Quantitative Finance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rquf20

Estimating correlations among elliptically distributed random variables under any form of heteroskedasticity

Matteo Pelagatti & Giacomo Sbrana

Quantitative Finance, 2024
Vol. 24, Nos. 3–4, 451–464, https://doi.org/10.1080/14697688.2023.2278502

Estimating correlations among elliptically distributed random variables under any form of heteroskedasticity

MATTEO PELAGATTI *† and GIACOMO SBRANA‡

†Department of Economics, Management and Statistics, University of Milano-Bicocca, Via Bicocca degli Arcimboldi 8, Milan, 20126, Italy
‡Department of Information Systems, Supply Chain & Decision Making, NEOMA Business School, 1 Rue du Marechal Juin, Mont-Saint-Aignan, 76130, France

(*Received 2 August 2022; accepted 26 October 2023; published online 4 January 2024*)

The paper introduces a semiparametric estimator of the correlations among elliptically distributed random variables invariant to any form of heteroscedasticity, robust to outliers, and asymptotically normal. Our estimator is particularly fit for financial applications as vectors of stock returns are generally well approximated by heteroskedastic processes with elliptical (conditional) distributions and heavy tails. The superiority of our estimator with respect to Pearson's sample correlation in financial applications is illustrated using simulated data and real high-frequency stock returns. Using simple exponentially weighted moving averages, we extend our estimator to the case of time-varying correlations and compare it to the popular GARCH-DCC model. We show that the two approaches have comparable performances through simulations and a simple application. However, our estimator is extremely fast to compute, computationally robust, and straightforward to implement.

Keywords: Correlation; Heteroscedasticity; Stochastic volatility; Elliptical distributions; Heavy tails; Quadrant correlation; Dynamic conditional correlation

JEL Classifications:: C13, C14, C32, C46, G1

1. Introduction

It is common to model financial returns as sequences of independent identically distributed random variables multiplied by volatility processes that govern the evolution of their standard deviations in time. The classes of GARCH (Engle 1982, Bollerslev 1986, Nelson 1991) and stochastic volatility (Hull and White 1987, Heston 1993, Harvey *et al.* 1994, Alizadeh *et al.* 2002) processes are popular examples of such models.

Moreover, for theoretical and empirical reasons, financial returns are generally assumed elliptically distributed (Owen and Rabinovitch 1983, Ingersoll 1987, Fiorentini *et al.* 2003, Pelagatti and Rondena 2004, Asai *et al.* 2006, Bauwens *et al.* 2006, Hamada and Valdez 2008, Diongue *et al.* 2009, Zhou *et al.* 2014, Dias *et al.* 2018).

We propose a semiparametric estimator of the correlation among elliptically distributed random variables that is invariant to any form of (deterministic or stochastic) heteroscedasticity and robust to outliers. We developed the idea of this statistic while looking for a two-step procedure to fit stochastic volatility (SV) models, such as those in Harvey *et al.* (1994) and Alizadeh *et al.* (2002), to very large vectors of returns. We needed the first step to be a reliable estimate of the correlation matrix of returns, and under the hypotheses of SV models, sample correlation is inconsistent.

We derive the large-sample behavior of our estimator, proving its consistency and asymptotic normality. We expect the asymptotic approximation to be rather accurate even in relatively small samples, as the random variables entering the central limit theorem are bounded and, therefore, possess moments of any order. Our simulation experiments confirm this conjecture.

We compare our estimator's behavior to sample correlation in a simulated environment and on real high-frequency

*Corresponding author. Email: matteo.pelagatti@unimib.it
The data and code used in the article are freely available in the GitHub repository https://github.com/matteopelagatti/signcor

returns. In particular, the simulation experiments are based on typical models for heteroskedastic returns used in finance (SV, GARCH, and Gaussian mixtures), while the estimation of correlations among high-frequency data is based on 5-minute returns of S&P100 constituents. Our estimator is reliable in both cases, while sample correlation behaves poorly.

Our proposal is related to Blomqvist's (1950) *quadrant correlation* (also *quadrant count ratio*), whose results we generalize in two directions. Indeed, we prove the estimator to be consistent for Pearson's population correlation under any elliptical distribution with arbitrary forms of heteroscedasticity, and we allow the observations to be centered using any consistent estimator of the population median. Indeed, Blomqvist (1950) proves consistency for Pearson's correlation only in the homoskedastic Gaussian case.

Finally, we show how to obtain estimates of time-varying correlations using (a transform of) exponentially weighted moving averages (EWMA) of the random quantities on which our estimator is built (i.e. the sign of the product of two random variables). The performance of our time-varying estimator is compared to that of the popular Dynamic Conditional Correlation (DCC) model of Engle (2002) both in simulation experiments and in a real-world application to global minimum variance portfolios. The performance of the two approaches is comparable: the DCC performs slightly better in simulations, while there is no clear winner in the application. However, our estimator is extremely fast to compute, computationally robust, and straightforward to implement. These features should make our estimator interesting to practitioners and financial software developers.

The rest of the paper is organized as follows. Section 2 introduces the estimator, derives its asymptotic properties (2.1), and illustrates its finite sample performance for three classes of models using simulation (2.2). Section 3 applies the estimator to high-frequency stock returns and compares it to the sample correlation and its trimmed version. Section 4 shows how to obtain a time-varying variant of our estimator using EWMA and compares its performance to the DCC. Section 5 draws some conclusions.

2. The sign-based correlation estimator

In this section, we show how to consistently estimate correlations among elliptical random variables (W_i, W_j) under any form of heteroscedasticity. Our estimator is based on a simple non-linear transformation of the mean sign of the product of the centered random variables:

$$\tilde{\rho}_{ij} := \sin\left(\frac{\pi}{2n}\sum_{t=1}^{n}\operatorname{sign}\left((W_{it}-\hat{\mu}_i)(W_{jt}-\hat{\mu}_j)\right)\right),$$

with $\hat{\mu}$ representing a consistent estimator of the median of the random variable and

$$\operatorname{sign}(x) := \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0\,. \\ 1 & \text{for } x > 0 \end{cases}$$

2.1. Definitions and asymptotics

The usual definition of population correlation assumes the existence of second-order moments. However, the class of elliptical distributions is an extension of the multivariate normal distribution, characterized by a nonnegative-definite scale matrix $\boldsymbol{\Sigma}$, proportional to the covariance matrix, when this exists. Thus, we can define the correlation matrix of an elliptically distributed vector as

$$\boldsymbol{R} := \boldsymbol{D}^{-1}\boldsymbol{\Sigma}\boldsymbol{D}^{-1},$$

where $\boldsymbol{D}$ is a diagonal matrix with values equal to the square root of the values on the diagonal of $\boldsymbol{\Sigma}$. The elements ρ_{ij} of $\boldsymbol{R}$ are well defined also when the moments of the distributions are not finite. We are interested in estimating ρ_{ij}, which coincides with the common definition of correlation when second moments exist and extends the definition to the case of non-existing moments.

ASSUMPTION 1 *Consider the vector process $\boldsymbol{x}_t$, whose i-th element is $X_{it} = Z_{it}\sigma_{it}$, for $i = \{1,\dots,d\}$, $t = \{1,\dots,n\}$, $0 < \sigma_{it} < \infty$, and let $\boldsymbol{z}_t = [Z_{1t},\dots,Z_{dt}]^\top$ be i.i.d. following a continuous d-dimensional elliptical distribution with zero median (and, thus, also zero mean when it exists) and correlation matrix $\boldsymbol{R} = \{\rho_{ij}\}$.† The sequences of scale factors σ_{it} can be either deterministic or stochastic.*

THEOREM 1 *Under Assumption* 1,

$$\nu_{ij} := \mathbb{E}\operatorname{sign}\left(X_{it}X_{jt}\right) = \frac{2}{\pi}\arcsin\rho_{ij} \quad \forall t. \tag{1}$$

REMARK 1 Notice that the evolution of the scale parameters σ_{it} can be deterministic or stochastic. The simple trick we exploit to get rid of the arbitrary scale parameters and estimate the correlations is

$$\begin{aligned}\operatorname{sign}(X_{it}X_{jt}) &= \frac{X_{it}X_{jt}}{|X_{it}||X_{jt}|} = \frac{(Z_{it}\sigma_{it})(Z_{jt}\sigma_{jt})}{|Z_{it}\sigma_{it}||Z_{jt}\sigma_{jt}|} \\ &= \frac{Z_{it}Z_{jt}}{|Z_{it}||Z_{jt}|} = \operatorname{sign}(Z_{it}Z_{jt}).\end{aligned}$$

REMARK 2 The reader acquainted with copula theory has probably noticed that the map between ρ_{ij} and ν_{ij} defined in equation (1) is the same as the relation that holds between the correlation coefficient of a bivariate elliptical copula and the population Kendall's tau of the two random variables

† The i.i.d. hypothesis can be relaxed provided that the marginal distribution of the processes Z_{it} is continuous and elliptical with correlation matrix $\boldsymbol{R}$ and that a law of large number (LLN) and a central limit theorem (CLT) apply. For example, Z_{it} could be a martingale difference sequence, and we could exploit the LLN and CLT for these kinds of processes (see Hall and Heyde 1980, for a complete treatment). Alternatively, Z_{it} could be a weakly dependent process, such as a stationary ergodic sequence, a mixing, or a near-epoch dependent process, as LLN and CLT also exist under these hypotheses (see Davidson 1994, in particular Chapters 20 and 24). Under weak dependence, the asymptotic variances of Theorem 2 would have to be adjusted to take into account the possible autocorrelation of the sequence $\operatorname{sign}(X_{it}X_{jt})$ but all the rest would remain valid.

(Lindskog *et al.* 2003). The population Kendall's tau can be expressed in terms of the copula function as

$$\tau_{ij} = 4\int_0^1\int_0^1 C_{ij}(uv)\mathrm{d}C_{ij}(u,v) - 1,$$

and it is often used to parametrize copula families, such as elliptical copulas. However, the correlation estimator based on Kendall's tau should not be used in the case of heteroscedasticity because it is not consistent (for an important counterexample, see the Appendix).

The following two theorems state the consistency and asymptotic normality of our estimator.

THEOREM 2 *Let us define the sample mean sign*

$$\hat{\nu}_{ijn} := \frac{1}{n}\sum_{t=1}^{n}\mathrm{sign}(X_{it}X_{jt}), \tag{2}$$

and its sine transform

$$\hat{\rho}_{ijn} = \sin\left(\frac{\pi}{2}\hat{\nu}_{ijn}\right). \tag{3}$$

Under Assumption 1,

- $\hat{\nu}_{ijn} \xrightarrow{a.s.} \nu_{ij}$;
- $\sqrt{n}(\hat{\nu}_{ijn} - \nu_{ij}) \xrightarrow{d} \mathsf{N}(0, 1-\nu_{ij}^2)$;
- $\hat{\rho}_{ijn} \xrightarrow{a.s.} \rho_{ij}$;
- $\sqrt{n}(\hat{\rho}_{ijn} - \rho_{ij}) \xrightarrow{d} \mathsf{N}(0, \sigma_{ij}^2)$ *with*

$$\begin{aligned}\sigma_{ij}^2 &= \left[\frac{\pi}{2}\cos\left(\frac{\pi}{2}\nu_{ij}\right)\right]^2(1-\nu_{ij}^2)\\ &= (1-\rho_{ij}^2)\left(\frac{\pi^2}{4} - \arcsin^2(\rho_{ij})\right).\end{aligned}$$

Thus, if two random variables are elliptically distributed, their correlation can be consistently estimated, regardless of any form of heteroscedasticity, using a transformation of the average sign of their product. This result is useful in financial applications because the sample correlation is an inconsistent estimator when the variances evolve as stochastic processes (as in GARCH and SV models).

THEOREM 3 *Under Assumption* 1, *define* $W_{it} = \mu_i + X_{it}$, *so that, conditionally on* $(\sigma_{1t}, \ldots \sigma_{dt})$, *the random vector* $(W_{1t}, \ldots, W_{dt})$ *is elliptically distributed with median vector* $(\mu_1, \ldots, \mu_d)$. *Moreover, assume that the standard deviations are bounded away from zero (i.e.* $0 < b_i < \sigma_{it}$*). If* $\hat{\mu}_{ni}$ *is a consistent estimator of* μ_i *such that* $\mathbb{E}|\hat{\mu}_{ni} - \mu_i| = O(n^{-1/2})$, *then*

$$\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\mathrm{sign}\left((W_{it}-\mu_i)(W_{jt}-\mu_j)\right) - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\mathrm{sign}\left((W_{it}-\hat{\mu}_{in})(W_{jt}-\hat{\mu}_{jn})\right)\right]^2 = 0.$$

The last results show that if we have a consistent estimator of the median, our sign-correlation enjoys the same properties proved in Theorem 2 if we subtract the estimated median from the observations. Natural candidates for $\hat{\mu}_{in}$ are the sample median, the sample mean (if the population mean exists), or robust location estimators. If in Theorem 3 $\hat{\mu}_{in}$ is the sample median, our estimator is consistent for the correlation (as redefined above) even if no moments exist.

REMARK 3 Standard estimators of the median of a random variable with a symmetric distribution are, in general, sufficiently robust against heteroscedasticity. For example, for the sample mean, a sufficient condition for the weak law of large numbers to apply in the case of heteroscedasticity is (see Theorem 4.4 in Shorack 2000) $n^{-2}\sum_{t=1}^{n}\sigma_t^2 \to 0$. Of course, this condition has to be supplemented with uniform integrability of the sequence of random variables since our theorem requires L_1 consistency and not just convergence in probability.

The necessary and sufficient condition for the sample median of non-identically distributed observations to converge in probability to their common population median is provided by Mizera and Wellner (1998, Theorem 1) and is extremely mild.

However, since Theorem 3 requires convergence in mean of the estimator, we can recall the following results that are sufficient for the convergence in mean of the sample mean and median. Let $X_1, \ldots, X_n$ be independent random variables with common median μ and variances $\sigma_1^2, \ldots, \sigma_n^2$, and let $\bar{X}_n$ and M_n the sample mean and median, then

$$\sqrt{n}\,\frac{\bar{X}_n - \mu}{\bar{\sigma}_n^2} \xrightarrow{d} \mathsf{N}(0,1), \qquad \sqrt{n}\,\frac{M_n - \mu}{\delta_n^2} \xrightarrow{d} \mathsf{N}(0,1),$$

where $\bar{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n}\sigma_t^2$, $\delta_n^2 = \frac{1}{4}(n^{-1}\sum_{t=1}^{n}f_t(\mu))^2$, and $f_t(\mu)$ density of X_t at the median. These conditions imply the L_1 consistency of the respective estimators, provided that $\bar{\sigma}_n^2$ and δ_n^2 are $o(n)$.

The last results can be found in Sen (1968, Section 3(II)), where it is also shown that under our Assumption 1, the asymptotic relative efficiency of the sample median with respect to the sample mean improves under heteroscedasticity. Sen (1968) refers to this property as *robust-efficiency of the sample median for heteroskedastic distributions*.

The above results are valid for independent processes with time-varying variances; however, using the theory of martingale differences, they can be extended to GARCH and stochastic variance processes (which are martingale differences possibly summed to a constant). If $X_1 - \mu, \ldots, X_n - \mu$ is martingale difference sequence and, for some $r \geq 1$, $\sum_{t=1}^{\infty}\mathbb{E}(X_t - \mu)^{2r}/t^{1+r} < \infty$, then the sample mean $\bar{X}_n$ converges almost surely to μ (Theorem 3.76 in White 2000). The Lindeberg-Feller central limit theorem applies if we assume that the sample variance is consistent for the asymptotic variance of the sample mean: using the same notation as above, it must hold (Theorem 5.24 in White 2000)

$$n^{-1}\sum_{t=1}^{n}(X_t - \mu)^2/\bar{\sigma}_n^2 - 1 \xrightarrow{p} 0.$$

As for the sample median of dependent processes, there is rich literature involving the sample quantiles under different types of dependence: m-dependent sequences (Sen 1968), φ-mixing processes (Sen 1972), functions of i.i.d. sequences (Wu 2005).

The above discussion let us conclude that the sample mean and median are consistent for common processes used in finance, such as GARCH and Stochastic Volatility. The sample median should be preferred for its robustness to heteroscedasticity and heavy tails.

REMARK 4 The square matrix $\hat{\boldsymbol{V}}_n = \{\hat{v}_{ijn}\}$ is positive semi-definite by construction, while $\hat{\boldsymbol{R}}_n = \{\hat{\rho}_{ijn}\}$ is only asymptotically so. Eventual negative definite estimates can be corrected using nearest correlation matrix methods such as those proposed by Knol (1989) or Higham (2002). Our application in section 4 uses the algorithm of Higham (2002).

2.2. *Simulation experiments*

We use Monte Carlo simulations to assess the accuracy of the asymptotic approximation for the distribution of our estimator by generating sample paths from three different models. We compare its performance with Pearson's and Kendall's sample correlations, which are inconsistent under heteroscedasticity. Correlations are estimated for sample sizes ranging from $n = 100$ to $n = 5000$ under the hypothesis of known and unknown median/mean. In the latter case, we apply our estimator on the deviations of the observations from their sample median; Pearson's correlation is computed the usual way (i.e. using the sample mean), while Kendall's correlation does not need any centering (i.e. it is invariant to location shifts).

EXPERIMENT 1 The first simulation is based on a stochastic volatility process (Harvey *et al.* 1994) with non-mean-reverting variance. The data-generating process is

$$\begin{aligned}
\boldsymbol{h}_t &= \boldsymbol{h}_{t-1} + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \text{i.i.d. } \mathsf{N}\left(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_2\right) \\
\boldsymbol{y}_t &= \begin{bmatrix} \exp(h_{1t}/2) & 0 \\ 0 & \exp(h_{2t}/2) \end{bmatrix} \boldsymbol{\varepsilon}_t, \\
\boldsymbol{\varepsilon}_t &\sim \text{i.i.d. } \mathsf{N}\left(\boldsymbol{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right).
\end{aligned}$$

We show results based on 10 000 sample paths of time series of lengths $n = 100, 101, \ldots, 5000$ with parameters $\sigma = 0.04$, $\rho = 0.7$ and $h_{1,0} = h_{2,0} = 0$.

Figure 1 compares the Monte Carlo percentiles (5% and 95%) of our sign-based estimator with those based on the asymptotic approximation for sample sizes up to $n = 5000$: the lines appear indistinguishable even in small samples. The same (Monte Carlo) percentiles are also plotted for the sample correlation, which is clearly not converging to any point.

EXPERIMENT 2 The second process we consider is the Constant Conditional Correlation GARCH (Bollerslev 1990), with conditional Student's t distribution. The data-generating process is

$$\begin{aligned}
\boldsymbol{y}_t &= \begin{bmatrix} \sigma_{1,t} & 0 \\ 0 & \sigma_{2,t} \end{bmatrix} \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \text{i.i.d. } \mathrm{t}_5\left(\boldsymbol{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right), \\
\sigma^2_{1,t+1} &= 0.002 + 0.05 y^2_{1,t} + 0.89 \sigma^2_{1,t} \\
\sigma^2_{2,t+1} &= 0.001 + 0.06 y^2_{2,t} + 0.90 \sigma^2_{2,t},
\end{aligned}$$

with $\rho = 0.8$.

Figure 2 illustrates the behavior of the estimators for sample sizes ranging from 100 to 5000 observations. Even though, in this example, the marginal variances of the two GARCH processes are finite, the behavior of sample correlation is not too dissimilar from that of the previous simulation. Again, the asymptotic approximation of the distribution of our estimator is rather accurate even in moderate-size samples, with a slight skewness towards 0 for the case of estimated medians when the sample is small.

EXPERIMENT 3 As the last example, we consider a simple Gaussian mixture with constant correlation $\rho = -0.9$, where the bivariate vector $\boldsymbol{x}_t$ is sampled from the following two distributions with equal probability:

$$\mathsf{N}\left(\boldsymbol{0}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right), \quad \mathsf{N}\left(\boldsymbol{0}, \begin{bmatrix} 1 & 5\rho \\ 5\rho & 25 \end{bmatrix}\right).$$

Figure 3 summarizes the results of this third experiment.

As the modulus of the population correlation increases (in this experiment, it is 0.9), the asymptotic approximation of the finite sample distribution becomes less accurate in moderate samples ($n < 500$), especially when the median is estimated. In this last case, there is also a slight bias of the SIGN COR estimator towards 0. Again, the asymptotic approximation of the distribution of our estimator is satisfactory even in moderate samples. It is simple to prove that the sample correlation is converging to -0.75 resulting inconsistent.

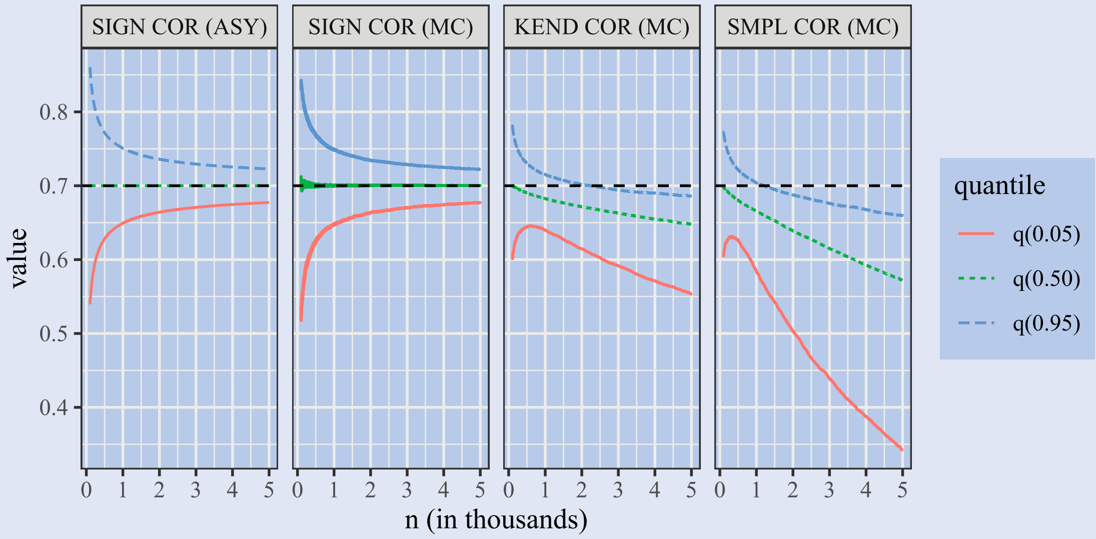
3. Application to high-frequency financial returns

Financial returns are well known to be heteroskedastic and leptokurtic, features that make Pearson's correlation unreliable. Moreover, elliptical distributions and copulas are often used to model vectors of financial returns. Under these conditions, our estimator is expected to be an excellent choice for estimating correlations.

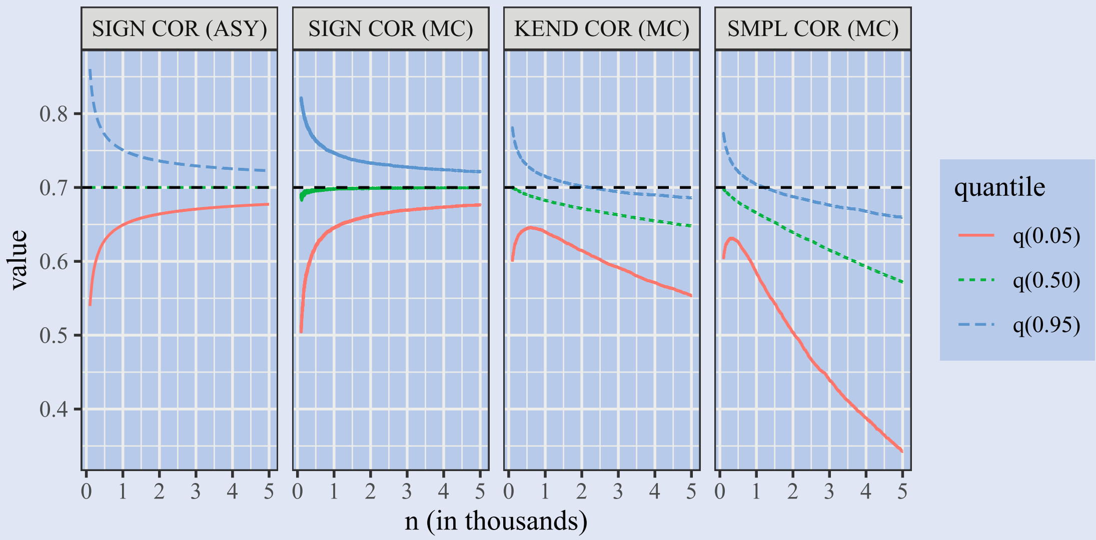
We use 5-minute financial returns of the S&P100 constituents ranging from 2020-09-01 to 2020-10-30 downloaded from FactSet®. Two months is a relatively short period and it is reasonable to believe that correlations among stock returns remained constant. On the other hand, intra-day volatility can show patterns linked to the time of the day and a small number of extreme values following unexpected news on a company or its economic environment. Figure 4 depicts the returns of the first five constituents (in alphabetical order) of S&P100 for the two months in our sample.

We compare the empirical distribution of the sample correlation and of our sign-based estimator using the following re-sampling scheme.

(1) In computing correlations, we keep only return pairs with non-zero values: both prices had to change in the 5-minute span.
(2) We draw random samples of $n/2$ time points, where n is the number of available return pairs.
(3) We estimate correlations both on the selected sample and on the complement of the selected sample.



(a) Known mean/median.



(b) Estimated mean/median: sample median for SIGN COR.

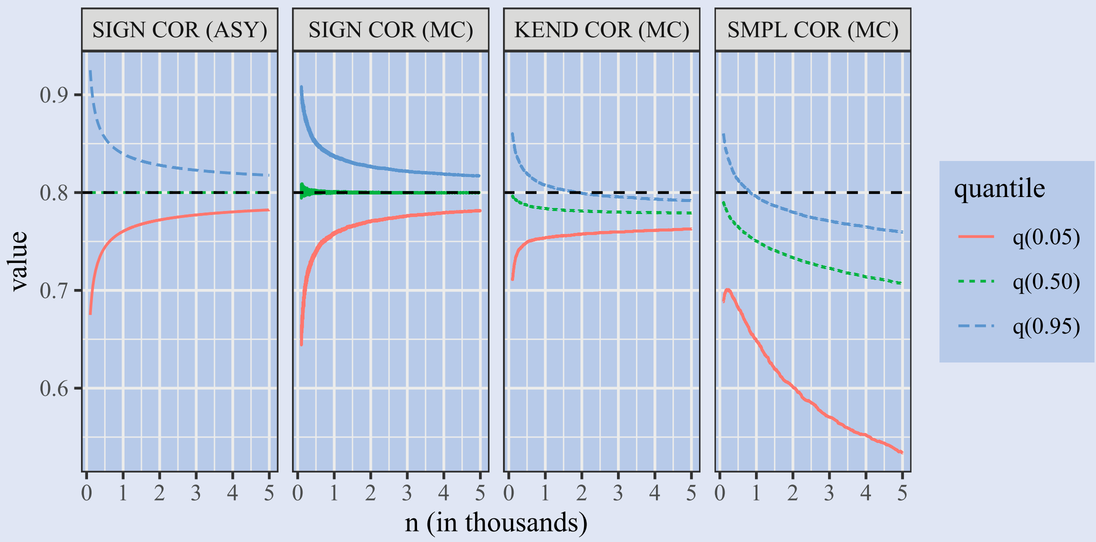
Figure 1. Percentiles (5%, 50%, 95%) of the asymptotic and Monte Carlo distribution of our estimator (SIGN COR) and Monte Carlo percentiles of the sample correlation (SMPL COR) and Kendall-based correlation (KEND COR) for the stochastic volatility model. (a) Known mean/median and (b) Estimated mean/median: sample median for SIGN COR.

Figure 5 depicts kernel densities for the sample correlation and for our sign-based estimator. Our estimates are centered at reasonable values and seem rather precise, while sample correlations are often multi-modal and spread over wide ranges of values. The case of GE-KO is rather odd: 50% of the sample correlations are below -0.50, with a mean of -0.49, while our estimates are centered at 0.30.

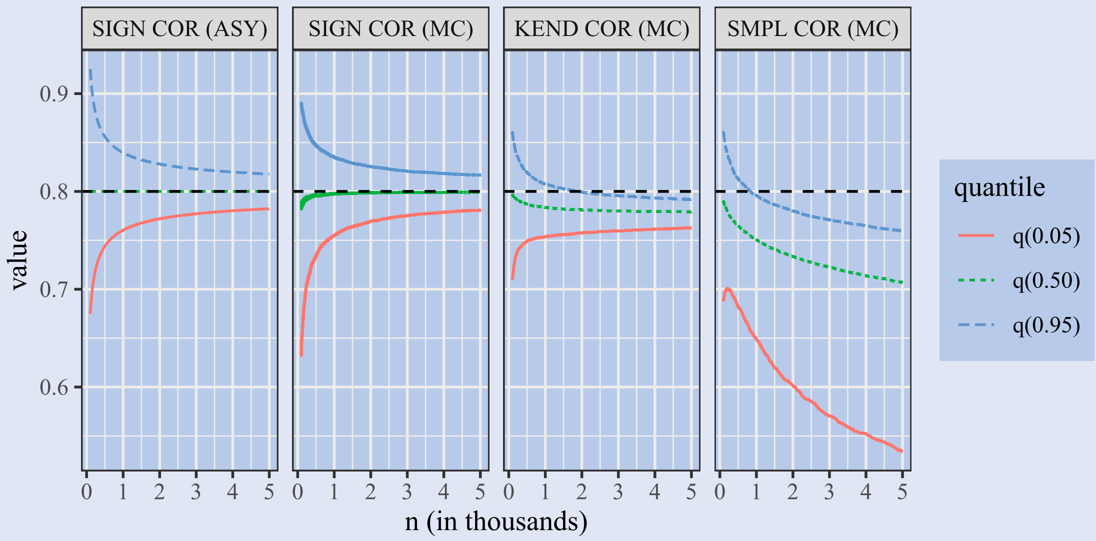
We also computed the difference between the estimate based on the selected sample and its complement (not reported here). The two-sample differences of our estimates are well concentrated around 0, while sample correlations tend to be multi-modal and spread over a wide range of values.

Since, as figure 4 highlights, there are some stocks for which few extreme returns dominate the sums in sample variances and covariances, we computed Pearson's correlations also on two trimmed samples, where absolute returns larger than 10% and 5% were removed.† Figure 5 plots kernel densities also for these two trimmed sample correlations. The more the returns are trimmed, the closer the sample correlation gets to our sign-based estimator. Not knowing the "real" population correlation, we can interpret the gradual approach of the trimmed correlation to our estimator as an indicator of the reliability of our method.

† Notice that in our notation the 5% trimmed correlation is more trimmed than the 10% correlation as percentages refer to returns and not to the share of omitted observations.



(a) Known mean/median.



(b) Estimated mean/median: sample median for SIGN COR.

Figure 2. Percentiles (5%, 50%, 95%) of the asymptotic and Monte Carlo distribution of our estimator (SIGN COR) and Monte Carlo percentiles of the sample correlation (SMPL COR) and Kendall-based correlation (KEND COR) for the Constant Conditional Correlation GARCH. (a) Known mean/median and (b) Estimated mean/median: sample median for SIGN COR.

4. Time-varying correlations

We can easily adapt the exponentially weighted moving average (EWMA or exponential smoothing) technique to make our correlation estimator time-varying. Let us consider the random quantities

$$S_{ij,t} = \text{sign}\left((X_{it} - \hat{\mu}_i)(X_{jt} - \hat{\mu}_j)\right), \quad P_{ij,t} = (S_{ij,t} + 1)/2$$

where $\hat{\mu}_k$ is the median of X_{kt} or a consistent estimator thereof, and t is the time index. If the random pair (X_{it}, X_{jt}) is absolutely continuous, then S_{ijt} takes values in $\{-1, 1\}$ and P_{ijt} in $\{0, 1\}$ and, thus, the latter follows a Bernoulli distribution with probability of success

$$p_{ij,t} = \mathbb{E}(P_{ij,t}).$$

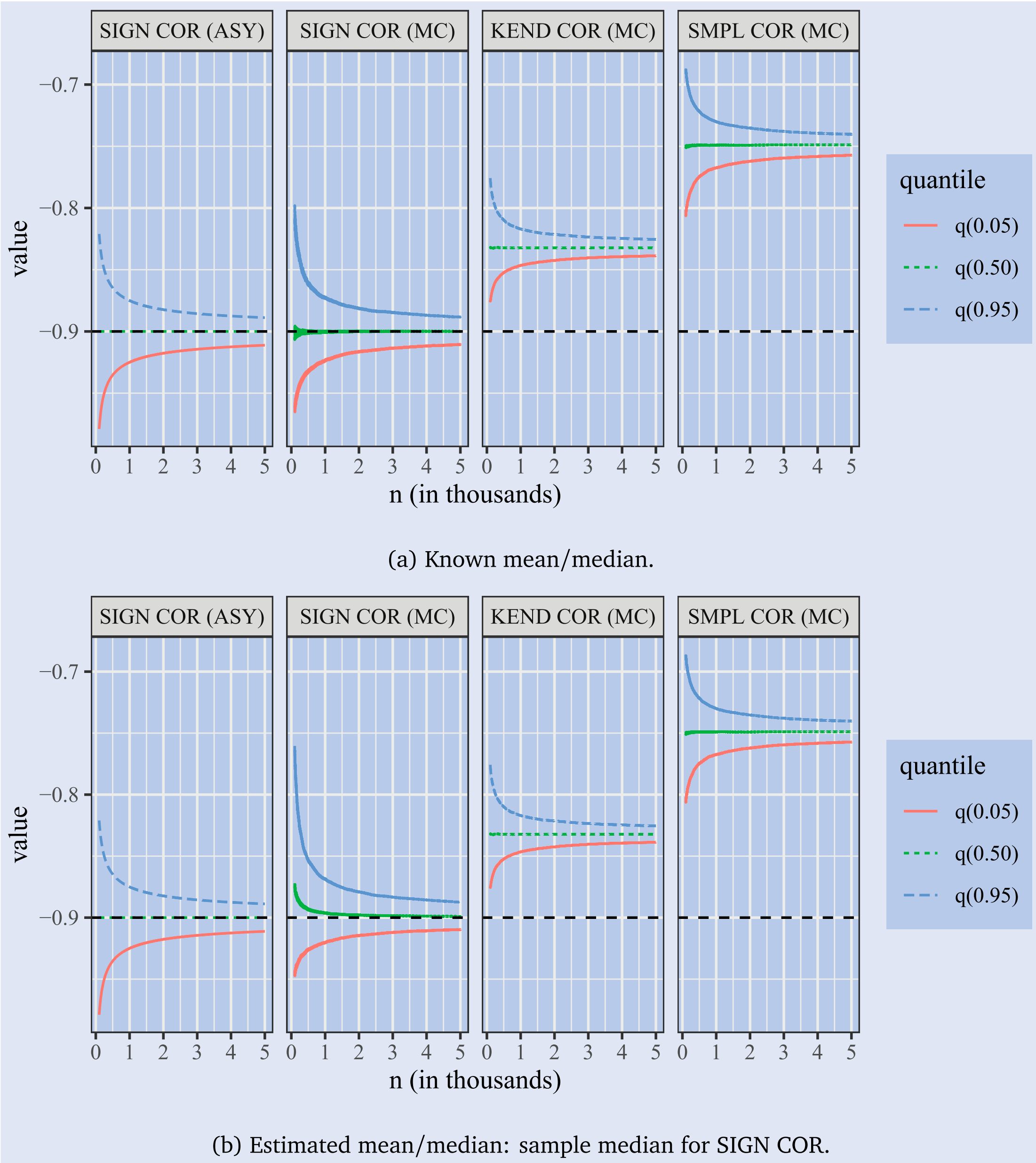
Assuming that (X_{it}, X_{jt}) follows an elliptical distribution, from Theorem 1, we have that the correlation among the two random variables is given by

$$\rho_{ij,t} = \sin\left(\frac{\pi}{2}\left(2p_{ij,t} - 1\right)\right).$$

The EWMA recursion to estimate a time-varying $p_{ij,t}$ is given by

$$\hat{p}_{ij,t} = \omega P_{ij,t} + (1 - \omega)\hat{p}_{ij,t-1}, \quad \forall i, j, t \tag{4}$$

and assuming $\hat{p}_{ij,t}$ to be a good predictor of the probability of success at time $t + 1$, we can estimate the unknown parameter



(a) Known mean/median.

(b) Estimated mean/median: sample median for SIGN COR.

Figure 3. Percentiles (5%, 50%, 95%) of the asymptotic and Monte Carlo distribution of our estimator (SIGN COR) and Monte Carlo percentiles of the sample correlation (SMPL COR) and Kendall-based correlation (KEND COR) for Gaussian mixture model. (a) Known mean/median and (b) Estimated mean/median: sample median for SIGN COR.

ω by maximizing the (quasi) log-likelihood

$$\ell(\omega) = \sum_{i,j,t} P_{ij,t} \log\left(\hat{p}_{ij,t|t-1}\right) + \left(1 - P_{ij,t}\right) \log\left(1 - \hat{p}_{ij,t|t-1}\right),$$

where we set $\hat{p}_{ij,t|t-1} = \hat{p}_{ij,t-1}$. A formal derivation of the recursion (4) can be obtained assuming that the probability $p_{ij,t}$ conditional on the observations $P_{ij,s}$, $s \leq t$, is distributed as Beta(a_t, b_t) and $p_{ij,t+1}$, conditional on the same information is Beta$(\omega a_t, \omega b_t)$ as in Harvey (1989, Section 6.6.2).

We implement four Monte Carlo experiments to assess how this filter behaves in extracting known time-varying correlations of arbitrary shapes. We simulate time series pairs whose correlation paths evolve in the following ways: for $t = 1, 2, \ldots, n$ and $n = 2000$,

rw $\rho_t = (1 + \exp(W_t))^{-1}$, with $W_t = \sum_{i=1}^{t} 0.05 \cdot Z_i$ and Z_i i.i.d. sequence of standard normal random variables;
sin $\rho_t = \sin(2\pi t/n)$;
step $\rho_t = -0.6\mathbb{I}(t < n/2) + 0.9\mathbb{I}(t \geq n/2)$;
sigmoid $\rho_t = (1 + \exp(-5[t/n - 0.5]))^{-1}$;

where $\mathbb{I}(\cdot)$ denotes the indicator function. We generate the random numbers

$$T_{0t} \sim t_d,$$
$$T_{1t} \sim t_d,$$
$$T_{2t} = \rho_t T_{1t} + \sqrt{1 - \rho_t^2}\, T_{0t},$$

where t_d denotes the Student's t distribution with d degrees of freedom. We also generate time-varying variances that evolve

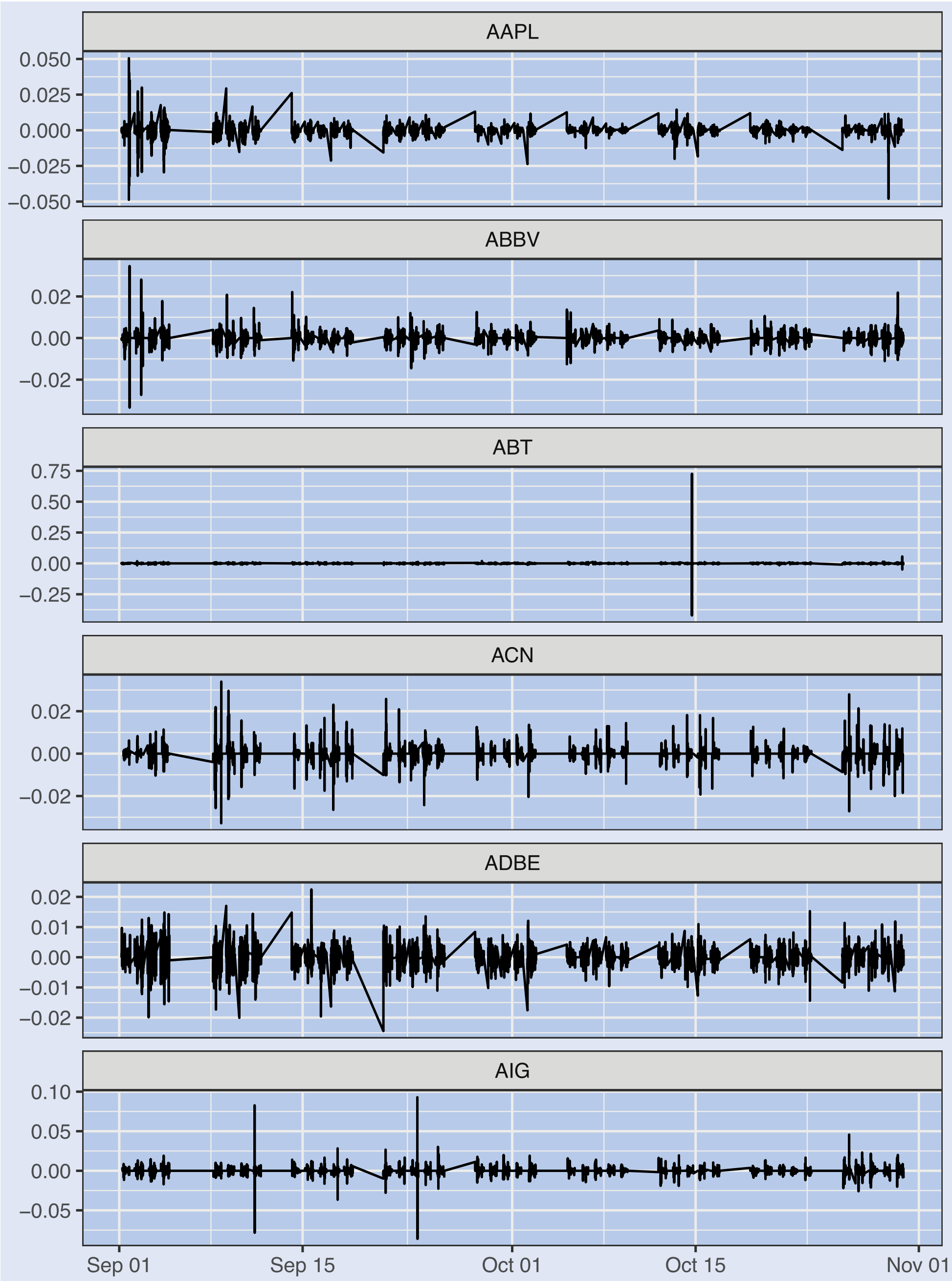


Figure 4. Five-minute returns of the first five constituents (in alphabetical order) of S&P100.

according to the following equation

$$\sigma_{i,t} = \exp(U_{it})/1000,$$

where U_{it} is a Gaussian random walk driven by white noise with a standard deviation of 0.01 and, finally, we compute the two time series

$$X_{1t} = \sigma_{1t}T_{1t}, \quad X_{2t} = \sigma_{2t}T_{2t},$$

from which we estimate the correlation paths.

We compare the behavior of our approach to that of the very popular GARCH-DCC model of Engle (2002). Neither our approach nor the DCC model is optimal under the above data-generating process: both are used as non-parametric alternatives to extract unobservable information about a time-varying correlation.

Figure 6 depicts one simulation for each correlation path and the estimates obtained with the two methods. Visual inspection suggests that the two estimates are comparable. However, table 1 reports mean absolute errors (MAE) and root mean square errors (RMSE) for a number of simulations.† Being very robust, our approach is generally somewhat less efficient than DCC unless the tails of the return distribution become extremely thick (bottom part of table 1). However, our filter is very simple to implement, very fast to compute,

† The DCC has been estimated using the *rmgarch* package (Ghalanos 2019) for *R* (R Core Team 2020).

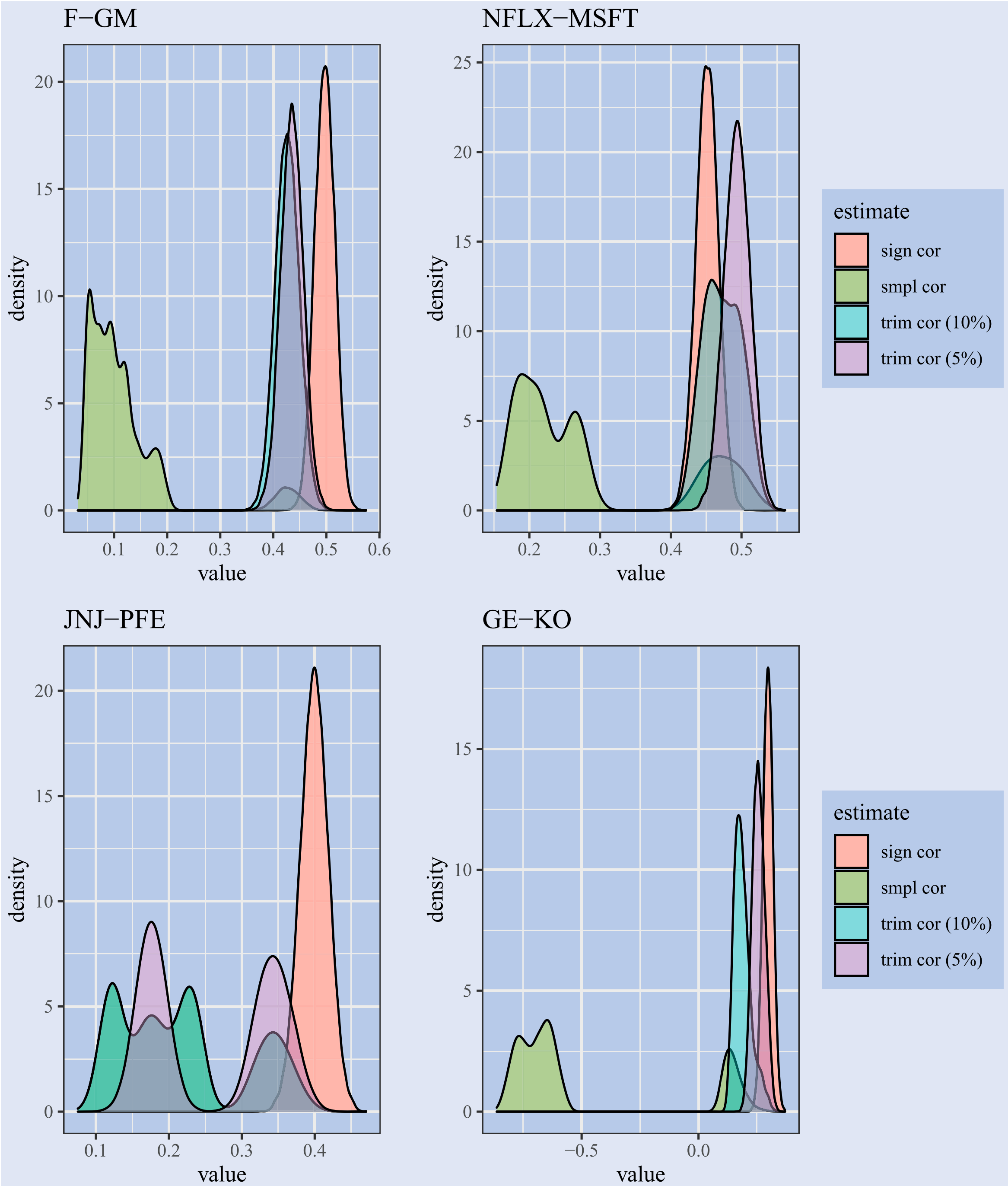


Figure 5. Empirical (kernel) densities of the correlation estimates for various stock pairs. The table with symbols and company names can be retrieved at https://en.wikipedia.org/wiki/S\&P_100.

and computationally robust, and this makes our approach feasible even when the size of the portfolio is too large for the DCC.

As an example application, we applied the DCC model and our approach coupled with univariate GARCH(1, 1) models to the first 10 stocks in our SP100 dataset† to build global minimum variance (GMV) portfolios. We used the daily returns of the full years 2010-2018 for estimating the models and built the GMV portfolios using daily one-step-ahead predictions of the covariance matrices. We consider both unconstrained portfolios and portfolios without short positions.

Figure 7 depicts the recursively computed annualized volatility (%) of the daily updated portfolios for the two approaches with and without non-negativity constraints. Overall, the DCC does slightly better, even though, until the jump in volatility in March 2020 due to the spread of the Covid-19 pandemic,‡ the portfolio based on our method showed lower volatility.

† The symbols of the selected stocks are RTX, LMT, WBA, DHR, SO, MSFT, UNH, JPM, KO, SBUX.

‡ The CBOE Volatility Index (VIX) peaks on day 2020-03-16.

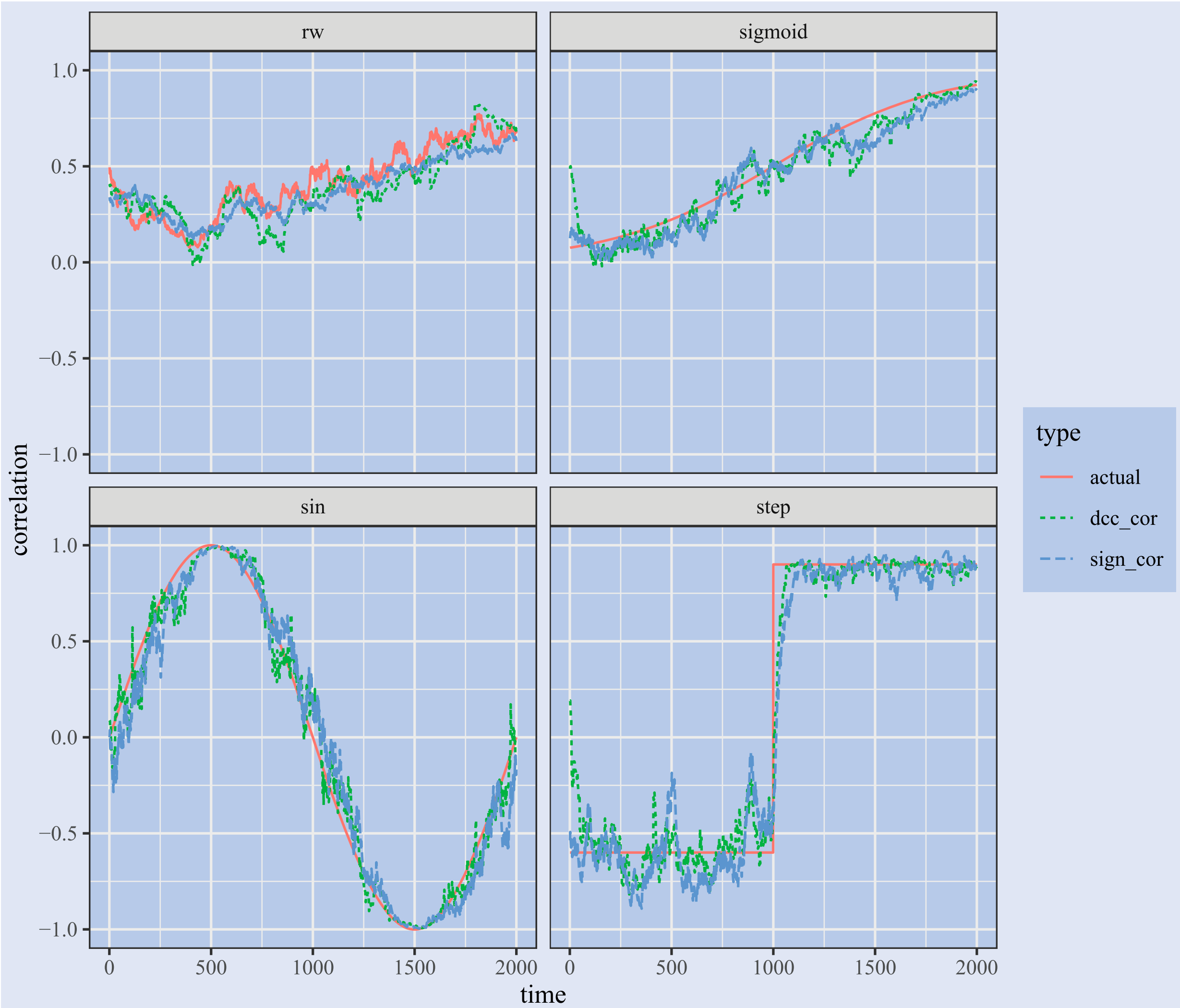


Figure 6. Estimates of time-varying correlation paths using our approach and the DCC model.

Table 1. Mean errors in estimating time-varying correlations.

		MAE			RMSE		
Correlation type	df	DCC	SCOR	RATIO	DCC	SCOR	RATIO
rw	8	0.06	0.08	1.31	0.09	0.11	1.27
sigmoid	8	0.07	0.07	1.06	0.10	0.09	0.96
sin	8	0.08	0.11	1.34	0.12	0.15	1.24
step	8	0.08	0.10	1.25	0.15	0.18	1.20
rw	4	0.08	0.09	1.16	0.10	0.11	1.11
sigmoid	4	0.08	0.08	0.99	0.11	0.11	0.95
sin	4	0.09	0.10	1.17	0.12	0.14	1.12
step	4	0.10	0.11	1.09	0.18	0.18	1.03
rw	2	0.17	0.09	0.57	0.22	0.12	0.56
sigmoid	2	0.16	0.08	0.53	0.20	0.11	0.53
sin	2	0.17	0.11	0.63	0.23	0.15	0.63
step	2	0.25	0.11	0.44	0.37	0.19	0.52

Figure 8 plots the annualized Sharpe ratios for the GMV portfolios built using the two methods. Again, the portfolio based on our method performed better until the jump in March 2020. After that date, the Sharpe ratios of the two portfolios become almost identical.

The no-short-selling constraint does not seem to significantly influence the portfolio volatility and Sharpe ratio under any model.

5. Conclusions

We proposed a simple and computationally efficient estimator for the correlation parameter of continuous elliptically distributed random variables and derived its asymptotic properties. Our estimator is scale-invariant and robust to outliers and heteroscedasticity of any kind.

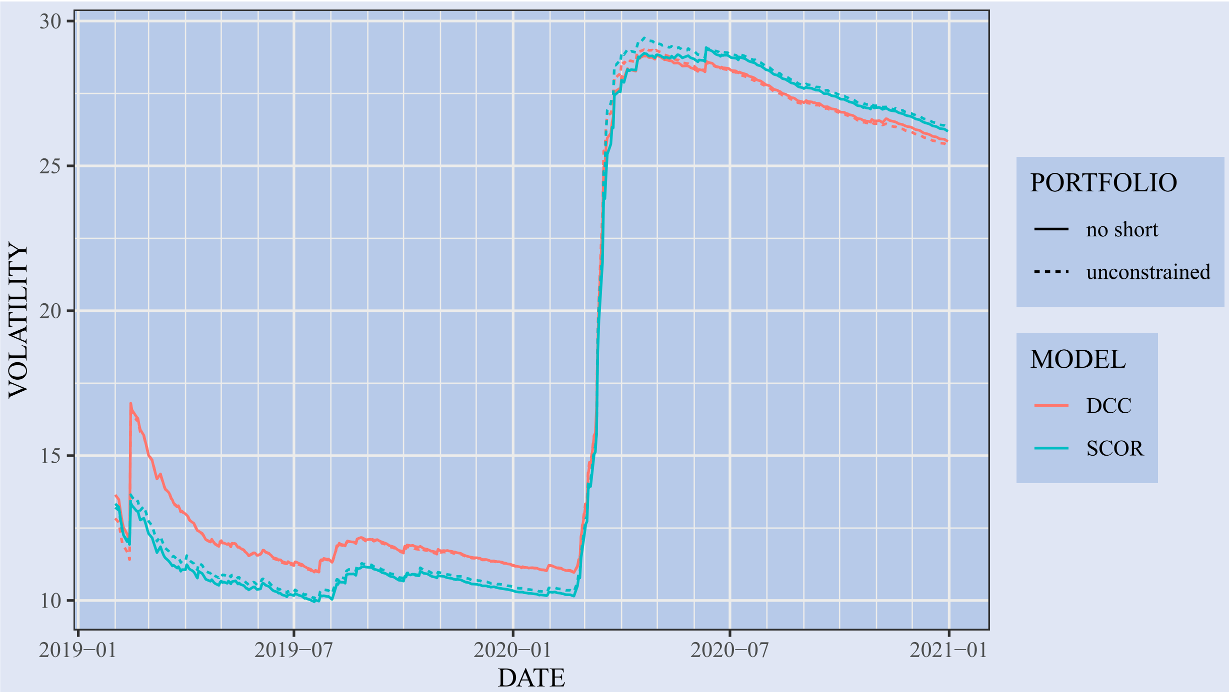


Figure 7. Recursively computed annualized volatility (%) for GMV portfolios based on DCC and our sign-correlation approach coupled with GARCH(1, 1) models.

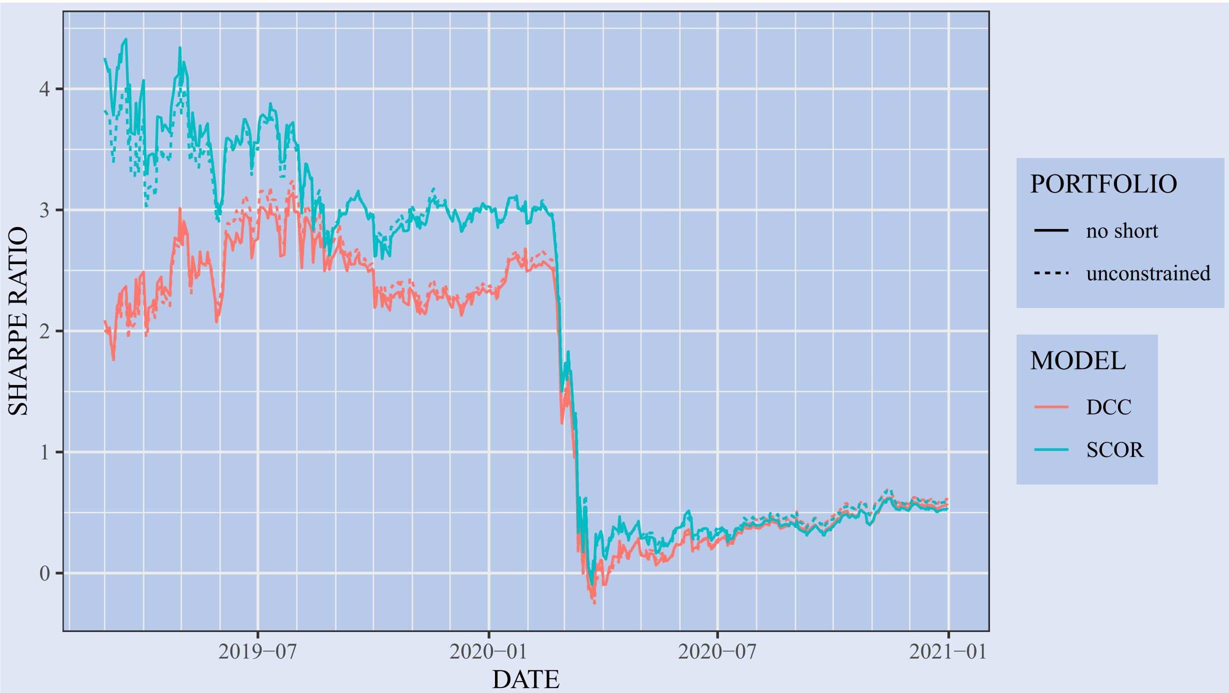


Figure 8. Recursively computed annualized Sharpe ratio for GMV portfolios based on DCC and our sign-correlation approach coupled with GARCH(1, 1) models.

Simulations show that the asymptotic distribution of our estimator is a good approximation of the small sample distribution, even for moderate n. This result is not surprising as asymptotic normality is obtained by averaging random variables (random signs) possessing moments of any order.

Under the hypothesis of homoscedasticity and elliptical copula, our estimator is consistent for the population Kendall's tau, even though the sample Kendall's tau is more efficient than our statistic under these conditions. However, enjoying a computational complexity of $O(n)$, our estimator can be a valid substitution for Kendall's tau when its computation becomes infeasible, or the data are heteroscedastic.

Financial time series analysis is the field of application for which we have developed this estimator; indeed, heteroscedasticity is a common feature of financial prices and returns. Through three simple simulation experiments, we show how poorly the sample correlation performs under GARCH and SV models and in scale mixtures, as opposed to our estimator, which preserves its consistency and approximate normality.

We apply our sign-based estimator and the sample correlation to intra-day financial returns, which display periodic volatility patterns and extreme outliers. The performance of the estimators is assessed by bootstrapping the five-minute returns. The bootstrap distribution of our estimator is well concentrated around reasonable correlation values, while sample correlations spread over a wide range of values and are often multi-modal. Trimming, by reducing heteroscedasticity and cutting outliers, makes the sample correlation's empirical distribution closer to that of our estimator, but, as for the case of scale mixtures, it seems to be somewhat biased towards zero.

Finally, we propose a simple way to obtain time-varying correlation estimates based on sign-correlations and EWMA and compare it to the popular DCC model of Engle (2002). The DCC model produces correlation estimates that are generally more efficient unless the return distributions are extremely thick-tailed. However, a simple application to global minimum variance portfolios shows that none of the two methods dominates the other. Our estimator's robustness, computational speed, and stability should make it particularly interesting to practitioners and financial software developers.

Acknowledgments

The author gratefully acknowledges the *Bicocca Data Science Lab* for providing computational resources.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Matteo Pelagatti http://orcid.org/0000-0002-1860-7535

References

Alizadeh, S., Brandt, M.W. and Diebold, F.X., Range-based estimation of stochastic volatility models. *J. Finance*, 2002, **57**(3), 1047–1091.

Asai, M., McAleer, M. and Yu, J., Multivariate stochastic volatility: A review. *Econom. Rev.*, 2006, **25**(2–3), 145–175.

Bauwens, L., Laurent, S. and Rombouts, J.V.K., Multivariate GARCH models: A survey. *J. Appl. Econom.*, 2006, **21**(1), 79–109.

Blomqvist, N., On a measure of dependence between two random variables. *Ann. Math. Stat.*, 1950, **21**(4), 593–600. http://www.jstor.org/stable/2236609

Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *J. Econom.*, 1986, **31**(3), 307–327.

Bollerslev, T., Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. *Rev. Econ. Stat.*, 1990, **72**(3), 498–505. http://www.jstor.org/stable/2109358

Cambanis, S., Huang, S. and Simons, G., On the theory of elliptically contoured distributions. *J. Multivar. Anal.*, 1981, **11**, 368–385.

Davidson, J., *Stochastic Limit Theory*, 1994 (Oxford University Press: Oxford).

Dias, A., Salmon, M. and Adcock, C., *Copulae and Multivariate Probability Distributions in Finance*, 2018 (Routledge: Abingdon).

Diongue, A.K., Guégan, D. and Wolff, R.C., BL-GARCH models with elliptical distributed innovations. *J. Stat. Comput. Simul.*, 2009, **80**(7), 775–791.

Engle, R.F., Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 1982, **50**(4), 987–1007.

Engle, R., Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econ. Stat.*, 2002, **20**(3), 339–350.

Fang, K.-T., Kotz, S. and Hg, K.-W., *Symmetric Multivariate and Related Distributions*, 1987 (Chapman & Hall: New York).

Fiorentini, G., Sentana, E. and Calzolari, G., Maximum likelihood estimation and inference in multivariate conditionally heteroscedastic dynamic regression models with StudenttInnovations. *J. Bus. Econ. Stat.*, 2003, **21**(4), 532–546.

Ghalanos, A., *rmgarch: Multivariate GARCH Models*. R package version 1.3-7, 2019.

Hall, P. and Heyde, C.C., *Martingale Limit Theory and Its Applications*, 1980 (Academic Press: New York).

Hamada, M. and Valdez, E.A., CAPM and option pricing with elliptically contoured distributions. *J. Risk Insur.*, 2008, **75**(2), 387–409.

Harvey, A., *Forecasting Structural Time Series and the Kalman Filter*, 1989 (Cambridge University Press: Cambridge).

Harvey, A., Ruiz, E. and Shephard, N., Multivariate stochastic variance models. *Rev. Econ. Stud.*, 1994, **61**(2), 247–264.

Heston, S.L., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 1993, **6**(2), 327–343.

Higham, N.J., Computing the nearest correlation matrix—A problem from finance. *IMA J. Numer. Anal.*, 2002, **22**(3), 329–343.

Hull, J. and White, A., Hedging the risks from writing foreign currency options. *J. Int. Money Finance*, 1987, **6**(2), 131–152.

Ingersoll, J.E., *Theory of Financial Decision Making*, 1987 (Rowman & Littlefield Pub Inc: Lanham).

Knol, D.L., Least-squares approximation of an improper correlation matrix by a proper one. *Psychometrika*, 1989, **54**(1), 53–61.

Lindskog, F., McNeil, A. and Schmock, U., Kendall's tau for elliptical distributions. In *Credit Risk: Measurement, Evaluation and Management*, edited by G. Bol, G. Nakhaeizadeh, S. T. Rachev, T. Ridder, K.-H. Vollmer, pp. 149–156, 2003 (Physica-Verlag: Heidelberg).

Mizera, I. and Wellner, J.A., Necessary and sufficient conditions for weak consistency of the median of independent but not identically distributed random variables. *Ann. Stat.*, 1998, **26**(2), 672–691.

Nadarajah, S. and Pogány, T.K., On the distribution of the product of correlated normal random variables. *C. R. Math.*, 2016, **354**(2), 201–204.

Nelson, D.B., Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 1991, **59**(2), 347–370.

Owen, J. and Rabinovitch, R., On the class of elliptical distributions and their applications to the theory of portfolio choice. *J. Finance*, 1983, **38**(3), 745–752.

Pelagatti, M.M. and Rondena, S., Dynamic conditional correlation with elliptical distributions. *SSRN Electron. J.*, 2004. Available at SSRN: https://ssrn.com/abstract=888732

R Core Team, *R: A Language and Environment for Statistical Computing*, 2020 (R Foundation for Statistical Computing: Vienna, Austria). https://www.R-project.org/.

Sen, P.K., Asymptotic normality of sample quantiles for m-dependent processes. *Ann. Math. Stat.*, 1968, **39**(5), 1724–1730.

Sen, P.K., On the bahadur representation of sample quantiles for sequences of φ-mixing random variables. *J. Multivar. Anal.*, 1972, **2**(1), 77–95.

Shorack, G.R., *Probability for Statisticians*, Springer Texts in Statistics, 2000 (Springer: New York).
White, H., *Asymptotic Theory for Econometricians*, revised edition edn, 2000 (Emerald Publishing Limited: Bingley).
Wu, W.B., On the bahadur representation of sample quantiles for dependent sequences. *Ann. Stat.*, 2005, **33**(4), 1934–1963.
Zhou, X., Malioutov, D., Fabozzi, F.J. and Rachev, S.T., Smooth monotone covariance for elliptical distributions and applications in finance. *Quant. Finance*, 2014, **14**(9), 1555–1571.

Appendix

Proof of Theorem 1

We need the following results for the proof of Theorem 1. The symbol $E_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ denotes a p-dimensional elliptical distribution with location vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and characteristic function ϕ.

THEOREM 4 *(Representation of an elliptically distributed random vector) Let $\boldsymbol{x}$ be a p-dimensional random vector; then $\boldsymbol{x}$ is distributed as $E_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ with $\operatorname{rank}(\boldsymbol{\Sigma}) = k \le p$ if and only if there is a random variable $R \ge 0$, independent of the k-dimensional random vector $\boldsymbol{u}$ uniformly distributed on the unit hypersphere $\{z \in \mathbb{R}^k | z^\top z = 1\}$ and a $p \times k$ matrix $\boldsymbol{A}$ with $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{\Sigma}$, such that*

$$\boldsymbol{x} \overset{d}{=} \boldsymbol{\mu} + R\boldsymbol{A}\boldsymbol{u}.$$

Proof For proof, see Cambanis *et al.* (1981) or the monograph by Fang *et al.* (1987). ∎

THEOREM 5 *(Distribution of the product of two standard normal variables) Let (X, Y) be a bivariate normal random vector with zero means, unit variances, and correlation coefficient ρ. Then, the probability density function of $Z = XY$ is*

$$f_Z(z) = \frac{1}{\pi\sqrt{1-\rho^2}} \exp\left[\frac{\rho z}{1-\rho^2}\right] K_0\left(\frac{|z|}{1-\rho^2}\right),$$

for $-\infty < z < \infty$, where $K_0(\cdot)$ denotes the modified Bessel function of the second kind of order zero:

$$K_0(x) = \int_0^\infty \cos(x \sinh t)\mathrm{d}t = \int_0^\infty \frac{\cos(xt)}{\sqrt{t^2+1}}\mathrm{d}t.$$

Proof See Nadarajah and Pogány (2016). ∎

COROLLARY 6 *If (X, Y) are jointly normally distributed with zero means, unit variances, and correlation ρ, then the probability $\Pr(Z > 0) = \Pr(X > 0 \wedge Y > 0) + \Pr(X < 0 \wedge y < 0)$ is given by*

$$\int_0^\infty f_Z(z)\mathrm{d}z = 1 - \frac{\arccos(\rho)}{\pi} = \frac{1}{2} + \frac{\arcsin(\rho)}{\pi}.$$

We can now prove Theorem 1.

THEOREM 7 *Let (X, Y) be jointly elliptically distributed with zero location and scale matrix given by*

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

If (X, Y) are jointly normal, then the result is just an application of Corollary 6:

$$\begin{aligned}\mathbb{E}[\operatorname{sign}(XY)] &= \Pr(XY > 0) - \Pr(XY < 0) = 2\Pr(XY > 0) - 1 \\ &= \frac{2}{\pi}\arcsin(\rho).\end{aligned}$$

Theorem 4 generalizes this result to all continuous elliptical random variables. Indeed, in our case (bivariate with zero location and unit scale), the representation of that theorem specializes to

$$\begin{bmatrix} X \\ Y \end{bmatrix} = R\boldsymbol{A}\boldsymbol{u} = R\begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}\begin{bmatrix} \cos(U) \\ \sin(U) \end{bmatrix},$$

with U uniformly distributed on $[0, 2\pi)$. Since the random variable R is positive (almost surely), the signs of X and Y do not depend on R:

$$\begin{aligned}\Pr(X > 0 \wedge Y > 0) &= \Pr(R\boldsymbol{A}\boldsymbol{u} > \boldsymbol{0}) = \Pr(\boldsymbol{A}\boldsymbol{u} > \boldsymbol{0}) \\ &= \Pr\left(\tilde{R}\boldsymbol{A}\boldsymbol{u} > \boldsymbol{0}\right) = \frac{1}{2} + \frac{\arcsin(\rho)}{\pi},\end{aligned}$$

where $\tilde{R}$ is the random variable that makes the elliptical distribution Gaussian ($\tilde{R}$ is chi-square with two degrees of freedom).

A counterexample for Kendall's tau inconsistency under heteroscedasticity

Let

$$\hat{\tau}_{ijn} := \frac{2}{n(n-1)} \sum_{t=2}^{n} \sum_{s=1}^{t-1} \operatorname{sign}\left((X_{it} - X_{is})(X_{jt} - X_{js})\right) \qquad \text{(A1)}$$

be the sample Kendall's tau.

Assume that the vector z_t of Assumption 1 is Gaussian and consider the generic summand of equation A1. For the properties of the normal distribution, conditionally on σ_{it} and σ_{jt} for $t = 1, \ldots, n$, we have

$$\begin{bmatrix} X_{it} - X_{is} \\ X_{jt} - X_{js} \end{bmatrix} \sim \mathsf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{it}^2 + \sigma_{is}^2 & \rho(\sigma_{it}\sigma_{jt} + \sigma_{is}\sigma_{js}) \\ \rho(\sigma_{it}\sigma_{jt} + \sigma_{is}\sigma_{js}) & \sigma_{jt}^2 + \sigma_{js}^2 \end{bmatrix}\right).$$

The correlation between the two Gaussian random variables is, then,

$$\mathsf{COR}\left(X_{it} - X_{is}, X_{jt} - X_{js}\right) = \rho\frac{\sigma_{it}\sigma_{jt} + \sigma_{is}\sigma_{js}}{\sqrt{(\sigma_{it}^2 + \sigma_{is}^2)(\sigma_{jt}^2 + \sigma_{js}^2)}}.$$

The above fraction takes values in the interval (0, 1]. Indeed, by taking the square of the denominator minus the square of the numerator we obtain

$$(\sigma_{it}^2 + \sigma_{is}^2)(\sigma_{jt}^2 + \sigma_{js}^2) - (\sigma_{it}\sigma_{jt} + \sigma_{is}\sigma_{js})^2 = (\sigma_{it}\sigma_{js} - \sigma_{is}\sigma_{jt})^2 \ge 0,$$

with equality holding only if $\sigma_{it}\sigma_{jt} = \sigma_{is}\sigma_{js}$. This last condition is a bit weaker than homoscedasticity, requiring only the product of the two variances to be constant over time. Let us name the above fraction $c_{ij,ts}$, by Theorem 1, for $0 < c_{ij,ts} < 1$ and $0 < \rho < 1$

$$\mathbb{E}\operatorname{sign}\left((X_{it} - X_{is})(X_{jt} - X_{js})\right) = \frac{2}{\pi}\arcsin(\rho_{ij}c_{ij,ts}) < \frac{2}{\pi}\arcsin(\rho_{ij}),$$

from which

$$\mathbb{E}\hat{\tau}_{ijn} < \frac{2}{\pi}\arcsin\left(\rho_{ij}\right),$$

for all n and for any deterministic or stochastic sequences of variances σ_{it} and σ_{jt} for which $\sigma_{it}\sigma_{jt} = \sigma_{is}\sigma_{js}$ does not hold. For a negative correlation coefficient ρ_{ij} the inequality reverses.

Thus, excluding ad-hoc selected sequences of variances for which $c_{ij,ts}$ is equal to one or converging to one, Kendall's tau is biased towards zero, and the bias does not reduce as n diverges.

If we consider the data generating process of Experiment 3 in section 2.2, where X_{it} has unitary variance and X_{jt} has variance equal to 1 or 5 with probability 0.5, with constant correlation equal to −0.9, we can easily compute the expected value of $c_{ij,ts}$:

$$\mathbb{E}c_{ij,ts} = \frac{1}{4}\left(\frac{1+1}{\sqrt{(2\times 2)}}\right) + \frac{1}{4}\left(\frac{1+5}{\sqrt{(2\times 26)}}\right)$$

$$+\frac{1}{4}\left(\frac{5+1}{\sqrt{(26\times 2)}}\right)+\frac{1}{4}\left(\frac{5+5}{\sqrt{(2\times 50)}}\right)$$
$$=\frac{1}{2}\left(1+\frac{3}{\sqrt{13}}\right)\approx 0.916.$$

If the law of large number applies, we can expect $\sin(\pi\hat{\tau}_{ijn}/2)$ to converge to $-0.9\cdot 0.916=-0.824$, as figure 3 seems to confirm.

Proof of Theorem 2

The first point of Theorem 2 is just an application of the strong of large numbers, since $\nu_{ij}=\mathbb{E}\,\mathrm{sign}(X_{it}X_{jt})$ is finite.

The second point is an application of the classical central limit theorem, as

$$\mathrm{Var}[\mathrm{sign}(X_{it}X_{jt})]=\mathbb{E}\,\mathrm{sign}(X_{it}X_{jt})^2-[\mathbb{E}\,\mathrm{sign}(X_{it}X_{jt})]^2=1-\nu_{ij}^2.$$

The third point applies the continuous mapping theorem, being the sine a continuous function.

The fourth point obtains using the delta method.

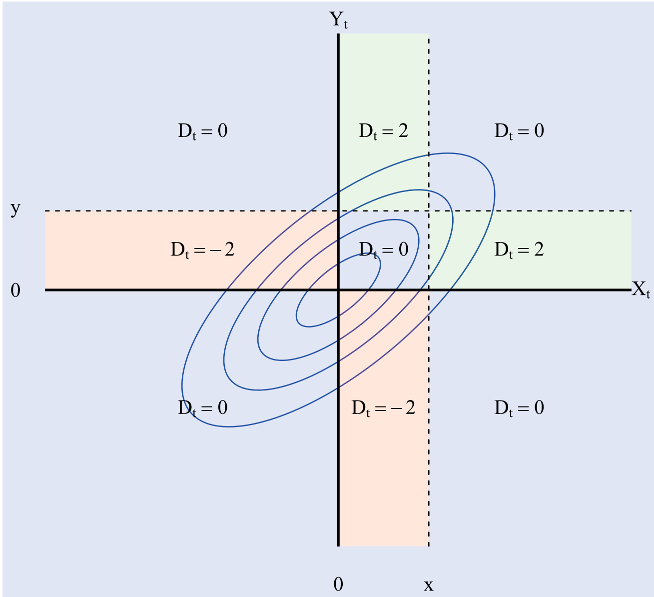


Figure A1. Values of D_t as a function of X_t, Y_t, $\bar{X}_n$ and $\bar{Y}_n$.

Proof of Theorem 3

First of all, we can assume without loss of generality that $\mu_i=0$, if the location estimator is equivariant to shifts (i.e. $loc(X+\mu)=loc(X)+\mu$) as the sample mean and median: if X_i has zero mean and $W_i=X_i+\mu_i$, then

$$W_{it}-\hat{\mu}_{Wi}=X_{it}+\mu_i-\hat{\mu}_{Wi}=X_{it}+\mu_i-\hat{\mu}_{Xi}-\mu_i=X_{it}-\hat{\mu}_{Xi}.$$

To simplify the notation in the proof and drop some indexes, let us rename $X_t=X_{it}$, $Y_t=X_{jt}$, $\bar{X}_n=\hat{\mu}_{Xi}$ and $\bar{Y}_n=\hat{\mu}_{Xj}$, and $X_t=U_t\sigma_t$, $Y_t=V_t\varsigma_t$ with (U_t,V_t) zero-mean i.i.d. process. Define, for any two scalars x, y, the random variable

$$D_t(x,y)=\mathrm{sign}(X_tY_t)-\mathrm{sign}\left((X_t-x)(Y_t-y)\right).$$

We are interested in the asymptotic behavior of $\bar{D}_n=\sum_{t=1}^n D_t(\bar{X}_n,\bar{Y}_n)/\sqrt{n}$. In particular, we want to study the convergence of $\mathbb{E}[\bar{D}_n^2]$ because if this quantity converges to zero, then the asymptotic distribution of the sign-correlation based on consistently estimated medians is the same as the one with known medians.

We can study the random variable $D_t(x,y)$ with the help of figure A1 where, given the values of x and y in a neighborhood of 0, all the outcomes of $D_t(x,y)$ as a function of X_t and Y_t are depicted. Since $D_t(x,y)$ can take only the values $-2,0,2$, $D_t(x,y)^2$ takes only the values 0 and 4.

Now, to keep the reasoning simple, let us assume that x and y are non-negative. We can express $D_t(x,y)$ as

$$\begin{aligned}D_t(x_t,y_t)&=2\mathbb{I}(0<X_t<x,Y<0)-2\mathbb{I}(0<X<x,Y>y)\\&\quad+2\mathbb{I}(X<0,0<Y<y)-2\mathbb{I}(X>x,0<Y<y)\\&=2\mathbb{I}(0<U_t<x_t,V_t<0)-2\mathbb{I}(0<U_t<x_t,V_t>y_t)\\&\quad+2\mathbb{I}(U_t<0,0<V_t<y_t)-2\mathbb{I}(U_t>x_t,0<V_t<y_t),\end{aligned}$$

where we set $x_t=x\sigma_t^{-1}$, $y_t=y\varsigma_t^{-1}$.

Conditioning on the standard deviations, we can take the expectation of D_t (we omit the conditioning in the formulae to keep the notation light):

$$\begin{aligned}\mathbb{E}D_t(x_t,y_t)&=-2\Pr(0<U_t<x_t,V_t<0)\\&\quad+2\Pr(0<U_t<x_t,V_t>y_t)\\&\quad-2\Pr(U_t<0,0<V_t<y_t)\\&\quad+2\Pr(U_t>x_t,0<V_t<y_t)\\&=-2\Pr(0<U_t<x_t,V_t<0)\\&\quad+2\Pr(-x_t<U_t<0,V_t<-y_t)\\&\quad-2\Pr(U_t<0,0<V_t<y_t)\\&\quad+2\Pr(U_t<-x_t,-y_t<V_t<0).\end{aligned}$$

Notice that the last equality is due to the elliptical symmetry of (W,Z). As the next step, we approximate each probability integral by its order-one Taylor expansion about the origin: let $F(x,y)$ be the distribution function of (W,Z), then

$$\Pr(0<U_t<x_t,V_t<0)=\frac{\partial F}{\partial x_t}(0,0)x_t+o(|x_t|)$$
$$\Pr(U_t<0,0<V_t<y_t)=\frac{\partial F}{\partial y_t}(0,0)y_t+o(|y_t|)$$
$$\Pr(-x_t<U_t<0,V_t<-y_t)=\frac{\partial F}{\partial x_t}(0,0)x_t+o(|x_t|+|y_t|)$$
$$\Pr(U_t<-x_t,-y_t<V_t<0)=\frac{\partial F}{\partial y_t}(0,0)y_t+o(|x_t|+|y_t|).$$

Thus,

$$\mathbb{E}D_t(x_t,y_t)=o(|x_t|+|y_t|).$$

By exploiting the assumption that the standard deviations are bounded away from zero, we can get rid of the conditioning on the standard deviation processes:

$$\mathbb{E}D_t(x_t,y_t)=o(|u|+|v|),$$

where we have set $u=xb_\sigma^{-1}$ and $v=yb_\varsigma^{-1}$, with b_σ and b_ς representing the standard deviations' lower bounds. Finally,

$$\begin{aligned}\mathbb{E}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n D_t(x_t,y_t)\right)^2&=\frac{1}{n}\sum_{s=1}^n\sum_{t=1}^n\mathbb{E}D_s(x_t,y_t)\mathbb{E}D_t(x_t,y_t)\\&=o\left(n(|u|+|v|)^2\right).\end{aligned}$$

Now, we have to plug in $\bar{X}_nb_\sigma^{-1}$ and $\bar{Y}_nb_\varsigma^{-1}$ in place of u and v.

$$E\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n D_t(\bar{X}_n,\bar{Y}_n)\right)^2=o\left(n[\mathbb{E}(|\bar{X}_n|+|\bar{Y}_n|)]^2\right)=o(1)$$

where the last identity obtains from the assumption $\mathbb{E}|\bar{X}_n|=O(n^{-1/2})=\mathbb{E}|\bar{Y}_n|$.