Michael W. Brandt
Duke University and the National Bureau of Economic Research

Francis X. Diebold
University of Pennsylvania and the National Bureau of Economic Research

A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations*

We extend range-based volatility estimation to the multivariate case. In particular, we propose a range-based covariance estimator motivated by a key financial economic consideration, the absence of arbitrage, in addition to statistical considerations. We show that this estimator is highly efficient yet robust to the market microstructure noise arising from bid-ask bounce and asynchronous trading.

I. Introduction

The price range, defined as the difference between the highest and lowest log asset prices over a fixed sampling interval (for concreteness, we focus on a 1-day interval), has a long, colorful, and distinguished history of use as a volatility estimator.[1] As emphasized most recently by Alizadeh, Brandt, and Diebold (2002), the range is a highly efficient volatility proxy, distilling volatility information from the entire intraday price path, in contrast to volatility proxies based on the daily return, such as the daily squared return, which use only the opening and closing prices. Moreover, data on the

* This work was supported by the National Science Foundation, the Wharton Institutions Center, and the Rodney L. White Center for Financial Research at the Wharton School. We thank Al Mandansky and an anonymous referee for insightful and forceful comments. We also thank Tim Bollerslev, Celso Brunetti, Rob Engle, Joel Hasbrouk, Peter Lildolt, Jeff Russell, and Niel Shephard, as well as seminar participants at the University of Pennsylvania and CIRANO for insightful discussions and comments. Clara Vega provided outstanding research assistance. Contact the corresponding author, Francis X. Diebold, at fdiebold@wharton.upenn.edu.

1. The relevant literature includes Garman and Klass (1980), Parkinson (1980), Beckers (1983), Ball and Torous (1984), Rogers and Satchell (1991), Kunitomo (1992), and Yang and Zhang (2000).

(*Journal of Business,* 2006, vol. 79, no. 1)

0021-9398/2006/7901-0003$10.00

range are widely available for individual stocks and exchange-traded futures contracts (including currencies, Treasury securities, and stock indices), not only presently but also, in many cases, over long historical spans. In fact, the range has been reported for many decades in business newspapers through so-called candlestick plots, showing the daily high, low, and close prices.

Despite these appealing properties of the range, one cannot help but notice a large and striking gap in the range-based volatility estimation literature: it is entirely univariate. That is, although range-based variance estimation has been discussed and refined extensively, range-based covariance estimation remains uncharted territory. The reason is that it is not at all obvious how to construct an appropriate range-based covariance estimator. Hence, the range would seem to join the ranks of other famously obvious and intuitive univariate statistics, such as the median, that have no similarly obvious or intuitive multivariate generalization.

The apparent failure of range-based volatility estimation to generalize to the multivariate case is particularly unfortunate, because financial economics is intimately concerned with multivariate interactions. Consider, for example, three pillars of modern finance: asset pricing, asset allocation, and risk management. Asset prices depend on covariance with the market and perhaps other risk factors. Similarly, optimal portfolio shares depend on the variances and covariances of asset returns, as does the portfolio value at risk.

We attempt to remedy the situation by proposing a simple and intuitive range-based covariance estimator. Our approach is not merely statistical; rather, it relies appealingly on a key financial economic consideration, the absence of arbitrage. In particular, we use no-arbitrage conditions to express covariances in terms of variances, which may then be estimated by standard range-based methods.

II. Range-Based Variance and Covariance Estimation

Before considering the range-based estimation of covariances, we must set the stage by considering certain aspects of univariate volatility estimation. Consider a univariate stochastic volatility diffusion for the log of an asset price p_t with instantaneous volatility σ_t. Suppose we sample this process discretely at m regular times throughout the day, which lasts from time t to $t+1$, to obtain the intraday returns $r_{(m),t+k/m} = p_{t+k/m} - p_{t+(k-1)/m}$, for $k = 1, \ldots, m$. Under conditions given in Andersen, Bollerslev and Diebold (2003a), the variance of the discrete-time returns over the 1-day interval conditional on the sample path $\{\sigma_{t+\tau}\}_{\tau=0}^{1}$ is

$$\bar{\sigma}_t^2 = \int_0^1 \sigma_{t+\tau}^2 d\tau. \tag{1}$$

The integrated volatility $\bar{\sigma}_t$ thus provides a canonical and natural measure of return volatility, and it features prominently in the financial economics literature (e.g., Hull and White 1987). Because the integrated volatility is inherently unobservable, several estimators have been proposed, including estimators based on daily returns (e.g., daily squared or absolute returns), high-frequency intraday returns (e.g., the "realized volatility" of Andersen et al. 2003b), and the daily range. In particular, Parkinson's (1980) celebrated range-based estimator of the daily integrated variance is given by

$$\gamma = 0.361\text{range}_t^2. \tag{2}$$

The univariate range-based volatility estimator has several appealing properties. First, of course, it is trivial to compute. Second, it is unbiased and highly efficient relative to competitors, such as the squared or absolute daily return (Andersen and Bollerslev 1998). Finally, it is robust to certain types of microstructure noise, such as bid-ask bounce (Alizadeh et al. 2002).

Now consider the multivariate case. In parallel with our univariate discussion, consider a stochastic volatility diffusion for a vector of log asset prices with diffusion matrix Σ, whose *ij*th element we denote σ_{ij}^2. Then, again, under the conditions given in Andersen et al. (2003a), the 1-day conditional covariance of the discrete-time returns on assets *i* and *j* is just the integrated instantaneous covariance,

$$\bar{\sigma}_{ij,t}^2 = \int_0^1 \sigma_{ij,t+\tau}^2 d\tau. \tag{3}$$

The attractive blend of convenience, efficiency, and robustness achieved by the range-based estimator in the univariate estimation of integrated volatility (1) makes one hungry for an extension to a range-based estimator of the integrated covariance (3) in the multivariate case. We now proceed to do so. The basic idea is very simple, and the implementation varies slightly, depending on whether the application is to foreign exchange, bonds, or stocks. We consider each in turn.

First, consider foreign exchange. In foreign exchange markets, the absence of triangular arbitrage implies a deterministic relationship between any pair of dollar rates and the corresponding cross rate. Consider two dollar exchange rates, denoted *A*/\$ and *B*/\$. Then, in the absence of triangular arbitrage, the cross rate is

$$\Delta\ln A/B = \Delta\ln A/\$ - \Delta\ln B/\$, \tag{4}$$

and hence the continuously compounded A/B return is $A/B = (A/\$)/(B/\$)$. Taking variances gives

$$\mathrm{Var}[\Delta \ln A/B] = \mathrm{Var}[\Delta \ln A/\$] + \mathrm{Var}[\Delta \ln B/\$] - 2\mathrm{Cov}[\Delta \ln A/\$, \Delta \ln B/\$], \tag{5}$$

and solving for the covariance yields

$$\mathrm{Cov}[\Delta \ln A/\$, \Delta \ln B/\$] = \frac{1}{2}(\mathrm{Var}[\Delta \ln A/\$] + \mathrm{Var}[\Delta \ln B/\$] - \mathrm{Var}[\Delta \ln A/B]). \tag{6}$$

This suggests a natural covariance estimator,

$$\widehat{\mathrm{Cov}}[\Delta \ln A/\$, \Delta \ln B/\$] = \frac{1}{2}\left(\widehat{\mathrm{Var}}[\Delta \ln A/\$] + \widehat{\mathrm{Var}}[\Delta \ln B/\$] - \widehat{\mathrm{Var}}[\Delta \ln A/B]\right), \tag{7}$$

where $\widehat{\mathrm{Var}}[\Delta \cdot]$ in principle can be any return variance estimator. Given the desirable properties of range-based volatility just estimation discussed, we advocate the use of Parkinson's (1980) range-based estimator in equation (2). We then assembled the estimated variance-covariance matrix as

$$\widehat{\Sigma} = \begin{bmatrix} \widehat{\mathrm{Var}}[\Delta \ln A/\$] & \widehat{\mathrm{Cov}}[\Delta \ln A/\$, \Delta \ln B/\$] \\ \widehat{\mathrm{Cov}}[\Delta \ln A/\$, \Delta \ln B/\$] & \widehat{\mathrm{Var}}[\Delta \ln B/\$] \end{bmatrix}. \tag{8}$$

In higher dimensional cases, we proceed in analogous fashion, estimating each pairwise covariance as in the preceding, then assembling the results into an estimated covariance matrix.

Now consider fixed-income markets, in which the absence of arbitrage implies a deterministic relationship among any two zero-coupon bond prices and the corresponding forward contract. Specifically, consider two bonds with maturities T_1 and T_2 and prices $P(T_1)$ and $P(T_2)$, with $T_1 < T_2$. The price of a forward contract between times T_1 and T_2 is $F(T_1, T_2) = P(T_2)/P(T_1)$. Taking logs gives $f(T_1, T_2) = p(T_2) - p(T_1)$; then, taking first differences gives $r_f(T_1, T_2) = r(T_2) - r(T_1)$; finally, taking variances gives

$$\mathrm{Var}[r_f(T_1, T_2)] = \mathrm{Var}[r(T_2)] + \mathrm{Var}[r(T_1)] - 2\mathrm{Cov}[r(T_1), r(T_2)]. \tag{9}$$

Hence, we can form the covariance estimator,

$$\widehat{\mathrm{Cov}}[r(T_1), r(T_2)] = \frac{1}{2}\left(\widehat{\mathrm{Var}}[r(T_2)] + \widehat{\mathrm{Var}}[r(T_1)] - \widehat{\mathrm{Var}}[r_f(T_1, T_2]\right), \tag{10}$$

and assemble the estimated variance-covariance matrix precisely as in the foreign exchange case.

Finally, consider equities. The return on a two-equity portfolio with shares λ and $1-\lambda$, denoted $r_p = \lambda r_1 + (1-\lambda) r_2$, has a variance of

$$\mathrm{Var}[r_p] = \lambda^2 \mathrm{Var}[r_1] + (1-\lambda)^2 \mathrm{Var}[r_2] + 2\lambda(1-\lambda)\mathrm{Cov}[r_1, r_2], \quad (11)$$

which suggests the covariance estimator,

$$\widehat{\mathrm{Cov}}[r_1, r_2] = \frac{1}{2\lambda(1-\lambda)} \left(\widehat{\mathrm{Var}}[r_p] - \lambda^2 \widehat{\mathrm{Var}}[r_1] - (1-\lambda)^2 \widehat{\mathrm{Var}}[r_2] \right). \quad (12)$$

This method of estimating the covariance via the range of the two-asset portfolio return is generally applicable to any two assets—not just equities—if data on the portfolio return range are available.

III. Discussion

Our no-arbitrage approach to range-based covariance estimation is widely applicable in the foreign exchange context because daily ranges of all legs of many currency triangles are available. For example, Datastream provides as much as 40 years of historical data on the daily high, low, and closing prices of 37 British-pound-denominated currencies and 14 Swiss-franc-denominated currencies. The International Monetary Market, a subsidiary of the Chicago Mercantile Exchange, recently introduced futures and options contracts on euro/British pound, euro/Swiss franc, and euro/Japanese yen cross rates. Finally, the New York Board of Trade offers futures contracts on 14 cross currencies, including seven euro-denominated contracts.

We hasten to add, however, that the practical applicability of our approach in other contexts is far more limited. For fixed-income securities, our approach is directly applicable to only select maturities, for which liquid bonds are aligned with liquid forward or futures contracts, such as the 3- and 6-month Eurodollar deposits and the 3-month Eurodollar futures. For equities, our approach rarely will be applicable, because historical data on the range of two-asset portfolios are typically not available.[2]

Thus far, we have said little about the theoretical properties of the range-based covariance estimator. One obvious point is that the covariance estimator is unbiased under the same conditions that deliver unbiasedness of Parkinson's (1980) variance estimator, because it is a linear combination of variances. Conversion to correlation, however, will introduce bias due to the nonlinearity of the transformation.

2. Some notable exceptions are the TSE 100, TSE 200, and TSE 300 indices of the Toronto Stock Exchange and the ASX 100, ASX 200, and ASX 300 indices of the Australian Stock Exchange.

A similarly obvious and related point is that $\widehat{\Sigma}$, in general, is not guaranteed to be positive definite. In our experience, however, positive definiteness is rarely violated in practice. If desired, positive definiteness can be imposed by estimating the Cholesky factor P of Σ, rather than Σ itself, where P is the unique lower-triangular matrix defined by $\Sigma = PP'$. Note that the elements of P are functions of the elements of Σ. Hence, we insert our range-based estimators of the relevant variances and covariances into P (computed analytically) to obtain an estimator of the Cholesky factor $\widehat{P}$, then form the estimator $\widehat{P}\widehat{P}'$ of the covariance matrix. Because the estimated Cholesky factor $\widehat{P}$ will be complex when $\widehat{\Sigma}$ is not positive definite, we define $\widehat{P}'$ as the conjugate transpose, which guarantees that $\widehat{P}\widehat{P}'$ is real.[3]

Ultimately, however, the interesting questions for financial economists center not on the theoretical properties of range-based covariance and correlation estimates under abstract conditions surely violated in practice, but rather on their performance in realistic situations involving small samples, discrete sampling, and market microstructure noise. As we argued previously, we have reason to suspect the good performance of the range-based approach, because of both its high efficiency due to the use of the information in the intraday sample path and its robustness to microstructure noise. We now turn to a brief Monte Carlo analysis designed to illuminate precisely those issues.

IV. Monte Carlo Exploration

We initially ignored market microstructure issues. We assumed that two dollar-denominated exchange rates P_1 and P_2 evolve as driftless diffusions with annualized volatilities σ of 15%, a covariance of 0.9, and hence a correlation ρ of 0.4. We further assumed that, at each instant, the cross rate P_3 is determined by the absence of triangular arbitrage as the ratio of the two dollar rates. Starting at $P_{1,0} = P_{2,0} = 1$, we simulated 24 hours worth of m regularly spaced intraday log price observations using

$$P_{i,t+k/m} = P_{i,t+(k-1)/m} + \sigma\sqrt{250/m}\,\varepsilon_{i,t+k/m},$$
$$\text{for } i = \{1, 2\} \text{ and } p_{3,t} = p_{1,t} - p_{2,t}, \qquad (13)$$

where $p_i = \ln P_i$, $[\varepsilon_1, \varepsilon_2]$ are standard normal innovations with correlation ρ, and there are 250 trading days per year. We considered sampling frequencies m ranging from $m = 18$ (one observation every 1 hour and 20 minutes) to $m = 1{,}440$ (one observation every minute) and used the resulting data to compute the daily range and intraday returns. We

3. Other ways to guarantee positive definiteness include the shrinkage approach of Ledoit and Wolf (2001) or the perturbation methods of Gill, Murray, and Wright (1981) and Schnabel and Eskow (1999).

TABLE 1 Range-Based and Realized Estimates in Merton's Utopia

	SD			Covariance			Correlation		
Sampling Frequency	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
Range-Based Estimates									
1 minute	14.099	4.279	4.373	.862	1.084	1.085	.371	.341	.342
5 minutes	13.746	4.277	4.457	.823	1.061	1.064	.369	.351	.352
10 minutes	13.477	4.274	4.537	.794	1.043	1.048	.368	.359	.360
20 minutes	13.090	4.266	4.674	.753	1.016	1.026	.366	.370	.372
40-minutes	12.525	4.255	4.923	.695	.977	.998	.363	.389	.391
1 hour 20 minutes	11.701	4.236	5.369	.615	.918	.961	.358	.420	.422
Realized Estimates with No-Arbitrage Condition									
1 minutes	14.997	.280	.280	.900	.064	.064	.400	.022	.022
5 minutes	14.985	.623	.624	.900	.143	.143	.400	.050	.050
10 minutes	14.971	.883	.883	.901	.202	.202	.399	.070	.070
20 minutes	14.943	1.249	1.250	.900	.285	.285	.398	.099	.099
40 minutes	14.888	1.758	1.762	.898	.404	.404	.395	.142	.142
1 hour 20 minutes	14.788	2.475	2.484	.896	.570	.570	.389	.203	.203
Realized Estimates with Cross Products									
1 minute	14.997	.280	.280	.900	.064	.064	.400	.022	.022
5 minutes	14.985	.623	.624	.900	.143	.143	.400	.050	.050
10 minutes	14.971	.883	.883	.901	.202	.202	.399	.070	.070
20 minutes	14.943	1.249	1.250	.900	.285	.285	.398	.099	.099
40 minutes	14.888	1.758	1.762	.898	.404	.404	.395	.142	.142
1 hour 20 minutes	14.788	2.475	2.484	.896	.570	.570	.389	.203	.203

NOTE.—Two dollar denominated exchange rates P_1 and P_2 evolve as driftless diffusions with an annualized volatility σ of 15%, covariance of 0.9, and correlation ρ of 0.4. At each instant, the cross-currency rate P_3 is given by the absence of triangular arbitrage as the ratio of the two base currencies. Starting at $P_{1,0} = P_{2,0} = 1$, we simulated 24 hours worth of m regularly spaced intraday log prices using $P_{i,t+k/m} = P_{i,t+(k-1)/m} + \sigma\sqrt{250/m}\,\varepsilon_{i,t+k/m}, i = \{1, 2\}$, and $P_{3,t} = P_{1,t} - P_{2,t}$, for $k = 1, \ldots, m$, where $p_i = \ln P_i$ and $[\varepsilon_1, \varepsilon_2]$ are standard-normal innovations with correlation ρ. The sampling frequency m ranges from 18 (one observation every hour and 20 minutes) to 1,440 (an observation every minute). We used this observed data to compute the daily range and intraday returns then construct three estimates of the volatilities, covariance, and correlation. We construct range-based covariance estimates using Parkinson's variance estimator and $\mathrm{Cov}[\Delta p_1, \Delta p_2] = 1/2(\mathrm{Var}[\Delta p_1] + \mathrm{Var}[\Delta p_2] - \mathrm{Var}[\Delta p_3])$. We constructed realized covariance estimates using either the realized variance estimator and the same expression for the covariance or the cross products of intraday returns. We repeated this procedure 10,000 times and report the mean, standard deviations (SD), and root mean squared errors (RMSE).

then constructed three estimates of the volatilities, covariance, and correlation of the two dollar rates. Specifically, we constructed range-based covariance matrix estimates using Parkinson's variance estimator (2) and equation (7), and for comparison, we computed the realized covariance matrix using two different approaches. First, in parallel fashion to the range-based estimator, we used the three realized variances constructed from the sum of squared intraday returns to obtain an estimate of the covariance. Second, we computed the realized covariance directly using the cross products of intraday returns. We repeated this procedure 10,000 times and report the means, standard deviations, and root mean squared errors of the resulting sampling distributions in table 1.

The results for the volatilities are familiar from Alizadeh et al. (2002).[4] The range-based estimates are downward biased because the range of the discretely sampled process is strictly less than the range of the underlying diffusion. The magnitude of the bias decreases as the sampling frequency increases. But, even in the limit as $m \to \infty$, the range is still only a noisy volatility proxy, which means that the standard deviation and RMSE of the range-based volatility estimator settle down to nonzero values. The realized volatility behaves quite differently because it converges not only in expectation but also in realization to the true volatility. The more frequently the underlying diffusion is sampled, the more precise the realized volatility gets, until, in the limit, the standard deviation and root mean squared error (RMSE) of the estimator are zero.

The results for the range-based covariance estimates follow from the properties of the range-based volatility estimates. The estimator $\text{Cov}[\Delta p_1, \Delta p_2] = 1/2(\text{Var}[\Delta p_1] + \text{Var}[\Delta p_2] - \text{Var}[\Delta p_3])$ involves three volatility estimates, each of which is downward biased by an amount that depends on the level of volatility (the higher is the volatility, the more likely that the true extremes are far from the observed extremes). Because the variance of p_3 is less than the variance of p_1 and p_2, due to the positive covariance, the covariance estimates are also downward biased because the downward bias of $\text{Var}[\Delta p_1] + \text{Var}[\Delta p_2]$ dominates the upward bias of $-\text{Var}[\Delta p_3]$. As with the volatility estimates, the bias vanishes as we increase the sampling frequency, and the standard deviation and RMSE stabilize. The realized covariances, computed either through the no-arbitrage condition or with return cross products, yield identical estimates that inherit the outstanding properties of the realized volatility estimates.

Finally, the range-based correlation is downward biased, although, by construction, the covariance in the numerator is less downward biased than the product of volatilities in the denominator (the correlation evaluated at the average covariance and volatilities with $m = 1{,}440$ is 0.4336). The source of this bias is the sampling variation of the covariance and volatility estimates through Jensen's inequality. Because the sampling variation does not vanish as $m \to \infty$, the range-based correlation estimator remains downward biased even in the limit. The realized correlation does not suffer from this bias.

Bid-ask bounce is a well-known reality of financial market data. To examine its effect on the covariance and correlation estimates, we augmented the Monte Carlo experiment with a simple model of bid-ask bounce and price discreteness taken from Hasbrouck (1999). Specifically, we took the dollar rates from the original experiment as the true prices and computed the bid and ask quotes $B_{i,t} = \text{floor}[P_{i,t} - 1/2 \text{ spread}, \text{ tick}]$ and $A_{i,t} = \text{ceiling}[P_{i,t} + 1/2 \text{ spread}, \text{ tick}]$ where floor[x, tick] and ceiling[x, tick] are

4. Since p_1 and p_2 follow the same stochastic process, we analyze the volatility estimates for only p_1.

functions that round x down or up to the nearest tick, respectively. For the cross rate, we computed the bid and ask quotes by imposing no arbitrage given the bid and ask quotes of the dollar rates. We then took the observed prices as $P_{i,t}^{\text{obs}} = q_{i,t}B_{i,t} + (1 - q_{i,t})A_{i,t}$, where $q_{i,t} \sim$ Bernoulli [1/2]. To capture the fact that the two base currencies are denominated in dollars, which means that the sale or purchase of the dollar might involve a simultaneous purchase or sale of the two currencies, we allow the buy-sell indicators $q_{1,t}$ and $q_{2,t}$ to be correlated with $\text{Corr}[q_{1,t}, q_{2,t}] = \eta$. The indicator $q_{3,t}$ is independent.

Table 2 presents the results for a bid-ask spread of 0.0005 and a tick size of 0.0001, which are realistic values for currencies (see Hasbrouck 1999). In panel A, the correlation η is set to 0, and in panel B, the correlation is 0.5. The effect of bid-ask bounce on the range-based estimates is relatively minor. In contrast, the effect on the realized volatilities, covariance, and correlation is striking. Consistent with the intuition outlined previously, the realized volatilities are upward biased when the data is sampled more frequently than once every 3 hours. By the time the data is sampled every minute, the bias inflates the true volatility by almost 100% (an average estimate of 29.7% as opposed to a true volatility of 15%). The results for the realized covariance depend on whether we construct the estimator using the no-arbitrage condition or return cross products and on the correlation of the bid-ask indicators. If we use the no-arbitrage condition, the realized covariance inherits the biases of the realized volatilities, to the point where, for 5-minute sampling, the average estimate is negative. In contrast, if we use return cross products and if the bid-ask indicators are independent (in panel A), the realized covariance is unbiased. The reason is that, if the bid-ask indicators are independent, then the expectation of the product of observed returns is equal to the expectation of the product of true returns. The bid-ask bounce, therefore, increases the variability of only the estimator. However, if the bid-ask indicators are correlated (in panel B), this argument no longer holds and the realized covariance is severely positively biased because each cross product of returns contains an upward bias due to the common component of the bid-ask indicators. Finally, the realized correlation, computed from the biased realized volatilities and biased covariance, is unreliable, ranging from -0.89 to 0.66.

Finally, asynchronous trading is another market microstructure effect that is likely to affect differently the range-based and realized covariance and correlation estimates. With infrequent trading, a security has a latent true price that is revealed only when a trade occurs. Between trades, the observed price is stale at the last traded price and therefore does not reflect the true price. In a univariate setting, infrequent trading induces positive serial correlation in the intraday returns, which in turn causes a downward bias in the realized volatility. In a bivariate setting, asynchronous infrequent trading, when the trades for the two assets do not

TABLE 2 Range-Based and Realized Estimates with Bid-Ask Bounce

	SD			Covariance			Correlation		
Sampling Frequency	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
Panel A. Independent Bid-Ask Bounce with $\eta = 0$									
Range-Based Estimates									
1 minute	14.512	4.278	4.306	.826	1.121	1.124	.327	.344	.352
5 minutes	14.006	4.274	4.388	.779	1.087	1.093	.327	.357	.365
10 minutes	13.671	4.272	4.474	.754	1.063	1.073	.331	.366	.373
20 minutes	13.228	4.263	4.617	.721	1.032	1.047	.335	.378	.384
40 minutes	12.622	4.256	4.875	.672	.989	1.015	.339	.397	.402
1 hour 20 minutes	11.767	4.236	5.328	.600	.928	.975	.340	.428	.432
Realized Estimates with No-Arbitrage Condition									
1 minutes	29.645	.490	14.653	−5.578	.462	6.495	−.636	.060	1.037
5 minutes	18.849	.760	3.924	−.395	.341	1.339	−.114	.100	.524
10 minutes	17.010	.990	2.241	.253	.351	.736	.082	.118	.339
20 minutes	15.994	1.327	1.658	.578	.396	.511	.217	.141	.231
40 minutes	15.422	1.820	1.868	.736	.486	.513	.296	.177	.206
1 hour 20 minutes	15.058	2.515	2.516	.815	.629	.635	.335	.232	.241
Realized Estimates with Cross Products									
1 minutes	29.645	.490	14.653	.900	.263	.263	.102	.030	.299
5 minutes	18.849	.760	3.924	.900	.223	.223	.253	.057	.158
10 minutes	17.010	.990	2.241	.901	.256	.256	.309	.076	.119
20 minutes	15.994	1.327	1.658	.901	322	.322	.347	.104	.117
40 minutes	15.422	1.820	1.868	.898	.429	.429	.368	.145	.149
1 hour 20 minutes	15.058	2.515	2.516	.896	.588	.588	.376	.205	.207
Panel B. Correlated Bid-Ask Bounce with $\eta = .5$									
Range-Based Estimates									
1 minutes	14.512	4.278	4.306	.810	1.123	1.127	.318	.347	.356
5 minutes	14.006	4.274	4.388	.764	1.089	1.097	.319	.360	.369
10 minutes	13.671	4.272	4.474	.741	1.065	1.077	.323	.369	.377
20 minutes	13.228	4.263	4.617	.710	1.033	1.051	.329	.380	.387
40 minutes	12.622	4.256	4.875	.664	.990	1.018	.333	.399	.405
1 hour 20 minutes	11.767	4.236	5.328	.595	.929	.978	.335	.430	.435
Realized Estimates with No-Arbitrage Condition									
1 minutes	29.645	.490	14.653	−7.812	.548	8.729	−.890	.073	1.292
5 minutes	18.849	.760	3.924	−.842	.375	1.782	−.241	.114	.651
10 minutes	17.010	.990	2.241	.029	.375	.949	.004	.130	.417
20 minutes	15.994	1.327	1.658	.465	.414	.601	.172	.151	.273
40 minutes	15.422	1.820	1.868	.679	.500	.547	.270	.185	.226
1 hour 20 minutes	15.058	2.515	2.516	.786	.640	.650	.321	.238	.251

TABLE 2 (*Continued*)

Sampling Frequency	SD			Covariance			Correlation		
	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
Realized Estimates with Cross Products									
1 minutes	29.645	.490	14.653	4.140	.265	3.251	.471	.024	.075
5 minutes	18.849	.760	3.924	1.549	.230	.688	.435	.050	.061
10 minutes	17.010	.990	2.241	1.225	.262	.418	.421	.070	.073
20 minutes	15.994	1.327	1.658	1.062	.328	.366	.410	.099	.099
40 minutes	15.422	1.820	1.868	.979	.433	.440	.401	.141	.141
1 hour 20 minutes	15.058	2.515	2.516	.937	.591	.592	.393	.202	.202

NOTE.—We simulated two currency prices as described in table 1 then computed the bid and ask quotes $B_{i,t} = \text{floor}[P_{i,t} - 1/2 \text{ spread, tick}]$ and $A_{i,t} = \text{ceilingr}[P_{i,t} + 1/2 \text{ spread, tick}]$, where floor[$x$, tick] and ceiling[$x$, tick] are functions that round x down or up to the nearest tick, respectively. The spread is set to 0.0005 and the tick size is 0.0001. For the cross currency, we computed the bid and ask quotes by imposing no-arbitrage, given the bid and ask quotes of the base currencies. We then took the observed prices as $P_{i,t}^{\text{obs}} = q_{i,t}B_{i,t} + (1 - q_{i,t})A_{i,t}$, where $q_{i,t} \sim$ Bernoulli [1/2]. The buy-sell indicators $q_{1,t}$ and $q_{2,t}$ are correlated with $\text{Corr}[q_{1,t}q_{2,t}] = \eta$, but the indicator $q_{3,t}$ is independent. In panel A, $\eta = 0$, and in panel B, $\eta = 0.5$. We used this observed data to compute the daily range and intraday returns then constructed three estimates of the volatilities, covariance, and correlation. We constructed range-based covariance estimates using Parkinson's variance estimator and $\text{Cov}[\Delta p_1, \Delta p_2] = 1/2(\text{Var}[\Delta p_1] + \text{Var}[\Delta p_2] - \text{Var}[\Delta p_3])$. We constructed realized covariance estimates using either the realized variance estimator and the same expression for the covariance or using the cross products of intraday returns. We repeated this procedure 10,000 times and report the means, standard deviations, and root mean squared errors.

take place at the same time, also creates a misalignment of the return cross products that may lead to a downward bias of the realized covariance.

To capture the effect of asynchronous infrequent trading in our Monte Carlo experiment, we used the discretization (13) with $m = 17{,}280$ (one observation per second) to simulate the latent "true" price processes. Then, for each process, we assigned n trade times randomly throughout the day and constructed stale price processes for which the price is equal to the price at the previous trade time until it is reset to the latent true price at the next trade time. Hence, the true prices look like continuous diffusions, while the stale prices look like discrete steps that occur at different times for the different currencies. Finally, we sampled these stale price processes at a regular frequency m, ranging again from 4 to 1,440, and proceeded just as in table 1 (i.e., there is no bid-ask bounce in this experiment).

We present the results for $n = 1{,}440$ (an average of one trade every minute) in table 3. The range-based estimates are slightly downward biased, because the infrequent trading magnifies the discretization bias. The realized volatilities are slightly downward biased due to the positive serial correlation induced by infrequent trading. Finally, when we computed the realized covariance and correlation using the no-arbitrage condition, the estimates inherited only the slight bias from infrequent trading, but when we instead used return cross products, the estimates were severely downward biased. In particular, the average realized covariance and correlation computed with return cross products are close to 0 in both panels. This extreme bias is due to the asynchronous price revelation.

TABLE 3 Range-Based and Realized Estimates with Asynchronous Trading

	SD			Covariance			Correlation		
Sampling Frequency	Mean	SD	RMSE	Mean	SD	RMSE	Mean	SD	RMSE
Range-Based Estimates									
1 minutes	14.037	4.333	4.436	.894	1.115	1.115	.382	.341	.341
5 minutes	13.743	4.335	4.511	.858	1.098	1.098	.379	.350	.351
10 minutes	13.472	4.315	4.575	.826	1.078	1.079	.377	.359	.359
20 minutes	13.096	4.308	4.708	.789	1.050	1.055	.377	.367	.368
40 minutes	12.531	4.280	4.939	.730	1.017	1.030	.374	.391	.392
1 hour 20 minutes	11.674	4.250	5.395	.646	.958	.991	.368	.428	.429
Realized Estimates with No-Arbitrage Condition									
1 minutes	14.989	.405	.405	.898	.097	.097	.399	.034	.034
5 minutes	14.952	.678	.680	.888	.163	.163	.395	.057	.057
10 minutes	14.940	.932	.934	.891	.222	.222	.396	.079	.079
20 minutes	14.923	1.292	1.294	.882	.315	.315	.390	.113	.113
40 minutes	14.898	1.773	1.775	.898	.433	.433	.394	.158	.158
1 hour 20 minutes	14.836	2.463	2.468	.913	.591	.591	.391	.220	.220
Realized Estimates with Cross Products									
1 minutes	14.989	.405	.405	.096	.096	.810	.043	.042	.360
5 minutes	14.952	.678	.680	.266	.212	.668	.119	.094	.296
10 minutes	14.940	.932	.934	.387	.257	.574	.173	.112	.253
20 minutes	14.923	1.292	1.294	.545	.317	.476	.243	.133	.206
40 minutes	14.898	1.773	1.775	.700	.402	.449	.310	.161	.185
1 hour 20 minutes	14.836	2.463	2.468	.824	.558	.563	.356	.210	.214

NOTE.—We simulated 24 hours worth of $m = 17{,}280$ regularly spaced intraday log prices (one price every second) for three currencies, as described in table 1. We then assigned to each log price process $n = 1{,}440$ trade times (an average of one trade every minute) randomly throughout the day and constructed stale price processes for which the price is equal to the price at the previous trade time until it is reset to the latent true price at the next trade time. We next sampled these stale price processes at a regular frequency m ranging from 18 (one observation every hour and 20 minutes) to 1,440 (one observation every minute). We used this observed data to compute the daily range and intraday returns and then construct three estimates of the volatilities, covariance, and correlation. We constructed range-based covariance estimates using Parkinson's variance estimator and $\mathrm{Cov}[\Delta p_1, \Delta p_2] = 1/2(\mathrm{Var}[\Delta p_1] + \mathrm{Var}[\Delta p_2] - \mathrm{Var}[\Delta p_3])$. We constructed realized covariance estimates using either the realized variance estimator and the same expression for the covariance or the cross products of intraday returns. We repeated this procedure 10,000 times and report the means, standard deviations, and root mean squared errors.

V. Conclusion

We extended the important idea of range-based volatility estimation to the multivariate case. In particular, we proposed a range-based covariance estimator motivated by financial economic considerations (the absence of arbitrage), in addition to statistical considerations. We showed that, unlike other univariate and multivariate volatility estimators, the range-based estimator is highly efficient yet robust to market microstructure noise arising from bid-ask bounce and asynchronous trading. Many extensions and applications of the ideas developed here are possible, and Brunetti and Lildolt (2002) take up several.

An intriguing application, which to the best of our knowledge has not yet been explored, involves constructing range-based volatility and covariance bets via a portfolio of lookback options. The payoff of a lookback straddle (a lookback call plus a lookback put) is equal to the range of the underlying asset over the life of the option. Therefore, lookback straddles are ideal for placing bets on the range-based volatility of an asset: their payoffs are high (low) when volatility as measured by the range is high (low). Our no-arbitrage approach to covariance estimation suggests an analogous way of placing bets on the covariance between two assets. Consider a portfolio of a long $A/\$$ lookback straddle, a long $B/\$$ lookback straddle, and a short A/B lookback straddle. Since each of the straddles is a variance bet, the payoffs of this portfolio are high (low) when covariance between the two dollar rates is high (low) over the life of the option.

References

Alizadeh, Sassan, Michael W. Brandt, and Francis X. Diebold. 2002. Range-based estimation of stochastic volatility models. *Journal of Finance* 57 (June):1047–92.

Andersen, Torben G., and Tim Bollerslev. 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39 (November):885–905.

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold. 2003. Parametric and nonparametric volatility measurement. In *Handbook of financial econometrics,* ed. L. P. Hansen and Y. Ait-Sahalia. Amsterdam: North-Holland, forthcoming.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold,and Paul Labys. 2003. Modeling and forecasting realized volatility. *Econometrica* 71 (March):579–626.

Ball, Clifford A., and Walter N. Torous. 1984. The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing, *Journal of Business* 57 (January):97–112.

Beckers, Stan. 1983. Variance of security price returns based on high, low, and closing prices. *Journal of Business* 56 (January):97–112.

Brunetti, Celso, and Peter Lildolt. 2002. Return-based and range-based (co)variance estimation, with an application to foreign exchange markets. Manuscript.

Garman, Mark B., and Michael J. Klass. 1980. On the estimation of price volatility from historical data. *Journal of Business* 53 (January):67–78.

Gill, Philip E., Walter Murray, and Margaret H. Wright, 1981. *Practical optimization*. New York: Academic Press.

Hasbrouck, Joel. 1999. Security bid/ask dynamics with discreteness and clustering: Simple strategies for modeling and estimation. *Journal of Financial Markets* 2:1–28.

Hull, John, and Alan White. 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42 (June):281–300.

Kunitomo, Naoto. 1992. Improving the Parkinson methods of estimating security price volatilities. *Journal of Business* 65 (April):295–302.

Ledoit, Olivier, and Michael W. Wolf. 2001. A well-conditioned estimator for large-dimensional covariance matrices. Working paper, Anderson School of Business, UCLA.

Parkinson, Michael. 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business* 53 (January):61–65.

Rogers, L. Christopher G., and Stephen E. Satchell. 1991. Estimating variance from high, low, and closing prices. *Annals of Applied Probability* 1 (November):504–12.

Schnabel, Robert B., and Elizabeth Eskow. 1999. A revised modified Cholesky factorization algorithm. *SIAM Journal of Optimization* 9 (September):1135–48.

Yang, Dennis, and Qiang Zhang. 2000. Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business* 73 (July):477–91.