 (0.72)	−0.072 (0.74)	0.141 (1.00)	−0.006 (0.37)	
$\widehat{M}_{jk,t-2:t-5}$	−0.044 (0.47)	−0.049 (0.42)	0.209 (0.97)	0.116 (0.49)	
$\widehat{M}_{jk,t-6:t-22}$	0.028 (0.37)	−0.020 (0.23)	0.409 (0.99)	0.417 (0.87)	
$\widehat{C}_{jk,t-1}$					0.184 (0.94)
$\widehat{C}_{jk,t-2:t-5}$					0.305 (0.96)
$\widehat{C}_{jk,t-6:t-22}$					0.304 (0.99)
R^2	0.397	0.376	0.354	0.313	0.284
R^2_{adj}	0.395	0.374	0.352	0.311	0.283

[a]The table reports the average parameter estimates for the vector HAR model in (14) averaged across 500 randomly selected pairs of stocks. The first three columns report results for the unrestricted models. The fourth column reports the estimates from a model that restricts the rows of $\Phi_{jk,\text{Day}}$, $\Phi_{jk,\text{Week}}$, and $\Phi_{jk,\text{Month}}$ to be the same, corresponding to a model for $\widehat{C}_{j,kt}$, while the final column reports the results of a standard HAR model on $\widehat{C}_{jk,t}$. The fraction of randomly selected pairs of stocks for which the coefficient is significant at the 5% level is given in brackets.

[18]Note that due to the linear nature of the HAR model and the fact that realized semicovariances sum exactly to the realized covariance, each coefficient in the fourth column is simply the sum of the corresponding coefficients in the first three columns.

components: the three lags of $\widehat{N}$ and the monthly lag of $\widehat{M}$ constitute the main drivers of the realized covariance $\widehat{C}$. Interestingly, the models based on the semicovariances also put a greater weight on more recent information compared to the standard HAR model reported in the last column: normalizing each of the explanatory variables by their sample means, the semicovariance-based HAR models effectively put a weight of 0.339 on lagged daily information, while the final column shows that a standard HAR model on average puts a weight of only 0.184 on the daily lag, implying a more muted reaction to new information. These differences are naturally associated with an improved fit of the semicovariance-based models, as shown by the R^2 values. In the next section, we investigate whether this improved in-sample fit is accompanied by a similar improvement in out-of-sample forecast performance for models that utilize the realized semicovariances.

4.2. *Portfolio Volatility Forecasting*

The realized variance of a portfolio depends on the realized semicovariances of the assets included in the portfolio. In particular, utilizing the result that $\widehat{C} = \widehat{P} + \widehat{N} + \widehat{M}$, the realized variance of a portfolio with portfolio weights w may be expressed as

$$\begin{aligned}\widehat{RV}^w &\equiv w^\top \widehat{C} w \\ &= w^\top \widehat{P} w + w^\top \widehat{N} w + w^\top \widehat{M} w \\ &\equiv \widehat{P}^w + \widehat{N}^w + \widehat{M}^w,\end{aligned}$$

where we use the superscript w to indicate the relevant (scalar-valued) portfolio quantities. These portfolio semicovariance measures are distinct from the portfolio semivariances of Barndorff-Nielsen, Kinnebrock, and Shephard (2010), which only depend on the high-frequency returns of the portfolio. The portfolio semicovariances, on the other hand, depend on the high-frequency returns for *all* of the individual assets included in the portfolio, and cannot be computed using only returns on the portfolio itself.

To explore whether portfolio semicovariances convey useful information beyond realized variances and semivariances, we extend the HAR model of Corsi (2009) to allow the forecasts to depend on each of the realized portfolio semicovariance components. Accordingly, the one-step forecast for the portfolio return variance is constructed from

$$\begin{aligned}RV^w_{t+1|t} = \phi_0 &+ \phi_{\text{Day},P}\widehat{P}^w_t + \phi_{\text{Week},P}\widehat{P}^w_{t-1:t-4} + \phi_{\text{Month},P}\widehat{P}^w_{t-5:t-21} \\ &+ \phi_{\text{Day},N}\widehat{N}^w_t + \phi_{\text{Week},N}\widehat{N}^w_{t-1:t-4} + \phi_{\text{Month},N}\widehat{N}^w_{t-5:t-21} \\ &+ \phi_{\text{Day},M}\widehat{M}^w_t + \phi_{\text{Week},M}\widehat{M}^w_{t-1:t-4} + \phi_{\text{Month},M}\widehat{M}^w_{t-5:t-21}.\end{aligned} \tag{15}$$

We will refer to this model as the SemiCovariance HAR (SCHAR) model. The general SCHAR model in (15) is obviously quite richly parameterized. Hence, motivated by the results in Table II, we also consider a restricted version, in which we only include the daily, weekly, and monthly lags of $\widehat{N}^w$, and the monthly lag of $\widehat{M}^w$. We will refer to this specification as the restricted SCHAR model, or SCHAR-r for short.

If the parameters associated with the lagged realized semicovariance component all coincide (i.e., $\phi_{j,P} = \phi_{j,N} = \phi_{j,M} = \phi_j$ for $j \in \{\text{Day, Week, Month}\}$), the SCHAR model trivially reduces to the basic HAR model and the corresponding forecasting scheme

$$RV^w_{t+1|t} = \phi_0 + \phi_{\text{Day}}\widehat{RV}^w_t + \phi_{\text{Week}}\widehat{RV}^w_{t-1:t-4} + \phi_{\text{Month}}\widehat{RV}^w_{t-5:t-21}.$$

This simple and easy-to-implement model has arguably emerged as the benchmark for judging alternative realized volatility-based forecasting procedures in the literature.

In addition to this commonly used benchmark, we also consider the forecasts from the Semivariance HAR (SHAR) model of Patton and Sheppard (2015). This model uses the portfolio realized semivariances defined as

$$\widehat{PSV} = \sum_{i=1}^{[T/\Delta_n]} p(w^\top \Delta_i^n X)^2, \qquad \widehat{NSV} = \sum_{i=1}^{[T/\Delta_n]} n(w^\top \Delta_i^n X)^2,$$

to decompose the daily realized variance into positive and negative semivariances, resulting in the forecasting scheme:

$$RV^w_{t+1|t} = \phi_0 + \phi_{\mathrm{Day},+}\widehat{PSV}_t + \phi_{\mathrm{Day},-}\widehat{NSV}_t + \phi_{\mathrm{Week}}\widehat{RV}^w_{t-1:t-4} + \phi_{\mathrm{Month}}\widehat{RV}^w_{t-5:t-21}.$$

This model has been found to perform particularly well from the perspective of portfolio variance forecasting, performing on par with or better than the forecasts from other HAR-style models, and as such constitutes another particularly challenging benchmark.[19]

We consider equally-weighted portfolios composed of $d = 10$ ("small") and 100 ("large") stocks randomly selected from the full set of 749 individual stocks. In each case, we ensure that the selected stocks contain an overlap of at least 1,100 daily observations. We then construct rolling out-of-sample forecasts based on each of the different models, with model parameters re-estimated daily using the most recent 1,000 daily observations. We rely on the commonly used mean-square-error (MSE) and QLIKE loss functions to evaluate the performance of the forecasts vis-a-vis the actual portfolio realized variances $\widehat{RV}^w_{t+1}$.[20] Table III reports the resulting losses averaged across 500 randomly selected portfolios. In addition, for each of the 500 portfolios, we also compute the ratio of each model's average loss relative to the benchmark HAR model, and report the average of these ratios over all of the random samples.[21]

Consistent with Patton and Sheppard (2015), the SHAR-based forecasts that utilize the portfolio realized semivariances result in fairly large relative gains vis-a-vis the benchmark HAR-based forecasts, especially for the large dimensional portfolios. The performance of the unrestricted SCHAR model, however, is mixed: it has smaller MSE loss than the HAR forecasts, but underperforms that same benchmark under QLIKE loss. This finding is hardly surprising. There is ample evidence in the forecasting literature emphasizing the importance of parsimony (see, e.g., Zellner (1992)), and the results in Table II clearly suggest that the unrestricted SCHAR model is "over-parameterized," and as such is likely to perform poorly in a forecasting context. Indeed, looking at the results for the restricted SCHAR-r model guided by the estimates in Table II, we see that the forecasts from this model unambiguously outperform those from the other models; the 13.8% improvement in terms of predictive accuracy (measured by MSE) for the large portfolio case

[19]A comprehensive comparison of these models with the numerous other specifications that have appeared in the volatility forecasting literature (e.g., Andersen, Bollerslev, and Diebold (2007), Corsi, Pirino, and Renò (2010), and Caporin, Kolokolov, and Renò (2017)) would be interesting. In the interest of space, we leave a more extensive empirical analysis along these lines for future research.

[20]The MSE and QLIKE loss functions may both be formally justified for the purpose of volatility forecast evaluation based on the use of imperfect ex post volatility proxies; see Patton (2011).

[21]Supplemental Material Section S6 reports corroborative evidence from a series of additional robustness checks, including the results from multivariate models designed to forecast the full $d \times d$ dimensional realized covariance matrix.

TABLE III

PERFORMANCE COMPARISON FOR PORTFOLIO VARIANCE FORECASTS[a]

Model	MSE Average	MSE Ratio	QLIKE Average	QLIKE Ratio
Panel A. Small Portfolio Case (d =10)				
HAR	1.849	1.000	0.141	1.000
SHAR	1.671	0.966	0.139	0.986
SCHAR	1.643	0.955	0.210	1.318
SCHAR-r	1.567	0.908	0.139	0.979
Panel B. Large Portfolio Case (d =100)				
HAR	0.048	1.000	0.119	1.000
SHAR	0.045	0.935	0.115	0.957
SCHAR	0.045	0.976	0.236	1.495
SCHAR-r	0.041	0.862	0.111	0.925

[a]The table reports the loss for forecasting the portfolio variance for portfolios of size $d = 10$ and 100 for each of the different forecasting models. The reported numbers are based on 500 randomly selected portfolios. The Average column provides the average loss over time and all portfolios. The Ratio column gives the time-average ratio of losses across all sets of portfolios relative to the HAR model.

is particularly impressive. As such, this clearly shows the benefit of utilizing the additional information inherent in the realized semicovariances, compared to both the HAR- and SHAR-based forecasts that only rely on the portfolio realized (semi)variances.

These gains in forecast accuracy are not unique to the two portfolio dimensions highlighted in Table III. Figure 5 plots the median loss ratios for the HAR, SHAR, and SCHAR-r models for all values of d ranging from 2 to 100, together with the 10% and 90% quantiles for the SCHAR-r forecasts computed across the 500 random portfolio-formations. As the figure shows, the median loss ratios for the SCHAR-r model are systematically below those of the other two models, the only exception being the QLIKE loss for $d = 2$. Also, the gains from using the information in the realized semicovariances accrue relatively quickly as the number of stocks in the portfolio increases, and for the QLIKE loss appear to reach somewhat of a plateau for $d \approx 40$ stocks.

The gains in forecast accuracy obtained from using realized semicovariances are also not specific to the one-day forecast horizon. Table S.VI in Section S7 of the Supplemental Material reports results for 5-day- and 22-day-ahead forecasts, and shows that the forecast improvements obtained by the SCHAR-r model relative to the standard HAR model appear even larger over longer forecast horizons.

To help understand the source of these forecast improvements, it is instructive to represent the SHAR and SCHAR models as HAR models with time-varying parameters. To illustrate the idea, consider the forecasts from a SCHAR-type model based only on the lagged daily realized semicovariances,

$$\begin{aligned} RV^w_{t+1|t} &= \phi_0 + \phi_{\mathrm{Day},P}\widehat{P}^w_t + \phi_{\mathrm{Day},N}\widehat{N}^w_t + \phi_{\mathrm{Day},M}\widehat{M}^w_t \\ &= \phi_0 + \left(\phi_{\mathrm{Day},P}\frac{\widehat{P}^w_t}{\widehat{RV}^w_t} + \phi_{\mathrm{Day},N}\frac{\widehat{N}^w_t}{\widehat{RV}^w_t} + \phi_{\mathrm{Day},M}\frac{\widehat{M}^w_t}{\widehat{RV}^w_t}\right)\widehat{RV}^w_t \\ &\equiv \phi_0 + \phi_{\mathrm{Day},t}\widehat{RV}^w_t. \end{aligned}$$

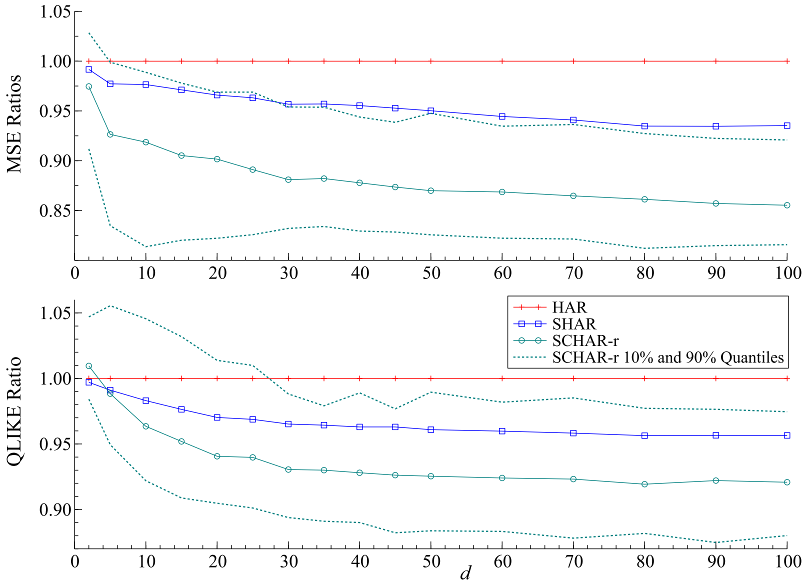

FIGURE 5.—Median loss ratios. *Note*: The graph plots the median loss ratios as a function of the number of stocks in the portfolio, d. The ratio is calculated as the average loss of the models divided by the average loss of the standard HAR, based on 500 random samples of d-stock portfolios.

As these equations show, even though the $\phi_{\mathrm{Day},P}$, $\phi_{\mathrm{Day},N}$, and $\phi_{\mathrm{Day},M}$ parameters used in the formulation of the model are all constant, the model may alternatively be interpreted as a first-order autoregression for $\widehat{RV}^{w}_{t+1}$ with a time-varying autoregressive parameter $\phi_{\mathrm{Day},t}$. This idea immediately extends to the SHAR and the more elaborate SCHAR-r forecasting models used in our empirical analysis, in which the parameters associated with the weekly and monthly lags in the implied HAR-type representations would be time-varying as well.

To study these effects, Figure 6 plots the daily, weekly, and monthly time-varying HAR parameters implied by the SHAR and SCHAR-r models for $d = 10$, averaged across 500 randomly selected ten-stock portfolios.[22] In addition to the implied daily, weekly, and monthly parameters, the last panel reports their sum as a measure of the overall persistence of the different models.

The daily, weekly, and monthly parameters for the HAR model are by definition all constant, with an average implied persistence of around 0.94. By comparison, the implied daily parameter estimates for the SHAR model vary slightly above the constant daily HAR parameter, while the constant weekly parameter for the SHAR model is slightly below that of the HAR model. As such, the overall persistence of the SHAR-based forecasts is generally close to that of the standard HAR-based forecasts.

[22]In contrast to the out-of-sample forecast results in Table III, which are based on a rolling estimation scheme, Figure 6 plots the implied parameter estimates obtained over the full sample period. Requiring observations to be available over the full sample reduces the number of stocks to 121. To avoid contaminating the results by a few influential outliers, we exclude any portfolios for which the maximum $\widehat{P}^w/\widehat{RV}^w$, $\widehat{N}^w/\widehat{RV}^w$, and $\widehat{M}^w/\widehat{RV}^w$ ratios exceed ten, and further smooth the implied parameters using a day $t-25$ to day $t+25$ moving average.

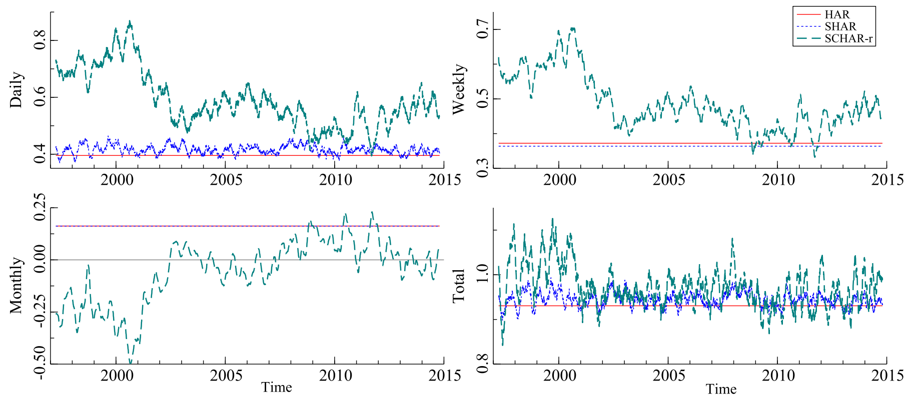

FIGURE 6.—Implied HAR parameters. *Note*: The figure plots the implied HAR-type parameters for predicting the variance of a ten-stock portfolio. The figure shows moving averages of the estimates, averaged across 500 randomly selected portfolios. The models are estimated over the full sample period.

By contrast, the implied time-varying daily and weekly HAR parameters for the SCHAR-r model both far exceed those of the standard HAR model, especially over the earlier part of the sample. On the other hand, the implied time-varying monthly parameters for the SCHAR-r model are typically less than the monthly parameters for the standard HAR and SHAR models. Meanwhile, the sum of the three implied parameters for the SCHAR-r model, which may be interpreted as an overall measure of persistence, are typically greater than those for the other two models. Thus, not only do the superior SCHAR-based forecasts respond more quickly to new information, the forecasts are typically also more persistent and slower to mean-revert than the benchmark HAR- and SHAR-based forecasts. This explains why the SCHAR-r model performs well not just over the short one-day forecast horizon, but also over weekly and monthly forecast horizons. It also highlights the usefulness of the information residing in the new realized semicovariance measures for volatility forecasting more generally.

5. CONCLUSION

We propose a decomposition of the realized covariance matrix based on the signs of the underlying high-frequency returns into positive, negative, and mixed-sign realized semicovariance components. Under a standard infill asymptotic setting for continuous-time Itô semimartingales, we derive the first- and second-order asymptotic properties of these new realized semicovariance measures. The asymptotic theory, taking the form of a non-central limit theorem, reveals the differential information carried by each of the realized semicovariance components, related to stochastic correlation, signed co-jumps, and notions of co-drifting and leverage effects. Using high-frequency data for a large cross-section of U.S. equities, we demonstrate how the asymptotic theory may be used to help understand key features of the realized semicovariance components. We further document distinctly different dynamic dependencies in the different realized semicovariance components. These differences in turn translate into superior forecast performance for models that utilize the realized semicovariance measures.

The infill asymptotic framework underlying much of the recent financial econometrics literature, including this paper, naturally suggests the use of high-frequency intraday data for the calculation of daily realized variation measures. However, the availability of reliable intraday data limits such analyses to relatively short and recent sample periods.

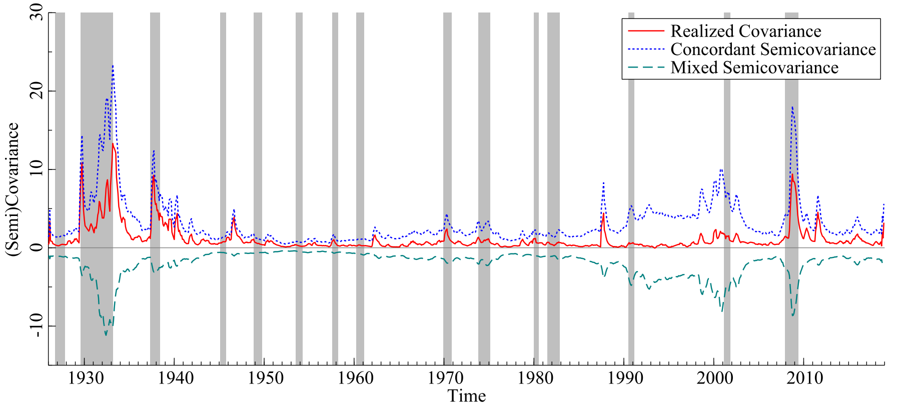

FIGURE 7.—Decomposition of monthly realized covariances. *Note*: The figure plots seven-month moving averages of the monthly time series of the concordant semicovariance ($\widehat{P}+\widehat{N}$), the mixed semicovariance ($\widehat{M}$), and the realized covariance ($\widehat{C}$). Each series is constructed as the average of the corresponding monthly realized measures averaged across 500 randomly selected pairs of CRSP stocks over the 1926–2018 period. The shaded regions represent NBER recession periods.

On the other hand, empirical work in the finance literature often employs daily data for the calculation of monthly or quarterly (co)variances over longer time periods (see, e.g., Schwert (1989)), which speak more directly to the longer holding periods that are commonplace in empirical finance. In this spirit, Figure 7 presents a lower-frequency analog to Figure 1, showing monthly covariances and semicovariances constructed from daily data spanning 1926 to 2018.[23] While the use of coarser daily returns in the construction of the monthly semicovariances introduces a wedge between the infill asymptotic theory and the empirical measures, thereby blurring the impact of co-jumps, short-lived co-drifts, and/or leverage effects for explaining the differences between the concordant semicovariance components, the figure still reveals notable differences through time in the relative importance of the concordant versus discordant components. In particular, the combined concordant semicovariance component generally, though not uniformly, increases in recessions more than the mixed component declines, leading to an overall increase in the total covariance. This is true for the Great Depression in the 1930s and the Great Recession in the late 2000s, as well as for the 1970s' oil-shock recession. However, it is not the case for the recessions of the early 1980s and 2000s. We leave further analyses of these intriguing features for future research.

APPENDIX A: REGULARITY CONDITIONS

The following regularity conditions are needed for our asymptotic theory.

ASSUMPTION 1: *The process X is an Itô semimartingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ of the form (3) with $J_t = \int_0^t \int_{\mathbb{R}} \delta(s,u)\mu(ds,du)$, where the process b is locally bounded, the process σ is càdlàg and takes value in $\mathbb{R}^{d\times d}$, δ is a predictable function, and μ is a Poisson random measure defined on $\mathbb{R}_+ \times \mathbb{R}$ with compensator $\nu(dt,du) = dt \otimes \lambda(du)$ for some finite measure λ on $\mathbb{R}$.*

[23]Analogously to Figure 1, each month we randomly select 500 pairs of stocks from that month's list of available stocks in the CRSP database, and average the monthly realized measures across all 500 pairs.

ASSUMPTION 2: *We have Assumption 1. Moreover, the process* σ *has the form* (5) *such that* (i) σ_t *is non-singular almost surely for all* t; (ii) $\tilde{b}$ *is locally bounded*; (iii) $\tilde{\sigma}$ *is* $d \times d \times d$ *càdlàg adapted process*; (iv) *the process* $\widetilde{M}$ *is a local martingale that is orthogonal to* W *with* $\|\Delta \widetilde{M}_t\| \leq \underline{\sigma}$ *for some constant* $\underline{\sigma} > 0$ *and its predictable quadratic covariation process has the form* $\langle \widetilde{M}, \widetilde{M} \rangle_t = \int_0^t \tilde{q}_s \, ds$ *for some locally bounded process* $\tilde{q}$; (v) *the compensator of the pure-jump process* $\sum_{s \leq t} \Delta \sigma_s 1_{\{\|\Delta \sigma_s\| > \underline{\sigma}\}}$ *has the form* $\int_0^t q_s \, ds$ *for some locally bounded process* q.

Overall, these assumptions are fairly mild and quite standard in the analysis of high-frequency data. In particular, they allow for price and volatility jumps and co-jumps, as well as the so-called leverage effect. The only notable restriction is the finite activity of the price jumps. As in Li, Todorov, and Tauchen (2017b), we purposely impose this condition because, in the current paper, the empirical interest vis-a-vis jumps mainly concerns "large" market-wide co-jumps, which occur relatively infrequently (and thus aligned with the finite-activity condition). Relaxing this condition would greatly complicate our technical exposition, without leading to any change in the actual numerical implementation.

APPENDIX B: PROOFS

B.1. *Notation and Preliminary Results*

We begin by defining some notation. Recall that for $j \in \{1, \ldots, d\}$, $\mathcal{T}_j = \{\tau : \Delta X_{j,\tau} \neq 0\}$ collects the jump times of asset j, and $\mathcal{T}_{j+} \equiv \{\tau \in \mathcal{T} : \Delta X_{j,\tau} > 0\}$ and $\mathcal{T}_{j-} \equiv \{\tau \in \mathcal{T} : \Delta X_{j,\tau} < 0\}$ collect the times at which asset j has positive and negative jumps, respectively. For each jump time τ, we denote by $i(\tau)$ the unique random integer i such that $(i-1)\Delta_n < \tau \leq i\Delta_n$. We then set

$$\mathcal{I}_j \equiv \{i(\tau) : \tau \in \mathcal{T}_j\} \quad \text{and} \quad \mathcal{I}_{j\pm} \equiv \{i(\tau) : \tau \in \mathcal{T}_{j\pm}\}.$$

Recall that $p(x) \equiv \max\{x, 0\}$ and $n(x) \equiv \min\{x, 0\}$. To simplify notation, we denote, for $x, y \in \mathbb{R}$, $f(x, y) \equiv p(x)p(y)$, $g(x, y) \equiv n(x)n(y)$, and $m(x, y) = p(x)n(y)$. Note that f and g are both continuously differentiable except on $\{(x, y) : x = 0 \text{ or } y = 0\}$, with gradients

$$\partial f(x, y) = \begin{pmatrix} 1_{\{x \geq 0\}} \max\{y, 0\} \\ \max\{x, 0\} 1_{\{y \geq 0\}} \end{pmatrix}, \qquad \partial g(x, y) = \begin{pmatrix} 1_{\{x \leq 0\}} \min\{y, 0\} \\ \min\{x, 0\} 1_{\{y \leq 0\}} \end{pmatrix}.$$

For generic functions h, h_1, and h_2 defined on $\mathbb{R}^d$, and an invertible $d \times d$ matrix a such that $c = aa^\top$, we define the following quantities:

$$\begin{cases} R_c(h) \equiv E_U[h(aU)], \qquad \gamma_c(h) \equiv E_U[h(aU)(aU)], \\ \hat{\gamma}_a(h) \equiv E_U[h(aU)U], \\ H_j \equiv h_j(aU) - R_c(h) - (c^{-1}\gamma_c(h_j))^\top aU, \quad j = 1, 2, \\ \bar{\gamma}_c(h_1, h_2) \equiv \mathrm{Cov}_U(H_1, H_2), \\ \Gamma_c(h_1, h_2) \equiv \mathrm{Cov}_U((c^{-1}\gamma_c(h_1))^\top aU, (c^{-1}\gamma_c(h_2))^\top aU), \\ \overline{\Gamma}_c(h_1, h_2) \equiv \mathrm{Cov}_U(h_1(aU), h_2(aU)), \end{cases} \tag{16}$$

where U is a generic d-dimensional standard normal variable, and E_U and Cov_U are the expectation and covariance operators with respect to U, respectively. We note that, except for $\hat{\gamma}_a(h)$, the expected values in (16) depend on a only through $c = aa^\top$, which explains our notation $R_c(\cdot)$, $\gamma_c(\cdot)$, etc.

We need to establish a few identities for these functionals. Observe that the variable H_j is the residual obtained from projecting the demeaned variable $h_j(aU) - E_U[h_j(aU)]$ onto aU, with $c^{-1}\gamma_c(h_j)$ being the corresponding projection coefficient. Since a is nonsingular, this residual may alternatively be written as $H_j = h_j(aU) - E_U[h_j(aU)] - \hat{\gamma}_a(h_j)^\top U$. Hence, the functional $\bar{\gamma}_c(h_1, h_2)$ can be rewritten as

$$\begin{aligned}\bar{\gamma}_c(h_1, h_2) = {} & E_U\big[\big(h_1(aU) - \hat{\gamma}_a(h_1)^\top U\big)\big(h_2(aU) - \hat{\gamma}_a(h_2)^\top U\big)\big] \\ & - E_U\big[h_1(aU)\big]E_U\big[h_2(aU)\big].\end{aligned} \tag{17}$$

We further note that $\overline{\Gamma}_c(h_1, h_2)$ computes the covariance of $h_1(aU) - E_U[h_1(aU)]$ and $h_2(aU) - E_U[h_2(aU)]$, while $\Gamma_c(h_1, h_2)$ computes the covariance of their projections. By a decomposition of covariance, we then deduce

$$\bar{\gamma}_c(h_1, h_2) = \overline{\Gamma}_c(h_1, h_2) - \Gamma_c(h_1, h_2). \tag{18}$$

In addition,

$$\Gamma_c(h_1, h_2) = \gamma_c(h_1)^\top c^{-1} \gamma_c(h_2). \tag{19}$$

Our analysis for the semicovariances relies on some explicit calculations of the expectations in (16). Lemma 1, below, provides the details for c taking the form

$$c = \begin{pmatrix} v_1^2 & \rho v_1 v_2 \\ \rho v_1 v_2 & v_2^2 \end{pmatrix}. \tag{20}$$

The proof is done by direct integration and is omitted for brevity.

LEMMA 1: *The following statements hold when the matrix c has the form* (20):

(a) $R_c(f) = R_c(g) = v_1 v_2 \psi(\rho)$ *and* $R_c(m) = -v_1 v_2 \psi(-\rho)$;

(b) $R_c(\partial f) = -R_c(\partial g) = (2\sqrt{2\pi})^{-1}(1+\rho)(v_2, v_1)^\top$;

(c) $\gamma_c(f) = -\gamma_c(g) = (2\sqrt{2\pi})^{-1}(1+\rho)^2 v_1 v_2 (v_1, v_2)^\top$;

(d) $\overline{\Gamma}_c(f, f) = \overline{\Gamma}_c(g, g) = v_1^2 v_2^2 (\Psi(\rho) - \psi(\rho)^2)$ *and* $\overline{\Gamma}_c(f, g) = -v_1^2 v_2^2 \psi(\rho)^2$.

B.2. *Proofs*

PROOF OF THEOREM 1: Let X' denote the continuous part of X, that is, $X'_t \equiv \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s$. We define $\widehat{P}'$ in the same way as $\widehat{P}$, but with X replaced by X'. It follows that

$$\widehat{P}_{jk} = \widehat{P}'_{jk} + \sum_{i \in \mathcal{I}_j \cup \mathcal{I}_k} p\big(\Delta_i^n X_j\big) p\big(\Delta_i^n X_k\big) - \sum_{i \in \mathcal{I}_j \cup \mathcal{I}_k} p\big(\Delta_i^n X'_j\big) p\big(\Delta_i^n X'_k\big).$$

By Theorem 3.4.1(b) in Jacod and Protter (2012) and Lemma 1(a),

$$\widehat{P}'_{jk} \xrightarrow{\mathbb{P}} \int_0^T v_{j,s} v_{k,s} \psi(\rho_{jk,s})\,ds. \tag{21}$$

We also note that $\sum_{i\in\mathcal{I}_j\cup\mathcal{I}_k} p(\Delta_i^n X_j)p(\Delta_i^n X_k) = \sum_{\tau\in\mathcal{T}_j\cup\mathcal{T}_k} p(\Delta_{i(\tau)}^n X_j)p(\Delta_{i(\tau)}^n X_k)$ and $\Delta_{i(\tau)}^n X \to \Delta X_\tau$ pathwise. Hence,

$$\sum_{i\in\mathcal{I}_j\cup\mathcal{I}_k} p\big(\Delta_i^n X_j\big)p\big(\Delta_i^n X_k\big) \xrightarrow{\mathbb{P}} P_{jk}^{\dagger}. \tag{22}$$

Finally, by a standard estimate for continuous Itô semimartingales and the fact that $\mathcal{I}_j\cup\mathcal{I}_k$ is almost surely finite, we deduce that $\sum_{i\in\mathcal{I}_j\cup\mathcal{I}_k} p(\Delta_i^n X_j')p(\Delta_i^n X_k') = O_p(\Delta_n)$. This estimate, combined with (21) and (22), implies the asserted convergence $\widehat{P}_{jk} \xrightarrow{\mathbb{P}} P_{jk}$. The proof for $\widehat{N}$ and $\widehat{M}$ follows essentially the same argument. *Q.E.D.*

PROOF OF THEOREM 2: Step 1. We begin by outlining the basic steps of the proof. First, we decompose $\Delta_n^{-1/2}(\widehat{P}_{12}-P_{12}) = \sum_{j=1}^5 \widetilde{P}^{(j)}$ and $\Delta_n^{-1/2}(\widehat{N}_{12}-N_{12}) = \sum_{j=1}^5 \widetilde{N}^{(j)}$, where

$$\begin{cases}
\widetilde{P}^{(1)} \equiv \Delta_n^{-1/2}\Bigg(\displaystyle\sum_{i\in\mathcal{I}_{1+}\cap\mathcal{I}_{2+}} p\big(\Delta_i^n X_1\big)p\big(\Delta_i^n X_2\big) - P_{12}^{\dagger}\Bigg),\\
\widetilde{P}^{(2)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{1+}\setminus\mathcal{I}_2} p\big(\Delta_i^n X_1\big)p\big(\Delta_i^n X_2\big),\\
\widetilde{P}^{(3)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{2+}\setminus\mathcal{I}_1} p\big(\Delta_i^n X_1\big)p\big(\Delta_i^n X_2\big),\\
\widetilde{P}^{(4)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{1-}\cup\mathcal{I}_{2-}} p\big(\Delta_i^n X_1\big)p\big(\Delta_i^n X_2\big),\\
\widetilde{P}^{(5)} \equiv \Delta_n^{-1/2}\Bigg(\displaystyle\sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} p\big(\Delta_i^n X_1\big)p\big(\Delta_i^n X_2\big) - P_{12}^{\star}\Bigg),
\end{cases}$$

and

$$\begin{cases}
\widetilde{N}^{(1)} \equiv \Delta_n^{-1/2}\Bigg(\displaystyle\sum_{i\in\mathcal{I}_{1-}\cap\mathcal{I}_{2-}} n\big(\Delta_i^n X_1\big)n\big(\Delta_i^n X_2\big) - N_{12}^{\dagger}\Bigg),\\
\widetilde{N}^{(2)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{1-}\setminus\mathcal{I}_2} n\big(\Delta_i^n X_1\big)n\big(\Delta_i^n X_2\big),\\
\widetilde{N}^{(3)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{2-}\setminus\mathcal{I}_1} n\big(\Delta_i^n X_1\big)n\big(\Delta_i^n X_2\big),\\
\widetilde{N}^{(4)} \equiv \Delta_n^{-1/2}\displaystyle\sum_{i\in\mathcal{I}_{1+}\cup\mathcal{I}_{2+}} n\big(\Delta_i^n X_1\big)n\big(\Delta_i^n X_2\big),\\
\widetilde{N}^{(5)} \equiv \Delta_n^{-1/2}\Bigg(\displaystyle\sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} n\big(\Delta_i^n X_1\big)n\big(\Delta_i^n X_2\big) - N_{12}^{\star}\Bigg).
\end{cases}$$

To prove the assertion of Theorem 2, it suffices to show the following joint stable convergence in law:

$$\Bigg(\sum_{j=1}^4 \widetilde{P}^{(j)}, \sum_{j=1}^4 \widetilde{N}^{(j)}\Bigg) \xrightarrow{\mathcal{L}\text{-}s} (\xi_P, \xi_N), \tag{23}$$

$$\begin{pmatrix} \widetilde{P}^{(5)} \\ \widetilde{N}^{(5)} \end{pmatrix} \xrightarrow{\mathcal{L}\text{-}s} \begin{pmatrix} B \\ -B \end{pmatrix} + \begin{pmatrix} L \\ -L \end{pmatrix} + \begin{pmatrix} \zeta \\ -\zeta \end{pmatrix} + \begin{pmatrix} \tilde{\zeta}_P \\ \tilde{\zeta}_N \end{pmatrix}. \tag{24}$$

In steps 2 and 3, below, we prove these claims in turn. We note that, by a standard argument, it is easy to show that the convergences in (23) and (24) hold jointly with $\mathcal{F}$-conditionally independent limits (as they involve non-overlapping Brownian increments). Therefore, the remaining task is to separately show (23) and (24).

Step 2. This step proves the claim in (23). Let $\mathcal{T} \equiv \mathcal{T}_1 \cup \mathcal{T}_2$ collect all jump times of the bivariate process X. We set, for each $\tau \in \mathcal{T}$, $\hat{\eta}_\tau = (\hat{\eta}_{1,\tau}, \hat{\eta}_{2,\tau})^\top \equiv \Delta_n^{-1/2}(\Delta_{i(\tau)}^n X - \Delta X_\tau)$. By Proposition 4.4.10 in Jacod and Protter (2012),

$$(\hat{\eta}_\tau)_{\tau\in\mathcal{T}} \xrightarrow{\mathcal{L}\text{-}s} (\tilde{\eta}_\tau)_{\tau\in\mathcal{T}}, \tag{25}$$

where $\tilde{\eta}_\tau = (\tilde{\eta}_{1,\tau}, \tilde{\eta}_{2,\tau})^\top \equiv \sqrt{\kappa_\tau}\tilde{\xi}_{\tau-} + \sqrt{1-\kappa_\tau}\tilde{\xi}_{\tau+}$ (recall the definitions in Section 2.2).

From (25), it is easy to derive the following representations uniformly for all $\tau \in \mathcal{T}$ (note that $\mathcal{T}$ is finite almost surely): for $j \in \{1, 2\}$,

$$p(\Delta_{i(\tau)}^n X_j) = \begin{cases} \Delta X_{j,\tau} + \Delta_n^{1/2}\hat{\eta}_{j,\tau} + o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T}_{j+}, \\ \Delta_n^{1/2} p(\hat{\eta}_{j,\tau}) + o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T} \setminus \mathcal{T}_j, \\ o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T}_{j-}, \end{cases}$$

and

$$n(\Delta_{i(\tau)}^n X_j) = \begin{cases} o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T}_{j+}, \\ \Delta_n^{1/2} n(\hat{\eta}_{j,\tau}) + o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T} \setminus \mathcal{T}_j, \\ \Delta X_{j,\tau} + \Delta_n^{1/2}\hat{\eta}_{j,\tau} + o_p(\Delta_n^{1/2}), & \text{if } \tau \in \mathcal{T}_{j-}. \end{cases}$$

Using these representations, we further deduce

$$\begin{cases} \widetilde{P}^{(1)} = \sum_{\tau\in\mathcal{T}_{1+}\cap\mathcal{T}_{2+}} (\Delta X_{1,\tau}\hat{\eta}_{2,\tau} + \Delta X_{2,\tau}\hat{\eta}_{1,\tau}) + o_p(1), \\ \widetilde{P}^{(2)} = \sum_{\tau\in\mathcal{T}_{1+}\setminus\mathcal{T}_2} \Delta X_{1,\tau} p(\hat{\eta}_{2,\tau}) + o_p(1), \\ \widetilde{P}^{(3)} = \sum_{\tau\in\mathcal{T}_{2+}\setminus\mathcal{T}_1} \Delta X_{2,\tau} p(\hat{\eta}_{1,\tau}) + o_p(1), \\ \widetilde{P}^{(4)} = o_p(1), \end{cases}$$

and

$$\begin{cases} \widetilde{N}^{(1)} = \sum_{\tau\in\mathcal{T}_{1-}\cap\mathcal{T}_{2-}} (\Delta X_{1,\tau}\hat{\eta}_{2,\tau} + \Delta X_{2,\tau}\hat{\eta}_{1,\tau}) + o_p(1), \\ \widetilde{N}^{(2)} = \sum_{\tau\in\mathcal{T}_{1-}\setminus\mathcal{T}_2} \Delta X_{1,\tau} n(\hat{\eta}_{2,\tau}) + o_p(1), \\ \widetilde{N}^{(3)} = \sum_{\tau\in\mathcal{T}_{2-}\setminus\mathcal{T}_1} \Delta X_{2,\tau} n(\hat{\eta}_{1,\tau}) + o_p(1), \\ \widetilde{N}^{(4)} = o_p(1). \end{cases}$$

These representations, together with the convergence (25), imply (23).

Step 3. This step proves the claim in (24). Let X' denote the continuous part of X, that is, $X'_t \equiv \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s$. We define $\widehat{P}'$ and $\widehat{N}'$ in the same way as $\widehat{P}$ and $\widehat{N}$, but with X replaced by X'. It is easy to see that $\widehat{P}' - \sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} p(\Delta_i^n X_1)p(\Delta_i^n X_2)$ and $\widehat{N}' -$

$\sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} n(\Delta_i^n X_1)n(\Delta_i^n X_2)$ are $O_p(\Delta_n)$. Hence, it suffices to prove the claim with $\widetilde{P}^{(5)}$ and $\widetilde{N}^{(5)}$ replaced by $\widetilde{P}' \equiv \Delta_n^{-1/2}(\widehat{P}'_{12} - P^\star_{12})$ and $\widetilde{N}' \equiv \Delta_n^{-1/2}(\widehat{N}'_{12} - N^\star_{12})$, respectively.

In order to derive the stable convergence in law of $(\widetilde{P}', \widetilde{N}')$, we apply Theorem 5.3.5 in Jacod and Protter (2012) to the test function $x \mapsto (p(x_1)p(x_2), n(x_1)n(x_2))$. To do so, it is enough to verify that the limiting variables $(B, -B)$, $(L, -L)$, $(\zeta, -\zeta)$, and $(\tilde{\zeta}_P, \tilde{\zeta}_N)$ coincide with Jacod and Protter's $\overline{A}$, $\overline{A}'$, $\overline{U}$, and $\overline{U}'$ variables. Turning to the details, we first note that, by Lemma 1(a), we can rewrite $P^\star_{12} = \int_0^T R_{c_s}(f)\,ds$ and $N^\star_{12} = \int_0^T R_{c_s}(g)\,ds$. By Lemma 1(b), we can rewrite the B term as $B = \int_0^T b_s^\top R_{c_s}(\partial f)\,ds = -\int_0^T b_s^\top R_{c_s}(\partial g)\,ds$. Given (6), no further rewriting is needed for the $(L, -L)$ term.

By Lemma 1(c), we can rewrite

$$\gamma_t = \gamma_{c_t}(f) \quad \text{and} \quad -\gamma_t = \gamma_{c_t}(g). \tag{26}$$

Hence, $\gamma_t = \sigma_t\hat{\gamma}_{\sigma_t}(f) = -\sigma_t\hat{\gamma}_{\sigma_t}(g)$ and, by (7), $\zeta = \int_0^T \hat{\gamma}_{\sigma_s}(f)^\top\,dW_s = -\int_0^T \hat{\gamma}_{\sigma_s}(g)^\top\,dW_s$.

By (26) and (19), we see that Γ_t defined in (8) can be written as

$$\Gamma_t = \begin{pmatrix} \Gamma_{c_t}(f,f) & \Gamma_{c_t}(f,g) \\ \Gamma_{c_t}(g,f) & \Gamma_{c_t}(g,g) \end{pmatrix}. \tag{27}$$

By Lemma 1(d), we see that $\overline{\Gamma}_t$ defined in (10) can be expressed equivalently as

$$\overline{\Gamma}_t = \begin{pmatrix} \overline{\Gamma}_{c_t}(f,f) & \overline{\Gamma}_{c_t}(f,g) \\ \overline{\Gamma}_{c_t}(g,f) & \overline{\Gamma}_{c_t}(g,g) \end{pmatrix}. \tag{28}$$

Recall that $\bar{\gamma}_t \equiv \overline{\Gamma}_t - \Gamma_t$. By (18), (27), and (28), it follows that

$$\bar{\gamma}_t = \begin{pmatrix} \bar{\gamma}_{c_t}(f,f) & \bar{\gamma}_{c_t}(f,g) \\ \bar{\gamma}_{c_t}(g,f) & \bar{\gamma}_{c_t}(g,g) \end{pmatrix}.$$

In view of (17), we verify that the local quadratic variation of the $(\tilde{\zeta}_P, \tilde{\zeta}_N)$ term coincides with that defined in (5.2.4) of Jacod and Protter (2012).

We are now ready to apply Theorem 5.3.5 in Jacod and Protter (2012) to finish the proof of (24), and hence, the assertion of Theorem 2. *Q.E.D.*

PROOF OF PROPOSITION 1: By Proposition 1 in Li, Todorov, and Tauchen (2017b), the set $\widehat{\mathcal{I}}$ coincides with $\mathcal{I}_1 \cup \mathcal{I}_2$ with probability approaching 1 and, in restriction to that sequence of events,

$$\widehat{P}^\star_{12} = \sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} p(\Delta_i^n X_1)p(\Delta_i^n X_2), \qquad \widehat{P}^\dagger_{12} = \sum_{i\in\mathcal{I}_1\cup\mathcal{I}_2} p(\Delta_i^n X_1)p(\Delta_i^n X_2),$$

$$\widehat{N}^\star_{12} = \sum_{i\notin\mathcal{I}_1\cup\mathcal{I}_2} n(\Delta_i^n X_1)n(\Delta_i^n X_2), \qquad \widehat{N}^\dagger_{12} = \sum_{i\in\mathcal{I}_1\cup\mathcal{I}_2} n(\Delta_i^n X_1)n(\Delta_i^n X_2).$$

The assertion then follows from (23) and (24) in the proof of Theorem 2. *Q.E.D.*

PROOF OF PROPOSITION 2: Let $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$. For notational simplicity, we denote, for each subset $\mathcal{S} \subseteq \mathcal{T}$,

$$\xi_P^*(\mathcal{S}) \equiv \Delta_n^{-1/2} \sum_{\tau \in \mathcal{S}} \Big(p\big(\Delta_{i(\tau)}^n X_1^* + \Delta_n^{1/2} \tilde{\eta}_{i(\tau),1}^*\big) p\big(\Delta_{i(\tau)}^n X_2^* + \Delta_n^{1/2} \tilde{\eta}_{i(\tau),2}^*\big) - p\big(\Delta_{i(\tau)}^n X_1^*\big) p\big(\Delta_{i(\tau)}^n X_2^*\big)\Big). \tag{29}$$

By Proposition 1 in Li, Todorov, and Tauchen (2017b), $\widehat{\mathcal{I}}$ coincides with $\mathcal{I}_1 \cup \mathcal{I}_2$ with probability approaching 1. Hence, we can restrict our calculations to that sequence of events without loss of generality. In particular, we can write ξ_P^* as $\xi_P^*(\mathcal{T})$ using the notation (29). We can then decompose ξ_P^* as

$$\xi_P^* = \xi_P^*(\mathcal{T}_{1+} \cap \mathcal{T}_{2+}) + \xi_P^*(\mathcal{T}_{1+} \setminus \mathcal{T}_2) + \xi_P^*(\mathcal{T}_{2+} \setminus \mathcal{T}_1) + \xi_P^*(\mathcal{T}_{1-} \cup \mathcal{T}_{2-}).$$

We note that $\Delta_{i(\tau)}^n X_j^* = 0$ for all $\tau \in \mathcal{T} \setminus \mathcal{T}_j$ with probability approaching 1. It is then easy to deduce that

$$p\big(\Delta_{i(\tau)}^n X_j^* + \Delta_n^{1/2} \tilde{\eta}_{i(\tau),j}^*\big) = \begin{cases} p\big(\Delta_{i(\tau)}^n X_j^*\big) + \Delta_n^{1/2} \tilde{\eta}_{i(\tau),j}^* + o_p\big(\Delta_n^{1/2}\big), & \text{if } \tau \in \mathcal{T}_{j+}, \\ \Delta_n^{1/2} p\big(\tilde{\eta}_{i(\tau),j}^*\big) + o_p\big(\Delta_n^{1/2}\big), & \text{if } \tau \in \mathcal{T} \setminus \mathcal{T}_j, \\ o_p\big(\Delta_n^{1/2}\big), & \text{if } \tau \in \mathcal{T}_{j-}. \end{cases}$$

From there, we readily deduce that

$$\begin{cases} \xi_P^*(\mathcal{T}_{1+} \cap \mathcal{T}_{2+}) = \displaystyle\sum_{\tau \in \mathcal{T}_{1+} \cap \mathcal{T}_{2+}} \big(p\big(\Delta_{i(\tau)}^n X_1^*\big) \tilde{\eta}_{i(\tau),2}^* + p\big(\Delta_{i(\tau)}^n X_2^*\big) \tilde{\eta}_{i(\tau),1}^* \big) + o_p(1), \\ \xi_P^*(\mathcal{T}_{1+} \setminus \mathcal{T}_2) = \displaystyle\sum_{\tau \in \mathcal{T}_{1+} \setminus \mathcal{T}_2} p\big(\Delta_{i(\tau)}^n X_1^*\big) p\big(\tilde{\eta}_{i(\tau),2}^*\big) + o_p(1), \\ \xi_P^*(\mathcal{T}_{2+} \setminus \mathcal{T}_1) = \displaystyle\sum_{\tau \in \mathcal{T}_{2+} \setminus \mathcal{T}_1} p\big(\Delta_{i(\tau)}^n X_2^*\big) p\big(\tilde{\eta}_{i(\tau),1}^*\big) + o_p(1), \\ \xi_P^*(\mathcal{T}_{1-} \cup \mathcal{T}_{2-}) = o_p(1). \end{cases} \tag{30}$$

By a standard result for spot covariance estimation (see, e.g., Theorem 9.3.2 in Jacod and Protter (2012)), we have $\hat{c}_{i(\tau)\pm} \xrightarrow{\mathbb{P}} c_{\tau\pm}$. Consequently, $(\tilde{\eta}_{i(\tau)}^*)_{\tau \in \mathcal{T}} \xrightarrow{\mathcal{L}|\mathcal{F}} (\tilde{\eta}_\tau)_{\tau \in \mathcal{T}}$, where $\xrightarrow{\mathcal{L}|\mathcal{F}}$ denotes the convergence in probability of $\mathcal{F}$-conditional laws under the uniform metric. In addition, we note that $p(\Delta_{i(\tau)}^n X_j^*) \xrightarrow{\mathbb{P}} \Delta X_{j,\tau}$ for all $\tau \in \mathcal{T}_{j+}$. Combining these convergence results with (30), we deduce that $\xi_P^* \xrightarrow{\mathcal{L}|\mathcal{F}} \xi_P$. Evidently, the same argument can be used to show that $\xi_N^* \xrightarrow{\mathcal{L}|\mathcal{F}} \xi_N$, jointly with $\xi_P^* \xrightarrow{\mathcal{L}|\mathcal{F}} \xi_P$, which implies the first assertion of Proposition 2 (i.e., $\xi_P^* - \xi_N^* \xrightarrow{\mathcal{L}|\mathcal{F}} \xi_P - \xi_N$). The size and power properties of the test readily follow from here. *Q.E.D.*

PROOF OF PROPOSITION 3: Let $\mathcal{G}$ be the σ-field generated by $(b_t, \sigma_t)_{0 \le t \le T}$. Since $\tilde{\sigma} = 0$, $L = 0$ by definition (recall (6)). Proposition 1 then implies that $\Delta_n^{-1/2}(\widehat{P}_{12}^\star - \widehat{N}_{12}^\star) - 2B \xrightarrow{\mathcal{L}\text{-}s} 2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N$. Since $\mathcal{G}$ and W are independent, and $\widetilde{W}$ in the definition of $\tilde{\zeta}$ is independent of $\mathcal{F}$, we see that $2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N$ is, conditional on $\mathcal{G}$, centered Gaussian with conditional variance $\Sigma^\star$ (recall (13)). By Theorem 3 of Li, Todorov, and Tauchen (2017a),

$\widehat{\Sigma}^\star \to \Sigma^\star$. By the property of stable convergence in law, we have

$$\frac{\Delta_n^{-1/2}(\widehat{P}^\star_{12} - \widehat{N}^\star_{12}) - 2B}{\sqrt{\widehat{\Sigma}^\star}} \xrightarrow{\mathcal{L}\text{-}s} \frac{2\zeta + \tilde{\zeta}_P - \tilde{\zeta}_N}{\sqrt{\Sigma^\star}} \equiv \mathcal{N},$$

where, by definition, $\mathcal{N}$ is a standard normal variable (defined on an extension of the original probability space) that is independent of $\mathcal{G}$. In restriction to $\Omega_0 = \{B = 0\}$, we have $\Delta_n^{-1/2}(\widehat{P}^\star_{12} - \widehat{N}^\star_{12})/\sqrt{\widehat{\Sigma}^\star} \xrightarrow{\mathcal{L}\text{-}s} \mathcal{N}$. Since $\Omega_0 \in \mathcal{G}$, this convergence readily implies the first assertion of the proposition concerning the null rejection probability.

To prove the second assertion, we fix some constant $R > 0$. We suppose that $\mathbb{P}(\Omega_a) > 0$ without loss of generality because, otherwise, the conditional probability $\mathbb{P}(C_n^+|\Omega_a)$ can be arbitrarily defined and there would be nothing to prove. By the property of stable convergence, we have

$$\begin{aligned}
&\mathbb{P}\left(\left\{\frac{\Delta_n^{-1/2}(\widehat{P}^\star_{12} - \widehat{N}^\star_{12})}{\sqrt{\widehat{\Sigma}^\star}} > z_\alpha\right\} \cap \Omega_a\right) \\
&\quad = \mathbb{P}\left(\left\{\frac{\Delta_n^{-1/2}(\widehat{P}^\star_{12} - \widehat{N}^\star_{12}) - 2B}{\sqrt{\widehat{\Sigma}^\star}} > z_\alpha - \frac{2B}{\sqrt{\widehat{\Sigma}^\star}}\right\} \cap \Omega_a\right) \\
&\quad = \mathbb{P}\left(\left\{\mathcal{N} > z_\alpha - \frac{2B}{\sqrt{\Sigma^\star}}\right\} \cap \Omega_a\right) + o(1) \\
&\quad \geq \mathbb{P}\big(\{\mathcal{N} > z_\alpha - R\} \cap \Omega_a\big) + o(1).
\end{aligned}$$

Note that $\Omega_a \in \mathcal{G}$. Since $\mathcal{N}$ is independent of $\mathcal{G}$, $\mathbb{P}(\{\mathcal{N} > z_\alpha - R\} \cap \Omega_a) = (1 - \Phi(z_\alpha - R))\mathbb{P}(\Omega_a)$. Hence,

$$\liminf_n \mathbb{P}\left(\left\{\frac{\Delta_n^{-1/2}(\widehat{P}^\star_{12} - \widehat{N}^\star_{12})}{\sqrt{\widehat{\Sigma}^\star}} > z_\alpha\right\} \cap \Omega_a\right) \geq \big(1 - \Phi(z_\alpha - R)\big)\mathbb{P}(\Omega_a).$$

Dividing both sides by $\mathbb{P}(\Omega_a)$, we finish the proof of the second assertion of this proposition. *Q.E.D.*

REFERENCES

AÏT-SAHALIA, Y., AND J. JACOD (2009): "Testing for Jumps in a Discretely Observed Process," *The Annals of Statistics*, 37, 184–222. [1530]

——— (2014): *High-Frequency Financial Econometrics*. Princeton, NJ: Princeton University Press. [1518, 1529,1530]

AÏT-SAHALIA, Y., AND D. XIU (2016): "Increased Correlation Among Asset Classes: Are Volatility or Jumps to Blame, or Both?" *Journal of Econometrics*, 194 (2), 205–219. [1519]

AMAYA, D., P. CHRISTOFFERSEN, K. JACOBS, AND A. VASQUEZ (2015): "Does Realized Skewness Predict the Cross-Section of Equity Returns?" *Journal of Financial Economics*, 118 (1), 135–167. [1519]

ANDERSEN, T. G., AND T. BOLLERSLEV (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158. [1522]

ANDERSEN, T. G., T. BOLLERLSEV, F. X. DIEBOLD, AND P. LABYS (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71 (2), 579–625. [1515,1516]

ANDERSEN, T. G., T. BOLLERSLEV, AND F. X. DIEBOLD (2007): "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89 (4), 701–720. [1531,1534,1537]
ANG, A., AND J. CHEN (2002): "Asymmetric Correlations of Equity Portfolios," *Journal of Financial Economics*, 63 (3), 443–494. [1519]
BACK, K. (1992): "Insider Trading in Continuous Time," *Review of Financial Studies*, 5 (3), 387–409. [1531]
BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2004a): "Power and Bipower Variation With Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2 (1), 1–37. [1531]
——— (2004b): "Econometric Analysis of Realized Covariation: High-Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72 (3), 885–925. [1516]
——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4 (1), 1–30. [1530,1531]
BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2008): "Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536. [1531]
——— (2011): "Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices With Noise and Non-Synchronous Trading," *Journal of Econometrics*, 162 (2), 149–169. [1516]
BARNDORFF-NIELSEN, O. E., S. KINNEBROCK, AND N. SHEPHARD (2010): "Measuring Downside Risk: Realised Semivariance," in *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, ed. by T. Bollerslev, J. R. Russell, and M. W. Watson. Oxford: Oxford University Press, 117–136. [1516,1518,1524, 1529,1530,1536]
BAUWENS, L., S. LAURENT, AND J. V. K. ROMBOUTS (2006): "Multivariate GARCH Models: A Survey," *Journal of Applied Econometrics*, 21 (1), 79–109. [1515]
BOLLERSLEV, T., AND V. TODOROV (2011a): "Estimation of Jump Tails," *Econometrica*, 79 (6), 1727–1783. [1531]
——— (2011b): "Tails, Fears, and Risk Premia," *Journal of Finance*, 66 (6), 2165–2211. [1531]
BOLLERSLEV, T., T. LAW, AND G. TAUCHEN (2008): "Risk, Jumps and Diversification," *Journal of Econometrics*, 144 (1), 234–256. [1519,1531]
BOLLERSLEV, T., J. LI, A. J. PATTON, AND R. QUAEDVLIEG (2020): "Supplement to 'Realized Semicovariances'," *Econometrica Supplemental Material*, 88, https://doi.org/10.3982/ECTA17056. [1520]
BOLLERSLEV, T., J. LI, AND Y. XUE (2018): "Volume, Volatility, and Public News Announcements," *Review of Economic Studies*, 85 (4), 2005–2041. [1522]
BOLLERSLEV, T., A. J. PATTON, AND R. QUAEDVLIEG (2020): "Multivariate Leverage Effects and Realized Semicovariance GARCH Models," *Journal of Econometrics*, 217 (2), 411–430. [1534]
CAPORIN, M., A. KOLOKOLOV, AND R. RENÒ (2017): "Systemic Co-Jumps," *Journal of Financial Economics*, 126 (3), 563–591. [1530,1531,1537]
CAPPIELLO, L., R. ENGLE, AND K. SHEPPARD (2006): "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4 (4), 537–572. [1519]
COLLIN-DUFRESNE, P., AND V. FOS (2016): "Insider Trading, Stochastic Liquidity, and Equilibrium Prices," *Econometrica*, 84 (4), 1441–1475. [1531]
CORSI, F. (2009): "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 7 (2), 174–196. [1519,1534,1536]
CORSI, F., D. PIRINO, AND R. RENÒ (2010): "Threshold Bipower Variation and the Impact of Jumps on Volatility Forecasting," *Journal of Econometrics*, 159 (2), 276–288. [1537]
DAS, S. R., AND R. UPPAL (2004): "Systemic Risk and International Portfolio Choice," *Journal of Finance*, 59 (6), 2809–2834. [1519]
DOVONON, P., S. GONCALVES, U. HOUNYO, AND N. MEDDAHI (2019): "Bootstrapping High-Frequency Jump Tests," *Journal of the American Statistical Association*, 114, 793–803. [1530]
ELTON, E. J., AND M. J. GRUBER (1973): "Estimating the Dependence Structure of Share Prices—Implications for Portfolio Selection," *Journal of Finance*, 28 (5), 1203–1232. [1515]
FISHBURN, P. C. (1977): "Mean-Risk Analysis With Risk Associated With Below-Target Returns," *American Economic Review*, 67 (2), 116–126. [1516]
GONÇALVES, S., AND N. MEDDAHI (2009): "Bootstrapping Realized Volatility," *Econometrica*, 77, 283–306. [1530]
HANSEN, P. R., AND A. LUNDE (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business and Economic Statistics*, 24 (2), 127–161. [1531]
——— (2014): "Estimating the Persistence and the Autocorrelation Function of a Time Series That Is Measured With Error," *Econometric Theory*, 30 (1), 60–93. [1534]

HANSEN, P. R., Z. HUANG, AND H. H. SHEK (2012): "Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27 (6), 877–906. [1533]

HOGAN, W. W., AND J. M. WARREN (1972): "Computation of the Efficient Boundary in the E-S Portfolio Selection Model," *Journal of Financial and Quantitative Analysis*, 7 (4), 1881–1896. [1516]

——— (1974): "Toward the Development of an Equilibrium Capital-Market Model Based on Semivariance," *Journal of Financial and Quantitative Analysis*, 9 (1), 1–11. [1516]

HONG, Y., J. TU, AND G. ZHOU (2007): "Asymmetries in Stock Returns: Statistical Tests and Economic Interpretation," *Review of Financial Studies*, 20 (5), 1547–1581. [1519]

JACOD, J. (1997): "On Continuous Conditional Gaussian Martingales and Stable Convergence in Law," in *Séminaire de Probabilités*. Berlin: Springer-Verlag, 232–246. [1530]

JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Berlin: Springer-Verlag. [1521,1524,1529,1530, 1543,1545-1547]

JACOD, J., AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes* (Second Ed.). New York: Springer-Verlag. [1522,1530]

JACOD, J., AND V. TODOROV (2009): "Testing for Common Arrivals of Jumps for Discretely Observed Multidimensional Processes," *The Annals of Statistics*, 37 (4), 1792–1838. [1519,1530]

JACOD, J., Y. LI, AND X. ZHENG (2017): "Statistical Properties of Microstructure Noise," *Econometrica*, 85 (4), 1133–1174. [1531]

KENDALL, M. G. (1953): "The Analysis of Economic Time-Series—Part I: Prices," *Journal of the Royal Statistical Society, Series A*, 116 (1), 11–34. [1515]

KINNEBROCK, S., AND M. PODOLSKIJ (2008): "A Note on the Central Limit Theorem for Bipower Variation of General Functions," *Stochastic Processes and Their Applications*, 118 (6), 1056–1070. [1524,1529,1530]

KRISTENSEN, D. (2010): "Nonparametric Filtering of the Realized Spot Volatility: A Kernel-Based Approach," *Econometric Theory*, 26 (1), 60–93. [1530]

KRONER, K. F., AND V. K. NG (1998): "Modeling Asymmetric Comovements of Asset Returns," *Review of Financial Studies*, 11 (4), 817–844. [1519]

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53 (6), 1315–1335. [1531]

LAHAYE, J., S. LAURENT, AND C. J. NEELY (2011): "Jumps, Cojumps and Macro Announcements," *Journal of Applied Econometrics*, 26 (6), 893–921. [1531]

LEE, S. S. (2012): "Jumps and Information Flow in Financial Markets," *Review of Financial Studies*, 25 (2), 439–479. [1531]

LEE, S. S., AND P. A. MYKLAND (2008): "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21 (6), 2535–2563. [1519,1531]

LI, J., AND D. XIU (2016): "Generalized Method of Integrated Moments for High-Frequency Data," *Econometrica*, 84 (4), 1613–1633. [1530]

LI, J., V. TODOROV, AND G. TAUCHEN (2017a): "Adaptive Estimation of Continuous-Time Regression Models Using High-Frequency Data," *Journal of Econometrics*, 200 (1), 36–47. [1528,1530,1547]

——— (2017b): "Jump Regressions," *Econometrica*, 85 (1), 173–195. [1519,1525,1542,1546,1547]

LI, J., V. TODOROV, G. TAUCHEN, AND R. CHEN (2017): "Mixed-Scale Jump Regressions With Bootstrap Inference," *Journal of Econometrics*, 201 (2), 417–432. [1531]

LI, Y., P. A. MYKLAND, E. RENAULT, L. ZHANG, AND X. ZHENG (2014): "Realized Volatility When Sampling Times Are Possibly Endogenous," *Econometric Theory*, 30 (3), 580–605. [1524,1530]

LONGIN, F., AND B. SOLNIK (2001): "Extreme Correlation of International Equity Markets," *Journal of Finance*, 56 (2), 649–676. [1519]

MAHEU, J. M., AND T. H. MCCURDY (2004): "News Arrival, Jump Dynamics, and Volatility Components for Indivdual Stock Returns," *Journal of Finance*, 59 (2), 755–793. [1534]

MANCINI, C. (2001): "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, LXIV, 19–47. [1519,1525]

——— (2009): "Non-Parametric Threshold Estimation for Models With Stochastic Diffusion Coefficient and Jumps," *Scandinavian Journal of Statistics*, 36 (2), 270–296. [1525]

MANCINI, C., AND F. GOBBI (2012): "Identifying the Brownian Covariation From the Co-Jumps Given Discrete Observations," *Econometric Theory*, 28 (2), 249–273. [1519]

MAO, J. C. T. (1970): "Survey of Capital Budgeting: Theory and Practice," *Journal of Finance*, 25 (2), 349–360. [1516]

MARKOWITZ, H. M. (1959): *Portfolio Selection*. New York: Wiley. [1516]

NEUBERGER, A. (2012): "Realized Skewness," *Review of Financial Studies*, 25 (11), 3423–3455. [1519]

NOURELDIN, D., N. SHEPHARD, AND K. SHEPPARD (2012): "Multivariate High-Frequency-Based Volatility (HEAVY) Models," *Journal of Applied Econometrics*, 27 (6), 907–933. [1533]

PATTON, A. J. (2004): "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation," *Journal of Financial Econometrics*, 2 (1), 130–168. [1519]

——— (2011): "Volatility Forecast Comparison Using Imperfect Volatility Proxies," *Journal of Econometrics*, 160 (1), 246–256. [1537]

PATTON, A. J., AND K. SHEPPARD (2015): "Good Volatility, Bad Volatility: Signed Jumps and the Persistence of Volatility," *Review of Economics and Statistics*, 97 (3), 683–697. [1519,1537]

POON, S.-H., M. ROCKINGER, AND J. TAWN (2004): "Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications," *Review of Financial Studies*, 17 (2), 581–610. [1519]

RENÒ, R. (2008): "Nonparametric Estimation of the Diffusion Coefficient of Stochastic Volatility Models," *Econometric Theory*, 24, 1174–1206. [1530]

SCHWERT, G. W. (1989): "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance*, 44 (5), 1115–1153. [1541]

TJØSTHEM, D., AND K. O. HUFTHAMMER (2013): "Local Gaussian Correlation: A New Measure of Dependence," *Journal of Econometrics*, 172 (1), 33–48. [1519]

ZELLNER, A. (1992): "Presidential Address: Statistics, Science and Public Policy," *Journal of the American Statistical Association*, 87 (417), 1–6. [1537]

ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [1531]

Co-editor Aviv Nevo handled this manuscript.

Manuscript received 8 February, 2019; final version accepted 18 February, 2020; available online 19 February, 2020.