Correlation Estimation Using Components of Japanese Candlesticks

V.Popov*†

†School of Mathematics and Statistics, University of St Andrews, KY16 9LZ St Andrews, UK

(*12 February 2016*)

Using the wick's difference from the classical Japanese candlestick representation of daily open, high, low, close prices brings efficiency when estimating the correlation in a bivariate Brownian motion. An interpretation of the correlation estimator in Rogers and Zhou (2008) in the light of wicks' difference allows us to suggest modifications, which lead to an increased efficiency and robustness against the baseline model. An empirical study on four major financial markets confirms the advantages of the modified estimator.

Keywords: Japanese Candlesticks; Correlation; Estimation; Brownian Motion; Jump Diffusions

JEL Classification: JEL C49, JEL G10

* Email: vmp@st-andrews.ac.uk. Tel +44 1334 46 1807

1.   Introduction

Under the assumption that the log-price process of a financial asset follows a Brownian motion, it is well-known that using the widely accessible information on daily high and low prices, additional to opening and closing prices, increases the efficiency of volatility estimation dramatically. Major references include Parkinson's development of a simple range-based volatility estimator (Parkinson (1980)), the derivation of an optimal volatility estimator among certain functionals of opening, high, low, and closing (OHLC) prices by Garman and Klass (1980), Rogers and Satchell (1991), correcting the downside bias due to discrete observation of the continuous price process, and Yang and Zhang (2000), extending the initial model framework of range-based volatility estimation from a driftless Brownian motion to the case of a non-zero drift rate. One drawback of the developed highly efficient OHLC-based volatility estimators, however, is their sensitivity with respect to violations of the Brownian motion assumption, which, in turn, allows their application for testing the specification of a Brownian motion, see Becker et al. (2007).

Until recently, the focus of using OHLC price data was on univariate volatility analysis. Using the argument that no arbitrage should be possible, Brandt and Diebold (2006) suggested a strikingly simple way to exploit the efficiency of OHLC-based variance estimators for covariance estimation by considering a two asset portfolio and expressing the covariance of the two-asset returns in terms of the portfolio return variance, additional to the two individual asset return variances. While the absence of triangular arbitrage in foreign exchange markets provides an immediate application, however, for stock returns the high and low prices of two-asset portfolios typically are not available, unless high frequency data is used, see Brunetti and Lildholdt (2008). A promising new strand of research on correlation estimation using OHLC price data has been initiated by an important result of Rogers and Shepp (2006), which gives the product moment of the extreme values of two correlated Brownian motions with correlation ρ as a function of the latter. Based on this result, Rogers and Zhou (2008) derive a correlation (covariance) estimator from OHLC data under the assumption of two correlated standard Brownian motions. In a first step, an estimator is constructed as a linear combination of the nine product moments between the open-high, open-low and open-close returns of two Brownian motions, which is unbiased if the two processes are either uncorrelated or perfectly correlated, and has minimum variance, when the processes are uncorrelated. In a second step, using the result by Rogers and Shepp, it is corrected for its bias in case of non-zero and non-perfect correlation.

In the present paper, we raise awareness of the concept of the Balanced Excess Return (BER), which is the difference (in absolute values) of the upper and lower wick in the Japanese-candlestick representation of OHLC prices, and use it to improve OHLC-based correlation estimators. Under the framework of a (standard) Brownian Motion we derive the univariate distribution of the BER. In the bivariate case, we calculate its correlation ρ_W as a function of the process correlation ρ using the Rogers-Shepp result. In the light of the newly introduced concept we interpret the Rogers-Zhou estimator as a weighted average of two estimators - one based on the close returns and a second one based on the BERs. The interpretation allows us to explain some of the shortcomings of the Rogers-Zhou estimator and suggest small but influential modifications. In a simulation study we study the properties unbiasedness and efficiency of the suggested estimators under the framework of the baseline model. We show that the modified Rogers-Zhou estimator is nearly unbiased and efficient even compared to the Brandt-Diebold estimator provided it uses the same amount of high-frequency data. In a next step we study the robustness of the estimators in the presence of jumps and excess kurtosis as in jump diffusion models. We back up our findings using real data from four major financial markets.

This paper is built as follows. In section 2 we define the BER and study its theoretical properties under the framework of a Brownian Motion. In section 3 we revise and improve correlation estimators based on the BERs. Section 4 contains a simulation study on the BER-based estimators, while an empirical study is provided in section 5. A summary of our results is given in section 6.

2.   Balanced Excess Returns under the Brownian Motion framework

Consider a log-price process. Assume that without loss of generality the opening price is normalized to one, i.e., the log-price process starts with zero at market opening. Let H be the highest, L be the lowest and C be the close log-price of this process. Due to the normalization of the opening price, H, L and C can equivalently be understood as the daily highest, lowest and close (log-)returns, respectively, and this is the way we will refer to these variables in the remaining part of the paper. In terms of the Japanese-candlestick representation of the OHLC prices, C is the body of the candlestick. Let

$$w_u := H - \max(0, C) \qquad \text{and} \qquad w_l := -(L - \min(0, C))$$

be the upper and, respectively, the lower wick of the candlestick. The difference of the wicks is given as

$$W := w_u - w_l = H - \max(0, C) + L - \min(0, C) = H + L - C$$

In the following we label the difference of the wicks as the Balanced Excess Return (BER), because it can be interpreted as the balance between the maximal possible profit $(H - \max(0, C))$ and the maximal possible loss $(L - \min(0, C))$ beyond the open-close band. Figure 1 illustrates the difference of the wicks along with the body of the candlestick and the range for a (discretized) Brownian Motion starting at zero.
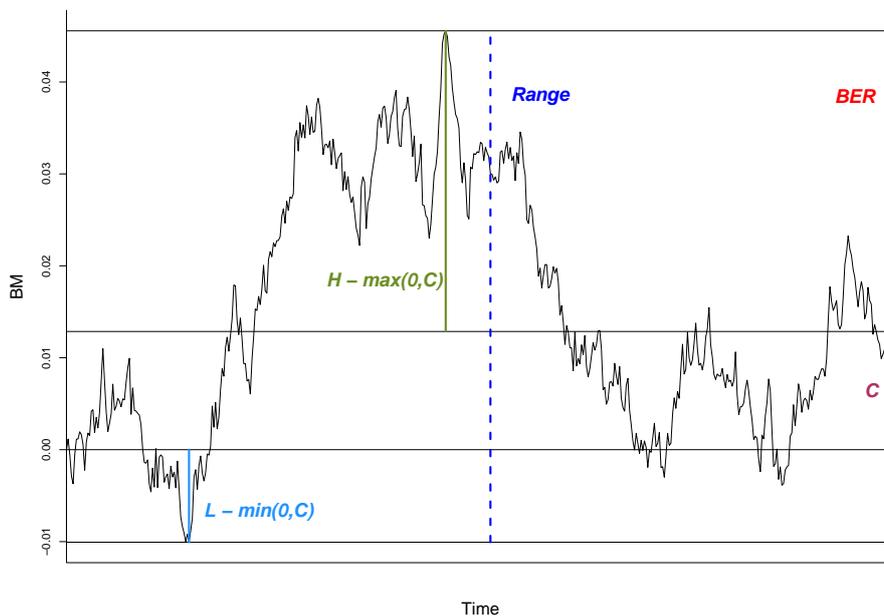

Figure 1. Candlestick body (C), range and wicks' difference (BER) of a log-price process following a Brownian Motion starting at 0.

Under a (univariate) standard Brownian Motion (BM), the distribution of the BER can be derived as (see Appendix A for more details):

$$\frac{\partial P(W < u)}{\partial u} = \sum_{k=1}^{\infty} \frac{16k^2}{(4k^2 - 1)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2ku)^2}{2}} = \sum_{k=1}^{\infty} \frac{8k}{(4k^2 - 1)^2} \phi_k(u), \tag{1}$$

where $\phi_k(u)$ is the density function of a normally distributed random variable with zero mean and variance $1/(2k)^2$. Consequently, the density of the BER is a weighted average (with weights summing to one) of scaled densities of normally distributed random variables. However, the BER is not normally distributed.

Using the density in (1) the first two moments of the BER can be easily obtained:

$$E(W) = 0$$

$$E(W^2) = \sum_{k=1}^{\infty} \frac{2}{k(4k^2 - 1)^2} = 3 - 4\ln 2$$

Alternatively one can use the Vasicek's generating function of high, low and close given in Appendix C, Garman and Klass (1980). By means of simple scaling one obtains the mean and the variance of the BER in the more general case when the volatility parameter σ, $\sigma > 0$, is not restricted to one:

$$E(W) = 0 \qquad \text{and} \qquad V(W) = \big(3 - 4\ln(2)\big)\sigma^2 \approx 0.22741\sigma^2 \tag{2}$$

Note that W has a much smaller variance than C. Furthermore, Klößner (2007) shows that for all Lévy processes (BM being a special case) it holds:

$$(W, C) \overset{\mathcal{D}}{=} (-W, C) \tag{3}$$

Klößner actually considers a version, in which the random variables are standardized by the range - see (15) on p.5 in Klößner (2007). The author first shows that (3) holds for intraday time-reversible processes and subsequently proves that Lévy processes are intraday time-reversible. Note that the underlying results for the proof of (3) can also be found in the proof of Lemma 1 in Klößner (2010). From (3) follows that the BER is symmetric around zero and has a zero mean independent of the value of the drift parameter under much more general conditions. An overview of moments and correlations between the BER, C and the range, denoted with l, and their squares is given in Table 1.

| | Mean/σ | Variance/σ^2 | Correlation | | |
			C	l	W
C	0	1	1	0	0
l	1.5958	0.2261	0	1	0
W	0	0.2274	0	0	1
	Mean/σ^2	Variance/σ^4	C^2	l^2	W^2
C^2	1	2	1	0.8406	-0.2156
l^2	2.7726	3.1316	0.8406	1	0.064
W^2	0.2274	0.1155	-0.2156	0.064	1

Table 1. Moments and correlations of (pairs of) C, l, W and their squares

Several results are worth pointing out:

- C, l, W are pairwise uncorrelated, i.e.:

$$Corr(C, l) = Corr(C, W) = Corr(l, W) = 0 \tag{4}$$

4

This property can be used in empirical studies to pin down deviations from the BM assumption. Note that the zero correlation between C and W holds for all Lévy processes - a result, which follows from (3).

- C, l, W are not independent (since their squares are correlated).
- C^2 and W^2 are negatively correlated, which implies negative dependence between them.

Next we consider a bivariate driftless BM. Although the bivariate distribution of the BER has not been derived yet, a recent result in Rogers and Shepp (2006) has made it possible to calculate the mean of the cross product of the two BERs. For notational purposes for the rest of this section we introduce an index i to denote the marginal log-price process ($i = 1, 2$).

Rogers and Shepp derive the mean of the product of the maxima in a bivariate BM as a function of the process correlation ρ $(-1 < \rho < 1)$:

$$E(H_1 H_2) = \sigma_1 \sigma_2 c^*(\rho), \qquad \text{where}$$

$$c^*(\rho) := \cos \alpha \int_0^\infty \frac{\cosh \nu \alpha}{\sinh \nu \pi / 2} \tanh \nu \gamma \mathrm{d}\nu \qquad (5)$$

with $\rho = \sin \alpha, \alpha \in (-\pi/2, \pi/2)$ and $2\gamma = \alpha + \pi/2$. For the values -0.5, 0 and 0.5 one obtains explicit result for the function $c^*(\rho)$. The results are summarized in Table 2 where we also include the well-known values of $E(H_1 H_2)/\sigma_1 \sigma_2$ at the boundary points of the parametric space of ρ, for which $c^*(\rho)$ is not defined.

ρ	-1	-0.5	0	0.5	1
$E(H_1 H_2)/\sigma_1 \sigma_2$	2ln2-1	0.5	$2/\pi$	$0.75\sqrt{3} - 0.5$	1

Table 2.   Exact values of $E(H_1 H_2)/\sigma_1 \sigma_2$

For values of ρ other than the ones listed above, $c^*(\rho)$ needs to be computed numerically.

Using Rogers and Shepp's result it can be shown that for the correlation between W_1 and W_2 (denoted by ρ_W) it holds:

$$\rho_W = \frac{E(W_1 W_2)}{\sigma_1 \sigma_2 (3 - 4\ln 2)} = \frac{2(c^*(\rho) - c^*(-\rho)) - \rho}{3 - 4\ln 2} =: g(\rho) \qquad (6)$$

Consequently the correlation of the BERs is a function of ρ.

In particular Figure 2 shows that ρ_W is an odd function for ρ and that it is invertible. Obviously, the nonlinearity of the dependence between ρ_W and ρ is very mild following the mild nonlinearity of the dependence between $c^*(\rho)$ and ρ as noted in Rogers and Shepp (2006). The largest absolute difference occurs in the vicinity of $\rho = \pm 0.64$. Moreover, it holds $|\rho_W| \leq |\rho|$, while equality holds only for the cases $\rho = -1, 0$ or 1.

Having outlined important properties of the BERs under the BM framework, we discuss their use for correlation estimation in the next section.

3.   Correlation estimation based on the BERs

The BERs were (implicitly) used for correlation estimation in Rogers and Zhou (2008) for the first time. Under the assumption of a *standard* bivariate BM, the authors derive the following correlation estimator as a solution to an optimization problem based on the nine cross-functionals of highs, lows and close returns:

$$\hat{\rho}_0 = \frac{1}{2} C_1 C_2 + \frac{1}{2} \cdot \frac{(H_1 + L_1 - C_1)(H_2 + L_2 - C_2)}{3 - 4\ln 2} = \frac{1}{2} C_1 C_2 + \frac{1}{2} \cdot \frac{W_1 W_2}{3 - 4\ln 2} \qquad (7)$$
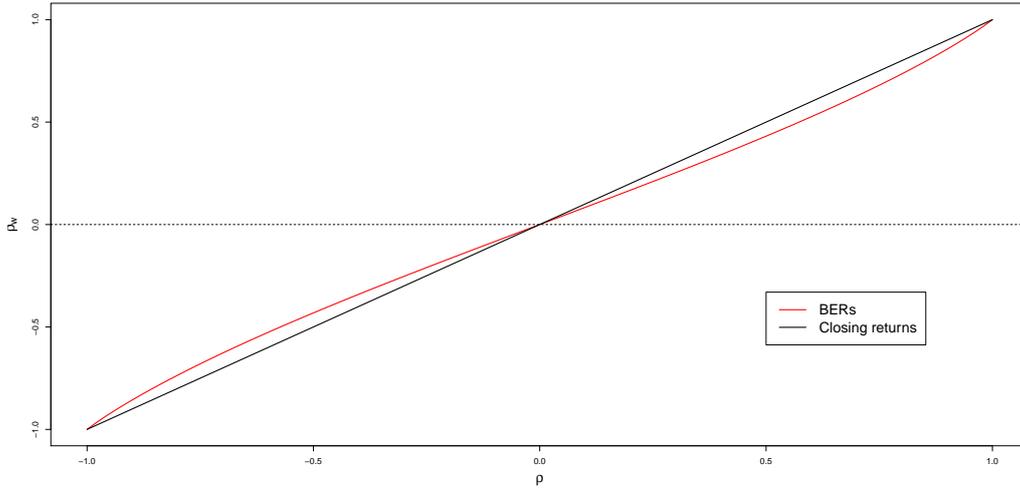
Figure 2. Correlations of the BERs and the close returns as functions of ρ

The estimator has a minimum variance $(1/2)$ for $\rho = 0$ and is unbiased for $\rho = -1, 0$ and 1. However, it is generally biased. To cope with this problem Rogers and Zhou suggest a multi-period estimator

$$\hat{\rho}_{RZ} := \phi^{-1}(\bar{\hat{\rho}}_0) \tag{8}$$

with

$$\bar{\hat{\rho}}_0 := \frac{1}{N} \sum_{j=1}^{N} \hat{\rho}_{0;j}, \tag{9}$$

where $\hat{\rho}_{0;j}$ is the estimator in (7) calculated for day j, $j = 1, ..., N$, $N \in \mathbb{N}$, and $\phi(\rho)$ is the mean of $\hat{\rho}_0$ as a function of ρ:

$$\phi(\rho) := E(\hat{\rho}_0) = \frac{1}{2}\rho + \frac{1}{2} \cdot \frac{2(c^*(\rho) - c^*(-\rho)) - \rho}{3 - 4\ln 2}$$

with $c^*(\cdot)$ given in (5). Note that $\phi^{-1}(\cdot)$ is used as a correction function for the bias of the estimator $\bar{\hat{\rho}}_0$.

The authors provide no interpretation of $\hat{\rho}_0$ other than the output of the formulated optimization problem. However, introducing the concept of BER allows us to gain valuable insights in this estimator, which in turn provide ideas for an improvement. Before we address this issue, we set the stage by defining a r_0-type of correlation estimators.

DEFINITION 1   Let (X, Y) be a two-dimensional random vector. Assume that the first two (central) moments of X and Y, say μ_X, μ_Y and respectively σ_X^2, σ_Y^2, exist and are known. Given a sample $(X_1, Y_1), (X_2, Y_2), ..., (X_N, Y_N)$ of length N, $N > 2$, the estimator of the correlation between X and Y given as

$$\frac{\frac{1}{N} \sum_{j=1}^{N} (X_j - \mu_X)(Y_j - \mu_Y)}{\sqrt{\sigma_X^2}\sqrt{\sigma_Y^2}}$$

6

is labeled an r_0-type correlation estimator.

Comparing this estimator with the classical Pearson correlation given as:

$$\frac{\frac{1}{N}\sum_{j=1}^{N}(X_j-\bar{X})(Y_j-\bar{Y})}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}(X_j-\bar{X})^2}\sqrt{\frac{1}{N}\sum_{j=1}^{N}(Y_j-\bar{Y})^2}} \quad \text{with} \quad \bar{X}=\frac{1}{N}\sum_{j=1}^{N}X_j \quad \text{and} \quad \bar{Y}=\frac{1}{N}\sum_{j=1}^{N}Y_j,$$

we observe that an r_0-type estimator makes use of the known means and variances by substituting them for their respective sample equivalents in the classical Pearson correlation estimator.
In terms of the Rogers-Zhou estimator we provide the following interpretation: $\bar{\hat{\rho}}_0$ (and respectively $\hat{\rho}_0$, which is a special case of $\bar{\hat{\rho}}_0$ when $N=1$) can be considered as an *equally weighted average of two r_0-type correlation estimators*:

$$\bar{\hat{\rho}}_0 = \frac{1}{2}\left(\frac{1}{N}\sum_{j=1}^{N}C_{1,j}C_{2,j} + \frac{1}{N}\sum_{j=1}^{N}\frac{W_{1,j}W_{2,j}}{3-4\ln2}\right) =: \frac{1}{2}(\bar{\hat{\rho}}_{0;C}+\bar{\hat{\rho}}_{0;W}) \tag{10}$$

The first one, $\bar{\hat{\rho}}_{0;C}$, is based only on the candlestick bodies and makes full use of the known zero means and unit variances. It is an unbiased estimator for the correlation parameter ρ. The second estimator, $\bar{\hat{\rho}}_{0;W}$ is an unbiased estimator for the correlation of the wicks' differences ρ_W (see (6)), which also makes full use of their known means and variances (see (2)). Because ρ and ρ_W differ except for the cases $\rho=-1,0$ and 1, the estimator $\bar{\hat{\rho}}_0$ is an average of unbiased estimators of two *different* correlations and consequently it is biased for either of them. This explains the need for the bias correction given in (8) when estimating ρ.

In the light of this interpretation we can identify and explain several problems plaguing $\bar{\hat{\rho}}_0$, which lies in the core of the Rogers-Zhou estimator.

- $\bar{\hat{\rho}}_0$ is drift-dependent

Rogers and Zhou acknowledge this problem but provide no solution to it. The reason for the drift-dependency is the fact the two r_0-type estimators assume zero mean of the close returns and, respectively, for the BERs. Whereas, as pointed out earlier, for the latter the zero-mean assumption is robust, this is not the case for the close returns.

- $\bar{\hat{\rho}}_0$ is generally not restricted to the admissible range [-1,1]

While the Cauchy-Bunyakoksky-Schwarz inequality restricts the Pearson correlation estimator to [-1,1], this is not true for r_0-type estimators. Consequently, especially for large absolute values of ρ and small N, $\bar{\hat{\rho}}_0$ can take values outside the admissible range. In such cases Rogers-Zhou estimator in $\hat{\rho}_{RZ}$ given in (8) is not defined. Applying a truncation on $\bar{\hat{\rho}}_0$ will introduce a downside bias.

- r_0-type estimators are generally not asymptotically efficient

To make this point clear, we will consider the case of correlation estimation in the bivariate normal model when the means and variances are known. This case is relevant for the Rogers-Zhou estimator since the close returns are normally distributed (although, as shown earlier, the BERs are not). Let $(X_1,Y_1),(X_2,Y_2),...,(X_N,Y_N)$ be a sample of length N, $N>2$, from a bivariate normal distribution with known means and variances and unknown correlation ρ. Let $\hat{\rho}_0^G$ be a r_0-type estimator and $\hat{\rho}_S^G$ be the classical Pearson estimator using sample variances to standardize the

covariance. We use the superscript G to indicate that the sample comes from a normal distribution, while the subscript S in the Pearson correlation estimator emphasizes on the fact that it relies on information from the sample for the means and the variances. Then the asymptotic distribution of the two estimators is:

$$\sqrt{N}(\hat{\rho}_0^G - \rho) \xrightarrow{\mathcal{D}} \mathrm{N}\left(0, 1 + \rho^2\right) \tag{11}$$

$$\sqrt{N}(\hat{\rho}_S^G - \rho) \xrightarrow{\mathcal{D}} \mathrm{N}\left(0, (1 - \rho^2)^2\right) \tag{12}$$

It is straightforward to show that $V(X_i Y_i) = 1 + \rho^2$ for all i and consequently $V(\hat{\rho}_0^G) = \dfrac{1 + \rho^2}{N}$. Applying the Central Limit Theorem then yields (11). See Cramér (1971) for (12). The results hold also for the case when the means are unknown and have to be estimated.

$\hat{\rho}_0^G$ is asymptotically inefficient compared to $\hat{\rho}_S^G$ since:

$$1 + \rho^2 \geq (1 - \rho^2)^2$$

with equality obtained for the case $\rho = 0$. This result can be viewed as a paradox since $\hat{\rho}_0^G$ uses more information compared to $\hat{\rho}_S^G$. A possible explanation is that the additional information of the known moments is not used in the correct way (as opposed, for example, to a maximum-likelihood estimator).

Another issue concerning the estimator in (8) is its complexity. In particular, $\phi^{-1}(\cdot)$ needs to be calculated numerically except for the cases listed in Table 2. This can be done using suitable statistical software but an approximation can simplify its practical application. First, note that since $\bar{\hat{\rho}}_{0;W}$ estimates the "wrong" correlation, a correction can be imposed only on it instead of correcting the combined estimator $\bar{\hat{\rho}}_0$. Consequently, we are interested in approximating $g^{-1}(\cdot)$ given in (6) rather than $\phi^{-1}(\cdot)$. The approximation should be an odd function in ρ_W, which is in line with Rogers and Zhou's observation that a quadratic approximation is not useful. In particular we seek a polynomial of power three such that:

$$\rho = \alpha_1 \rho_W^3 + \alpha_2 \rho_W^2 + \alpha_3 \rho_W + \alpha_4$$

The results in Table 2 imply the following relations:

ρ	-1	0	0.5	1
ρ_W	-1	0	$\frac{3\sqrt{3}-5}{6-8\ln 2}$	1

and thus the vector with the polynomial coefficients $\boldsymbol{\alpha}$ should solve the following system:

$$B\boldsymbol{\alpha} = z,$$

where (setting $\frac{3\sqrt{3}-5}{6-8\ln 2} =: k$)

$$B = \begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ k^3 & k^2 & k & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad z = \begin{pmatrix} -1 \\ 0 \\ 0.5 \\ 1 \end{pmatrix}$$

One easily obtains $\boldsymbol{\alpha} \approx (-0.1958, 0, 1.1958, 0)'$. Setting $\alpha^* = 1.1958$, the approximation can be summarized as follows:

$$\rho \approx \alpha^* \rho_W + (1 - \alpha^*)\rho_W^3 = 1.1958\rho_W - 0.1958\rho_W^3, \tag{13}$$

which is indeed an odd function of ρ_W.

To cope with the problems listed above we propose the following slight but influential modifications to the Rogers-Zhou estimator:

1) Substitute the inefficient r_0-type estimators with Pearson-type estimators using sample means and variances

2) Correct the estimator for ρ_W with the approximated correction function given in (13)

In that way we propose the following BER-based correlation estimator:

$$\hat{\rho}_J := \frac{1}{2}(\hat{\rho}_C + 1.1958\hat{\rho}_{\mu;W} - 0.1958\hat{\rho}_{\mu;W}^3), \qquad \text{where} \tag{14}$$

$$\hat{\rho}_C := \frac{\sum_{j=1}^N (C_{1,j} - \bar{C}_1)(C_{2,j} - \bar{C}_2)}{\sqrt{\sum_{j=1}^N (C_{1,j} - \bar{C}_1)^2 \sum_{j=1}^N (C_{2,j} - \bar{C}_2)^2}} \qquad \text{and}$$

$$\hat{\rho}_{\mu;W} := \frac{\sum_{j=1}^N W_{1,j} W_{2,j}}{\sqrt{\sum_{j=1}^N W_{1,j}^2 \sum_{j=1}^N W_{2,j}^2}} \tag{15}$$

with $\bar{C}_i = \frac{1}{N}\sum_{j=1}^N C_{i,j}$, $i = 1, 2$. The subscript μ in $\hat{\rho}_{\mu;W}$ indicates that the estimator uses the known (zero) means. There is no need to demean the BERs since their zero-mean property holds under the general class of Lévy processes.

Some of the asymptotic properties of the proposed estimator can be derived using already available results for the (multivariate) asymptotic distribution of the Pearson correlation coefficient. First note that the latter is consistent and asymptotically normal - see (Lehmann 1999, p.317) for example. Thus $\hat{\rho}_C$ and $\hat{\rho}_{\mu;W}$ are consistent for ρ and ρ_W, respectively. Due to the approximation, however, $\hat{\rho}_J$ is not consistent for ρ. In fact, it is consistent for ($\alpha = 1.1958$)

$$\rho_\alpha = \frac{1}{2}\left(\rho + \alpha\rho_W + (1 - \alpha)\rho_W^3\right),$$

which in general differs from ρ except for the cases stated in Table 2. This is not a major drawback. First, note that for small to medium sized finite samples, the MSE is dominated by the variance and the contribution of the approximation error is negligible. We note that the difference between the true value and the approximation, evaluated numerically on a fine grid, is not larger than 0.001554991 in absolute value. Second, a consistent estimator can be obtained using the true correction function. The estimator:

$$\hat{\rho}_{J;g} = \frac{1}{2}(\hat{\rho}_C + g^{-1}(\hat{\rho}_{\mu;W}))$$

with $g(\cdot)$ given in (6) is consistent for ρ. The result follows from Slutsky's theorem and the fact that $g^{-1}(\cdot)$ is continuous.

The asymptotic distribution of $\hat{\rho}_J$ can be obtained using a result in Neudecker and Wesselman (1990). According to Theorem 2 on p.597, the asymptotic distribution of the sample correlation matrix (i.e. the matrix containing the Pearson correlation estimators) of the vector (C_1, W_1, C_2, W_2) is normal. Note that the theorem only requires that the moments of the random vector up to the fourth order exist. Since the distribution of the vector $(\hat{\rho}_C, \hat{\rho}_{\mu;W})$ is marginal to the one of the sample correlation matrix, it follows that:

$$\sqrt{N}\left(\begin{array}{c} \hat{\rho}_C - \rho \\ \hat{\rho}_{\mu;W} - \rho_W \end{array}\right) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{\Sigma}),$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution and $\boldsymbol{\Sigma}$ is a certain variance-covariance matrix. A precise expression of $\boldsymbol{\Sigma}$ would require, among other results, moments of higher order of cross-products of BER. These have not been derived yet. For $\mathbf{x} := (x_1, x_2)'$ define

$$h(\mathbf{x}) = \frac{1}{2}\left(x_1 + \alpha x_2 + (1 - \alpha)x_2^3\right)$$

Applying the delta method then gives

$$\sqrt{N}(\hat{\rho}_J - \rho_\alpha) = \sqrt{N}\left[h\left((\hat{\rho}_C, \hat{\rho}_{\mu;W})'\right) - h\left((\rho, \rho_W)'\right)\right] \xrightarrow{\mathcal{D}} N(\mathbf{0}, \boldsymbol{\delta\Sigma\delta}')$$

with $\boldsymbol{\delta} = \frac{1}{2}(1, \alpha + 3(1-\alpha)\rho_W^2)$. Consequently, $\hat{\rho}_J$ is asymptotically normal. Using similar arguments, it can be shown that $\hat{\rho}_{J;g}$ is also asymptotically normal.

In the next section we provide insights on the properties unbiasedness, efficiency and robustness of $\hat{\rho}_J$ in comparison with other estimators discussed in the present paper. We note at this stage that provided there is no censoring, $\hat{\rho}_{RZ}$ is nearly unbiased because, as noted above, the non-linearity of the correction function is only mild. On the other hand, $\hat{\rho}_J$ has a small sample bias. The reason for that lies in the fact that $\hat{\rho}_J$ can be considered as a weighted average of two Pearson-type correlation estimators and it is well-known that the Pearson correlation estimator is biased of order $O(1/N)$ - see Cramér (1971) for example. Note that for small samples a bias correction for the Pearson correlation estimator exists - see Olkin and Pratt (1958). For reasonably sized samples, however, this bias is negligible. In terms of efficiency, there is little theoretical support on the variances of the estimators (except for $\hat{\rho}_0$ when $\rho = 0$) for reasons stated above. Moreover, it is important to know how sensitive are the estimators to the model assumption. Such problems are addressed in a simulation study described in the next section.

Before we close this section we add a comment on the estimator proposed by Rogers and Zhou. In financial applications the assumption of a standard BM (which implies that the drift and the volatility parameters are known) is rather unrealistic. Usually the drift and the volatility parameters need to be estimated from the sample data. Although Rogers and Zhou do not address the problem, a feasible version of their estimator is straightforward to derive using a volatility estimator based on the squared close returns to standardize $\bar{\bar{\rho}}_0$:

$$\hat{\rho}_{RZ}^a := \phi^{-1}\left(\frac{\bar{\bar{\rho}}_0}{\sqrt{\frac{1}{N}\sum_{j=1}^N C_{1,j}^2 \frac{1}{N}\sum_{j=1}^N C_{2,j}^2}}\right) = \tag{16}$$

$$= \phi^{-1}\left(\frac{1}{2}\left(\frac{\frac{1}{N}\sum_{j=1}^N C_{1,j}C_{2,j}}{\sqrt{\frac{1}{N}\sum_{j=1}^N C_{1,j}^2 \frac{1}{N}\sum_{j=1}^N C_{2,j}^2}} + \frac{\frac{1}{N}\sum_{j=1}^N W_{1,j}W_{2,j}}{\sqrt{\frac{1}{N}\sum_{j=1}^N C_{1,j}^2 \frac{1}{N}\sum_{j=1}^N C_{2,j}^2}(3 - 4\ln2)}\right)\right)$$

One cannot help but notice that the estimator for ρ_W employs a peculiar mix of theoretical $(3 - 4\ln2)$ and sample variances $(1/N\sum_{j=1}^N C_{i,j}^2)$ when estimating the variance of the BERs. Note that the variance of the BERs standardized by the estimated scaling parameter is only approximately $3 - 4\ln2$. Thus, even though $\hat{\rho}_{RZ}^a$ can be further adjusted to account for a non-zero drift, the core difference between the feasible Rogers-Zhou estimator and $\hat{\rho}_J$ lies in the estimation of the BER variance. For finite samples:

$$(3 - 4\ln2)\frac{1}{N}\sum_{j=1}^N C_{i,j}^2 \neq \frac{1}{N}\sum_{j=1}^N W_{i,j}^2, \qquad i = 1, 2$$

10

and thus the distribution of $\frac{\frac{1}{N}\sum_{j=1}^{N}W_{1,j}W_{2,j}}{\sqrt{\frac{1}{N}\sum_{j=1}^{N}C_{1,j}^2\frac{1}{N}\sum_{j=1}^{N}C_{2,j}^2}(3-4\ln2)}$ and $\hat{\rho}_{\mu;W}$ may differ substantially. We investigate the effect of this difference on the unbiasedness and the efficiency of the estimators in the next section.

4.   Simulation study

First we start by studying the properties unbiasedness and efficiency of the correlation estimators of interest under the framework of a standard BM. In a next step we study the performance of the estimators in the presence of jumps and excess kurtosis.

The competing estimators are:

- the classical estimator $\hat{\rho}_C$, which is based on the candlestick body and neither uses additional information on highs and lows, nor relies on known moments (serving as a benchmark). In this way the simulation study differs from the one in Rogers and Zhou (2008), where the inefficient $\bar{\hat{\rho}}_{0;C}$ - see (10) - is used as a benchmark.
- Rogers-Zhou's $\hat{\rho}_{RZ}$, which additionally incorporates information on the wick's difference and relies on known means and/or variances
- its feasible version $\hat{\rho}_{RZ}^a$
- the modified Rogers-Zhou estimator $\hat{\rho}_J$ proposed in the present paper and
- the Brand-Diebold correlation estimator proposed in Brandt and Diebold (2006)

The latter estimator is defined as:

$$\hat{\rho}_{BD} := \frac{\frac{1}{2\lambda_1\lambda_2}\sum_{j=1}^{N}(l_{w,j}^2 - \lambda_1^2 l_{1,j}^2 - \lambda_2^2 l_{2,j}^2)}{\sqrt{\sum_{j=1}^{N}l_{1,j}^2\sum_{j=1}^{N}l_{2,j}^2}}, \tag{17}$$

where λ_1, λ_2 are real and $l_{w,j}$ is the range of a portfolio from the two assets with weights λ_1 and λ_2. Obviously, $\hat{\rho}_{BD}$ requires additional information on the portfolio range, which is not readily available in general. Brandt and Diebold (2006) list several exceptions, the most notable being the foreign exchange market in the absence of triangular arbitrage. In the general case, however, in order to obtain the portfolio range one needs to reconstruct the path of the portfolio return using high-frequency data. For this reason it is debatable if the estimator can be classified as a correlation estimator based on *daily* OHLC prices and thus the comparison with the other estimators can be regarded as "unfair". Nevertheless, we include $\hat{\rho}_{BD}$ in the present simulation study for completeness and emphasize that the study provides useful insights in terms of the efficiency and robustness of this estimator.

4.1.   *Comparison of the correlation estimators under the BM assumption*

To compare the estimators under the original framework we simulate random samples of the following tuple $(C_1, l_1, W_1, C_2, l_2, W_2, C_{1-2}, l_{1-2}, W_{1-2})$ for sample sizes $N = 25$ and $N = 250$ (roughly corresponding to the amount of one month respectively one year of daily data). Each sextuple $(C_1, l_1, W_1, C_2, l_2, W_2)$ is obtained by simulations of discretizations of correlated BMs with $f = 500$ steps, which, in case of a 24-hour trading day, corresponds to a price observation every 2.88 minutes. The triple $(C_{1-2}, l_{1-2}, W_{1-2})$ gives the close return, the range and the BER of a portfolio with weights $\lambda_1 = 1$ and $\lambda_2 = -1$. Each sample is replicated $M = 10,000$ times. In that way the simulation study in the present work differs from the simulation study in Rogers and Zhou (2008). In the latter N is set to one while $M = 20,000$ and the efficiency of $\hat{\rho}_0$ rather than that of $\hat{\rho}_{RZ}$ is reported. However, for the evaluation of the bias the authors calculate $\hat{\rho}_{RZ}$ applied on the full

ρ	$\hat{\rho}_C$	$\hat{\rho}_{BD}$	$\hat{\rho}_{RZ}$	$\hat{\rho}_{RZ}^a$	$\hat{\rho}_J$
			Mean		
0.9	0.897	0.899	0.869	0.880	0.895
0.8	0.793	0.798	0.782	0.801	0.791
0.7	0.693	0.699	0.694	0.716	0.691
0.6	0.591	0.598	0.594	0.613	0.588
0.5	0.492	0.499	0.494	0.508	0.488
0.4	0.395	0.399	0.399	0.410	0.392
0.3	0.292	0.298	0.297	0.305	0.291
0.2	0.195	0.199	0.199	0.205	0.195
0.1	0.095	0.099	0.098	0.100	0.096
0.0	-0.003	-0.000	-0.000	-0.000	-0.001
			$N\cdot$Variance		
0.9	0.046	0.008	0.362	0.302	0.029
0.8	0.157	0.029	0.487	0.408	0.097
0.7	0.304	0.056	0.553	0.490	0.182
0.6	0.461	0.089	0.585	0.532	0.280
0.5	0.617	0.121	0.589	0.536	0.374
0.4	0.750	0.155	0.597	0.581	0.452
0.3	0.905	0.186	0.585	0.600	0.519
0.2	0.971	0.206	0.586	0.627	0.562
0.1	1.035	0.224	0.608	0.669	0.609
0.0	1.020	0.227	0.583	0.641	0.597
			$\sqrt{N\cdot\text{MSE}}$		
0.9	0.214	0.090	0.622	0.558	0.172
0.8	0.398	0.170	0.703	0.639	0.314
0.7	0.552	0.237	0.744	0.704	0.429
0.6	0.680	0.299	0.765	0.732	0.532
0.5	0.786	0.348	0.768	0.733	0.615
0.4	0.866	0.394	0.773	0.764	0.673
0.3	0.952	0.432	0.765	0.775	0.722
0.2	0.986	0.453	0.766	0.792	0.750
0.1	1.017	0.473	0.780	0.818	0.781
0.0	1.010	0.476	0.763	0.801	0.773

Table 3.   Bivariate standard BM, $N = 25$, M=10,000 replications

"sample" size $M = 20,000$. We report results only for (the more realistic for financial markets) positive correlations ($\rho \in \{0, 0.1, ..., 0.9\}$), the results for negative correlations are similar.

Estimates lying outside the admissible range are censored to the corresponding lower or upper bound. We calculate the mean, the variance and $\sqrt{N \cdot \text{MSE}}$ (to which we refer as root mean squared error (RMSE)) of the 5 competing estimators over the M replications. The results for $N = 25$ and $N = 250$ are reported in Table 3 and Table 4 respectively. In the following we dispose of the "hat" when denoting the estimated variance of the estimators to ease the notation.

4.1.1. **Small sample bias.**

Most estimators show no noticeable bias when $N = 25$. The only exceptions are $\hat{\rho}_{RZ}$ and $\hat{\rho}_{RZ}^a$. For larger (absolute) values of ρ both estimators are biased towards zero due to the imposed censoring. For $0.1 < \rho < 0.8$, the feasible Rogers-Zhou estimator suffers from an upward bias, the reason for which seems to be the variance estimator of the BER. For $\rho > 0.8$ this upside bias cancels out to some extent with the downside bias introduced through the censoring. The Pearson-based estimators ($\hat{\rho}_C$ and $\hat{\rho}_J$) show a small downside bias which is in line with the theory. When the sample size is increased to 250, all estimators are virtually unbiased.

It is interesting to note that Brandt-Diebold estimator $\hat{\rho}_{BD}$ is also unbiased. This may seem surprising since the range-based variance estimators, which build $\hat{\rho}_{BD}$, suffer from disretization bias. The explanation is that the bias in the numerator and in the denominator cancel out on average. In fact it is straightforward to show that for a driftless BM $\hat{\rho}_{BD}$ is a consistent estimator for the true correlation parameter *independent of the sampling frequency*, provided that the sampling frequency is the same for the two assets and the portfolio.

Table 3 shows that even for $N = 25$ the difference between the true value of ρ and the estimated

ρ	$\hat{\rho}_C$	$\hat{\rho}_{BD}$	$\hat{\rho}_{RZ}$	$\hat{\rho}_{RZ}^a$	$\hat{\rho}_J$
			Mean		
0.9	0.899	0.900	0.897	0.901	0.900
0.8	0.799	0.800	0.798	0.801	0.800
0.7	0.699	0.700	0.699	0.700	0.699
0.6	0.599	0.600	0.598	0.600	0.598
0.5	0.499	0.500	0.499	0.501	0.499
0.4	0.399	0.400	0.398	0.399	0.398
0.3	0.301	0.300	0.300	0.301	0.300
0.2	0.200	0.200	0.200	0.200	0.199
0.1	0.100	0.100	0.100	0.100	0.099
0.0	0.000	-0.000	-0.000	-0.000	-0.000
			N·Variance		
0.9	0.038	0.008	0.558	0.616	0.023
0.8	0.134	0.028	0.594	0.573	0.080
0.7	0.264	0.058	0.612	0.517	0.164
0.6	0.413	0.090	0.613	0.479	0.255
0.5	0.554	0.122	0.598	0.495	0.347
0.4	0.720	0.156	0.601	0.496	0.429
0.3	0.815	0.180	0.601	0.538	0.506
0.2	0.945	0.214	0.596	0.572	0.560
0.1	0.996	0.227	0.605	0.601	0.607
0.0	1.015	0.237	0.610	0.617	0.625
			$\sqrt{N}\cdot\text{MSE}$		
0.9	0.194	0.090	0.749	0.785	0.151
0.8	0.366	0.169	0.771	0.757	0.284
0.7	0.514	0.240	0.782	0.719	0.405
0.6	0.643	0.299	0.784	0.692	0.506
0.5	0.745	0.350	0.774	0.703	0.589
0.4	0.849	0.394	0.776	0.704	0.656
0.3	0.903	0.425	0.776	0.733	0.712
0.2	0.972	0.462	0.772	0.756	0.748
0.1	0.998	0.476	0.778	0.775	0.779
0.0	1.007	0.486	0.781	0.785	0.790

Table 4.   Bivariate standard BM, $N = 250$, M=10,000 replications

mean of $\hat{\rho}_{BD}$ is never larger than 0.002 in absolute value. However, when the sample size is much smaller, the effect of the sampling variation is larger and the estimator is biased. This explains the downside bias of $\hat{\rho}_{BD}$ reported in Table 1 (Brandt and Diebold 2006, p.7) when $N = 1$.

4.1.2.   Efficiency. In the following section, similar to Garman and Klass (1980), we define efficiency of an estimator B compared to an estimator A as the variance ratio $V(A)/V(B)$. An efficient estimator then is the estimator with the smallest variance within a set of (nearly) unbiased estimators. Efficiency gains of B over A, in turn, are defined as the percentage increase of $V(A)$ over $V(B)$.

Both Tables 3 and 4 show that it pays to use Pearson-based rather than r_0-type estimators for the close returns and BERs. The results from the simulation study are in line with the theoretical results from the bivariate normal model - $\hat{\rho}_J$ outperforms $\hat{\rho}_{RZ}$ (except for the vicinity of zero) with the efficiency gains increasing with increasing absolute value of ρ. In particular, the efficiency gains for $\rho = 0.9$ amount to 1248%. Thus, paradoxically as it may sound, it pays off not to use any a priori information on the variance of the excess returns. More precisely, such information should not be used *appropriately*.

The feasible estimator $\hat{\rho}_{RZ}^a$ performs slightly better than $\hat{\rho}_{RZ}$ in terms of efficiency but it is still outperformed by $\hat{\rho}_J$. Consequently, there are indications that the estimators $\hat{\rho}_{\mu;W}$ and $\frac{\frac{1}{N}\sum_{j=1}^N W_{1,j}W_{2,j}}{\sqrt{\frac{1}{N}\sum_{j=1}^N C_{1,j}^2 \frac{1}{N}\sum_{j=1}^N C_{2,j}^2(3-4\ln2)}}$ indeed differ in distribution.

For both sample sizes $\hat{\rho}_J$ completely outperforms the benchmark estimator with efficiency gains around 65%. On the other hand, $\hat{\rho}_C$ is more efficient than $\hat{\rho}_{RZ}$ for $|\rho| > 0.5$ although it makes

no use of information on the wicks. For smaller absolute values of ρ the latter estimator is more efficient but its efficiency gains are not as impressive as when compared to $\bar{\bar{\rho}}_{0;C}$ (as in Rogers and Zhou (2008)). Consequently, it pays to include BERs for the estimation of ρ but this should be done in the "right" way.

When data on the portfolio OHLC is available, $\hat{\rho}_{BD}$ clearly outperforms all other competitors for both sample sizes. Its efficiency gains with respect to $\hat{\rho}^a_{RZ}$ for the case $\rho = 0.9$ are impressive - 4514% ($N = 25$) and 6975% ($N = 250$). The results in Table 3 and Table 4 provide interesting insights on the inefficiency of the estimator $\hat{\rho}_{BD}$ when information on the entire path of the two BMs is used in order to derive the daily highs and lows of the composite process (as is the case in the simulation study). To make this point clear consider dividing the day in 10 subintervals with OHLC data available for each subinterval. Consequently, for 25 days there will be 250 pairs of close returns and BERs, correlated with ρ respectively ρ_W, available, i.e. $\hat{\rho}_J$ can be calculated on 250-observation-basis. This means that the variance of this estimator will roughly correspond to the variance of $\hat{\rho}_J$ implied in Table 4. Then in terms of efficiency of the two estimators one can compare $250 \cdot V(\hat{\rho}_J)$ (reported in Table 4) with $25 \cdot V(\hat{\rho}_{BD})$ (reported in Table 3) multiplied by 10. Then it becomes clear that $\hat{\rho}_J$ completely outperforms $\hat{\rho}_{BD}$ even though it uses even less information than the latter estimator (only OHLC data for the subperiods as opposed to the whole (discretized) path). It is fair to note that in this case the BER will be calculated from 50 rather than 500 observations per interval but $\hat{\rho}_J$ will not be seriously affected by this discretization problem because on the one side, as pointed out in Rogers and Zhou (2008), the BERs suffer less than the range from the discretization errors and, on the other side, the estimator based on BERs enters $\hat{\rho}_J$ (only) with weight 0.5.

To highlight these arguments we conduct an additional simulation, which marks the situation described in the previous paragraph. In particular we divide each of 25 days into 10 subintervals of equal length and simulate discretizations of correlated BMs with $f = 50$ steps for each subinterval. In that way we obtain quadruples (C_1, W_1, C_2, W_2) for each subinterval. Next we use the close returns from the subintervals to construct the path of the portfolio with weights, as before, $\lambda_1 = 1$ and $\lambda_2 = -1$. From the path containing 10 observations we obtain the daily range of the portfolio. The daily range of the two assets is calculated in a similar way, i.e. using paths constructed from subintervals' close returns. Although the observed daily range for the two assets is available, this approach allows the discretization bias of the portfolio's volatility estimator to be the same as the discretization bias of the respective estimators for the two assets. This is done to the advantage of $\hat{\rho}_{BD}$ because it could still profit from the canceling out effect of the bias. We calculate $\hat{\rho}_J$ using the $25 \cdot 10 = 250$ pairs of close and excess returns, while $\hat{\rho}_{BD}$ is calculated from the 25 daily triples of ranges.

The simulation is replicated $M = 10,000$ times and the (sample) means, variances (scaled by 250) and RMSE (scaled by $\sqrt{250}$) are reported in Table 5.

	Mean		$250\cdot$ Variance		$\sqrt{250}\cdot$ RMSE	
ρ	$\hat{\rho}_J$	$\hat{\rho}_{BD}$	$\hat{\rho}_J$	$\hat{\rho}_{BD}$	$\hat{\rho}_J$	$\hat{\rho}_{BD}$
0.9	0.899	0.899	0.023	0.126	0.152	0.355
0.8	0.799	0.798	0.083	0.441	0.290	0.665
0.7	0.698	0.698	0.167	0.901	0.410	0.950
0.6	0.597	0.597	0.260	1.374	0.512	1.173
0.5	0.496	0.497	0.358	1.945	0.601	1.395
0.4	0.397	0.397	0.436	2.426	0.662	1.558
0.3	0.298	0.298	0.506	2.759	0.712	1.662
0.2	0.199	0.199	0.562	3.141	0.750	1.772
0.1	0.099	0.097	0.597	3.404	0.773	1.845
0.0	-0.000	0.000	0.595	3.445	0.771	1.856

Table 5.   Comparison of $\hat{\rho}_J$ with Brandt-Diebold type of estimators using the same amount of information.

The results are expected. Reducing the sampling frequency by 10 times has little effect on the unbiasedness of $\hat{\rho}_J$. As expected the estimator $\hat{\rho}_{BD}$ also remains unbiased. However, in terms of efficiency the Brandt-Diebold estimator is completely outperformed by the modified Rogers-Zhou one with $\hat{\rho}_J$ being around 5 times more efficient for every ρ in the table. Consequently, when $\hat{\rho}_J$ is calculated using the same high-frequency data as $\hat{\rho}_{BD}$, i.e. when the two estimators are compared on a fair basis, then the one proposed in this paper performs better.

We conclude that as long as the assumption of a bivariate BM is satisfied and no additional information is available for the calculation of $\hat{\rho}_{BD}$, the estimator proposed in the present paper $\hat{\rho}_J$ is nearly unbiased and efficient in the set of studied estimators. Rogers-Zhou estimator $\hat{\rho}_{RZ}$ is not only limited in application but is also for larger absolute values of ρ the worst performing estimator even compared to the benchmark.

Before we close this section, we provide numerical arguments supporting the theoretical results regarding the asymptotic distribution of $\hat{\rho}_J$. In particular, for $\rho = 0, 0.1, ..., 0.9$ we simulate $M = 1000$ replications of samples (C_1, W_1, C_2, W_2) of length $N = 1500$. For each sample we calculate $\hat{\rho}_J$ and apply a Jarque-Bera test on the M estimates. The p-values are provided in Table 6. The results are in line with the theory - all p-values are greater than any standard significant level, which means that the Null hypothesis of a normal distribution cannot be rejected.

ρ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
p-value	0.553	0.306	0.953	0.298	0.496	0.260	0.377	0.358	0.160	0.151

Table 6. p-values from the Jarque-Bera test on a sample of $M = 1000$ $\hat{\rho}_J$ estimates, each calculated from a sample of $N = 1500$ pairs of close returns and BERs.

4.2.   *Robustness check*

In this section we consider the effects of jumps and excess kurtosis on the studied estimators under the framework of Merton Jump Diffusions introduced in Press (1967).

DEFINITION 2   Let $(W_t)_{t\in[0,1]}$ be a standard BM, $\mu \in \mathrm{R}, \sigma > 0$, N_t a Poisson process with intensity $\lambda > 0$ and $\sum_{i=1}^{N_t} Y_i$ an independent from $(W_t)_{t\in[0,1]}$ Compound Poisson process with $Y_i \sim N(\mu_J, \sigma_J^2)$. Then the stochastic process $X = (X_t)_{t\in[0,1]}$ with

$$X_t = \mu t + \sigma W_t + \sum_{j=1}^{N_t} Y_j$$

is a Merton Jump Diffusion (MJD).

According to this definition $(X_t)_{t\in[0,1]}$ is a BM with drift μ and scaling parameter σ, subjected to a finite number of random jump discontinuities occurring with exponential interarrival times (see (Madan and Seneta 1987, p.154)). In finance, such jumps are caused by the arrival of news with influence on the whole market (change in the interest rate for example) or specifically on the individual assets (earnings announcements for example).

In case $\mu = 0$, the characteristic function of the return X_t, $t \in [0, 1]$ is:

$$\psi_{X_t}(x) = \exp\{-\frac{tx^2\sigma^2}{2} + \lambda t(\exp\{ix\mu_J - \frac{x^2\sigma_J^2}{2}\} - 1)\},$$

where i denotes the imaginary unit. Using $\psi_{X_t}(x)$ it can be shown that the tail of X_t is heavier than the tail of a Gaussian return, whereas its distribution is symmetric when $\mu_J = 0$.

Next we construct correlated MJDs using the following procedure. Let Z_0, Z_1 and Z_2 be three independent MJD with (the same) parameters $\mu = \mu_J = 0$, σ, σ_J and λ. For $0 \leq \rho \leq 1$ the two processes

$$X_i = \sqrt{\rho}Z_0 + \sqrt{1-\rho}Z_i \qquad i = 1, 2 \tag{18}$$

are jump diffusions with characteristic functions of the returns

$$\psi_{X_{1,t}}(x) = \psi_{X_{2,t}}(x) = \exp\{-\frac{tx^2\sigma^2}{2} +$$

$$+ \lambda t(\exp\{-\frac{x^2\rho\sigma_J^2}{2}\} - 1) + \lambda t(\exp\{-\frac{x^2(1-\rho)\sigma_J^2}{2}\} - 1)\},$$

correlated with ρ. We provide the following interpretation of the construction in (18). $\sqrt{\rho}Z_0$ is the (scaled) market component with systematic shocks, occurring at the same time and with the same size for the individual asset price processes. As opposed to it, $\sqrt{1-\rho}Z_1$ and $\sqrt{1-\rho}Z_2$ represent the (scaled) idiosyncratic components, whose shocks are independent from one another and from Z_0. Please note that X_1 and X_2 differ from the classical definition of a MJD in that their jump part is a sum of two compound Poisson processes with generally different variances ($\rho\sigma_J^2$ and $(1-\rho)\sigma_J^2$), although the intensity parameter, the parametric family for the distribution of the jumps and the respective means are identical.

In the present work we simulate discretizations of the jump diffusions given above with $\mu = 0$, $\mu_J = 0$, $\sigma = \sqrt{0.5}$, $\sigma_J = \sqrt{0.5}$ and $\lambda = 1$ so that the means and the variances of the close returns are respectively zero and unity, whereas the excess kurtosis varies between 0.375 and 0.75 depending on ρ. The simulation technique for a discretized MJD is the following.

- *Step 1:* Simulate $f = 500$ normal rvs with zero mean and variance $1/f$.
- *Step 2:* Simulate $f = 500$ exponentially distributed rvs with parameter $\lambda = 1$. Cumulate the resulting series in order to determine the points of time, at which the jumps occur. Leave out values greater than unity.
- *Step 3:* For each of the points of time from Step 2 simulate a normally distributed rv with zero mean and variance 0.5 in order to determine the size of the jump.
- *Step 4:* Add the jumps to the increments of the random walk for the corresponding period so that the increments of a MJD are obtained (we use the convention that the jumps occur at the end of the time interval). These increments have zero means and variances $\frac{1}{f}(\sigma^2 + \lambda\sigma_J^2) = 0.002$.

Conduct Step 1 to Step 4 for each of the three processes Z_0, Z_1 and Z_2. Finally

- *Step 5:* Obtain the increments of X_1 and X_2 in accordance with (18).
- *Step 6:* Build the cumulative sums of the increments from Step 5.

The rest of the simulation is analog to the procedure described in section 4.1 with sample size $N = 250$ and number of replications $M = 10,000$. For each $\rho \in \{0, 0.1, ..., 0.9\}$ the means and the RMSE of the four estimators (we drop the estimator $\hat{\rho}_{RZ}$ and consider only its feasible version) of interest over the M replications are reported in Table 7.

The classical estimator is robust against the change in the underlying process and remains nearly unbiased. Its variance, however, has slightly increased, which leads to an increased RMSE compared to Table 4. The reason for that is the fact that the close returns are no longer normally distributed. $\hat{\rho}_{BD}$ is mildly affected by the change in the underlying process, which reflects in a small bias towards zero (especially in the vicinity of 0). As opposed to the Brandt-Diebold estimator, $\hat{\rho}_{RZ}^a$ has a considerable bias towards zero. The source of the bias is twofold. On the one hand, the variance of the excess returns is no longer $3 - 4\ln2$. On the other hand, the relation between ρ and ρ_W given in (6) and consequently the correction in (8) are no longer valid. For these reasons $\hat{\rho}_{RZ}^a$ is very sensitive to the change of the underlying process. $\hat{\rho}_J$ still depends on the relation between ρ and

ρ	$\hat{\rho}_C$	$\hat{\rho}_{BD}$	$\hat{\rho}_{RZ}^a$	$\hat{\rho}_J$	$\hat{\rho}_C$	$\hat{\rho}_{BD}$	$\hat{\rho}_{RZ}^a$	$\hat{\rho}_J$
		Mean				$\sqrt{N} \cdot \mathrm{MSE}$		
0.9	0.900	0.898	0.821	0.889	0.215	0.121	1.476	0.255
0.8	0.800	0.798	0.732	0.785	0.399	0.218	1.298	0.416
0.7	0.699	0.699	0.643	0.685	0.555	0.295	1.129	0.527
0.6	0.599	0.599	0.553	0.587	0.695	0.363	0.994	0.603
0.5	0.498	0.500	0.463	0.491	0.804	0.415	0.891	0.656
0.4	0.399	0.399	0.372	0.395	0.882	0.453	0.814	0.700
0.3	0.300	0.297	0.282	0.301	0.939	0.479	0.757	0.735
0.2	0.199	0.193	0.190	0.205	0.982	0.504	0.734	0.775
0.1	0.100	0.084	0.096	0.106	1.005	0.555	0.712	0.786
0.0	0.001	-0.033	0.001	0.001	1.003	0.718	0.692	0.774

Table 7. Correlated bivariate jump diffusions, parameters $\mu = \mu_J = 0$, $\sigma^2 = \sigma_J^2 = 0.5$, $\lambda = 1$, $N = 250$, M=10,000 replications

ρ_W but does not employ any assumption of the theoretical variances of the (standardized) BERs. Consequently, it is more robust than $\hat{\rho}_{RZ}^a$, which reflects in a smaller bias. In terms of RMSE, none of the estimators shows a clear advantage over its competitors for all studied values of ρ. The choice of a correlation estimator in this case is, therefore, debatable.

One possibility to improve on the performance of $\hat{\rho}_J$ in practical applications is to use Monte Carlo simulations to estimate the relation between ρ and ρ_W. Then the estimated relation can be used to correct $\bar{\hat{\rho}}_{0;W}$. This can be done, however, under the condition that a suitable multivariate model has been fit to the data.

5.   Empirical study

To back up the results from the simulation study with empirical evidence, we conduct an empirical study, which is centered around the modified Rogers-Zhou estimator

$$\hat{\rho}_J := \frac{1}{2} \left(\hat{\rho}_C + g^{-1}(\hat{\rho}_{\mu;W}) \right),$$

where, in order to avoid any inaccuracy due to the approximation of $g^{-1}(\cdot)$, we use the exact correction rather than its approximation. We want to know if there is a statistically significant difference in the mean of $\hat{\rho}_J$ and the classical estimator $\hat{\rho}_C$, which is the Pearson's sample correlation coefficient applied on the candlestick bodies. This would implicitly answer the question if the correction $g^{-1}(\cdot)$ of the BER correlation is adequate. If it is not, we want to learn more about the relation between ρ, the correlation of the candlestick bodies, and ρ_W. Furthermore, we want to check if $\hat{\rho}_J$ retains the efficiency gains over $\hat{\rho}_C$ promised by the simulation study in section 4.1 when the estimators are applied on real data. Our empirical study follows the set-up in Rogers and Zhou (2008), where the authors conduct a small-scale empirical study on four assets from the American stock market. Our goal is to extend the number of asset pairs and to include other major stock markets. We note also that the authors compare the r_0-type estimators, while we concentrate on the more efficient Pearson-based correlation estimators.

Throughout this section, returns (high, low, close) are related to the open price rather than to the previous close. Correlation estimators are calculated from (non-overlapping) small-size samples of length $N = 25$. This ensures that the sample of correlation pairs is large enough, while the assumption of a constant correlation is still plausible. We note that in the empirical study in Rogers and Zhou (2008) the authors also use small-size samples with $N = 1$.

5.1. *Data*

The empirical study in this section concentrates on four major stock markets - the American, the German, the French and the UK one. In particular we study pairs of constituents of the following benchmark indices - the Dow Jones Industrial Average (DJI) for the American, the "Deutscher Aktienindex" DAX for the German, the CAC 40 (or simply CAC) for the French and the FTSE 100 (or simply FTSE) index for the UK market. The first two indices consist of 30 large publicly traded companies trading on the local stock exchange markets, whereas the third consists of 40 and the fourth consists of 100 such companies.

The data source is Thompson Reuters' statistical database DataStream. The data used for the study is filtered from:

- *days with "not annotated" (NA) entries.* Mostly weekends and bank holidays.
- *days on which no trading takes place*, i.e. open price equals low, high and close price.
- *days with irregularities in the data* such as high price being lower than open and/or close price, respectively low price being higher than open and/or close price.

For the data cleaning we assume the following approach. Days with NA entries and days without trading are deleted from the data so that only days with actual, simultaneous (for all studied assets on the particular market) trading are considered. With respect to the irregularities we set 1) the high price equal the maximum of open, high and close price and 2) low price equal the minimum of open, low and close price.

For the American market we consider 28 constituents of DJI dropping Kraft foods and JPMorgan Chase due to incomplete data. This results in $\binom{28}{2} = 378$ pairs of assets. OHLC data is collected from 1993-10-01 until 2012-02-23. In that way we obtain $4625 = 25 \cdot 185$ complete records of daily OHLC prices for this time period. For the German market we consider 28 out of 30 constituents of DAX with the exceptions being Deutsche Brse AG and BASF. The former is omitted due to incomplete data, while the latter is left out due to a large amount of irregularities. The time series span from 2000-12-19 until 2012-02-23, i.e. $2825 = 25 \cdot 113$ days with actual, simultaneous for *all* 28 assets trading. In terms of the French market we consider 37 constituents of CAC (The missing assets are EDF (DataStream code F:EDF), GDF SUEZ (F:GDF) and LEGRAND (F:LRRS)), resulting in $\binom{37}{2} = 666$ pairs of assets. The time series span from 2002-01-28 until 2012-02-23, which leads to $2450 = 25 \cdot 98$ complete OHLC records for all assets. Last but not least, 83 out of 102 FTSE assets are used for the study on the UK market (19 are left out because of incomplete records). The number of pairs is then $\binom{83}{2} = 3403$. Starting from 2002-08-29 we collect until 2012-02-23 altogether $2050 = 25 \cdot 82$ complete "clean" OHLC records for all considered assets.

5.2. *Set up of the study*

The four markets are studied separately and the procedure for all of them is analogous. First, from the daily OHLC data for each asset on each day the pair of close returns and BERs is calculated. Then we divide the data for each asset in M *non-overlapping* samples of length 25, $M = 185$ for the American, $M = 113$ for the German, $M = 98$ for the French and $M = 82$ for the UK markets. The number of pairs is denoted by N^* with $N^* = 378$ for the American and the German markets, while $N^* = 666$ for the French and $N^* = 3403$ for the UK markets. For each sample k, $k = 1, ..., M$, the pair $(\hat{\rho}_{i;k;C}, \hat{\rho}_{i;k;J})$ is calculated, where i indicates the asset pair, $i = 1, ..., N^*$. Thus on each market we obtain $M \cdot N^*$ pairs of Pearson-based correlation coefficients. Note that in the following we use $\hat{\rho}_{i;k;\mu;W}$ to denote the correlation estimator of the BERs given in (15) calculated for the asset pair i from sample k.

The methodology we employ has been suggested in Rogers and Zhou (2008). In particular, for each asset pair we calculate the sample means and the sample standard deviations of the two correlation estimators. In a next step we calculate the respective confidence intervals with confidence level approximately 95 % (mean $\pm 2 \cdot \frac{1}{\sqrt{M}}$ standard deviation). For the construction of the intervals

we use the asymptotic distribution of the correlation estimators. To demonstrate that this is a reasonable approach we conduct a small empirical analysis on the estimates. To gain a feeling of their distribution, we randomly choose two pairs of assets from the American market (3M - AT & T and GE - Travelers Companies) and plot histograms of the empirical distribution of $\hat{\rho}_{i;k;C}$ and $\hat{\rho}_{i;k;J}$ for each asset pair in Figure 3.
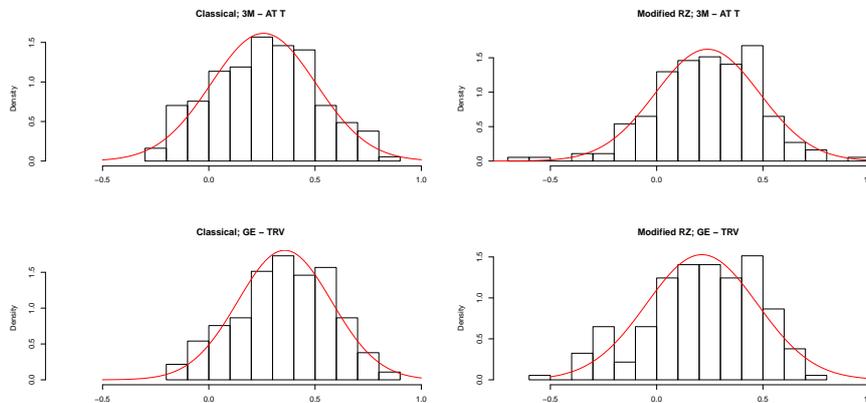

Figure 3. Histograms of estimates from $\hat{\rho}_{i;k;C}$ (Classical) and $\hat{\rho}_{i;k;J}$ (Modified RZ) for the asset pairs 3M - AT & T and GE - TRV. The solid red line is the density of a normal distribution with mean and variance equal the sample counterparts for each case.

The histogram analysis confirms that for the chosen estimates the normal assumption is not unjustified. To generalise this result we conduct a Jarque-Bera test on *every* sample of correlation estimates (from both classical and modified Rogers-Zhou estimators) for each market. We summarise the results in Table 8, which reports the absolute and percentage number of cases, for which the hypothesis of normal distribution of the estimates has been rejected at 5 % significance level.

	#	%
DJI	103	0.136
DAX	117	0.155
CAC	141	0.106
FTSE	284	0.042

Table 8. Summary of p-values in absolute (#) and percentage (%) numbers from the Jarque-Bera test, which lie below 5% for each market.

These results confirm the findings from the histogram analysis, i.e. that the normal approximation is reasonable.
When the CI are constructed, we inspect if the means of the estimators lie in each other confidence intervals. Furthermore we compare the variance ratios defined as $\widehat{Var}(\hat{\rho}_C)/\widehat{Var}(\hat{\rho}_J)$.

5.3. *Results and discussion*

5.3.1. Mean comparison of $\hat{\rho}_C$ and $\hat{\rho}_J$. We provide a summary of the mean comparison of $\hat{\rho}_C$ and $\hat{\rho}_J$ in Table 9. The first column gives the number of asset pairs for the particular market. In the following two columns the number (in absolute and percentage terms) of asset pairs is given, for which the sample mean of $\hat{\rho}_J$, say $\bar{\hat{\rho}}_J$, lies in the confidence interval derived from $\hat{\rho}_C$, say $CI(\bar{\hat{\rho}}_C)$. For the next two columns it is the other way round. The 6th and the 7th columns give the number of pairs, for which both sample means lie in the overlapping part of the two intervals. In the last two columns the number of pairs is given, for which the sample mean of $\hat{\rho}_J$ lies under the lower

		(a)		(b)		$(a) \wedge (b)$		$\bar{\hat{\rho}}_J < \hat{\rho}_C^L$	
	N^*	#	%	#	%	#	%	#	%
	378	184	0.487	116	0.307	116	0.307	194	0.513
DAX	378	145	0.384	102	0.270	102	0.270	233	0.616
CAC	666	259	0.389	208	0.312	208	0.312	407	0.611
FTSE	3403	2044	0.601	1596	0.469	1596	0.469	1359	0.399

Table 9.   Mean comparison of $\hat{\rho}_C$ and $\hat{\rho}_J$, (a) gives the number - absolute (#) and percentage (%) - of cases for which $\bar{\hat{\rho}}_J \in CI(\bar{\hat{\rho}}_C)$; (b) gives the number - absolute (#) and percentage (%) - of cases for which $\bar{\hat{\rho}}_C \in CI(\bar{\hat{\rho}}_J)$. $\hat{\rho}_C^L$ denotes the lower bound of $CI(\bar{\hat{\rho}}_C)$.

	N^*	Min	LH	Median	UH	Max	#	%
DJI	378	1.067	1.347	1.451	1.524	1.884	0	0.000
DAX	378	0.991	1.284	1.392	1.530	2.309	1	0.003
CAC	666	0.672	1.198	1.298	1.408	2.003	17	0.026
FTSE	3403	0.739	1.312	1.438	1.584	2.370	32	0.009

Table 10.   Statistics of the variance ratios $\widehat{Var}(\hat{\rho}_C)/\widehat{Var}(\hat{\rho}_J)$, LH stands for lower-hinge, UH stands for upper-hinge.

bound of the confidence interval derived from $\hat{\rho}_C$, say $\hat{\rho}_C^L$. The rows correspond to the studied market with the benchmark index labeling the market.

Several things should be pointed out. First, the percentage pairs, for which the mean differences of the two estimators are well within sampling error, is not negligible but still low. It is around 30 % for the American, the German and French market and slightly higher - slightly less than 50% - for the UK one. If the sample mean of $\hat{\rho}_C$ lies in the confidence interval from $\hat{\rho}_J$, then the opposite is also true. Moreover, the percentage pairs for which the confidence interval derived from $\hat{\rho}_C$ covers the $\bar{\hat{\rho}}_J$ is higher than the other way round. This indicates that $\hat{\rho}_J$ is generally less volatile than $\hat{\rho}_C$ (see below). Last but not least, in all cases, in which the estimated mean of $\hat{\rho}_J$ is not in the confidence interval derived from $\hat{\rho}_C$, it lies below the lower bound.

We make the following conclusions. The bias correction $g^{-1}(\cdot)$ is not a bad approximation of the relation between ρ_W and ρ_r, but there is room for improvement. There is a clear indication that for a greater part of the studied asset pairs, the gap between ρ_r and ρ_W is much larger than is the case under the BM model.

5.3.2.   Variance comparison of $\hat{\rho}_J$ and $\hat{\rho}_C$. Next we compare the variance ratios of the estimators $\hat{\rho}_J$ and $\hat{\rho}_C$ in Table 10. In the first five columns (not counting the column with N^*) we give the minimum, the lower-hinge, the median, the upper-hinge and the maximum (i.e. the five-number summary) of the N^* variance ratios for each market. In the last two columns we give the number - the absolute (#) and the percentage (%) one - of the asset pairs, for which the variance ratio is less than one. For completeness we plot in Figure 4 the variance ratios of the pairs from each market ordered by size.

The results are reassuring. For all markets the efficiency gains through the additional information on the BER correlation is around 40% on average (slightly higher for DJI and FTSE, and a little lower for CAC). This is comparable with the results from the simulation study in section 4.1 (see Table 3), where the efficiency gains of $\hat{\rho}_J$ over $\hat{\rho}_C$ are around 65% (depending on ρ). Note that for some asset pairs the efficiency gains even exceed 100%. Moreover, depending on the market there are no or very few cases, for which the estimated variance of $\hat{\rho}_C$ is smaller than that of $\hat{\rho}_J$. Therefore, we conclude that $\hat{\rho}_J$ retains its advantages in terms of efficiency when applied on empirical data.

Before we close this section we provide an interesting insight. It builds on the idea that market turmoil causes prices to jump, which in turn diverges ρ_W from ρ further than postulated by a BM (see Section 4.2). In that respect we suggest a simple test for indicating "calm" periods on the market. These could be months or years, for example. The approach is cross-sectional. In particular
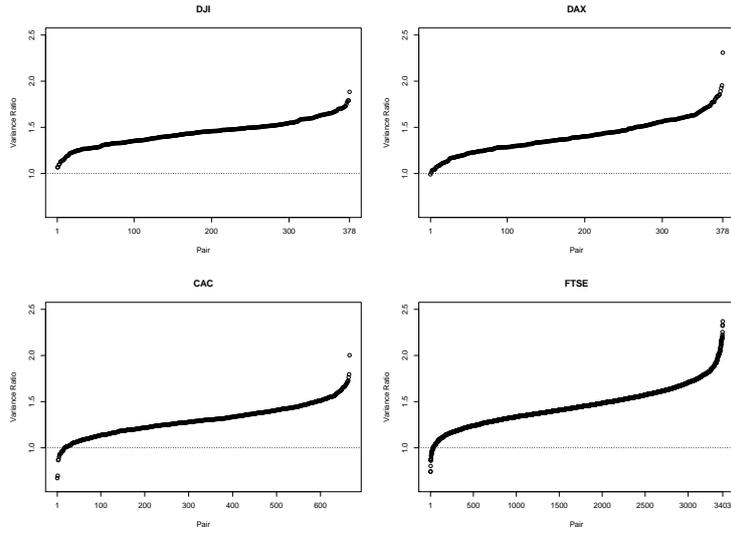
Figure 4. Variance ratios of different asset pairs from the four major stock markets

for each period we apply a paired t-test on the N^* observations $(\hat{\rho}_{i;k;C}, \hat{\rho}_{i;k;J})$ corresponding to the N^* pairs of asset from the particular market. More precisely we test the hypothesis that there is no significant difference in the estimators' means:

$$H_0 : E(\hat{\rho}_{i;k;C} - \hat{\rho}_{i;k;J}) = 0 \quad \text{for all } i\text{=}1,....,N^*$$

with k fixed. Non-rejection of the null-hypothesis can be regarded as an indication for a "calm" period. Notice that

$$\hat{\rho}_{i;k;C} - \hat{\rho}_{i;k;J} = \frac{1}{2} \left(\hat{\rho}_{i;k;C} - g^{-1}(\hat{\rho}_{i;k;\mu;W}) \right),$$

i.e. in other words we test if the correction $g^{-1}(\cdot)$ is adequate.

We conduct a small empirical study on this problem and briefly sketch the results. We are interested in pinning down "calm" *years* on the four aforementioned markets and set $N = 250$, i.e. we calculate $\hat{\rho}_C$ and $\hat{\rho}_J$ from samples of size 250. The sample years roughly correspond to calender years. In that way we study the period 1994-2011 for the American, 2001-2011 for the German, 2003-2011 for the French and the UK markets. We pin down two "calm" years on the American market (1995 and 1996), one on the French (2005) and none on the German and the UK ones. For the rest there is a significant difference in the mean of the two estimators, which we interprete as an indicator for market turmoils.

6.   Conclusion

In the present paper we introduced the concept of the BER, which is the difference of the wicks in a Japanese candlestick, and derived its distribution under a standard BM. In a next step we used it to improve correlation estimators based on components of daily-data Japanese candlesticks. In particular, we showed that using Pearson-type correlation estimators for the close returns and the BERs rather than r_0- types as suggested in Rogers and Zhou (2008) increases the efficiency and the robustness of the Rogers-Zhou estimator. Another contribution of the present paper is a simple approximation of the correction function for the latter estimator, which eases its application. A large-scale empirical study confirmed that considering BERs as in $\hat{\rho}_J$ leads to an increased efficiency

compared to the classical estimator based only on the candlestick body.

Future research can concentrate on transferring the BER-based estimators in the high-frequency field, as Bannouh et al. (2009) did in terms of the Brandt-Diebold estimator. On the one side, it will be interesting to study the properties of $\hat{\rho}_J$ in the presence of microstructure noise and non-synchronuous trading. Note that in this case the estimator will be applied on interval OHLC data with frequency lower than a day. On the other side, in the framework of jump diffusions with finite activity, if the interval length is sufficiently small, there will be only one jump in each interval. The latter will affect the body but not the (balanced) wicks. Consequently, in a jump diffusion where the correlation of the BM components differs from the correlation of the jump components, estimators based on the BER could potentially filter the later and provide information on the former. Such estimators would then be "jump-robust". Finally, the assumption of constant correlation in the BM model could be relaxed. An example of models that allow for dynamic correlation can be found in Engle (2002). Extending this framework to accommodate the dynamics of the correlation of BERs poses a challenging task for future research.

Appendix A

The joint density of the random vector $(H, L, C)'$ is given for $a \leq z \leq b$, $a \leq 0$, $b \geq 0$ as (see 1.15.8, p.174 in Borodin and Salminen (2002)):

$$\frac{\partial P(L < a, H < b, C < z)}{\partial a \partial b \partial z} =$$

$$= \frac{2}{\sqrt{2\pi}\sigma^3} \sum_{k=-\infty}^{\infty} \left(2k^2 e^{-\frac{(z+2k(b-a))^2}{2\sigma^2}} \left(\frac{(z+2k(b-a))^2}{\sigma^2} - 1 \right) - \right.$$

$$\left. 2k(k+1) e^{-\frac{(z-2a+2k(b-a))^2}{2\sigma^2}} \left(\frac{(z-2a+2k(b-a))^2}{\sigma^2} - 1 \right) \right)$$

We set $\sigma = 1$. The joint density of the close returns, the range and the BER is obtained using the transformation

$$
\begin{array}{lll}
W = H + L - C & & u = b + a - z \\
l = H - L & \text{implying} & s = b - a \\
C = C & & z = z
\end{array}
$$

Denote the system of linear equations defined by the equations above as $(u, s, z) = \mathbf{g}((a, b, z))$. Further denote $h(a, b, z) = \frac{\partial P(L < a, H < b, C < z)}{\partial a \partial b \partial z}$ (i.e. the trivariate density of high, low and close) and $f(u, s, z) = \frac{\partial P(C < z, l < s, W < u)}{\partial z \partial s \partial u}$ (i.e. the trivariate density of close, range and BER). According to the density transformation rule:

$$f(u, s, z) = h(\mathbf{g}^{-1}(a, b, z))|J|,$$

where J is the Jacobian of the transformation, i.e. the determinant of the matrix with first order partial derivatives of $\mathbf{g}^{-1}(\cdot)$. In particular this matrix is:

$$
\begin{pmatrix}
\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\[2mm]
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\[2mm]
0 & 0 & 1
\end{pmatrix}
$$

with $|J|$ equal $\frac{1}{2}$.

Applying the density transformation rule above we obtain the following joint density function

$$\frac{\partial P(C < z, l < s, W < u)}{\partial z \partial s \partial u} =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \left(2k^2 e^{-\frac{(z+2ks)^2}{2}} \left((z+2ks)^2 - 1\right) - \right.$$

$$\left. 2k(k+1) e^{-\frac{((2k+1)s-u)^2}{2}} \left(((2k+1)s-u)^2 - 1\right) \right) \qquad (19)$$

for $|z| + |u| \leq s < \infty$. Please note that this joint density function can also be used to obtain the marginal densities of C and l.

Integrating over the range yields the following joint density function of C and W:

$$\frac{\partial P(C < z, W < u)}{\partial z \partial u} =$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \left(k e^{-\frac{(z+2k(|u|+|z|))}{2}} (z+2k(|u|+|z|)) - \right.$$

$$\left. \frac{2k(k+1)}{2k+1} e^{-\frac{((2k+1)(|u|+|z|)-u)}{2}} \left((2k+1)(|u|+|z|)-u\right) \right) \qquad (20)$$

for u and z both $\in (-\infty, +\infty)$. It remains to integrate the bivariate density function over the close return. The result is the (marginal) density function of the BER

$$\frac{\partial P(W < u)}{\partial u} = \frac{4}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \left(\frac{k^2}{4k^2-1} e^{-\frac{(2ku)^2}{2}} - \frac{k(k+1)}{(2k+1)^2} e^{-\frac{((2k+1)|u|-u)^2}{2}} \right)$$

for $u \in (-\infty, +\infty)$. It can be shown that the density above can be further simplified to the following expression:

$$\frac{\partial P(W < u)}{\partial u} = \sum_{k=1}^{\infty} \frac{16k^2}{(4k^2-1)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2ku)^2}{2}} = \sum_{k=1}^{\infty} \frac{8k}{(4k^2-1)^2} \phi_k(u), \qquad (21)$$

Acknowledgment

References

Bannouh, K., D. van Dijk, and M. Martens (2009) Range-Based Covariance Estimation Using High-Frequency Data: The Realized Co-Range, *Journal of Financial Econometrics*, 7(4), pp. 341–372.

Becker, M., R. Friedmann, S. Klößner, and W. Sanddorf-Köhle (2007) A Hausman test for Brownian motion, *AStA - Advances in Statistical Analysis*, 91(1), pp. 3–21.

Borodin, A. N. and P. Salminen (2002) *Handbook of Brownian Motion - Facts and Formulae*, Probability and its Applications, Birkhäuser, Basel et al., 2. edition.

Brandt, M. W. and F. X. Diebold (2006) A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations, *Journal of Business*, 79(1), pp. 61–74.

Brunetti, C. and M. P. Lildholdt (2008) Range-based covariance estimation: the Co-range, Johns Hopkins University.

Cramér, H. (1971) *Mathematical Methods of Statistics*, Princeton University Press.

Engle, R. F. (2002) Dynamic Conditional Correlation: A simple Class of Multivariate Generalized Autoregressive Conditional Hetereroskedasticity Models, *Journal of Business and Economic Statistics*, 20(3), pp. 339–350.

Garman, M. B. and M. J. Klass (1980) On the Estimation of Security Price Volatilities from Historical Data, *Journal of Business*, 53(1), pp. 67–78.

Klößner, S. (2007) On Intraday Time-Reversibility of Return Processes, Paper presented at Statistics under one umbrella, Bielefeld, 2007, March 27-30.

Klößner, S. (2010) A high-low-based omnibus test for symmetry, the Lévy property, and other hypotheses on intraday returns, *Finance & Stochastics*, 14(1), pp. 1–12.

Lehmann, E. L. (1999) *Elements of Large-Sample Theory*, Springer-Verlag New York, Inc.

Madan, D. B. and E. Seneta (1987) Simulation of Estimates Using the Empirical Characteristic Function, *International Statistical Review*, 55(2), pp. 153–161.

Neudecker, H. and A. M. Wesselman (1990) The Asymptotic Matrix of the Sample Correlation Matrix, *Linear Algebra and its Applications*, 127, pp. 589–599.

Olkin, I. and J. W. Pratt (1958) Unbiased estimation of certain correlation coefficients, *Annals of Mathematical Statistics*, 29, pp. 201–211.

Parkinson, M. (1980) The Extreme Value Method for Estimating the Variance of the Rate of Return, *Journal of Business*, 53(1), pp. 61–65.

Press, J. S. (1967) A compound events model for security prices, *Journal of Business*, 40(3), pp. 317–335.

Rogers, L. C. G. and S. E. Satchell (1991) Estimating variance from high, low and closing prices, *The Annals of Applied Probability*, 1(4), pp. 504–512.

Rogers, L. C. G. and L. A. Shepp (2006) The correlation of the maxima of correlated Brownian motions, *Journal of Applied Probability*, 43, pp. 880–883.

Rogers, L. C. G. and F. Zhou (2008) Estimating correlation from high, low, opening and closing prices, *The Annals of Applied Probability*, 18(2), pp. 813–823.

Yang, D. and Q. Zhang (2000) Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices, *Journal of Business*, 73(3), pp. 477–491.