High-Frequency Covariance Estimates With Noisy and Asynchronous Financial Data

Yacine AÏT-SAHALIA, Jianqing FAN, and Dacheng XIU

This article proposes a consistent and efficient estimator of the high-frequency covariance (quadratic covariation) of two arbitrary assets, observed asynchronously with market microstructure noise. This estimator is built on the marriage of the quasi–maximum likelihood estimator of the quadratic variation and the proposed generalized synchronization scheme and thus is not influenced by the Epps effect. Moreover, the estimation procedure is free of tuning parameters or bandwidths and is readily implementable. Monte Carlo simulations show the advantage of this estimator by comparing it with a variety of estimators with specific synchronization methods. The empirical studies of six foreign exchange future contracts illustrate the time-varying correlations of the currencies during the 2008 global financial crisis, demonstrating the similarities and differences in their roles as key currencies in the global market.

KEY WORDS: Covariance; Generalized synchronization method; Market microstructure noise; Quasi-Maximum Likelihood Estimator; Refresh Time.

1. INTRODUCTION

The covariation between asset returns plays a crucial role in modern finance. For instance, the covariance matrix and its inverse are the key statistics in portfolio optimization and risk management. Many recent financial innovations involve complex derivatives, like exotic options written on the minimum, maximum or difference of two assets, or some structured financial products, such as CDOs. All of these innovations are built upon, or in order to exploit, the correlation structure of two or more assets. As technological developments make high frequency data commonly available, much effort has been put into developing statistical inference methodologies for continuous time models with intra-day data, enabling us to capture the daily variation of some interesting statistics that were otherwise unobservable from daily or weekly data.

Realized variance estimation is an example of such statistics. Unfortunately, unlike those low frequency time series that are homogeneously spaced, tick-by-tick transactions of different assets usually occur randomly and asynchronously; in addition, with high frequency data comes market microstructure noise. These factors make it difficult to employ a Realized Covariance (RC) estimator directly. Popular estimators in the univariate variance case include Two Scales Realized Volatility (TSRV) of Zhang, Mykland, and Aït-Sahalia (2005), the first consistent estimator for integrated volatility in the presence of noise, Multi-Scale Realized Volatility (MSRV), a modification of TSRV which achieves the best possible rate of convergence proposed by Zhang (2006), Realized Kernels (RK) by Barndorff-Nielsen et al. (2008a) and the Pre-Averaging (PA) approach by Jacod et al. (2009), both of which contain sets of nonparametric estimators that can also achieve the best convergence rate. In contrast with these nonparametric estimators, Xiu (2010) has extended the Maximum-Likelihood Estimator of Aït-Sahalia, Mykland, and Zhang (2005) designed for parametric volatility to a Quasi-Maximum Likelihood Estimator (QMLE), as a misspecified maximum likelihood estimator in the setting of stochastic volatility. This parametric estimator has also proved to be consistent and efficient without any tuning parameters. Related work includes Bandi and Russell (2003), Delattre and Jacod (1997), Fan and Wang (2007), Gatheral and Oomen (2009), Hansen and Lunde (2006), Kalnina and Linton (2008), Li and Mykland (2007), Aït-Sahalia, Mykland, and Zhang (2009), Zhang, Mykland, and Aït-Sahalia (2009), and Li et al. (2010).

These advances in variance estimation pave the way for the efficient estimation of covariance with noisy data. Studies of correlation estimated from asynchronous high frequency stock price returns date back to at least Epps (1979), who documented the Epps effect, i.e., the fact that the sample correlation tends to have a strong bias towards zero as the sampling interval progressively shrinks. The same effect has been documented for exchange rates, see, e.g., Guillaume et al. (1997) and Muthuswamy et al. (2001). Since then, researchers have been trying to resolve this puzzle. There are at least two possible approaches. First, dealing with asynchronous data can be achieved with subsampling, using previous tick or other interpolation methods. This procedure may induce a potential bias. Hayashi and Yoshida (2005) have proposed a modification of the Realized Covariance (RC) estimator (HY), which is consistent and immune to this bias. Also, it is reasonable to conjecture that microstructure noise might be at least partly responsible for the documented bias as well, based on the fact that the magnitude of the noise relative to that of the price signal will increase the realized variance estimator, which serves as the denominator in the correlation calculation: see Large (2007) and Griffin and Oomen (2008), for example. Voev and Lunde (2007) provided a bias correction procedure for the HY estimator to correct for the effect of the noise, but the estimator does not achieve consistency. Zhang (2009) has demonstrated theoretically that there may be a bias associated with

Yacine Aït-Sahalia is the Otto A. Hack 1903 Professor of Finance and Economics, Department of Economics, Jianqing Fan is the Frederick L. Moore Professor of Finance and Statistics, Department of Operation Research and Financial Engineering, and Dacheng Xiu is a Ph.D. Student, Bendheim Center for Finance (E-mail: *dachengx@princeton.edu*), Princeton University, Princeton, NJ 08544. Aït-Sahalia's research was supported by National Science Foundation (NSF) grant SES-0850533. Fan's research was supported by NSF grants DMS-0714554 and DMS-0704337. The authors are very grateful for the helpful comments of two referees, the associate editor, and the editor.

Journal of the American Statistical Association
December 2010, Vol. 105, No. 492, Theory and Methods
DOI: 10.1198/jasa.2010.tm10163

the RC estimator with the previous tick method due to asynchronocity of the data. Zhang (2009) has put forward a consistent Two Scales Realized Covariance estimator (TSCV) using the previous tick method, which is capable of dealing with asynchronous and noisy data. Recently, Barndorff-Nielsen et al. (2008b) suggested to synchronize the high frequency prices using a Refresh Time scheme, and implemented Multivariate Realized Kernels (MRK) to provide a consistent and semi-definite estimator of the covariance matrix, while Christensen, Kinnebrock, and Podolskij (2010) proposed a multivariate PA estimator. These nonparametric estimators involve the selection of bandwidths or other tuning parameters, and implementation can therefore raise subtle issues in practice.

This article proposes another consistent and rate-efficient estimator based on the QMLE and generalized time synchronization method. Unlike some of the alternatives, it involves no tuning parameters to be set and is very easy to implement. The article is organized as follows. Section 2 summarizes the univariate QMLE of integrated volatility. Section 3 details the proposed covariance estimator and the synchronization method. Section 4 includes Monte Carlo simulations to compare different covariance estimators. Section 5 provides a detailed empirical study of six foreign exchange futures contracts. Section 6 concludes. The Appendix contains the mathematical proofs.

2. QMLE REVIEW

Assume that the logarithm of the transaction price $\tilde{X}$ is observed at $0 = \tau_0 < \tau_1 < \cdots < \tau_n = T$ and is related to the latent true asset log-price X and the microstructure noise U in an additive way, $\tilde{X}_t = X_t + U_t$. This simple model resembles the model of Roll (1984), where the efficient price is a random walk. See also Hasbrouck (2007) for many extensions of this basic reduced-form microstructure model. Unlike these discrete time models, we assume that the latent log-price follows a continuous time Itô process. More precisely,

Assumption 1. The true log-price process satisfies

$$dX_t = \mu_t\,dt + \sigma_t\,dW_t$$

with the volatility process a positive and locally bounded Itô semimartingale, and the drift a locally bounded and progressively measurable process. The noise U_t is iid, and independent of price and volatility processes, with mean 0, variance a^2 and a finite fourth moment. The observations are equally spaced, denoted by Δ.

Unfortunately, the classical maximum likelihood estimation is not feasible in this general setting, since the parameter of interest $\int_0^T \sigma_t^2\,dt$ is a random variable, and the volatility process is left unspecified. Xiu (2010) has suggested to analyze model misspecification, where volatility is stochastic, in the context of the parametric MLE, which assumes that volatility is a fixed parameter. For this purpose, consider the following (potentially misspecified) model:

Assumption 2. The volatility of the observed log return process is constant, that is,

$$dX_t = \sigma\,dW_t.$$

The noise U_t is iid Gaussian, with mean 0 and variance a^2 and independent of the stock price.

The simple Assumption 2 is of course most certainly unrealistic, but its main advantage is that the parameters σ^2 and a^2 can easily be estimated by using maximum-likelihood, as proposed by Aït-Sahalia, Mykland, and Zhang (2005). Specifically, under this model, the observed log-return $Y_i = \tilde{X}_{\tau_i} - \tilde{X}_{\tau_{i-1}}$ follows an MA(1) process, and the log likelihood function for $\mathbf{Y}$'s is

$$l(a^2,\sigma^2) = -\frac{1}{2}\log\det(\mathbf{\Omega}) - \frac{n}{2}\log(2\pi) - \frac{1}{2}\mathbf{Y}'\mathbf{\Omega}^{-1}\mathbf{Y}, \quad (1)$$

where

$$\mathbf{\Omega} = \begin{pmatrix} \sigma^2\Delta + 2a^2 & -a^2 & 0 & \cdots & 0 \\ -a^2 & \sigma^2\Delta + 2a^2 & -a^2 & \ddots & \vdots \\ 0 & -a^2 & \sigma^2\Delta + 2a^2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -a^2 \\ 0 & \cdots & 0 & -a^2 & \sigma^2\Delta + 2a^2 \end{pmatrix}.$$

The optimization algorithm can be efficiently implemented since the matrix inverse of $\mathbf{\Omega}$ is explicit and hence (1) is explicit.

The QMLE is obtained by maximizing the above quasi-likelihood function, even when the data generating process follows Assumption 1 rather than the more restrictive Assumption 2. Aït-Sahalia, Mykland, and Zhang (2005) showed that this estimation approach is robust to departure from the Gaussian noise assumption. But the other part of Assumption 2 that is questionable is the constancy of volatility. In practice, volatility is expected to be stochastic, but Xiu (2010) has shown that even in that situation, the QMLE $(\hat{\sigma}^2, \hat{a}^2)$ remains consistent and optimal in terms of its rate of convergence under the fairly general Assumption 1. This provides a simple and easily implementable method for estimating the integrated volatility, with the following Central Limit Theorem:

Theorem 1. Given Assumption 1, we have

$$\begin{pmatrix} n^{1/4}(\hat{\sigma}^2 - \frac{1}{T}\int_0^T \sigma_t^2\,dt) \\ n^{1/2}(\hat{a}^2 - a_0^2) \end{pmatrix} \xrightarrow{\mathcal{L}_X} \mathrm{MN}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{5a_0\int_0^T \sigma_t^4 dt}{T(\int_0^T \sigma_t^2\,dt)^{1/2}} + \frac{3a_0(\int_0^T \sigma_t^2\,dt)^{3/2}}{T^2} & 0 \\ 0 & 2a_0^4 + \mathrm{Cum}_4[U] \end{pmatrix}\right),$$

where MN is a mixed Normal, $\mathcal{L}_X$ denotes $\sigma(X)$-stable convergence in law, and $\mathrm{Cum}_4[U] = EU_t^4 - 3(EU_t^2)^2$.

Remark 1. If the log-price process is driven by two or more Brownian factors, i.e.,

$$dX_t = \sigma_{1t}\,dW_{1t} + \sigma_{2t}\,dW_{2t},$$

where $E(dW_{1t}\cdot dW_{2t}) = \rho_t\,dt$, then Theorem 1 continues to hold, namely, with $\sigma_t^2 = \sigma_{1t}^2 + \sigma_{2t}^2 + 2\rho_t\sigma_{1t}\sigma_{2t}$, we have

$$n^{1/4}\left(\hat{\sigma}^2 - \frac{1}{T}\int_0^T \sigma_t^2\,dt\right) \xrightarrow{\mathcal{L}_X} \mathrm{MN}\left(0, \frac{5a_0\int_0^T \sigma_t^4\,dt}{T(\int_0^T \sigma_t^2\,dt)^{1/2}} + \frac{3a_0(\int_0^T \sigma_t^2\,dt)^{3/2}}{T^2}\right).$$

The distribution of the time intervals between observations would not affect the consistency, if the time intervals remain iid and independent of the prices. However, the sequence of log-returns does matter since it determines the autocorrelation structure of the observations.

3. COVARIANCE/CORRELATION ESTIMATION VIA THE QMLE

3.1 Model Setup

We now extend the previous results to incorporate a two-dimensional log-price process $\mathbf{X}_t = (X_{1t}, X_{2t})$, discretely observed over the interval $[0, T]$. Suppose that the observations are recorded at times $0 = t_{i,0} \le t_{i,1} \le t_{i,2} \le \cdots \le t_{i,n_i} = T$, respectively, where $i = 1, 2$. As in the univariate case, one can only observe $\tilde{X}_{i,t}$, contaminated by an additive error $U_{i,t}$, associated at each observation point. Further, we make the following assumption:

Assumption 3. The latent log-price process satisfies

$$dX_{it} = \mu_{it}\,dt + \sigma_{it}\,dW_{it}$$

with $E(dW_{1t} \cdot dW_{2t}) = \rho_t\,dt$, and the volatility processes positive and locally bounded Itô semimartingales, and the drifts locally bounded and progressively measurable processes. The noise $\mathbf{U}_t$ is an iid 2-dimensional vector with mean 0, diagonal covariance matrix $\mathbf{\Theta}$ and has a finite fourth moment.

3.2 Covariance/Correlation Estimator for Synchronized Data

The estimator is based on the following identity:

$$\mathrm{Cov}(X_1, X_2) = \tfrac{1}{4}\big(\mathrm{Var}(X_1 + X_2) - \mathrm{Var}(X_1 - X_2)\big).$$

Therefore, we propose

$$\widehat{\mathrm{Cov}}(\tilde{X}_1, \tilde{X}_2) = \tfrac{1}{4}\big(\widehat{\mathrm{Var}}(\tilde{X}_1 + \tilde{X}_2) - \widehat{\mathrm{Var}}(\tilde{X}_1 - \tilde{X}_2)\big), \qquad (2)$$

where $\widehat{\mathrm{Cov}}(\cdot, \cdot)$ is our estimator, $\widehat{\mathrm{Var}}(\cdot, \cdot)$ denotes the QMLE of the quadratic variation, and "$\cdot$" indicates the data we are actually using. However, in order to compute the prices $\tilde{X}_1 + \tilde{X}_2$ and $\tilde{X}_1 - \tilde{X}_2$, we need the two assets be synchronically traded. We will discuss in details the next subsection how to deal with nonsynchronized trading.

Straightforwardly, the correlation estimator is given by:

$$\widehat{\mathrm{Corr}}(\tilde{X}_1, \tilde{X}_2) = \frac{\widehat{\mathrm{Cov}}(\tilde{X}_1, \tilde{X}_2)}{\sqrt{\widehat{\mathrm{Var}}(\tilde{X}_1)}\sqrt{\widehat{\mathrm{Var}}(\tilde{X}_2)}}. \qquad (3)$$

Remark 2. It may be more efficient to consider linear combinations

$$\widehat{\mathrm{Cov}}(\tilde{X}_1, \tilde{X}_2) = \frac{1}{4\gamma(1-\gamma)}\big(\widehat{\mathrm{Var}}(\gamma\tilde{X}_1 + (1-\gamma)\tilde{X}_2) - \widehat{\mathrm{Var}}(\gamma\tilde{X}_1 - (1-\gamma)\tilde{X}_2)\big), \qquad (4)$$

where γ can be selected to minimize the asymptotic variance. If $\gamma = 1/2$, (4) recurs to (2). One alternative γ may be $\widehat{\mathrm{Var}}(\tilde{X}_2)/(\widehat{\mathrm{Var}}(\tilde{X}_1) + \widehat{\mathrm{Var}}(\tilde{X}_2))$.

The constructed covariance matrix estimates may not be positive semi-definite. This property is nevertheless essential to many applications in practice. To enforce it, one possibility consists in projecting the resulting symmetric matrix onto the space of positive semi-definite matrices: see for example an application to portfolio allocation in Fan, Li, and Yu (2010).

In Section 2, we have shown consistency and a central limit theorem for the QMLE of the variance. Based on Theorem 1 and Remark 1, we can show that this estimator is $n^{1/4}$-consistent for the covariance and correlation. This rate is the optimal one. The following theorem provides the central limit results for the covariance estimator under an idealized data observation scheme where the two assets are observed at synchronized times:

Theorem 2. Given Assumption 3, and that the data are synchronized and equally spaced, that is, $n := n_1 = n_2$ and $\tau_j := t_{1,j} = t_{2,j}$, and that $\Delta = \tau_j - \tau_{j-1}$, for $1 \le j \le n$, then the following Central Limit Theorem holds:

$$n^{1/4}\left(\widehat{\mathrm{Cov}}(\tilde{X}_1, \tilde{X}_2) - \frac{1}{T}\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt\right) \xrightarrow{\mathcal{L}_X} \mathrm{MN}(0, V), \qquad (5)$$

$$n^{1/4}\left(\widehat{\mathrm{Corr}}(\tilde{X}_1, \tilde{X}_2) - \frac{\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt}{\sqrt{\int_0^T \sigma_{1t}^2\,dt}\sqrt{\int_0^T \sigma_{2t}^2\,dt}}\right) \xrightarrow{\mathcal{L}_X} \mathrm{MN}(0, \tilde{V}), \qquad (6)$$

where V and $\tilde{V}$ are given in the Appendix.

3.3 Data Synchronization

We would stop here if the data were synchronized, meaning that the prices of the two assets were observed at the same times. However, this is not the case in practice, at least for high frequency financial data. In most cases, high frequency transactions for two assets occur at times that are not synchrone. This practical issue may induce a large bias for the estimation, and may be (at least partly) responsible for the Epps effect. We note further that both terms on the right side of (2) need synchronized data. The remaining question is what kind of data synchronization procedure one should use. Clearly, if we apply the QMLE to estimate the diagonal elements in the covariance matrix, it would be better to cross out a small number of data points rather than adding more through an interpolation method, because the former strategy may suffer from efficiency loss, while the latter one may result in inconsistency due to the change in the autocorrelation structure. We define the following concept, which we then use to propose a general synchronization scheme.

Definition 1. A sequence of time points $\{\tau_0, \tau_1, \tau_2, \ldots, \tau_n\}$ is said to be the Generalized Sampling Time for a collection of M assets, if

1. $0 = \tau_0 < \tau_1 < \cdots < \tau_{n-1} < \tau_n = T$.
2. There exists at least one observation for each asset between consecutive τ_i's.
3. The time intervals, $\{\Delta_j = \tau_j - \tau_{j-1}, 1 \le j \le n\}$, satisfies $\sup_i \Delta_i \xrightarrow{\mathrm{P}} 0$.

The Generalized Synchronization method is built upon the Generalized Sampling Time by selecting an arbitrary observation $\tilde{X}_{i,\check{t}_j}$ for the ith asset between the time interval $(\tau_{j-1}, \tau_j]$. The synchronized data sets are, therefore, $\{\tilde{X}^{\tau}_{i,\tau_j}, 1 \le i \le M, 1 \le j \le n\}$ such that $\tilde{X}^{\tau}_{i,\tau_j} = \tilde{X}_{i,\check{t}_j}$.

The concept of Generalized Synchronization method is more general than that of the Previous Tick approach discussed in Zhang (2009), and the Refresh Time scheme proposed by Barndorff-Nielsen et al. (2008b), namely, the Replace All scheme in deB. Harris et al. (1995).

More precisely, if we require $\{\tau_j\}$ to be equally spaced between $[0, T]$, the previous tick for each asset before τ_j to be selected, we are back to the Previous Tick approach. Or, if we choose τ_j recursively as

$$\tau_{j+1} = \max_{1\le i\le M}\{t_{i,N_i(\tau_j)+1}\},$$

where $\tau_1 = \max\{t_{1,1}, t_{2,1}, \ldots, t_{M,1}\}$ and $N_i(t)$ measures the number of observations for asset i before time t, and if we select those ticks that occur right before or at τ_j's, we return to the Refresh Time scheme. In both cases, the previous ticks of the assets, if needed, are regarded as if they were observed at the sampling time τ_j's. By contrast, we advocate choosing an arbitrary tick for each asset within each interval. In practice, it may happen that the order of consecutive ticks is not recorded correctly. Because our synchronization method has no requirement on tick selection, the estimator is robust to data misplacement error, as long as these misplaced data points are within the same sampling intervals.

It is apparent that the Refresh Time scheme is highly dependent on the relatively illiquid asset. On the one hand, the number of the synchronized pairs are smaller than the number of the observations of this asset, inducing an inevitable loss of data for the other asset. More importantly, it is very likely that the Refresh Time points are determined by the occurrence of the relatively more illiquid asset, rendering the selected observations of the other asset always ahead of the corresponding illiquid asset. This hidden effect may induce some additional bias in the estimation.

Alternatively, we can design the synchronization scheme requiring each asset to lead in turn. Take two assets, for example. If we require the first asset to lead, we choose $\tau_1 = t_{2,N_2(t_{1,1})+1}$. Recursively,

$$\tau_i = t_{2,N_2(t_{1,N_1(\tau_{i-1})+1})+1}.$$

Literally, it means that right after τ_{i-1}, we find the first observation of $\tilde{X}_{1t}$, which should happen at $t_{1,N_1(\tau_{i-1})+1}$, and then the next Generalized Sampling Time is defined to be the point when the first $\tilde{X}_{2t}$ is observed right after $t_{1,N_1(\tau_{i-1})+1}$. In this case, at all sampling time points, the second asset would always have records. The previous tick of the first asset, if needed, is regarded as if it were observed a bit later at the sampling time. Hence, in the synchronized pairs, the first asset always leads the second one.

Remark 3. Refresh Time includes the largest amount of data among all Generalized Sampling Time.

Combining the synchronization scheme with our estimator, we make the following assumption:

Assumption 4. The Generalized Sampling Time $\{\tau_j\}$ is independent of the price process, the volatility process and the noise. The time intervals, $\{\Delta_j = \tau_j - \tau_{j-1}, 1 \le j \le n\}$, are iid with mean $\bar{\Delta}$. The number of observations n is therefore random, of order $O_P(1/\bar{\Delta})$.

Replacing the idealized data with the products of the Generalized Sampling Time, we obtain:

Theorem 3. Given Assumptions 3 and 4, the QMLE of the quadratic variation for $\tilde{X}^{\tau}_1 + \tilde{X}^{\tau}_2$ is consistent, that is,

$$\widehat{\mathrm{Var}}(\tilde{X}^{\tau}_1 + \tilde{X}^{\tau}_2) - \frac{1}{T}\int_0^T \sigma^2_{1t} + \sigma^2_{2t} + 2\rho_t\sigma_{1t}\sigma_{2t}\,dt = O_P\big(\bar{\Delta}^{1/4}\big).$$

Therefore,

$$\widehat{\mathrm{Cov}}(\tilde{X}^{\tau}_1, \tilde{X}^{\tau}_2) - \frac{1}{T}\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt = O_P\big(\bar{\Delta}^{1/4}\big),$$

$$\widehat{\mathrm{Corr}}(\tilde{X}^{\tau}_1, \tilde{X}^{\tau}_2) - \frac{\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt}{\sqrt{\int_0^T \sigma^2_{1t}\,dt}\sqrt{\int_0^T \sigma^2_{2t}\,dt}} = O_P\big(\bar{\Delta}^{1/4}\big).$$

In other words, the rate of convergence of the estimators are the same as those given in Theorem 2.

3.4 Synchronization Comparison With the HY Estimator

It is interesting to compare the synchronization method embedded in the HY estimator and the synchronization scheme proposed here. Recall that Hayashi and Yoshida (2005) proposed

$$\langle X_1, X_2\rangle_{HY} = \sum_{i,j}\big(X_{1,t_{1,j}} - X_{1,t_{1,j-1}}\big)\big(Y_{1,t_{2,i}} - Y_{1,t_{2,i-1}}\big) \times 1_{\{(t_{1,j-1},t_{1,j}]\cap(t_{2,i-1},t_{2,i}]\neq\emptyset\}}, \quad (7)$$

where X_1 and X_2 are the observations, in an estimator that assumes no noise and is therefore infeasible in our setting. Nevertheless, the synchronization method proposed therein may be applicable. A priori, the HY method has advantages over the Refresh Time scheme in that the former utilizes all possible data. However, this ignores the fact that the HY method effectively deletes data through some cancelation in the calculation as well. If we fix index j and sum over i at first, we can rewrite the formula (7) in the following way:

$$\langle X_1, X_2\rangle_{HY} = \sum_j\big(X_{1,t_{1,j}} - X_{1,t_{1,j-1}}\big)\big(Y_{1,t_{2,j+}} - Y_{1,t_{2,(j-1)-}}\big), \quad (8)$$

where $t_{2,j+} = \inf\{t_{2,k} : t_{2,k} \ge t_{1,j}\}$ and $t_{2,j-} = \sup\{t_{2,k} : t_{2,k} \le t_{1,j}\}$.

It follows from (8) and Figure 1 that at least any records for the second asset that occur in $(t_{2,(j-1)+}, t_{2,j-})$ will not play a role in the calculation due to cancelation. Similarly, if three consecutive observations of the first asset form two intervals which share the same corresponding interval of the second asset, then the middle observation of the first asset will not be used either. In the simulation and empirical studies, we will compare the effective sample size of the HY method and the Refresh Time scheme.

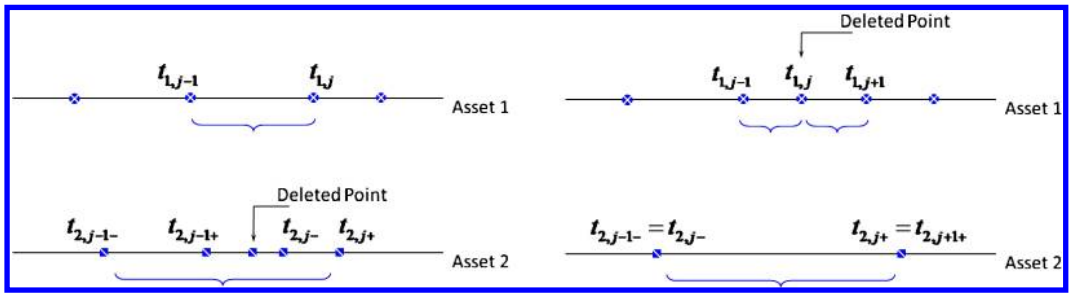

Figure 1. These two graphs illustrate how the HY synchronization method delete data points. The online version of this figure is in color.

4. SIMULATION STUDY

We now conduct Monte Carlo simulations to verify the applicability of the above theoretical results in realistic settings. The data is generated using stochastic volatility models, and for this purpose we consider mainly the Heston Model. For $i = 1, 2$, the data generating process is

$$dX_{it} = \sigma_{it}\, dW_{it},$$
$$d\sigma_{it}^2 = \kappa_i(\bar{\sigma_i}^2 - \sigma_{it}^2) + s_i\sigma_{it}\, dB_{it} + \sigma_{it-}J_{it}^V\, dN_{it},$$

where $E(dW_{it} \cdot dB_{jt}) = \delta_{ij}\rho_i\, dt$ and $E(dW_{1t} \cdot dW_{2t}) = \rho\, dt$.

We generate sample paths using the Euler method, where the first observation for volatility process σ_{i0}^2 is sampled from a Gamma distribution $\Gamma(2\kappa_i\bar{\sigma_i}^2/s_i^2,\ s_i^2/2\kappa_i)$. The jump size in volatility is $J_t^V = \exp(z)$, where $z \sim N(\theta_i, \mu_i)$, and N_{it} is a Poisson Processes independent of Brownian motions with intensity λ_i. The noise of each asset is iid $N(0, a_i^2)$. The parameter values are reported in Table 1.

First, we generate N random intervals from the exponential distribution for both processes. The data are equally spaced and therefore, there are no issues with synchronization and we can easily estimate the true covariance using the RC method. Then we randomly select about $N/2$ observations for X_1 and $2N/3$ for X_2 from the whole sample using Bernoulli trials, and contaminate the data with noise. Next we synchronize the data using the Refresh Time procedure and apply the QMLE estimator to $\tilde{X}_1^\tau \pm \tilde{X}_2^\tau$, $\tilde{X}_1^\tau$, and $\tilde{X}_2^\tau$, respectively, and construct the estimator for the covariance. We also include a modified version QMLE* as defined in (4).

Alternatively, we also implement the TSCV estimator by Zhang (2009) for comparison. Suppose the number of subsamples are K and J, respectively. Define

$$[\tilde{X}_1^\tau, \tilde{X}_2^\tau]_T^S = \frac{1}{S}\sum_{i=S}^{n}\big(\tilde{X}_{1,\tau_i}^\tau - \tilde{X}_{1,\tau_{i-S}}^\tau\big)\big(\tilde{X}_{2,\tau_i}^\tau - \tilde{X}_{2,\tau_{i-S}}^\tau\big)$$

and then the TSCV estimator is given by

$$\langle\tilde{X}_1^\tau, \tilde{X}_2^\tau\rangle_{TSCV} = \left([\tilde{X}_1^\tau, \tilde{X}_2^\tau]_T^K - \frac{\bar{n}_K}{\bar{n}_J}[\tilde{X}_1^\tau, \tilde{X}_2^\tau]_T^J\right),$$

where $1 \le J \ll K = O(n^{2/3})$ and $\bar{n}_S = (n - S + 1)/S$, $S = K, J$.

We also implement the MRK proposed in Barndorff-Nielsen et al. (2008b) for comparison. Suppose we have n number of observations for each asset after synchronization. First we redefine the initial and final time points to control the edge effect associated with realized kernels. Fix $m \in \mathbb{N}$, with $N = n + 1 - 2m$. Define $\tilde{\mathbf{X}}_{\tau_j} = (\tilde{X}_{1,\tau_j}^\tau, \tilde{X}_{2,\tau_j}^\tau)'$,

$$\check{\mathbf{X}}_0 = \frac{1}{m}\sum_{j=1}^{m}\tilde{\mathbf{X}}_{\tau_j},$$

$$\check{\mathbf{X}}_N = \frac{1}{m}\sum_{j=1}^{m}\tilde{\mathbf{X}}_{\tau_{n-m+j}},$$

and $\check{\mathbf{X}}_j = \tilde{\mathbf{X}}_{\tau_{j+m}}$, where $j = 1, 2, \ldots, N-1$. Let $\mathbf{x}_j = \check{\mathbf{X}}_j - \check{\mathbf{X}}_{j-1}$, $j = 1, 2, \ldots, N$.

The multivariate realized kernel estimator is

$$\mathbf{K}(X) = \sum_{h=-H}^{H} k\left(\frac{|h|}{H+1}\right)\boldsymbol{\Gamma}_h,$$

where

$$\boldsymbol{\Gamma}_h = \begin{cases} \displaystyle\sum_{j=|h|+1}^{N}\mathbf{x}_j\mathbf{x}_{j-h}', & h \ge 0 \\ \displaystyle\sum_{j=|h|+1}^{N}\mathbf{x}_{j+h}\mathbf{x}_j', & h < 0, \end{cases}$$

and the kernel function is chosen to be the Parzen kernel:

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3, & 0 \le x \le 1/2 \\ 2(1-x)^3, & 1/2 \le x \le 1 \\ 0, & x > 1 \text{ or } x < 0. \end{cases}$$

Also, we make the choice of bandwidth $H = \max(H_1, H_2)$, with $H_i = c^*\xi_i^{4/5}n^{3/5}$, where $\xi_i^2 = a_i^2/\sqrt{T\int_0^T\sigma_{it}^4\, dt}$ and $c^* = 3.5134$. Here, in order to optimize the behavior of the MRK, we try the infeasible case, i.e., H_i is calculated using the true values.

The benchmark is the RC estimator using the original complete synchronized data set without being contaminated by the

Table 1. Summary of the parameter values in the Monte Carlo simulation

Asset	X_{i0}	κ_i	s_i	$\bar{\sigma}_i^2$	ρ_i	λ_i	θ_i	μ_i	a_i	ρ
$i = 1$	log(100)	6	0.5	0.16	−0.6	12	−5	0.8	0.005	0.5
$i = 2$	log(40)	4	0.3	0.09	−0.75	36	−6	1.2	0.001	—

Table 2. Simulation comparisons of the estimates

		$\hat{\Sigma}_{11}$	$\hat{\Sigma}_{12}$	$\hat{\Sigma}_{22}$
$\bar{\Delta} = 2$ s	RC	0.1603	0.0583	0.0904
$n = 11{,}700$		(0.0032)	(0.0018)	(0.0018)
	QMLE	0.1603	0.0584	0.0906
$\bar{n}_1 = 5853$		(0.0242)	(0.0112)	(0.0074)
$\bar{n}_2 = 7801$	MRK	0.1967	0.0584	0.0920
$\bar{n} = 5573$		(0.0466)	(0.0147)	(0.0147)
	TSCV	0.1565	0.0569	0.0883
		(0.0306)	(0.0148)	(0.0146)
$m_1 = 5151$	HY		0.0584	
$m_2 = 5307$			(0.0028)	
Span $= 1.92$ s	QMLE*		0.0582	
			(0.0108)	
$\bar{\Delta} = 30$ s	RC	0.1600	0.0581	0.0901
$n = 781$		(0.0121)	(0.0068)	(0.0068)
	QMLE	0.1606	0.0582	0.0904
$\bar{n}_1 = 392$		(0.0509)	(0.0241)	(0.0178)
$\bar{n}_2 = 522$	MRK	0.2116	0.0575	0.0923
$\bar{n} = 372$		(0.0732)	(0.0270)	(0.0264)
	TSCV	0.1504	0.0543	0.0847
		(0.0536)	(0.0256)	(0.0247)
$m_1 = 345$	HY		0.0581	
$m_2 = 355$			(0.0106)	
Span$= 28.7$ s	QMLE*		0.0575	
			(0.0232)	

NOTE: This table reports the summary statistics for the simulation results. $\bar{\Delta}$ is the sampling frequency for the original data. n is the size of the whole sample. $\bar{n}_i$ is the average number of observations for the ith asset. $\bar{n}$ is the average number of synchronized observations using the Refresh Time scheme. m_1 and m_2 are the effective sample sizes for two series, respectively, using the HY method. Span measures the average difference between the refresh time and the real observation time. $\hat{\Sigma}_{ij}$ is the component of the estimated covariance matrix. The data in parenthesis are the RMSE. RC is the benchmark method applied to the no noise and the whole data (Synchronized). The Monte Carlo sample size is 10,000.

market microstructure noise. This ideal estimator is only available in simulations, and is unfeasible in practice. This benchmark estimator is only computed here so that other estimators can be compared to it. We also implement the HY estimator with asynchronous data but no noise. The comparison results with various sampling frequencies are shown in Table 2.

It is clear from the RMSE reported here that the QMLE, the TSCV and the MRK behave quite well from the simulation studies and that neither of them suffers from a synchronization problem. But in some case, the bias cannot be ignored. One common drawback for the two nonparametric estimators is the choice of bandwidth, in that the selected bandwidth is suboptimal when trying to estimate the diagonal elements of the covariance matrix.

By contrast, the QMLE has no tuning parameters, and is asymptotically unbiased at least for synchronous data, and is designed to have a smaller asymptotic variance since the rate is higher than that of the $n^{1/5}$-consistent MRK and the $n^{1/6}$-consistent TSCV. The comparison of the TSCV and the MRK indicates that the asymptotic bias associated with the MRK estimator overcomes its edge from its higher convergence rate in finite sample.

As to the synchronization comparison, Table 2 provides an example where the Refresh Time scheme may actually outperform the HY method in terms of data inclusion. Empirically, there is not a wide gap between the amount of data retained by each method, as shown in Table 3. The HY method does not account for the noise and consequently may interfere with the autocorrelation structure introduced by the microstructure noise; hence it is not surprising that is may be dominated in our setting where the data are noisy.

Table 3. Comparison between Refresh Time and HY synchronization scheme

	AUD	CAD	EUR	GBP	JPY	CHF
Refresh Time						
% of AUD		56.9%	87.2%	73.7%	70.0%	58.6%
% of CAD	59.2%		86.6%	74.3%	74.9%	58.8%
% of EUR	31.3%	21.7%		52.2%	53.7%	34.4%
% of GBP	43.1%	41.7%	84.9%		64.3%	47.3%
% of JPY	40.5%	37.8%	78.6%	57.9%		42.7%
% of CHF	57.3%	55.1%	93.6%	79.1%	79.4%	
HY						
% of AUD		59.1%	85.5%	73.5%	76.5%	60.8%
% of CAD	61.3%		84.7%	74.0%	74.5%	60.9%
% of EUR	35.0%	33.2%		53.4%	55.1%	37.6%
% of GBP	46.4%	44.9%	81.6%		64.5%	50.0%
% of JPY	44.3%	41.3%	76.8%	58.9%		46.0%
% of CHF	59.7%	57.7%	90.7%	77.7%	78.1%	

NOTE: In this table, we compare the Refresh Time and HY synchronization scheme by calculating the percentage of data reserved.

Figure 2 verifies in simulations the central limit distribution given in Theorem 2. We estimate the covariance and correlation using synchronized assets and plot the histogram of the standardized estimates. The histogram matches well with the standard Gaussian distribution predicted by the theorem.

5. EMPIRICAL ESTIMATION WITH FOREIGN EXCHANGE FUTURES PRICES

5.1 Data Description

Foreign exchange future contracts are traded on the Chicago Mercantile Exchange (CME) on a 24-hour clock. These

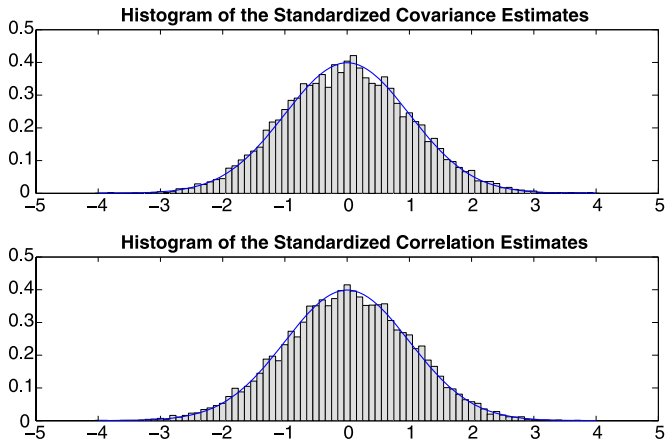

Figure 2. In this figure, we plot the histograms of the covariance and correlation estimator. Here, the data are synchronized, and the average length of the sampling intervals is $\bar{\Delta} = 1$ s. The online version of this figure is in color.

Table 4. Summary statistics of the log returns of the foreign exchange futures

FX	Avg. no.	Avg. freq.	Mean	Std. err.	1st lag	2nd lag	3rd lag
AUD	6142	14.07 s	9.97e−08	1.53e−04	−0.0769	0.0082	0.0031
CAD	5898	14.65 s	2.25e−08	1.24e−04	−0.1186	0.0033	0.0007
EUR	17,097	5.05 s	2.16e−08	5.69e−05	−0.1141	0.0099	0.0014
GBP	10,508	8.22 s	1.35e−08	7.62e−05	−0.0651	0.0123	0.0037
JPY	11,685	7.39 s	−2.18e−08	8.64e−05	−0.1066	0.0089	0.0005
CHF	6284	13.75 s	2.35e−08	1.07e−04	−0.0846	0.0067	−0.0025

NOTE: The second and third columns are, respectively, the average numbers of trades and average trading frequencies per day. The fourth and fifth column are the average log-returns and their association standards. The auto-correlation coefficients of lags 1, 2, and 3 are depicted at the sixth, seventh, and eighth columns.

marked-to-market futures contracts are very liquid. We are interested in estimating the covariations among pairs of currencies; we focus on the Australian dollar, Canadian dollar, Euro, British pound, Japanese yen, and Swiss franc futures contracts. The contracts are quoted in terms of the unit value of the foreign currency measured in US dollars. We use currency futures instead of currency spot prices because the former contracts are traded in an exchange setting. By contrast, the currency spot markets, while extremely liquid, are over-the-counter, and may therefore suffer from additional microstructure issues, e.g., the financial stability of the counterparties and the quality of the execution they provide. Another advantage with futures data is that the interest rate differentials are embedded in them, which facilitates the calculation of the carry trade returns, so that we can avoid calculating them in the same way as in Brunnermeier, Nagel, and Pedersen (2009). The cleaned data are available from Tick Data, Inc., among other data vendors.

Timing is a critical issue in the FX future market. Our sample period runs from January 1, 2007 to June 30, 2009, covering the most critical stage of the recent financial crisis. The future contracts are automatically rolled over to provide continuous price records. The tick-by-tick transaction prices are recorded in exchange time (Chicago Time). Therefore, in a given day, trading activity starts from Asia and Europe, followed by North America and then Australia. There is no overlap between the opening hours of the Asian market and the US market. The advantage of using the Exchange Time is that we do not need to deal with different daylight saving rules for different continents in that the traders on the CME are likely to adjust their trading behavior according to the US time (where most of the activity takes place) rather than their local time. The electronic trading starts at 5 p.m. on the previous day and ends at 4 p.m. on current day. In between, there is a one-hour gap. Therefore, we redefine the day in accordance with the electronic trading system, so that we can avoid including potential jumps from the market close price to its open price.

Our data preprocessing eliminates the transactions on Saturdays and Sundays, US federal holidays, the day after Thanksgiving, December 24 to 26, and December 31 to January 2, because of the relatively smaller volume of activity. We are left with 618 days in the sample. For each particular FX contract, we take the average price for any multiple transactions that happen exactly at the same time stamp. The summary statistics for each FX futures are listed in Table 4. It appears that there exists an MA(1) structure for each futures contract, indicating that

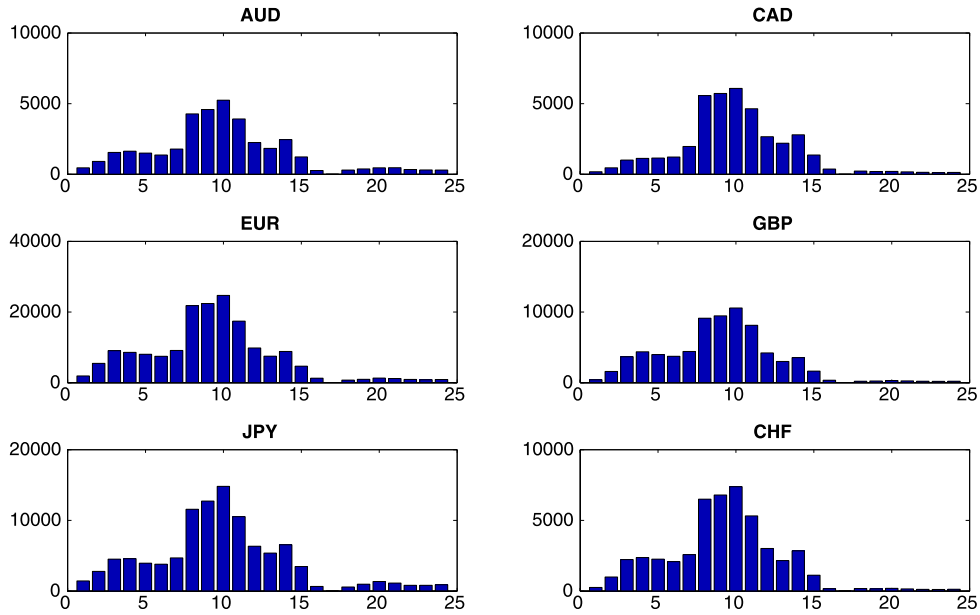

Figure 3. In this figure, we plot the average hourly trading volume of the foreign exchange future contracts from Jan 2007 to Jun 2009. The x-axis is measured in Chicago time. The online version of this figure is in color.

Table 5. Summary of the correlation estimates

	AUD	CAD	EUR	GBP	JPY	CHF
AUD		0.2592	0.3783	0.3044	0.0061	0.2994
CAD	0.2584		0.2708	0.2169	0.0203	0.2282
EUR	0.3774	0.2715		0.4703	0.2138	0.6409
GBP	0.3011	0.2173	0.4724		0.1079	0.4051
JPY	0.0064	0.0201	0.2131	0.1083		0.2656
CHF	0.2955	0.2272	0.6475	0.4032	0.2635	

NOTE: This table reports the average correlations among the six currencies. The numbers on the upper diagonal are based on Refresh Time scheme, while the numbers on the lower diagonal are the average of two synchronization schemes, which in addition, require one asset to lead the other one.

our model assumptions based on market microstructure noise are plausible.

Figure 3 plots the average hourly trading volume for each contract over the whole sample period, indicating that traders on CME tend to trade these currency futures simultaneously, and that there is no clear evidence of home bias in our sample, that is, e.g., traders tend to trade JPY instead of EUR, during their active trading hours. Therefore, it would be reasonable to apply the Refresh Time scheme to synchronize the data sets.

5.2 Empirical Findings

The estimation results are reported in Table 5. We find that the correlations are all positive, which is consistent with the findings in the literature based on lower frequency data, see, e.g., Campbell, de Medeiros, and Viceira (2009). As shown in Figure 4, the relationship between the Swiss franc and the Euro exhibits a large correlation, followed by the British Pound, consistently with the high degree of integration among the European economies as well as the exchange rate policies followed by the Swiss Central Bank. The correlation between the Swiss franc and the Euro slightly decreased in the middle of the crisis, but remained at a high level. As one of the world's major commodity currency, the Australian dollar has lower correlations with the other currencies, while the Canadian dollar is even further detached from them, owing perhaps to its close dependence on the American economy rather than the European or Asian economies. The large impact of the carry trade demise at various points during the financial crisis on the Australian dollar (expecially in relation to the Japanese yen) is partly responsible for large swings in that exchange rate: high interest rates in Australia and low interest rates in Japan combine to make this pair of currencies an attractive target for carry traders.

Figure 4 also describes very similar patterns for those pairs that include the Japanese yen: correlations decreased from their respective normal levels during the period of highest uncertainty. This is probably due to the characteristics of the Japanese economy and its role in the financial crisis. The Japanese economy is strongly dependent on exports, and the yen played the role of a reserve currency during the crisis since the Japanese banking system emerged relatively unscathed (as did the Cana-

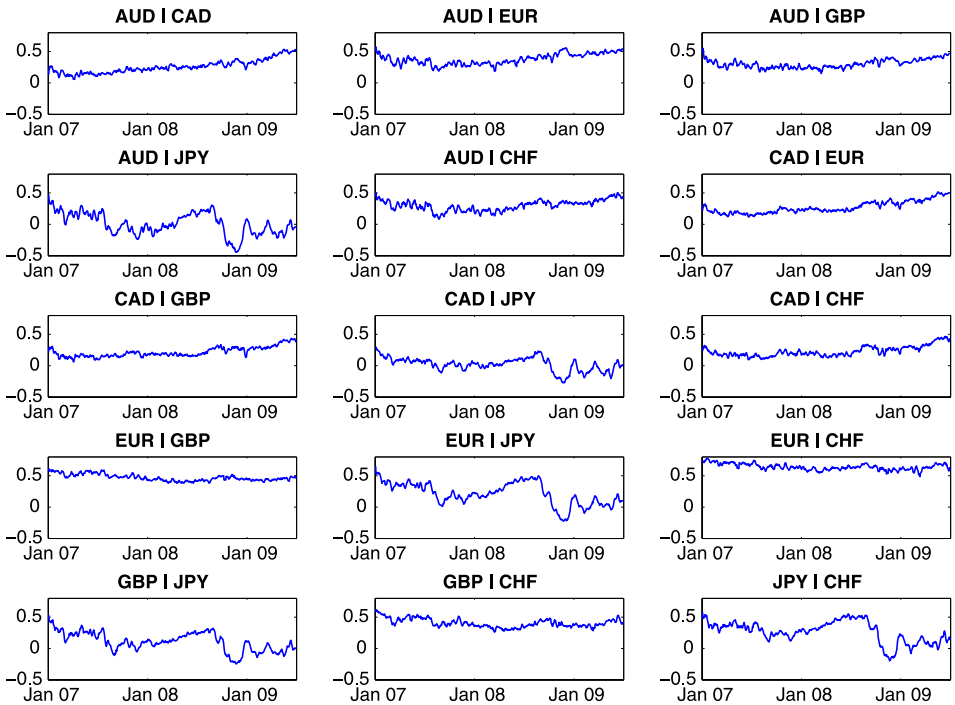

Figure 4. In this figure, we report the time series of correlation estimates of all six FX futures. Each curve is a 5-day moving average of the daily correlations. The online version of this figure is in color.

dian one). An additional factor was likely the return of yen deposits previously borrowed by hedge funds and similar actors in the pursuit of carry trades that were the victims of heavy deleveraging at the height of the crisis.

5.3 Robustness Checks

Unlike stock prices that mainly respond to individual company events, the FX markets is driven primarily by macroeconomic events such as interest rate policy changes, for instance. The correlation estimates may therefore be sensitive to specific news announcement. The most active trading hours, as can be seen from Figure 3, are from 7 a.m. to 11 a.m. Chicago Time when both the Europe and US trading desks are open. In addition, most of the US macro news announcements are released during this period, see, e.g., Andersen et al. (2003). To check for the effect of these announcements, we divide the intraday observations into three trading sessions: 5 p.m.–7 a.m., 7 a.m.–11 p.m., and 11 p.m.–4 p.m. In the first session, the Asian and Australian markets are open, with European markets joining later. The second session is the most active period within the day. Only the North American markets are open in the last session. The estimates are reported in Table 6. The correlation estimates in the second session are mostly higher than the other two, but the differences between the first session and last session are not significant.

Refresh Time tends to be selected when less liquid assets trade, hence the price of the liquid asset in the synchronized pairs are more likely to be stale. The average time difference between the refresh time and the real time is a measure of staleness. To check the robustness of the synchronization procedure, we require in addition that one asset in each pair has to lead the other one, and then taking the average estimates given by the two synchronization scheme. The comparison results are reported in Table 5 (lower diagonal). The estimates do not differ significantly from the estimates given by the Refresh Time scheme.

6. CONCLUSIONS

This article proposes a quasi-likelihood based estimator of high frequency covariance/correlation with noisy and asynchronous tick-by-tick data as well as a generalized synchronization scheme. This is so far the only estimator that can deal with both noise and asynchronism while maintaining optimal convergence rates. Moreover, the feature of being bandwidth-free makes this estimator advantageous in practice over other nonparametric estimators, by removing an essentially arbitrary degree of freedom. The proposed synchronization scheme is conservative in deleting data. Empirically, the article documents the time-varying correlations between six major currencies during the recent financial crisis.

APPENDIX

A.1 Proof of Theorem 2

Without loss of generality, we assume that the drift terms are zero. Consider two new processes $\tilde{X}_1 \pm \tilde{X}_2$. Apparently, each of them is a one dimensional Itô process with additive iid noise. According to Theorem 1, $\widehat{\mathrm{Var}}(\tilde{X}_1 + \tilde{X}_2)$, for example, is a consistent and efficient estimator of $\frac{1}{T}\int_0^T \sigma_{1t}^2 + \sigma_{2t}^2 + 2\rho_t\sigma_{1t}\sigma_{2t}\,dt$ with convergence rate $n^{1/4}$. Therefore, the covariance and correlation estimators are $n^{1/4}$-consistent, that is,

$$\widehat{\mathrm{Cov}}(\tilde{X}_1, \tilde{X}_2) - \frac{1}{T}\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt = O_P\big(n^{-1/4}\big),$$

$$\widehat{\mathrm{Corr}}(\tilde{X}_1, \tilde{X}_2) - \frac{\int_0^T \rho_t\sigma_{1t}\sigma_{2t}\,dt}{\sqrt{\int_0^T \sigma_{1t}^2\,dt}\sqrt{\int_0^T \sigma_{2t}^2\,dt}} = O_P\big(n^{-1/4}\big).$$

Also, the central limit theorem for $\widehat{\mathrm{Var}}(\tilde{X}_1 + \tilde{X}_2)$ clearly holds and the asymptotic variance is given by Remark 1 with a_0 replaced by $(a_{1,0}^2 + 2\theta a_{1,0}a_{2,0} + a_{2,0}^2)^{1/2}$, where θ is the correlation between U_{1t} and U_{2t}. In the synchronous case, we can identify the correlation between the noises. Hence, we can work with a more general case here.

Therefore, in order to obtain the central limit theorem for the covariance and correlation estimators, it is sufficient to derive the asymptotic covariances of the four QMLE estimators:

$$V^{\alpha,\beta} = \mathrm{Acov}\big(\widehat{\mathrm{Var}}(\alpha\tilde{X}_1 + \tilde{X}_2), \widehat{\mathrm{Var}}(\tilde{X}_1 - \beta\tilde{X}_2)\big),$$

Table 6. Comparisons of the correlation estimates

Active markets	Asia, Australia, Europe	Europe, US, Canada	US, Canada
Chicago Time	5 p.m.–7 a.m.	7 a.m.–11 a.m.	11 a.m.–4 p.m.
AUD–CAD	0.2476	0.2766	0.2342
AUD–EUR	0.3735	0.4042	0.3153
AUD–GBP	0.3001	0.3233	0.2534
AUD–JPY	−0.0041	0.0486	−0.0321
AUD–CHF	0.3146	0.3114	0.2042
CAD–EUR	0.2618	0.2885	0.2314
CAD–GBP	0.2125	0.2294	0.1858
CAD–JPY	0.0122	0.0469	−0.0156
CAD–CHF	0.2298	0.2391	0.1747
EUR–GBP	0.4555	0.4948	0.4648
EUR–JPY	0.1801	0.2663	0.1939
EUR–CHF	0.6685	0.6341	0.5754
GBP–JPY	0.0692	0.1746	0.0929
GBP–CHF	0.3996	0.4306	0.3582
JPY–CHF	0.2198	0.3213	0.2641

NOTE: This table compares the correlations of each FX pairs in three sessions within a day.

where α and β can be 0, 1, and -1.

We start with $\alpha = \beta = 1$. Suppose (σ_1^2, a_1^2) and (σ_2^2, a_2^2) are the parameters in two respective quasi-likelihood functions. It can be shown that the QMLE for the volatility and the noise variance are asymptotically independent, so we only need to consider the score functions with respect to σ_k^2, denoted as $\Psi_{k,n}$, where $k = 1, 2$. Let the inverse of the covariance matrix in each likelihood (1) be $\mathbf{\Omega}_k^{-1} = (\omega_k^{i,j})$. For any process A, denote $\Delta_i^n A = A_{\tau_i} - A_{\tau_{i-1}}$. Define

$$\Psi_{k,n}(\omega) = \frac{1}{2\sqrt{n}}\left\{\frac{\partial \log(\det \mathbf{\Omega}_k)}{\partial \sigma_k^2} + \mathbf{Y}_k' \frac{\partial \mathbf{\Omega}_k^{-1}}{\partial \sigma_k^2}\mathbf{Y}_k\right\},$$

$$\bar{\Psi}_{k,n}(\omega) = \frac{1}{2\sqrt{n}}\left\{\frac{\partial \log(\det \mathbf{\Omega}_k)}{\partial \sigma_k^2} + \operatorname{tr}\left(\frac{\partial \mathbf{\Omega}_k^{-1}}{\partial \sigma_k^2}\mathbf{\Sigma}_k\right)\right\},$$

where the ith components of $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are

$$Y_{1,i} = \Delta_i^n \tilde{X}_1 + \Delta_i^n \tilde{X}_2,$$

$$Y_{2,i} = \Delta_i^n \tilde{X}_1 - \Delta_i^n \tilde{X}_2.$$

$\mathbf{\Sigma}_1 = (\Sigma_{1,i,j})$ and $\mathbf{\Sigma}_2 = (\Sigma_{2,i,j})$ are specified in the following way:

$$\Sigma_{1,i,i} = \int_{\tau_{i-1}}^{\tau_i} \sigma_{1,t}^2 + \sigma_{2,t}^2 + 2\rho_t \sigma_{1t}\sigma_{2t}\,dt + 2a_1^{*2},$$

$$\Sigma_{2,i,i} = \int_{\tau_{i-1}}^{\tau_i} \sigma_{1,t}^2 + \sigma_{2,t}^2 - 2\rho_t \sigma_{1t}\sigma_{2t}\,dt + 2a_2^{*2},$$

$$\Sigma_{1,i,i-1} = \Sigma_{1,i,i+1} = -a_1^{*2},$$

$$\Sigma_{2,i,i-1} = \Sigma_{2,i,i+1} = -a_2^{*2},$$

where $a_1^{*2} = a_{1,0}^2 + a_{2,0}^2 + 2\theta a_{1,0}a_{2,0}$ and $a_2^{*2} = a_{1,0}^2 + a_{2,0}^2 - 2\theta a_{1,0}a_{2,0}$.

Also, let $\epsilon_{1,i} = \Delta_i^n U_1 + \Delta_i^n U_2$ and $\epsilon_{2,i} = \Delta_i^n U_1 - \Delta_i^n U_2$. Thus, the difference between $\Psi_{1,n}$ and $\bar{\Psi}_{1,n}$ is

$$\begin{aligned}
\Psi_{1,n} - \bar{\Psi}_{1,n} &= \frac{1}{2\sqrt{n}}\left\{\mathbf{Y}_1' \frac{\partial \mathbf{\Omega}_1^{-1}}{\partial \sigma_1^2}\mathbf{Y}_1 - \operatorname{tr}\left(\frac{\partial \mathbf{\Omega}_1^{-1}}{\partial \sigma_1^2}\mathbf{\Sigma}_1\right)\right\} \\
&= \frac{1}{2\sqrt{n}}\left\{\sum_{i=1}^n \frac{\partial \omega_1^{ii}}{\partial \sigma_1^2}\left((\Delta_i^n X_1 + \Delta_i^n X_2)^2 \right.\right. \\
&\quad \left. - \int_{\tau_{i-1}}^{\tau_i} \sigma_{1t}^2 + \sigma_{2t}^2 + 2\rho_t\sigma_{1t}\sigma_{2t}\,dt\right) \\
&\quad + \sum_{i=1}^n \sum_{j\neq i}^n \frac{\partial \omega_1^{ij}}{\partial \sigma_1^2}(\Delta_i^n X_1 + \Delta_i^n X_2)(\Delta_j^n X_1 + \Delta_j^n X_2) \\
&\quad + 2\sum_{i=1}^n \sum_{j=1}^n \frac{\partial \omega_1^{ij}}{\partial \sigma_1^2}\epsilon_{1,j}(\Delta_i^n X_1 + \Delta_i^n X_2) \\
&\quad \left. + \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \omega_1^{ij}}{\partial \sigma_1^2}(\epsilon_{1,i}\epsilon_{1,j} - E\epsilon_{1,i}\epsilon_{1,j})\right\} \\
&:= \frac{1}{2\sqrt{n}}(M_1^1 + M_2^1 + M_3^1 + M_4^1), \qquad \text{(A.1)}
\end{aligned}$$

$\Psi_{2,n} - \bar{\Psi}_{2,n}$ can be decomposed in a similar way. It follows from theorem 7.1 in Jacod (2007) that

$$\operatorname{Acov}\left(n^{-1/4}M_2^1, n^{-1/4}M_2^2\right) = \frac{a_2^2\sigma_1^2 + a_1^2\sigma_2^2 + 3a_1a_2\sigma_1\sigma_2}{2\sqrt{T}\sigma_1^3\sigma_2^3(\sigma_1 a_2 + \sigma_2 a_1)^3}\int_0^T (\sigma_{1t}^2 - \sigma_{2t}^2)^2\,dt.$$

Further, we can show that $n^{-1/4}(M_1^1 + M_2^1)$ and $n^{-1/4}(M_1^2 + M_2^2)$ jointly $\sigma(X)$-stable convergence in law, due to the fact that $n^{-1/4}M_1^1$ and $n^{-1/4}M_1^2$ are asymptotically negligible.

Note that

$$\sum_{i=1}^n (\Delta_i^n X_1 + \Delta_i^n X_2)(\Delta_i^n X_1 - \Delta_i^n X_2) \xrightarrow{\mathcal{P}} \int_0^T (\sigma_{1t}^2 - \sigma_{2t}^2)\,dt.$$

By conditioning on the filtration $\sigma(X)$ and standard central limit theorem, it follows that

$$\operatorname{Acov}\left(n^{-1/4}M_3^1, n^{-1/4}M_3^2\right) = \frac{(a_{1,0}^2 - a_{2,0}^2)}{\sqrt{T}\sigma_1\sigma_2(\sigma_1 a_2 + \sigma_2 a_1)^3}\int_0^T (\sigma_{1t}^2 - \sigma_{2t}^2)\,dt.$$

Last, the contribution of the noise terms is

$$\begin{aligned}
&\operatorname{Acov}\left(n^{-1/4}M_4^1, n^{-1/4}M_4^2\right) \\
&= \lim_{n\to\infty} \sum_{i,j=1}^n \sum_{k,l=1}^n \frac{\partial \omega_1^{ij}}{\partial \sigma_1^2}\frac{\partial \omega_2^{kl}}{\partial \sigma_2^2} \\
&\qquad \times \left(\operatorname{cum}(\epsilon_{1,i}, \epsilon_{1,j}, \epsilon_{2,k}, \epsilon_{2,l}) + 2\operatorname{cov}(\epsilon_{1,i}, \epsilon_{1,j})\operatorname{cov}(\epsilon_{2,k}, \epsilon_{2,l})\right) \\
&:= \lim_{n\to\infty} V_1\left(\frac{\partial \omega_1}{\partial \sigma_1^2}, \frac{\partial \omega_2}{\partial \sigma_2^2}\right) + V_2\left(\frac{\partial \omega_1}{\partial \sigma_1^2}, \frac{\partial \omega_2}{\partial \sigma_2^2}\right),
\end{aligned}$$

where

$$V_1(v, \omega) = \sum_{i,j,k,l=1}^n v^{ij}\omega^{kl}\operatorname{cum}(\epsilon_{1,i}, \epsilon_{1,j}, \epsilon_{2,k}, \epsilon_{2,l}),$$

$$\begin{aligned}
V_2(v, \omega) &= \sum_{i,j,k,l=1}^n v^{ij}\omega^{kl}2\operatorname{cov}(\epsilon_{1,i}, \epsilon_{1,j})\operatorname{cov}(\epsilon_{2,k}, \epsilon_{2,l}) \\
&= 2(a_{1,0}^2 - a_{2,0}^2)^2 \\
&\quad \times \sum_{i=1}^N \sum_{j=1}^N \{v^{ij}(\omega^{j-1,i-1} + \omega^{j-1,i+1} - 2\omega^{j-1,i} + \omega^{j+1,i-1} \\
&\quad + \omega^{j+1,i+1} - 2\omega^{j+1,i} - 2(\omega^{j,i-1} + \omega^{j,i+1} - 2\omega^{j,i}))\}.
\end{aligned}$$

And it follows from a similar proof of lemma 1 in Aït-Sahalia, Mykland, and Zhang (2005) that,

$$\operatorname{cum}(\epsilon_{1,i}, \epsilon_{1,j}, \epsilon_{2,k}, \epsilon_{2,l}) = \begin{cases} 2S(U_1, U_2), & \text{if } i = j = k = l \\ (-1)^{s(i,j,k,l)}S(U_1, U_2), & \text{if } \max(i,j,k,l) = \min(i,j,k,l) + 1 \\ 0, & \text{otherwise,} \end{cases}$$

where $S(U_1, U_2) = \operatorname{cum}_4[U_1] + \operatorname{cum}_4[U_2] - 2\operatorname{cov}(U_1^2, U_2^2) + 4\operatorname{cov}^2(U_1, U_2)$ and $s(i,j,k,l)$ denotes the number of indices among (i,j,k,l) that are equal to $\min(i,j,k,l)$.

Note that for $k = 1, 2$, we have

$$\begin{aligned}
&V_k\left(\frac{\partial \omega_1}{\partial \sigma_1^2}, \frac{\partial \omega_2}{\partial \sigma_2^2}\right) \\
&= \left(\frac{\partial \eta_1}{\partial \sigma_1^2}\right)\left(\frac{\partial \eta_2}{\partial \sigma_2^2}\right)V_k\left(\frac{\partial \omega_1}{\partial \eta_1}, \frac{\partial \omega_2}{\partial \eta_2}\right) \\
&\quad + \frac{1}{\gamma_1^2\gamma_2^2}\left(\frac{\partial \gamma_1^2}{\partial \sigma_1^2}\frac{\partial \gamma_2^2}{\partial \sigma_2^2}\right)V_k(\omega_1, \omega_2) \\
&\quad - \frac{1}{\gamma_1^2}\frac{\partial \eta_2}{\partial \sigma_2^2}\frac{\partial \gamma_1^2}{\partial \sigma_1^2}V_k\left(\omega_1, \frac{\partial \omega_2}{\partial \eta_2}\right) - \frac{1}{\gamma_2^2}\frac{\partial \eta_1}{\partial \sigma_1^2}\frac{\partial \gamma_2^2}{\partial \sigma_2^2}V_k\left(\frac{\partial \omega_1}{\partial \eta_1}, \omega_2\right),
\end{aligned}$$

where η_i and γ_i are given in (A.2) and (A.3), with σ^2 and a^2 replaced by σ_i^2 and a_i^2. Following a similar calculation as in lemma 2 of Xiu (2010), we can obtain

$$V_1\left(\frac{\partial\omega_1}{\partial\sigma_1^2},\frac{\partial\omega_2}{\partial\sigma_2^2}\right)=O(1),$$

$$V_2\left(\frac{\partial\omega_1}{\partial\sigma_1^2},\frac{\partial\omega_2}{\partial\sigma_2^2}\right)=\frac{(a_{1,0}^2-a_{2,0}^2)^2\sqrt{T}}{2a_1a_2(a_2\sigma_1+a_1\sigma_2)^3}n^{1/2}+o\big(n^{1/2}\big).$$

Therefore,

$$\operatorname{Acov}\big(n^{-1/4}M_4^1,n^{-1/4}M_4^2\big)=\frac{(a_{1,0}^2-a_{2,0}^2)^2\sqrt{T}}{2a_1a_2(a_2\sigma_1+a_1\sigma_2)^3}.$$

Combining the above results with Remark 1 and applying lemma 1 and proposition 5 in Barndorff-Nielsen et al. (2008b), we can obtain the central limit theorem for two-dimensional case:

$$\begin{pmatrix}n^{1/4}(\Psi_{1,n}-\bar\Psi_{1,n})\\ n^{1/4}(\Psi_{2,n}-\bar\Psi_{2,n})\end{pmatrix}\xrightarrow{\mathcal{L}_X}\mathrm{MN}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}U_{11}&U_{12}\\U_{12}&U_{22}\end{pmatrix}\right),$$

where

$$\begin{aligned}U_{12}&=\frac{a_2^2\sigma_1^2+a_1^2\sigma_2^2+3a_1a_2\sigma_1\sigma_2}{8\sqrt{T}\sigma_1^3\sigma_2^3(\sigma_1a_2+\sigma_2a_1)^3}\int_0^T(\sigma_{1t}^2-\sigma_{2t}^2)^2\,dt\\&\quad+\frac{(a_{1,0}^2-a_{2,0}^2)}{4\sqrt{T}\sigma_1\sigma_2(\sigma_1a_2+\sigma_2a_1)^3}\int_0^T(\sigma_{1t}^2-\sigma_{2t}^2)\,dt\\&\quad+\frac{(a_{1,0}^2-a_{2,0}^2)^2\sqrt{T}}{8a_1a_2(a_2\sigma_1+a_1\sigma_2)^3},\\U_{11}&=\frac{5\int_0^T(\sigma_{1t}^2+2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2)^2\,dt}{64a_1\sigma_1^7\sqrt{T}}\\&\quad+\frac{(a_{1,0}^2+a_{2,0}^2+2\theta a_{1,0}a_{2,0})^2\sqrt{T}}{64a_1^5\sigma_1^3}\\&\quad+\frac{(a_{1,0}^2+a_{2,0}^2+2\theta a_{1,0}a_{2,0})\int_0^T\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t}\,dt}{32\sigma_1^5a_1^3\sqrt{T}},\\U_{22}&=\frac{5\int_0^T(\sigma_{1t}^2-2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2)^2\,dt}{64a_2\sigma_2^7\sqrt{T}}\\&\quad+\frac{(a_{1,0}^2+a_{2,0}^2-2\theta a_{1,0}a_{2,0})^2\sqrt{T}}{64a_2^5\sigma_2^3}\\&\quad+\frac{(a_{1,0}^2+a_{2,0}^2-2\theta a_{1,0}a_{2,0})\int_0^T\sigma_{1t}^2+\sigma_{2t}^2-2\rho_t\sigma_{1t}\sigma_{2t}\,dt}{32\sigma_2^5a_2^3\sqrt{T}}.\end{aligned}$$

Therefore, a direct application of the Delta method yields

$$\begin{pmatrix}n^{1/4}(\hat\sigma_{1,n}^2-\frac{1}{T}\int_0^T\sigma_{1t}^2+2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2\,dt)\\ n^{1/4}(\hat\sigma_{2,n}^2-\frac{1}{T}\int_0^T\sigma_{1t}^2-2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2\,dt)\end{pmatrix}\xrightarrow{\mathcal{L}_X}\mathrm{MN}\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}V^{1,-1}&V^{1,1}\\V^{1,1}&V^{-1,1}\end{pmatrix}\right),$$

where

$$\begin{aligned}V^{1,-1}&=T^{-1/2}\left(5\frac{q_1^*}{\sigma_1^*}+3\sigma_1^{*3}\right)a_1^*,\\V^{-1,1}&=T^{-1/2}\left(5\frac{q_2^*}{\sigma_2^*}+3\sigma_2^{*3}\right)a_2^*,\\V^{1,1}&=\frac{16a_1^*a_2^*}{T(a_2^*\sigma_1^*+a_1^*\sigma_2^*)^3}\\&\quad\times\left(\frac{a_2^{*2}\sigma_1^{*2}+a_1^{*2}\sigma_2^{*2}+3a_1^*a_2^*\sigma_1^*\sigma_2^*}{2\sqrt{T}}\int_0^T(\sigma_{1t}^2-\sigma_{2t}^2)^2\,dt\right.\\&\quad+\frac{(a_{1,0}^2-a_{2,0}^2)\sigma_1^{*2}\sigma_2^{*2}}{\sqrt{T}}\int_0^T(\sigma_{1t}^2-\sigma_{2t}^2)\,dt\\&\quad\left.+\frac{(a_{1,0}^2-a_{2,0}^2)^2\sigma_1^{*3}\sigma_2^{*3}\sqrt{T}}{2a_1^*a_2^*}\right)\end{aligned}$$

with $\sigma_1^{*2}=\frac{1}{T}\int_0^T\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t}\,dt$, $\sigma_2^{*2}=\frac{1}{T}\int_0^T\sigma_{1t}^2+\sigma_{2t}^2-2\rho_t\sigma_{1t}\sigma_{2t}\,dt$, $q_1^*=\frac{1}{T}\int_0^T(\sigma_{1t}^2+2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2)^2\,dt$, and $q_2^*=\frac{1}{T}\times\int_0^T(\sigma_{1t}^2-2\rho_t\sigma_{1t}\sigma_{2t}+\sigma_{2t}^2)^2\,dt$.

Therefore,

$$n^{1/4}\left(\widehat{\operatorname{Cov}}(\tilde X_1,\tilde X_2)-\frac{1}{T}\int_0^T\rho_t\sigma_{1t}\sigma_{2t}\,dt\right)\xrightarrow{\mathcal{L}_X}\mathrm{MN}(0,V),$$

where

$$V=\begin{pmatrix}\frac14&-\frac14\end{pmatrix}\begin{pmatrix}V^{1,-1}&V^{1,1}\\V^{1,1}&V^{-1,1}\end{pmatrix}\begin{pmatrix}\frac14\\-\frac14\end{pmatrix}.$$

Similarly,

$$\begin{aligned}V^{\alpha,\beta}&=\frac{16a_1^\alpha a_2^\beta}{2T^{3/2}(a_2^\beta\sigma_1^\alpha+a_1^\alpha\sigma_2^\beta)^3}\\&\quad\times\left(\big((a_2^\beta)^2(\sigma_1^\alpha)^2+(a_1^\alpha)^2(\sigma_2^\beta)^2+3a_1^\alpha a_2^\beta\sigma_1^\alpha\sigma_2^\beta\big)\right.\\&\quad\times\int_0^T(\alpha\sigma_{1t}^2-\beta\sigma_{2t}^2+(1-\alpha\beta)\rho_t\sigma_{1t}\sigma_{2t})^2\,dt\\&\quad+2(\alpha a_{1,0}^2-\beta a_{2,0}^2+(1-\alpha\beta)\theta a_{1,0}a_{2,0})(\sigma_1^\alpha)^2(\sigma_2^\beta)^2\\&\quad\times\int_0^T(\alpha\sigma_{1t}^2-\beta\sigma_{2t}^2+(1-\alpha\beta)\rho_t\sigma_{1t}\sigma_{2t})\,dt\\&\quad+(a_1^\alpha a_2^\beta)^{-1}(\alpha a_{1,0}^2-\beta a_{2,0}^2+(1-\alpha\beta)\theta a_{1,0}a_{2,0})^2\\&\quad\left.\times(\sigma_1^\alpha)^3(\sigma_2^\beta)^3T\right),\end{aligned}$$

where $a_1^\alpha=(\alpha^2a_{1,0}^2+a_{2,0}^2+2\alpha\theta a_{1,0}a_{2,0})^{1/2}$, $a_2^\beta=(a_{1,0}^2+\beta^2a_{2,0}^2-2\beta\theta a_{1,0}a_{2,0})^{1/2}$, $\sigma_1^\alpha=(\frac{1}{T}\int_0^T\alpha^2\sigma_{1t}^2+\sigma_{2t}^2+2\alpha\rho_t\sigma_{1t}\sigma_{2t}\,dt)^{1/2}$, and $\sigma_2^\beta=(\frac{1}{T}\int_0^T\sigma_{1t}^2+\beta^2\sigma_{2t}^2-2\beta\rho_t\sigma_{1t}\sigma_{2t}\,dt)^{1/2}$.

Denote the asymptotic variances of the $\widehat{\operatorname{Var}}(\tilde X_i)$ as V^i, for $i=1,2$. According to Theorem 1,

$$V^i=\frac{5a_{i,0}\int_0^T\sigma_{i,t}^4\,dt}{T(\int_0^T\sigma_{i,t}^2\,dt)^{1/2}}+\frac{3a_{i,0}(\int_0^T\sigma_{i,t}^2\,dt)^{3/2}}{T^2}.$$

Therefore, using the Delta method again, we obtain

$$n^{1/4}\left(\widehat{\operatorname{Corr}}(\tilde X_1,\tilde X_2)-\frac{\int_0^T\rho_t\sigma_{1t}\sigma_{2t}\,dt}{\sqrt{\int_0^T\sigma_{1t}^2\,dt}\sqrt{\int_0^T\sigma_{2t}^2\,dt}}\right)\xrightarrow{\mathcal{L}_X}\mathrm{MN}(0,\tilde V)$$

and $\tilde V=\mathbf{e}\mathbf{\Lambda}\mathbf{e}'$, where

$$\mathbf{\Lambda}=\begin{pmatrix}V^1&.&.&.\\V^{1,0}&V^{1,-1}&.&.\\V^{-1,0}&V^{1,1}&V^{-1,1}&.\\V^{0,0}&V^{0,-1}&V^{0,1}&V^2\end{pmatrix}$$

and

$$\mathbf{e} = \Bigg(-\frac{T\int_0^T \rho_t \sigma_{1t}\sigma_{2t}\,dt}{2(\int_0^T \sigma_{1t}^2\,dt)^{3/2}(\int_0^T \sigma_{2t}^2\,dt)^{1/2}}, \frac{T}{4(\int_0^T \sigma_{1t}^2\,dt)^{1/2}(\int_0^T \sigma_{2t}^2\,dt)^{1/2}},$$
$$-\frac{T}{4(\int_0^T \sigma_{1t}^2\,dt)^{1/2}(\int_0^T \sigma_{2t}^2\,dt)^{1/2}}, -\frac{T\int_0^T \rho_t \sigma_{1t}\sigma_{2t}\,dt}{2(\int_0^T \sigma_{1t}^2\,dt)^{1/2}(\int_0^T \sigma_{2t}^2\,dt)^{3/2}} \Bigg).$$

A.2 Proof of Theorem 3

In order to simplify the notation and without loss of generality, we assume that $\tilde{X}_{it}$ is observed at $\{t_{i,j}, 0 \le j \le n, i = 1, 2\}$, where $t_{1,j-1} \le t_{2,j-1} \le \tau_{j-1} < t_{1,j} \le t_{2,j}$, for each $0 \le j \le n$. Otherwise, we can just swap the subscripts 1 and 2, if needed. Denote $\bar{n} = T/\bar{\Delta}$. Let $\Omega^{-1} = (\omega^{ij})$. Of course, we can add one more condition: $K^{-1} \le \sigma_{it}^2 \le K, \forall t \in [0, T]$, since one can always relax the constraint by localization scheme.

$$\begin{aligned}\boldsymbol{\Psi}_n &= (\Psi_n^1(\omega), \Psi_n^2(\omega))' \\ &= \Bigg(\frac{1}{2\sqrt{\bar{n}}} \Bigg\{ \frac{\partial \log(\det \boldsymbol{\Omega})}{\partial \sigma^2} + \mathbf{Y}' \frac{\partial \boldsymbol{\Omega}^{-1}}{\partial \sigma^2} \mathbf{Y} \Bigg\}, \\ &\qquad \frac{1}{2\bar{n}} \Bigg\{ \frac{\partial \log(\det \boldsymbol{\Omega})}{\partial a^2} + \mathbf{Y}' \frac{\partial \boldsymbol{\Omega}^{-1}}{\partial a^2} \mathbf{Y} \Bigg\} \Bigg)',\end{aligned}$$

where $Y_i = X_{1t_{1,i}} + X_{2t_{2,i}} - X_{1t_{1,i-1}} - X_{2t_{2,i-1}} + \epsilon_i$ and $\epsilon_i = U_{1t_{1,i}} + U_{2t_{2,i}} - U_{1t_{1,i-1}} - U_{2t_{2,i-1}}$. Also,

$$\bar{\boldsymbol{\Psi}}_n = \Bigg(\frac{1}{2\sqrt{\bar{n}}} \Bigg\{ \frac{\partial \log(\det \boldsymbol{\Omega})}{\partial \sigma^2} + \operatorname{tr}\Bigg(\frac{\partial \boldsymbol{\Omega}^{-1}}{\partial \sigma^2} \boldsymbol{\Sigma}_0 \Bigg) \Bigg\}, \frac{1}{2\bar{n}} \Bigg\{ \frac{\partial \log(\det \boldsymbol{\Omega})}{\partial a^2} + \operatorname{tr}\Bigg(\frac{\partial \boldsymbol{\Omega}^{-1}}{\partial a^2} \boldsymbol{\Sigma}_0 \Bigg) \Bigg\} \Bigg)',$$

where $\boldsymbol{\Sigma}_0 = (\Sigma_{0,i,j})$ is given by

$$\Sigma_{0,i,i} = \int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}^2\,dt + \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}^2\,dt + \int_{t_{2,i-1}}^{t_{1,i}} 2\rho_t \sigma_{1t}\sigma_{2t}\,dt + 2u_1^2,$$

$$\Sigma_{0,i,i-1} = \Sigma_{0,i-1,i} = -u_1^2 + \int_{t_{1,i-1}}^{t_{2,i-1}} \rho_t \sigma_{1t}\sigma_{2t}\,dt$$

and $u_1^2 = a_{1,0}^2 + a_{2,0}^2$. The other coefficients are 0.

On the other hand, we can write

$$\omega^{ij} = \frac{(-\eta)^{|i-j|} - (-\eta)^{i+j} - (-\eta)^{2n-i-j+2} + (-\eta)^{2n-|i-j|+2}}{\gamma^2(1-\eta^2)(1-\eta^{2n+2})},$$

where

$$\eta = \frac{1}{2a^2}\{-2a^2 - \sigma^2\bar{\Delta} + \sqrt{\sigma^2\bar{\Delta}(4a^2 + \sigma^2\bar{\Delta})}\}, \tag{A.2}$$

$$\gamma^2 = \frac{1}{2}\{2a^2 + \sigma^2\bar{\Delta} + \sqrt{\sigma^2\bar{\Delta}(4a^2 + \sigma^2\bar{\Delta})}\}. \tag{A.3}$$

We consider the same decomposition of $\boldsymbol{\Psi}_n - \bar{\boldsymbol{\Psi}}_n$ as in formula (A.1). Obviously, data synchronization has no effect on the order of the variances for the last two terms; hence we are left with the following two asymptotic results that are in need of verification:

$$\sum_{i=1}^n \omega^{ii} \Bigg\{ \Bigg(\int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}\,dW_{1t} + \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}\,dW_{2t} \Bigg)^2 - \int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}^2\,dt - \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}^2\,dt - \int_{t_{2,i-1}}^{t_{1,i}} 2\rho_t \sigma_{1t}\sigma_{2t}\,dt \Bigg\} = O_p(1),$$

$$\sum_{i=1}^n \sum_{j \ne i}^n \omega^{ij} \Bigg(\int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}\,dW_{1t} + \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}\,dW_{2t} \Bigg) \times \Bigg(\int_{t_{1,j-1}}^{t_{1,j}} \sigma_{1t}\,dW_{1t} + \int_{t_{2,j-1}}^{t_{2,j}} \sigma_{2t}\,dW_{2t} \Bigg) - \sum_{i=1}^n (\omega^{i,i-1} + \omega^{i-1,i}) \int_{t_{1,i-1}}^{t_{2,i-1}} \rho_t \sigma_{1t}\sigma_{2t}\,dt = O_p(\bar{n}^{1/4}).$$

By lemma 1 in Xiu (2010) and using the tower property of the conditional expectation (conditioning on $\{\Delta_i\}$), we have for any $k = 1, 2$,

$$\sum_{i=1}^n \omega^{ii} \Bigg\{ \Bigg(\int_{t_{k,i-1}}^{t_{k,i}} \sigma_{kt}\,dW_{kt} \Bigg)^2 - \int_{t_{k,i-1}}^{t_{k,i}} \sigma_{kt}^2\,dt \Bigg\} = O_p(1),$$

$$\sum_{i=1}^n \sum_{j \ne i}^n \omega^{ij} \Bigg(\int_{t_{k,i-1}}^{t_{k,i}} \sigma_{kt}\,dW_{kt} \int_{t_{k,j-1}}^{t_{k,j}} \sigma_{kt}\,dW_{kt} \Bigg) = O_p(\bar{n}^{1/4}).$$

Hence, it is sufficient to show that

$$\sum_{i=1}^n \omega^{ii} \Bigg\{ \int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}\,dW_{1t} \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}\,dW_{2t} - \int_{t_{2,i-1}}^{t_{1,i}} \rho_t \sigma_{1t}\sigma_{2t}\,dt \Bigg\} = O_p(1),$$

$$\sum_{i=1}^n \omega^{i,i-1} \int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}\,dW_{1t} \int_{t_{2,i-2}}^{t_{2,i-1}} \sigma_{2t}\,dW_{2t} - \sum_{i=1}^n \omega^{i,i-1} \int_{t_{1,i-1}}^{t_{2,i-1}} \rho_t \sigma_{1t}\sigma_{2t}\,dt = O_p(1).$$

These two equalities are in fact similar, since $\omega_{i,i-1} = \omega_{i,i} + o(1)$. It can be proved by noting that $\int_{t_{1,i-1}}^{t_{2,i-1}} \rho_t \sigma_{1t}\sigma_{2t}\,dt$ is the quadratic covariation of $\int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}\,dW_{1t}$ and $\int_{t_{2,i-2}}^{t_{2,i-1}} \sigma_{2t}\,dW_{2t}$.

So far, we have obtained the point-wise convergence of $\boldsymbol{\Psi}_n - \bar{\boldsymbol{\Psi}}_n$ to 0. Following the same reasoning in the proofs of theorems 2 and 4 in Xiu (2010), we can prove the stochastic equicontinuity of $\boldsymbol{\Psi}_n - \bar{\boldsymbol{\Psi}}_n$, which guarantees that

$$\sup_{\sigma^2, a^2} \|\boldsymbol{\Psi}_n(\sigma^2, a^2) - \bar{\boldsymbol{\Psi}}_n(\sigma^2, a^2)\| \xrightarrow{\mathrm{P}} 0.$$

Next, we verify the identifiability condition and find the roots of $\bar{\boldsymbol{\Psi}}_n$.

$$\begin{aligned}\bar{\Psi}_n^2 &= \frac{1}{2\bar{n}} \Bigg\{ \operatorname{tr}\Bigg(\boldsymbol{\Omega}^{-1} \frac{\partial \boldsymbol{\Omega}}{\partial a^2} \Bigg) + \operatorname{tr}\Bigg(\frac{\partial \boldsymbol{\Omega}^{-1}}{\partial a^2} \boldsymbol{\Sigma}_0 \Bigg) \Bigg\} \\ &= \frac{1}{\bar{n}} \Bigg\{ (\operatorname{tr} \boldsymbol{\Omega}^{-1} - \operatorname{tr} \boldsymbol{\Omega}^{-1}\mathbf{J}) + u_1^2 \Bigg(\operatorname{tr} \frac{\partial \boldsymbol{\Omega}^{-1}}{\partial a^2} - \operatorname{tr} \frac{\partial \boldsymbol{\Omega}^{-1}}{\partial a^2}\mathbf{J} \Bigg) \\ &\quad + \frac{1}{2} \sum_{i=1}^n \frac{\partial \omega^{ii}}{\partial a^2} \Bigg(\int_{t_{1,i-1}}^{t_{1,i}} \sigma_{1t}^2\,dt + \int_{t_{2,i-1}}^{t_{2,i}} \sigma_{2t}^2\,dt + \int_{t_{2,i-1}}^{t_{1,i}} 2\rho_t \sigma_{1t}\sigma_{2t}\,dt \Bigg) dt \Bigg\} \\ &\quad + \frac{1}{2} \sum_{i=1}^n \Bigg(\frac{\partial \omega^{i-1,i}}{\partial a^2} + \frac{\partial \omega^{i,i-1}}{\partial a^2} \Bigg) \int_{t_{1,i-1}}^{t_{2,i-1}} \rho_t \sigma_{1t}\sigma_{2t}\,dt,\end{aligned}$$

where $\mathbf{J} = (J_{ij})$, where $J_{i-1,i} = 1$, and the other components of $\mathbf{J}$ is 0, and

$$\begin{aligned}&\operatorname{tr} \boldsymbol{\Omega}^{-1} - \operatorname{tr}(\boldsymbol{\Omega}^{-1}\mathbf{J}) \\ &\quad = \frac{n(1+\eta)(1+\eta^{2n+1}) - \eta(1+\eta^{2n}) - 2\eta^2(1-\eta^{2n-1})/(1-\eta)}{\gamma^2(1-\eta^2)(1-\eta^{2n+2})} \\ &\quad = \frac{n}{2a^2} - \frac{\sqrt{Ta^2\sigma^2}}{4a^4} n^{1/2} + O(1).\end{aligned} \tag{A.4}$$

Let $\omega^m = \omega^{n^{2/3},n^{2/3}}$. Note that $\omega^{i,j} = \omega^{j,i}$, and for any $n^{2/3} \le i \le n - n^{2/3}$,

$$\omega^{ii} = \omega^m(1+o(1)) = \omega^{i,i-1}(1+o(1)),$$
$$\frac{\partial\omega^{ii}}{\partial a^2} = \frac{\partial\omega^{i,i-1}}{\partial a^2}(1+o(1)) = -\frac{1}{2a^2}\omega^{ii}(1+o(1)),$$
$$\frac{\partial\omega^{ii}}{\partial\sigma^2} = \frac{\partial\omega^{i,i-1}}{\partial\sigma^2}(1+o(1)) = -\frac{1}{2\sigma^2}\omega^{ii}(1+o(1)).$$

For $i \le n^{2/3}$ and $i \ge n - n^{2/3}$, they are dominated by ω^m, and the integration is over an interval which shrinks at the rate $n^{-1/3}$. So

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{n}\frac{\partial\omega^{ii}}{\partial a^2}\left(\int_{t_{1,i-1}}^{t_{1,i}}\sigma_{1t}^2\,dt + \int_{t_{2,i-1}}^{t_{2,i}}\sigma_{2t}^2\,dt + \int_{t_{2,i-1}}^{t_{1,i}}2\rho_t\sigma_{1t}\sigma_{2t}\,dt\right)\\
&\quad+\frac{1}{2}\sum_{i=1}^{n}\left(\frac{\partial\omega^{i-1,i}}{\partial a^2}+\frac{\partial\omega^{i,i-1}}{\partial a^2}\right)\int_{t_{1,i-1}}^{t_{2,i-1}}\rho_t\sigma_{1t}\sigma_{2t}\,dt\\
&=\frac{\partial\omega^m}{\partial a^2}\int_0^T(\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t})\,dt\,(1+o_p(1))\\
&=-\frac{\bar n^{1/2}(a^2)^{-3/2}}{4\sqrt{\sigma^2T}}\\
&\quad\times\int_0^T(\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t})\,dt\,(1+o_p(1)). \qquad (A.5)
\end{aligned}$$

By calculation, we can also obtain

$$\begin{aligned}
\bar\Psi_n^2 = &\left(\frac{1}{2a^2}-\frac{u_1^2}{2a^4}\right)+\left(\frac{3u_1^2\sqrt{\sigma^2T}}{8a^5}-\frac{\sqrt{\sigma^2T}}{4a^3}\right.\\
&\left.-\frac{\int_0^T(\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t})\,dt}{8a^3\sqrt{\sigma^2T}}\right)\bar n^{-1/2}+o_p(\bar n^{-1/2}).
\end{aligned}$$

Hence, if we solve $\bar\Psi_n^2 = 0$, we get

$$\begin{aligned}
a_n^{*2} = &u_1^2+\left(\frac{3u_1^2\sqrt{\sigma_n^{*2}T}}{4a_n^*}-\frac{\sqrt{\sigma_n^{*2}T}a_n^*}{2}\right.\\
&\left.-\frac{a_n^*\int_0^T(\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t})\,dt}{4\sqrt{\sigma_n^{*2}T}}\right)n^{-1/2}\\
&+o_p(\bar n^{-1/2}). \qquad (A.6)
\end{aligned}$$

On the other hand, let $\mathbf{\Gamma}$ be the symmetric trigonal matrix, whose ith element on the diagonal is $\Gamma_i := \int_{t_{1,i-1}}^{t_{1,i}}\sigma_{1t}^2\,dt + \int_{t_{2,i-1}}^{t_{2,i}}\sigma_t^2\,dt + \int_{t_{2,i-1}}^{t_{1,i}}2\rho_t\sigma_{1t}\sigma_{2t}\,dt - \sigma^2\bar\Delta$, and the element off the diagonal is $\int_{t_{1,i-1}}^{t_{2,i-1}}\rho_t\sigma_{1t}\sigma_{2t}\,dt$ then

$$\begin{aligned}
\bar\Psi_n^1(\sigma^2,a^2) &= \frac{1}{2\sqrt{\bar n}}\left\{\operatorname{tr}\left(\mathbf{\Omega}^{-1}\frac{\partial\mathbf{\Omega}}{\partial\sigma^2}\right)+\frac{\partial\operatorname{tr}(\mathbf{\Omega}^{-1}\mathbf{\Sigma}_0)}{\partial\sigma^2}\right\}\\
&= \frac{1}{2\sqrt{\bar n}}\left\{\operatorname{tr}\left(\mathbf{\Omega}^{-1}\frac{\partial\mathbf{\Omega}}{\partial\sigma^2}\right)\right.\\
&\quad\left.+\frac{\partial\operatorname{tr}(\mathbf{\Omega}^{-1}(\mathbf{\Omega}+(2\mathbf{I}-\mathbf{J}-\mathbf{J}')(a^2-u_1^2)+\mathbf{\Gamma}))}{\partial\sigma^2}\right\}\\
&= \frac{1}{2\sqrt{\bar n}}\left\{\sum_{i=1}^{n}\left(\frac{\partial\omega^{i,i-1}}{\partial\sigma^2}+\frac{\partial\omega^{i,i-1}}{\partial\sigma^2}\right)\Gamma_{i,i-1}+\sum_{i=1}^{n}\frac{\partial\omega^{ii}}{\partial\sigma^2}\Gamma_{ii}\right\}\\
&\quad-\frac{\sqrt{T}}{8a^3\sigma}(a^2-u_1^2)+O_p(\bar n^{-1/2})\\
&= -\frac{1}{8a\sigma^3\sqrt{T}}\left(\int_0^T\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t}\,dt-\sigma^2T\right)\\
&\quad-\frac{\sqrt{T}}{8a^3\sigma}(a^2-u_1^2)+O_p(\bar n^{-1/2}). \qquad (A.7)
\end{aligned}$$

Set $\bar\Psi_n^1(\sigma^2,a^2) = 0$, that is,

$$\begin{aligned}
&\int_0^T\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t}\,dt-\sigma_n^{*2}T\\
&\quad=-\frac{\sigma_n^{*2}T}{a_n^{*2}}(a_n^{*2}-u_1^2)+O_p(\bar n^{-1/2})\\
&\quad=O_p(\bar n^{-1/2}). \qquad (A.8)
\end{aligned}$$

Therefore, all the desired equalities remain as in Xiu (2010), except that we replace a_0^2 with u_1^2 and $\int_0^T\sigma_t^2\,dt$ with $\int_0^T\sigma_{1t}^2+\sigma_{2t}^2+2\rho_t\sigma_{1t}\sigma_{2t}\,dt$. Thus, the identifiability condition holds, and hereby we have $\hat\sigma^2-\sigma_n^{*2}=o_p(1)$ and $\hat a^2-a_n^{*2}=o_p(1)$. The desired convergence rates follow from Taylor expansion of $\mathbf{\Psi}_n-\bar{\mathbf{\Psi}}_n$ as in the proof of Theorem 2.

[Received March 2010. Revised June 2010.]

REFERENCES

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005), "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, 18, 351–416. [1504,1505,1513]

——— (2009), "Ultra High Frequency Volatility Estimation With Dependent Microstructure Noise," *Journal of Econometrics*, to appear. [1504]

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38–62. [1512]

Bandi, F. M., and Russell, J. R. (2003), "Microstructure Noise, Realized Volatility and Optimal Sampling," technical report, University of Chicago Graduate School of Business. [1504]

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008a), "Designing Realized Kernels to Measure ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536. [1504]

——— (2008b), "Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices With Noise and Non-Synchronous Trading," technical report, Department of Mathematical Sciences, University of Aarhus. [1505,1507,1508,1514]

Brunnermeier, M. K., Nagel, S., and Pedersen, L. H. (2009), "Carry Trades and Currency Crashes," *NBER Macroeconomics Annual 2008*, 23, 313–347. [1510]

Campbell, J. Y., de Medeiros, K. S., and Viceira, L. M. (2009), "Global Currency Hedging," *Journal of Finance*, to appear. [1511]

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010), "Pre-Averaging Estimators of the ex-post Covariance Matrix in Noisy Diffusion Models With Non-Synchronous Data," *Journal of Econometrics*, to appear. [1505]

deB. Harris, F. H., McInish, T. H., Shoesmith, G. L., and Wood, R. A. (1995), "Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets," *Journal of Financial and Quantitative Analysis*, 30, 563–581. [1507]

Delattre, S., and Jacod, J. (1997), "A Central Limit Theorem for Normalized Functions of the Increments of a Diffusion Process, in the Presence of Round-Off Errors," *Bernoulli*, 3, 1–28. [1504]

Epps, T. W. (1979), "Comovements in Stock Prices in the Very Short Run," *Journal of the American Statistical Association*, 74, 291–296. [1504]

Fan, J., and Wang, Y. (2007), "Multi-Scale Jump and Volatility Analysis for High-Frequency Financial Data," *Journal of the American Statistical Association*, 102, 1349–1362. [1504]

Fan, J., Li, Y., and Yu, K. (2010), "Vast Volatility Matrix Estimation Using High Frequency Data for Portfolio Selection," technical report, Princeton University. [1506]

Gatheral, J., and Oomen, R. C. (2009), "Zero-Intelligence Realized Variance Estimation," *Finance and Stochastics*, to appear. [1504]

Griffin, J., and Oomen, R. (2008), "Sampling Returns for Realized Variance Calculations: Tick Time or Transaction Time?" *Econometric Reviews*, 27, 230–253. [1504]

Guillaume, D. M., Dacorogna, M. M., Dave, R., Müller, U. A., Olsen, R. B., and Pictet, O. V. (1997), "From the Bird's Eye to the Microscope: A Survey of New Stylized Facts of the Intradaily Foreign Exchange Markets," *Finance and Stochastics*, 1, 95–129. [1504]

Hansen, P. R., and Lunde, A. (2006), "Realized Variance and Market Microstructure Noise," *Journal of Business & Economic Statistics*, 24, 127–161. [1504]

Hasbrouck, J. (2007), *Empirical Market Microstructure*, New York: Oxford University Press. [1505]

Hayashi, T., and Yoshida, N. (2005), "On Covariance Estimation of Non-Synchronously Observed Diffusion Processes," *Bernoulli*, 11, 359–379. [1504,1507]

Jacod, J. (2007), "Statistics and High Frequency Data: SEMSTAT Seminar," technical report, Université de Paris-6. [1513]

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009), "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach," *Stochastic Processes and Their Applications*, 119, 2249–2276. [1504]

Kalnina, I., and Linton, O. (2008), "Estimating Quadratic Variation Consistently in the Presence of Endogenous and Diurnal Measurement Error," *Journal of Econometrics*, 147, 47–59. [1504]

Large, J. (2007), "Accounting for the Epps Effect: Realized Covariation, Cointegration and Common Factors," technical report, Oxford-Man Institute, University of Oxford. [1504]

Li, Y., and Mykland, P. A. (2007), "Are Volatility Estimators Robust With Respect to Modeling Assumptions?" *Bernoulli*, 13, 601–622. [1504]

Li, Y., Mykland, P., Renault, E., Zhang, L., and Zheng, X. (2010), "Realized Volatility When Sampling Times Are Possibly Endogenous," technical report, Hong Kong University of Science and Technology. [1504]

Muthuswamy, J., Sarkar, S., Low, A., and Terry, E. (2001), "Time Variation in the Correlation Structure of Exchange Rates: High Frequency Analysis," *Journal of Futures Markets*, 21, 127–144. [1504]

Roll, R. (1984), "A Simple Model of the Implicit Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127–1139. [1505]

Voev, V., and Lunde, A. (2007), "Integrated Covariance Estimation Using High-Frequency Data in the Presence of Noise," *Journal of Financial Econometrics*, 5, 68–104. [1504]

Xiu, D. (2010), "Quasi-Maximum Likelihood Estimation of Volatility With High Frequency Data," *Journal of Econometrics*, 159, 235–250. [1504, 1505,1514-1516]

Zhang, L. (2006), "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach," *Bernoulli*, 12, 1019–1043. [1504]

——— (2009), "Estimating Covariation: Epps Effect and Microstructure Noise," *Journal of Econometrics*, to appear. [1504,1505,1507,1508]

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [1504]

——— (2009), "Edgeworth Expansions for Realized Rolatility and Related Estimators," *Journal of Econometrics*, to appear. [1504]