A New Tail-based Correlation

Jinjing Liu *

McGill University

December 16, 2016

Abstract

The dependence between assets tends to increase when the market declines. This paper develops a correlation measure focusing on market declines using the expected shortfall (ES), referred to as the ES-implied correlation, to improve the existing value at risk (VaR)-implied correlation. Simulations which define period-by-period true correlations show that the ES-implied correlation is much closer to true correlations than is the VaR-implied correlation with respect to average bias, standard deviation and root-mean-squared error. More importantly, this paper develops a series of test statistics to measure and test correlation asymmetries, as well as to evaluate the impact of weights on the VaR-implied correlation and the ES-implied correlation. The test statistics indicate that the linear correlation significantly underestimates correlations between US and other G7 countries during market downturns, and the choice of weights does not have significant impact on the VaR-implied correlation or the ES-implied correlation.

Key words: Tail Dependence, Correlation asymmetries, Test Statistics

JEL classification number: C12, C15, G15

*I am indebted to my PhD advisor John Galbraith for his invaluable guidance and support. I am also grateful for comments from Jean-Marie Dufour, Victoria Zinde-Walsh and my discussant Maarten van Oordt at the 50th Annual Conference of the Canadian Economics Association (CEA), as well as participants at the 50th CEA annual meeting, 2015 Annual Meeting of the Eastern Finance Association (EFA), 11th Centre Interuniversitaire de Recherche en Economie Quantitative (CIREQ) PhD Students' Conference, CIREQ Seminar at McGill University. All errors or omissions are mine. Please address correspondence to jinjing.liu@mail.mcgill.ca.

1 Introduction

It is a core principle of portfolio theory that diversification can reduce risk. Risk diversification depends on assets being less correlated, so that a fall in one investment can be offset by a rise in another investment. The correlation between assets is traditionally estimated by the linear correlation. However, a number of empirical studies have found that assets tend to fall together when the market falls. As a correlation measure not distinguishing among market situations, the linear correlation tends to underestimate the dependence between assets when the market falls. As McNeil, Frey, and Embrechts (2005) pointed out, the linear correlation is only natural in the context of elliptical models[1] since only elliptical models can be fully characterized by a mean vector and a covariance matrix. On the other hand, it is exactly during market downturns that wealth decreases and protection is most needed. Underestimating the dependence leads investors to overestimate their risk diversification and can not protect their wealth when the market declines.

As a consequence, the literature proposes the value at risk (VaR)-implied correlation to estimate asset dependence under different conditions. The VaR-implied correlation equals the linear correlation when asset returns follow multivariate normal distribution, but captures the increased correlation between assets during market downturns when asset returns are not from normal distribution. However, the VaR-implied correlation has a number of disadvantages. First, the risk measure VaR, on which the VaR-implied correlation is based, does not consider losses beyond VaR. VaR is a quantile of loss distribution. Although the probability for events in the tails to happen is very small, these events cause large losses once they happen. Disregarding losses beyond VaR may cause tail risk, the risk that arises when the possibility of extreme losses is greater than expected. Yamai and Yoshiba (2005) illustrated several cases where the tail risk of VaR causes serious problems. Second, VaR is not coherent. Coherence requires the risk of a combination of individual assets not exceeding the sum of the individual risks, i.e., risk can be reduced by diversification. According to Artzner, Delbaen, Eber, and Heath (1997, 1999), reasonable risk measures should be coherent. Third, the VaR-implied correlation is untrustworthy, or even inaccessible around the centre of asset distributions because the demeaned VaR, the denominator in the formula of the VaR-implied correlation, is close to 0 around the centre.

Consequently, this paper suggests a development of this style of tail-based measure, the expected shortfall (ES)-implied correlation, which repairs the above shortcomings of the VaR-implied correlation. The expected shortfall is the average of losses falling beyond VaR and thus avoids tail risk. In addition, it is a coherent measure. See Artzner et al. (1999), Acerbi and Tasche (2002) and Tasche (2002). Inui and Kijima (2005) even showed that expected shortfall is a basic coherent

[1] The multivariate normal and T distributions are special cases of elliptical distributions. More information about elliptical distributions could be found in chapter 3 of McNeil et al. (2005).

measure because it gives the minimum value among the class of plausible coherent risk measures, and any coherent risk measure is a convex combination of expected shortfalls.

In order to study the possible gains and losses for using the ES-implied correlations, I design four cases in the simulations. In cases 1 and 2, the linear correlation is appropriate and is used as a benchmark to judge whether the ES-implied correlation, which allows extra generality, embodies a large sacrifice when correlation is constant. Cases 3 and 4 illustrate what gains may be possible to use the ES-implied correlations when the linear correlation is inappropriate. Simulation results show that the ES-implied correlation does not cause significant sacrifice when the linear correlation is appropriate, but produces substantial gains when the linear correlation is appropriate. Among all the cases, the ES-implied correlation is much closer to the true correlation than is the VaR-implied correlation.

In the empirical analysis, I investigate the relation between equity returns of G7 countries since previous studies have discovered that correlations between international equity returns increase in bear markets. See Campbell, Koedijk, and Kofman (2002), Longin and Solnik (2001), and Garcia and Tsafack (2011). Using the ES-implied correlation, I also find that correlations between US and other G7 countries are higher during market declines than during normal time.

To measure and test the amount that correlation deviates from the linear correlation during market downturns and upturns, I develop a series of H-statistics based on the VaR- and ES-implied correlations. The statistics from the ES-implied correlation clearly demonstrate that on average, dependence during market downturns increases significantly, while the test statistics from the VaR-implied correlation do not point to a clear pattern. The empirical analysis also shows that correlations tend to decrease during market upturns, but the amount decreased is generally less economically and statistically significant than the amount increased during market downturns.

Notice that portfolio weights are given as exogenous when computing VaR-implied and ES-implied correlations. I further develop test statistics to measure and test the impact of the choice of weights on the implied correlations. Although Cotter and Longin (2011) found little difference in the VaR-implied correlations from using different weights by eyes, they did not provide a method to test the significance of the difference. In addition, their paper is limited to estimating correlations using the pairwise method. According to whether to include a third asset in the portfolio or not, estimation methods can be classified into pairwise estimation and joint estimation, where pairwise estimation assigns weights only to the two assets between whom correlation is examined, yet joint estimation also assigns weights to other assets. Using the test statistics, this paper finds that the ES-implied correlation is less affected by the choice of weights than is the VaR-implied correlation even when excluding the centre of distribution, where the VaR-implied correlation is known to be unstable.

The ES-implied correlation is developed on the literature of the VaR-implied correlation. Campbell et al. (2002) pioneered the VaR-implied correlation and used it to test whether real data follow the normal distribution. Cotter and Longin (2011) investigated the impact of the portfolio weights, the type of position, the frequency of data and the probability level on VaR-implied correlations by using US and UK equity indices. Mittnik (2014) extended the pairwise method used in these papers to joint estimation.

Besides the VaR-implied correlation, there also exist another two tail-based dependence measures: the exceedence correlation and the tail dependence coefficient. The exceedance correlation was pioneered by Longin and Solnik (2001) and studied by Ang and Chen (2002) and Campbell, Forbes, Koedijk, and Kofman (2008). The exceedance correlation estimates the correlation between assets conditional on asset returns falling above or below a pre-specified level.[2] Although the exceedance correlation is easy to understand and simple to calculate, Ang and Chen (2002), Campbell et al. (2002), and Longin and Solnik (2001) noticed that the exceedance correlation has a conditioning bias. For example, when asset returns follow a multivariate normal distribution with a given linear correlation, the exceedence correlation does not equal the linear correlation. As Artavanis (2014) explained, the measure of co-movement depends on the relative performance that matches good and bad states of assets. Truncating data would leave only one state, thus inducing a conditioning bias in the exceedance correlation. Figure 1 plots exceedance correlations conditional on the quantiles across the distribution when returns are drawn from a bivariate normal distribution with the actual correlation of 0.5. It appears that exceedance correlations deviate a lot from the actual correlation. When moving into tails, the exceedance correlation converges to 0. Because of the conditioning bias, the exceedance correlation needs to be adjusted before measuring correlation asymmetries.

The tail dependence coefficient calculates the asymptotic probability that one asset provides extremely small or large returns given another asset provides extreme returns.[3] See Garcia and Tsafack (2011), Patton (2006), and Fortin and Kuzmics (2002) for example. One advantage of the tail dependence coefficient is that it does not need to choose a threshold as other conditional

[2] A general form of the exceedance correlation between two variables X and Y at thresholds δ_1 and δ_2 is

$$\rho(\delta_1,\delta_2) = \begin{cases} Corr(X,Y|X \leq \delta_1, Y \leq \delta_2), \delta_1 < 0, \delta_2 < 0 \\ Corr(X,Y|X \geq \delta_1, Y \geq \delta_2), \delta_1 \geq 0, \delta_2 \geq 0 \end{cases}$$

[3] The coefficient of upper tail dependence is

$$\tau_U = \lim_{\alpha \to 0} Pr[F_X(x) \geq \alpha | F_Y(y) \geq \alpha]$$

and the coefficient of lower tail dependence is

$$\tau_L = \lim_{\alpha \to 0} Pr[F_X(x) \leq \alpha | F_Y(y) \leq \alpha].$$

correlations do. However, this also induces one drawback: as a measure of dependence under very extreme circumstances, the tail dependence coefficient is realized infrequently. Similar to the exceedence correlation, the tail dependence coefficient also deviates from the linear correlation even when returns follow multivariate normal distribution.

Different with the exceedance correlation and tail dependence coefficient, the VaR-implied correlation does not have a conditioning bias. Although the VaR-implied correlation is also defined on only downside state, it is conditional not only on individual assets' returns falling beyond a given threshold, but also on returns of a portfolio composed of the assets falling beyond the threshold. The condition on the portfolio counteracts the conditioning bias from truncating individual asset returns and thus makes the VaR-implied correlation free of conditioning bias. This paper will focus on comparing the performance of the ES-implied correlation with the VaR-implied correlation.

The paper is organized as follows. Section 2 presents the estimation of the ES-implied correlation in two-asset and multi-asset environments, as well as the construction of H-statistics. Section 3 reports the results of simulations, which are designed to evaluate the performance of the ES-implied correlation in comparison with the linear correlation and the VaR-implied correlation. Section 4 analyzes the dependence between US and other G7 countries conditional on different market situations and illustrates how to apply the ES-implied correlation in risk management and asset allocation. Section 5 concludes.

2 Method

2.1 Pairwise method of the VaR-implied and ES-implied correlations

This section presents a pairwise method of the ES-implied correlation, where I only consider two individual assets and their correlation. As a background to the discussion, I first introduce the VaR-implied correlation.

VaR is a function of losses. The loss L is usually given as the negative of returns. The VaR at confidence level α is defined as the minimum value such that the probability of not exceeding this value at least equals α. Formally,

$$VaR(L)_\alpha = \inf\{l | P(L \leq l) \geq \alpha\}, \tag{1}$$

In other words, VaR is the α-quantile of the loss distribution. To simplify the notation, q_α is used to denote VaR_α in the following.

Let r_1 and r_2 be the returns of two assets and r_p be the return of a portfolio which is composed of the two assets with weights w_1 and w_2, where $w_1 + w_2 = 1$. Assume that the loss distribution of asset i belongs to a location-scale family and is characterized by a location parameter μ_i, a scale parameter σ_i and a zero-location, unit-scale distribution F_{Z_i}, referred to as the standard distribution, then

$$L_i = \mu_i + \sigma_i Z_i, \tag{2}$$

where Z_i follows the standard distribution F_{Z_i}.

The VaR of asset i is

$$VaR_{i,\alpha} = \mu_i + \sigma_i VaR(Z_i)_\alpha \tag{3}$$

where $VaR(Z_i)_\alpha$is the VaR of the noise variable Z_i at α confidence level.

Substituting (3) into

$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho\sigma_1\sigma_2, \tag{4}$$

leads to the standardized VaR-implied correlation

$$\rho_{VaR,\alpha} = \frac{(\frac{q_{p,\alpha}-\mu_p}{q(Z_p)_\alpha})^2 - w_1^2(\frac{q_{1,\alpha}-\mu_1}{q(Z_1)_\alpha})^2 - w_2^2(\frac{q_{2,\alpha}-\mu_2}{q(Z_2)_\alpha})^2}{2w_1w_2\frac{(q_{1,\alpha}-\mu_1)(q_{2,\alpha}-\mu_2)}{q(Z_1)_\alpha q(Z_2)_\alpha}}. \tag{5}$$

Campbell et al. (2002) removed the standard distributions in equation (5) and pioneered the VaR-implied correlation:

$$\rho_{VaR,\alpha} = \frac{(q_{p,\alpha}-\mu_p)^2 - w_1^2(q_{1,\alpha}-\mu_1)^2 - w_2^2(q_{2,\alpha}-\mu_2)^2}{2w_1w_2(q_{1,\alpha}-\mu_1)(q_{2,\alpha}-\mu_2)}. \tag{6}$$

Campbell et al. (2002) used the VaR-implied correlation to examine whether returns are from normal distribution or not. Mittnik (2014) demonstrated that the VaR-implied correlation is also able to capture the change in the correlation in the tails.

Since the VaR is well known for not considering losses beyond it and not being coherent, while the ES remedies these problems, it is natural to develop a correlation measure using the ES. The ES at confidence level α is defined as the average of losses beyond the VaR, i.e.,

$$ES(L)_\alpha = E(L|L \geq VaR(L)_\alpha) \tag{7}$$

when the distribution of L is continuous.[4]

[4]When the distribution is discontinuous, $ES(L)_\alpha = \frac{E(L;L\geq VaR(L)_\alpha)+VaR(L)_\alpha(1-\alpha-Pr(L\geq VaR(L)_\alpha))}{1-\alpha}$. See Acerbi and Tasche (2002) and McNeil et al. (2005).

Similarly, given the loss $L_i = \mu_i + \sigma_i Z_i$,

$$ES_{i,\alpha} = \mu_i + \sigma_i ES(Z_i)_\alpha, \tag{8}$$

where $ES(Z_i)_\alpha$ is the ES of the noise variable Z_i at α confidence level.

Substituting (8) into equation (4) leads to the standardized ES-implied correlation,

$$\rho_{ES,\alpha} = \frac{(\frac{ES_{p,\alpha}-\mu_p}{ES(Z_p)_\alpha})^2 - w_1^2(\frac{ES_{1,\alpha}-\mu_1}{ES(Z_1)_\alpha})^2 - w_2^2(\frac{ES_{2,\alpha}-\mu_2}{ES(Z_2)_\alpha})^2}{2w_1 w_2 \frac{(ES_{1,\alpha}-\mu_1)(ES_{2,\alpha}-\mu_2)}{ES(Z_1)_\alpha ES(Z_2)_\alpha}}, \tag{9}$$

The standardized ES-implied correlation only reflects the relation between assets in the second moment and is equivalent to the linear correlation.

Following Campbell et al. (2002), Cotter and Longin (2011), and Mittnik (2014), this paper defines the ES-implied correlation as the correlation that removes $ES(Z_i)_\alpha$, $i = 1, 2, p$ in equation (9), i.e., the ES-implied correlation is

$$\rho_{ES,\alpha} = \frac{(ES_{p,\alpha} - \mu_p)^2 - w_1^2(ES_{1,\alpha} - \mu_1)^2 - w_2^2(ES_{2,\alpha} - \mu_2)^2}{2w_1 w_2 (ES_{1,\alpha} - \mu_1)(ES_{2,\alpha} - \mu_2)}. \tag{10}$$

Under the assumption of individual assets and the portfolio having the same standard distribution, equation (10) is equivalent to equation (9). For example, when returns follow normal distribution, the standardized ES-implied correlation and the ES-implied correlation both equal the linear correlation. When this assumption does not hold, contrary to the standardized ES-implied correlation, the ES-implied correlation defined in equation (10) reflects the information in the standard distribution and thereby is referred to as non-standardized.

2.2 Modification of the ES-implied correlation

The traditional definition of the ES is the average of losses falling beyond the corresponding VaR. However, first, $ES_0 = \mu$, implying that the denominator is close to 0 in equations (9) and (10) when α is very small, and second, ES only considers the returns beyond the given quantile level, thus making ES not consider the extreme values of positive returns. Analyzing the dependence between positive returns is also useful. Thus, this paper defines the ES in two parts. When $\alpha < 0.5$, the ES is defined the same as ES, dented by ES^-. When $\alpha \geq 0.5$, the ES, denoted by ES^+, is modified to be

$$ES^+_\alpha = E(L|L < VaR(L)_\alpha). \tag{11}$$

The standardized ES-implied correlation is therefore modified to be

$$\rho_{ES,\alpha} = \begin{cases} \dfrac{(\frac{ES^+_{p,\alpha}-\mu_p}{ES(Z_p)^+_\alpha})^2 - w_1^2(\frac{ES^+_{1,\alpha}-\mu_1}{ES(Z_1)^+_\alpha})^2 - w_2^2(\frac{ES^+_{2,\alpha}-\mu_2}{ES(Z_2)^+_\alpha})^2}{2w_1w_2\frac{(ES^+_{1,\alpha}-\mu_1)(ES^+_{2,\alpha}-\mu_2)}{ES(Z_1)^+_\alpha ES(Z_2)^+_\alpha}}, \alpha \geq 0.5, \\ \\ \dfrac{(\frac{ES^-_{p,\alpha}-\mu_p}{ES(Z_p)^-_\alpha})^2 - w_1^2(\frac{ES^-_{1,\alpha}-\mu_1}{ES(Z_1)^-_\alpha})^2 - w_2^2(\frac{ES^-_{2,\alpha}-\mu_2}{ES(Z_2)^-_\alpha})^2}{2w_1w_2\frac{(ES^-_{1,\alpha}-\mu_1)(ES^-_{2,\alpha}-\mu_2)}{ES(Z_1)^-_\alpha ES(Z_2)^-_\alpha}}, \alpha < 0.5. \end{cases} \tag{12}$$

The non-standardized ES-implied correlation is modified to be

$$\rho_{ES,\alpha} = \begin{cases} \dfrac{(ES^+_{p,\alpha}-\mu_p)^2 - w_1^2(ES^+_{1,\alpha}-\mu_1)^2 - w_2^2(ES^+_{2,\alpha}-\mu_2)^2}{2w_1w_2(ES^+_{1,\alpha}-\mu_1)(ES^+_{2,\alpha}-\mu_2)}, \alpha \geq 0.5, \\ \\ \dfrac{(ES^-_{p,\alpha}-\mu_p)^2 - w_1^2(ES^-_{1,\alpha}-\mu_1)^2 - w_2^2(ES^-_{2,\alpha}-\mu_2)^2}{2w_1w_2(ES^-_{1,\alpha}-\mu_1)(ES^-_{2,\alpha}-\mu_2)}, \alpha < 0.5. \end{cases} \tag{13}$$

The following proves that the modified ES-implied correlations calculated from the left and the right are the same at $\alpha = 0.5$.

Proposition 1. *The standardized and non-standardized ES-implied correlations are continuous at* $\alpha = 0.5$.

Proof. From Corollary 3.3 in Acerbi and Tasche (2002), we know that ES is continuous. Thus $\lim_{\alpha \to 0.5^-} ES^+ = ES^+_{0.5}$. Since $ES^-_{0.5} + ES^+_{0.5} = \frac{\int_0^1 VaR_u du}{0.5} = 2\mu$,

$$ES^-_{0.5} - \mu = -(ES^+_{0.5} - \mu). \tag{14}$$

Substituting equation (14) into equations (12) and (13), we can see that $\lim_{\alpha \to 0.5^+} \rho_{ES,\alpha} = \lim_{\alpha \to 0.5^+} \rho_{ES,\alpha} = \rho_{ES,0.5}$ for both the standardized and non-standardized correlations. □

In the empirical analysis, the VaR-implied correlation violates the [-1,1] correlation interval frequently. However, due to the fact that the ES is coherent, the ES-implied correlation has the following good property:

Proposition 2. *The ES-implied correlation does not exceed* 1 *when short selling is not allowed.*

Proof. Recall that a risk measure ζ is coherent if it is: 1) subadditive, meaning $\zeta(L_1 + L_2) \leq \zeta(L_1) + \zeta(L_2)$; 2) positive homogeneous, meaning $\zeta(wL) = w\zeta(L)$ for every $w > 0$; 3) monotonic, meaning $\zeta(L_1) \leq \zeta(L_2)$ for $L_1 \leq L_2$; and 4) translation invariant, meaning $\zeta(L + l) = \zeta(L) + l$ for every $l \in R$.

When short selling is not allowed,

$$\begin{aligned} ES_{p,\alpha} &= ES(w_1L_1 + w_2L_2)_\alpha \\ &\leq ES(w_1L_1)_\alpha + ES(w_2L_2)_\alpha \\ &= w_1ES(L_1)_\alpha + w_2ES(L_2)_\alpha \\ &= w_1ES_{1,\alpha} + w_2ES_{2,\alpha}, \end{aligned}$$

where the inequality holds because of subadditivity and the second equality holds due to positive homogeneity. Not allowing short selling is for applying positive homogeneity.

Since expected shortfall is a monotonic risk measure, $ES_\alpha \geq ES_0$, where the latter equals $\int_0^1 VaR_u du = \mu$. Hence,

$$0 \leq ES_{p,\alpha} - \mu \leq w_1(ES_{1,\alpha} - \mu) + w_2(ES_{2,\alpha} - \mu). \tag{15}$$

Thus under the assumption of no short selling, $\rho_{ES,\alpha} \leq 1$ when $\alpha < 0.5$.

To prove $\rho_{ES,\alpha} \leq 1$ when $\alpha \geq 0.5$, I express ES^+_α as the traditional expected shortfall ES_α:

$$ES^+_\alpha - \mu = \frac{\int_0^1 VaR_u du - \int_\alpha^1 VaR_u du}{\alpha} - \mu = -\frac{1-\alpha}{\alpha}(ES_\alpha - \mu). \tag{16}$$

The following holds after substituting equation (16) into equation (15):

$$0 \geq ES^+_{p,\alpha} - \mu \geq w_1(ES^+_{1,\alpha} - \mu) + w_2(ES^+_{2,\alpha} - \mu). \tag{17}$$

Thus $\rho_{ES,\alpha} \leq 1$ also holds when $\alpha \geq 0.5$. □

Equation (16) also implies that the ES-implied correlation is symmetric around $\alpha = 0.5$ when the return distribution is symmetric.

2.3 Estimation and consistency

To estimate the risk measure-implied correlations, I need to compute the VaR and ES of the individual assets and the portfolio. There are three main methods to estimate the VaR and ES: the Gaussian approach, the extreme value theory (EVT) approach and the empirical approach.

The Gaussian approach assumes that returns follow normal distributions. The VaR and ES are then functions of the mean and the standard deviation. See Castellacci and Siclari (2003) for an application of this approach. VaR and ES are very easy to compute using this approach. However, the assumption of normality has been proved unrealistic by many empirical studies. The estimated VaR and ES are thus inaccurate. More importantly, it is easy to observe that the implied correlations estimated by the Gaussian approach actually equal the Pearson linear correlation. So it is unnecessary to use this approach to estimate the implied correlations since they can be obtained easily by calculating the linear correlation.

Extreme value theory focuses on the study of the tail behaviour and is used widely to estimate VaR and ES. See for example Fernandez (2010). However, it is accurate only in the tails. Thus Danielsson and De Vries (2000) used it along with the empirical method. Under the assumption that returns follow extreme value distributions, VaR and ES are functions of the parameters of the extreme value distributions. Generally, there are two methods to estimate the extreme value distribution, block maxima method (BMM) and peak over threshold (POT). Fitting the extreme value distribution requires specifying either the size of the block or the threshold. An inappropriate choice of block size or threshold will cause inaccurate estimation. Thus this paper does not employ this approach to estimate VaR or ES.

The empirical approach uses the empirical distribution of the data to approximate the actual distribution. Fernandez (2010) and Danielsson and De Vries (2000) found that this approach generates smaller errors than the Gaussian approach. More importantly, the empirical approach is parameter free and easy to implement. If even the VaR- and ES-implied correlations computed by using this simple approach can capture the change in the dependence, the implied correlations computed by using sophisticated approach would perform better. Thus, this paper chooses the empirical approach to estimate VaR and ES.

The rest of this section presents the estimation and convergence of the implied correlations. I start from the estimation of the VaR-implied correlation. Let $L_{j:T}$ be the jth largest value in the historical losses $L_t, t = 1, 2, ..., T$, F be the cumulative distribution function of losses and F_T be the empirical distribution function. The empirical estimation of VaR at confidence level α is $L_{[\alpha T]:T}$, where $[\alpha T]$ is the integer part of αT.

Therefore, the empirical estimate of the VaR-implied correlation

$$\hat{\rho}_{VaR,\alpha} = \frac{(\hat{q}_{p,\alpha} - \hat{\mu}_p)^2 - w_1^2(\hat{q}_{1,\alpha} - \hat{\mu}_1)^2 - w_2^2(\hat{q}_{2,\alpha} - \hat{\mu}_2)^2}{2w_1 w_2(\hat{q}_{1,\alpha} - \hat{\mu}_1)(\hat{q}_{2,\alpha} - \hat{\mu}_2)}$$

where $\hat{q}_{i,\alpha} = L^i_{[\alpha T]:T}, i = 1, 2, p$ approximates the α-quantiles and the sample mean $\hat{\mu}_i, i = 1, 2, p$ approximates population means of assets 1, 2, and the portfolio. Shorack and Wellner (2009) have proved that $L_{[\alpha T]:T}$ converges to $F^{-1}(\alpha)$ almost surely. Thus $\hat{\rho}_{VaR,\alpha}$ converges to $\rho_{VaR,\alpha}$ in probability when weak law of large number holds and almost surely when strong law holds. When returns are multivariate normally distributed or they have the same standard distribution F_Z, the VaR-implied correlation is the linear correlation.

For the estimation of the ES-implied correlation, the ES at confidence level less than 0.5 can be estimated by $\frac{\sum_{j=[\alpha T]+1}^{T} L_{j:T}}{T-[\alpha T]}$ and the ES at confidence level exceeding 0.5 can be estimated by $\frac{\sum_{j=1}^{[\alpha T]} L_{j:T}}{[\alpha T]}$. Acerbi and Tasche (2002) proved that the estimate of the traditional ES converges to the actual expected shortfall almost surely. Similarly, the estimate of the modified ES also converges to the actual value of the modified expected shortfall. Therefore, the empirical estimate of the ES-implied correlation converges to the ES-implied correlation in probability when weak law of large number holds and almost surely when strong law holds. When standard distributions of the individual assets and the portfolio are the same, the empirical estimate converges to the linear correlation.

2.4 Joint estimation of the ES-implied correlation

The above sections have presented the pairwise method of estimating the correlation between two assets. For $n \geq 1$ assets, there exist $C(n,2) = \frac{n(n-1)}{2}$ correlations, where $C(n,k)$ denotes the number of k combinations from n elements. The $C(n,2)$ correlations could either be estimated one by one, using the pairwise method, or be estimated together. Estimating correlations one by one may result in a loss of information since other assets in the portfolio may have an impact on the correlation. Mittnik (2014) found that assigning weights to other assets could improve efficiency and reduce the frequency with which the VaR-implied correlation violates the [-1,1] interval. Following Mittnik (2014), this section introduces a closed-form solution for estimating the $C(n,2)$ ES-implied correlations jointly.

Given a portfolio composed of n assets with weights $w_i, i = 1,2,...,n, \sum_{i=1}^{n} w_i = 1$,

$$(ES_{\alpha,p})^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j (ES_{\alpha,i})(ES_{\alpha,j})\rho_{ES,\alpha,ij} \tag{18}$$

holds for demeaned ES.

Denoting the correlation matrix by R, equation (18) can be rewritten as

$$ES_p^2 = (ES \diamond w)' R (ES \diamond w),$$

where α is dropped in order to simplify notations, ES is a $n \times 1$ vector composed of expected shortfalls of all assets, w is a $n \times 1$ vector of weights and $\diamond$ is the Schur product, i.e., $ES \diamond w = \sum_{i=1}^{n} w_i ES_i$. Bring all the known $\rho_{ii} = 1, i = 1,2,...,n$ to the left, then

$$\tilde{ES}_p^2 = ES_p^2 - \sum_{j=1}^{n} w_j^2 ES_j^2 = (ES \diamond w)'(R - I)(ES \diamond w),$$

where I is the identity matrix and $\tilde{ES}_p^2$ represents excess squared quantiles.

Employing the formula $vec(ABC) = (C' \otimes A)vec(B)$, where $\otimes$ is the Kronecker product and vec is the conventional vectorization operator, the above equation equals

$$\tilde{ES}_p^2 = (ES \diamond w)' \otimes (ES \diamond w)' vec(R - I).$$

There exists a unique $n^2 \times \frac{n(n-1)}{2}$ matrix D composed of zeros and ones such that the vectorization of $R - I$ equals

$$\begin{aligned} vec(R-I) &= (0 \quad \rho_{12} \quad ... \quad \rho_{1n} \quad \rho_{12} \quad 0 \quad ... \quad \rho_{2n} \quad ... \quad \rho_{n-1,n} \quad 0)' \\ &= D \quad (\rho_{12} \quad ... \quad \rho_{1n} \quad \rho_{23} \quad ... \quad \rho_{2n} \quad ... \quad \rho_{n-1,n})'. \end{aligned}$$

The $\frac{n(n-1)}{2} \times 1$ vector $\rho = (\rho_{12} \quad ... \quad \rho_{1n} \quad \rho_{23} \quad ... \quad \rho_{2n} \quad ... \quad \rho_{n-1,n})'$ includes all the correlations that need to be estimated.

Therefore, given a weight vector w,

$$\tilde{ES}_p^2 = (ES \diamond w)' \otimes (ES \diamond w)' D\rho. \tag{19}$$

To exactly estimate correlations, $n(n-1)/2$ equations are needed. When there are $m = \frac{n(n-1)}{2}$ vectors of weights $(w_1, ...w_m)$,

$$\begin{pmatrix} \tilde{ES}_{p_1}^2 \\ ... \\ \tilde{ES}_{p_m}^2 \end{pmatrix} = \begin{bmatrix} (ES \diamond w_1)' \otimes (ES \diamond w_1)' \\ ... \\ (ES \diamond w_m)' \otimes (ES \diamond w_m)' \end{bmatrix} D\rho.$$

Denote $\tilde{Q} = (\tilde{ES}_{p_1}^2 ... \tilde{ES}_{p_n}^2)'$ and $X = \begin{bmatrix} (ES \diamond w_1)' \otimes (ES \diamond w_1)' \\ ... \\ (ES \diamond w_m)' \otimes (ES \diamond w_m)' \end{bmatrix} D$, then the correlation vector is obtained by

$$\rho = X^{-1}\tilde{Q}. \tag{20}$$

Equation (20) is referred to as exact identification since every equation in the formula is satisfied exactly. When assigning weights only to every two assets, correlations calculated from the exact identification equal correlations calculated from the pairwise method in section 2.1.

When adding the number of weight vectors more than unknown correlations, X would not be a square matrix any more. The estimator can be obtained by least squares:

$$\rho = (X'X)^{-1}X'\tilde{Q}. \tag{21}$$

Equation (21) is referred to as overidentification, where an error term exists so that $\tilde{Q} = X\rho + u$ instead of $\tilde{Q} = X\rho$ in the case of exact identification.

2.5 Quantitative measures of correlation asymmetries

Since correlations are found to be asymmetric rather than constant, this section develops a series of H statistics for measuring and testing the amount that correlation deviates from the linear correlation during market downturns and upturns. The following session presents the downside H, the upside H, the downside AH, and the upside AH statistics. The downside and upside H statistics measure the maximum degree to which implied correlations deviate from the linear correlation, while the downside and upside AH statistics evaluate the average of correlation deviations. Because the VaR-implied correlation is very unsteady when the probability level is around 0.5 and the linear correlation is only inappropriate in the tails, the downside and upside statistics are constructed to assess correlation asymmetries in intervals of (0, 0.3) and (0.7, 1), respectively.

The downside H statistic is defined as the supremum of deviations of the linear correlation from tail-based correlations, i.e.,

$$H^{-}_{VaR} = \sup_{\alpha \in (0,0.3)} (\hat{\rho}_{VaR,\alpha} - \hat{\rho}) \tag{22}$$

for the VaR-implied correlation and the downside H statistic for the ES-implied correlation is

$$H^{-}_{ES} = \sup_{\alpha \in (0,0.3)} (\hat{\rho}_{ES,\alpha} - \hat{\rho}), \tag{23}$$

where $\hat{\rho}$ denotes the empirical linear correlation, $\hat{\rho}_{VaR,\alpha}$ denotes the empirical VaR-implied correlation, and $\hat{\rho}_{ES,\alpha}$ denotes the empirical ES-implied correlation.

The upside H statistic evaluates the highest degree to which correlation is overestimated by the linear correlation in the right tail of return distribution, i.e.,

$$H^{+}_{VaR} = \sup_{\alpha \in (0.7,1)} (\hat{\rho} - \hat{\rho}_{VaR,\alpha}) \tag{24}$$

for the VaR-implied correlation and

$$H^{+}_{ES} = \sup_{\alpha \in (0.7,1)} (\hat{\rho} - \hat{\rho}_{ES,\alpha}) \tag{25}$$

for the ES-implied correlation.

The downside AH statistic measures the average of correlation asymmetries in the left tail. The downside AH statistic using the VaR-implied correlation is

$$AH^{-}_{VaR} = \frac{1}{0.3T}\int_{0}^{0.3}(\hat{\rho}_{VaR,\alpha} - \hat{\rho})d\alpha, \tag{26}$$

where $0.3T$ is the sample size locating in the $\alpha \in [0, 0.3]$ interval. The downside AH statistic using the ES-implied correlation is

$$AH^{-}_{ES} = \frac{1}{0.3T}\int_{0}^{0.3}(\hat{\rho}_{ES,\alpha} - \hat{\rho})d\alpha. \tag{27}$$

The upside AH statistic measures average upside correlation asymmetries. The upside AH statistic using the VaR-implied correlation is

$$AH^{+}_{VaR} = \frac{1}{0.3T}\int_{0.7}^{1}(\hat{\rho} - \hat{\rho}_{VaR,\alpha})d\alpha \tag{28}$$

and the upside AH statistic using the ES-implied correlation is

$$AH^{+}_{ES} = \frac{1}{0.3T}\int_{0.7}^{1}(\hat{\rho} - \hat{\rho}_{ES,\alpha})d\alpha. \tag{29}$$

The probability level α is assumed to be uniformly distributed between 0 and 1. When α follows other distributions, the difference between implied correlations and linear correlations is assigned different weights at different probability levels. For example, Ang and Chen (2002) chose weights proportional to the number of observations.

Besides evaluating correlation asymmetries numerically and comparing the degree of correlation asymmetries across assets, the largest advantage of the H-statistics is to test the significance of the difference between correlations under extreme circumstances and normal time. The null hypothesis of the test is that returns follow multivariate normal distribution, under which condition, there is no difference between linear correlation and tail-based correlations. The alternative hypothesis is that linear correlation underestimates (overestimates) correlations during market downturns (upturns). Thus when the test statistic is too large, the null hypothesis is rejected.

I employ the Monte-Carlo (MC) method to test the significance of the H statistic and the AH statistic. The MC method provides a simple way to test statistics whose finite distribution is unknown, but can be simulated. The idea of the MC method is to obtain the empirical distribution of the test statistic from simulation and uses it to estimate the cumulative distribution function of the test statistic. Since originated by Dwass (1957), the MC method has been widely studied. See

Dufour (2006), Hastings (1970), and Gilks (2005) for example. Dufour and Kurz-Kim (2010) and Beaulieu, Dufour, and Khalaf (2014) used it to test parameters in the stable distribution.

Given two assets with sample size n, the following procedures are designed to test the significance of correlation asymmetries:

Step A: Estimate test statistics $\hat{H}$, $\hat{AH}$, and $\hat{\rho}$ from the empirical data.

Step B: Draw n pairs of data from the bivariate normal distribution with correlation $\hat{\rho}$.

Step C: Compute the test statistic, named $H^{(1)}$ for H statistic and $AH^{(1)}$ for AH, from the simulated data.

Step D: Repeat step B and step C M times and get a sequence of test statistics, $H^{(1)},...,H^{(M)}$ and $AH^{(1)},...,AH^{(M)}$.

Step E: Calculate the p-values, $\hat{p}_H = \frac{1}{M+1}\sum_{m=1}^{M}(I(H^{(m)} \geq \hat{H})+1)$ for H, and $\hat{p}_{AH} = \frac{1}{M+1}\sum_{m=1}^{M}(I(AH^{(m)} \geq \hat{AH})+1)$ for AH, where $I(\cdot)$ is known as the indicator function. The hypothesis is rejected at level α if the MC p-value is less than or equal to α.

Notice that the H statistics in this paper differ from the statistic in Ang and Chen (2002) in several ways. First, the thresholds considered in the paper are continuous, while the thresholds in the paper of Ang and Chen (2002) are a number of discrete points and are given as a priori. Second, while Ang and Chen (2002) used the quadratic deviation and the sum of deviation between the linear correlation and the exceedance correlation, this paper measures the maximum and the average of the deviation of the linear correlation from implied correlations. It is common to construct statistics using maximum. See Hansen (1996), Davies (1977) and Davies (1987). Third, this paper uses the risk measure-implied correlations, which are free of conditioning bias, to measure correlation asymmetries, while Ang and Chen (2002) have to adjust conditioning bias of the exceedance correlation. Fourth, the significance of the test statistic is examined by Monte-Carlo in this paper, while Ang and Chen (2002) used GMM and the delta method to get the standard deviation of test statistic first and then calculate the p-value. Compared to the estimation of GMM, the Monte-Carlo method is parameter-free and is easier to implement.

The H and AH statistics also can evaluate the impact of weights on risk measure-implied correlations. The sign of the difference is not important any more since the interest here is to test whether using different weights could lead to different implied correlations. Thus I use the absolute value of the difference to construct statistics. The new H statistics do not distinguish between downside and upside, and take the supremum across all probability levels except the interval $(0.3, 0.7)$ to avoid the unstable VaR-implied correlations. That is,

$$H_{VaR} = \left(\sup_{\alpha \in (0,0.3)\cup(0.7,1)} |\hat{\rho}^{(1)}_{VaR,\alpha} - \hat{\rho}^{(2)}_{VaR,\alpha}| \right) \tag{30}$$

and

$$H_{ES}=\left(\sup_{\alpha\in(0,0.3)\cup(0.7,1)}|\hat{\rho}^{(1)}_{ES,\alpha}-\hat{\rho}^{(2)}_{ES,\alpha}|\right),\tag{31}$$

where $\hat{\rho}^{(1)}_{VaR,\alpha}$ and $\hat{\rho}^{(1)}_{ES,\alpha}$ are implied correlations using a choice of weights, while $\hat{\rho}^{(2)}_{VaR,\alpha}$ and $\hat{\rho}^{(2)}_{ES,\alpha}$ are implied correlations using a different choice of weights. The corresponding AH are

$$AH_{VaR}=\frac{1}{0.6}\left(\int_0^{0.3}|\hat{\rho}^{(1)}_{VaR,\alpha}-\hat{\rho}^{(2)}_{VaR,\alpha}|d\alpha+\int_{0.7}^{1}|\hat{\rho}^{(1)}_{VaR,\alpha}-\hat{\rho}^{(2)}_{VaR,\alpha}|d\alpha\right)\tag{32}$$

and

$$AH_{ES}=\frac{1}{0.6}\left(\int_0^{0.3}|\hat{\rho}^{(1)}_{ES,\alpha}-\hat{\rho}^{(2)}_{ES,\alpha}|d\alpha+\int_{0.7}^{1}|\hat{\rho}^{(1)}_{ES,\alpha}-\hat{\rho}^{(2)}_{ES,\alpha}|d\alpha\right).\tag{33}$$

Since the VaR- and ES-implied correlations are invariant with respect to weights for elliptical distributions,[5] the difference in implied correlations from choosing different weights should be insignificant when data are from the multivariate normal distribution. Thus this paper simulates data from the multivariate normal distribution and calculates the statistics using simulated data to test the significance of the statistics from the empirical data. The process is similar to the test of correlation asymmetries and thus is not repeated.

3 Simulation

In order to study the possible sacrifice and gains for using the ES-implied correlation, I design four cases in the simulation. In the first two cases, correlation is constant and the linear correlation is appropriate in order to judge whether allowing extra generality embodies a large sacrifice; in the other two cases, correlation changes in the tails and the linear correlation is inappropriate in order to judge what gains may be possible by using the expected shortfall-based measure.

Case 1: correlation is constant $\rho = 0.5$ and the data are from a multivariate normal distribution with mean 0 and covariance matrix $\Sigma = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}$. The estimated linear correlation is the maximum-likelihood estimation (MLE), and therefore has the desirable asymptotic properties of maximum likelihood, including consistency and efficiency. Fisher (1915) and Gayen (1951) have discussed the distribution of the estimated linear correlation.

Case 2: correlation is still constant and the data follow a multivariate T distribution with degrees of freedom 3 and covariance matrix $\frac{T-2}{T}\Sigma$, where T is the sample size. The estimated linear

[5] See Mittnik (2014) and Campbell et al. (2002)

correlation is no longer the MLE.

Case 3: correlation is non-constant and changes at some designed probability levels. Since the correlation is non-constant, better performance from methods which allow for the change in the correlation is expected.

Only non-standardized implied correlations are reported in the above cases since standardized implied correlations are actually linear correlation and are thus unnecessary to compute, especially when data follow multivariate normal distribution. Besides, the results of non-standardized and standardized implied correlations are very similar when data are from multivariate normal and multivariate T distribution.

The empirical analysis has found that the distribution of returns is asymmetric. Therefore, case 4 simulates the data from a skewed T distribution. The results of both non-standardized and standardized implied correlations are reported. While the standardized implied correlations always converge to the linear correlation because the effect from the standard distributions of individual assets and the portfolio is eliminated, the non-standardized implied correlations do not necessarily converge to the linear correlation.

Case 4: correlation is again constant, but data are drawn from a skewed bivariate T distribution with degrees of freedom 3 and a skewness parameter 5. See Fernández and Steel (1998) and von Rohr, Hoeschele, et al. (2002) for details of this distribution.

The average bias, standard deviation and root-mean-square error (RMSE) are calculated to evaluate the performance of different correlations. The squared RMSE is equal to the sum of the squared standard deviation and the squared average bias.

The average bias is the average of the bias through m simulations, i.e.,

$$\frac{1}{m}\sum_{j=1}^{m}(\hat{\rho}_j-\rho),$$

the standard deviation is

$$\sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{\rho}_j-\frac{1}{m}\sum_{j=1}^{m}\hat{\rho}_j)^2},$$

and the RMSE is

$$\sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{\rho}_j-\rho)^2}.$$

The estimator which makes these values close to 0 is regarded as a good indicator.

Throughout the simulation, I use the pairwise method and the weight vector $(0.5, 0.5)$ to estimate the correlations. The effect of the choice of weight vectors on estimated implied correlations will be studied in the empirical analysis.

Figure 2 plots the average bias and standard deviation of the estimated correlations from the multivariate normal distribution with sample size $T = 10^3, 10^4$ and 10^5 through $m = 10^4$ replications. The RMSE is very close to standard deviation since average bias is very small in this case and thus is not reported. Throughout the paper, the black, red and green points correspond to the estimated linear correlation, VaR-implied correlation and ES-implied correlation, respectively.

Consistent with the theory that the estimated linear correlation is best unbiased when the data are from multivariate normal distribution, the estimated linear correlation coefficient is consistent and efficient in case 1.

The estimated ES-implied correlation is as good as the estimated linear correlation almost everywhere. When the vertical axis is set to be (-0.010,0.010) in the plot of average bias and (0,1) in the plot of standard deviation, it is easy to notice that the implied correlations deviate from the actual correlation in the tails when sample size is small. However, the deviation is very small compared to the value of the actual correlation and this problem decreases as sample size increases. The problem may be due to few observations in the tails when sample size is small.

The estimated VaR-implied correlation has two symmetric U shapes. The deviation in the center of the distribution is because the denominator in the formula of VaR-implied correlation is close to 0 around the center. The deviation in the tails is probably also due to not enough observations. The estimated VaR-implied correlation deviates a lot when sample size is small. As sample size increases, its performance gets closer to the actual correlation and the estimated ES-implied correlation. When T reaches 10^5, there is little difference between the estimated VaR-implied correlation and the other two estimated correlations.

Figure 3 plots the result of data simulated from the multivariate T distribution with the actual correlation $\rho = 0.5$. Again, three sample sizes are considered, $T = 10^3, 10^4$ and 10^5, and the process is repeated $m = 10^4$ times. In this case, the estimated linear correlation is not MLE any more. The estimated ES-implied correlation has both the smallest bias and the smallest standard deviation among the three correlation estimates. The standard deviations of implied correlations show a U shape when sample size is small. For both the estimated VaR-implied correlation and the estimated ES-implied correlation, their performances in the tails improve as sample size increases. Again, the poor performances in the tails may be due to few observations.

Case 3 mimics the correlation between US and Canada and assumes two breaks, which divide the whole space $(-\infty, \infty)$ into 3 regions. The first region $(-\infty, -1] \times (-\infty, -1]$ mimics the bear market, where the actual correlation is 0.77. The second region $(-1, 1] \times (-1, 1]$ mimics the moderate market situation, where the actual correlation is 0.74. The third region $(1, \infty) \times (1, \infty)$ mimics the bull market, where the actual correlation is 0.66.

The data following such a distribution are generated by truncation. Ang and Bekaert (2002) and Campbell et al. (2008) documented that truncation changes the correlation between assets. Given the correlation after truncation, one needs to decide the correlation before truncation. Employing the MC technique in a wide search, I find that the data for regions 1, 2, and 3 can be generated by truncating a bivariate normal distribution with marginal distributions $N(0,1)$ and correlation coefficients $\rho^1_{before} = 0.94$, $\rho^2_{before} = 0.9$ and $\rho^3_{before} = 0.88$, respectively.

To be exact, the following steps are used to determine the correlations that are used before truncation.

Step A: Choose a possible value for the correlation before truncation, and generate 10^5 random variables from a bivariate normal distribution with marginal distribution $N(0,1)$ and this correlation.

Step B: Truncate the simulated data to the region that we concern and calculate the correlation in that region.

Step C: Repeat this process 10^4 times and record the correlation every time.

Step D: Calculate the average of the 10^4 truncated correlations. If the average correlation equals the targeted correlation, the correlation chosen in step A is the right correlation before truncation. If the average is greater than the targeted correlation, reduce the correlation and repeat step A, B and C till it makes the average of the correlations across 10^4 replications equal the targeted correlation. If the average is smaller than the targeted correlation, increase the correlation and repeat step A, B and C till the average of the correlations and the targeted correlation are the same.

Generate a proportion of $p_1 = 12\%$ of random variables truncating from a bivariate normal distribution with correlation ρ^1_{before} in region 1, $p_3 = 6\%$ of random variables truncating from a bivariate normal distribution with correlation ρ^3_{before} in region 3, and $p_2 = 1 - p_1 - p_3 = 82\%$ of random variables truncating from a bivariate normal distribution with correlation ρ^2_{before} in region 2, then the actual correlation is

$$\rho_\alpha = \begin{cases} 0.77, & \alpha \leq 12\%, \\ 0.74, & 12\% < \alpha \leq 94\%, \\ 0.66, & \alpha > 94\%. \end{cases}$$

Figure 4 plots the average estimated correlations across $m = 10^3$ simulations with sample size $T = 10^4$. In comparison with the large deviation of the estimated linear correlation from the actual linear correlation, the estimated implied correlations are more trustworthy. The estimated VaR-implied correlation is still unstable around the center of each region.

We are naturally interested in how far a deviation needs to be from constancy in order that the estimated ES-implied correlation provides an improvement. Table 1 reports the summary statistics of the RMSEs in four situations: $p_1 = 0, p_2 = 1, p_3 = 0$; $p_1 = 2\%, p_2 = 96\%, p_3 = 2\%$; $p_1 = 12\%, p_2 = 82\%, p_3 = 6\%$; and $p_1 = 12\%, p_2 = 76\%, p_3 = 12\%$. Situation 1 assumes no break point, in which case the estimated linear correlation is the best unbiased estimator. Situation 2 assumes that 2% of the data are from another multivariate normal distributions in the left tail and right tails. Situation 3 increases this proportion and makes the proportion different in the left and right tails in order to evaluate the impact of asymmetry. Situation 4 then increases the proportion in the right tail to the same level of the proportion in the left tail.

The estimated linear correlation generates the least RMSE in situation 1. However, even when only 2% data from other distributions are included in the tails, the RMSE of the estimated linear correlation increases sharply. The RMSE keeps growing when the proportion increases in the tails, but not that much from situation 3 to situation 4.

Compared to the estimated linear correlation, the estimated risk measure-implied correlation is less affected by the change of the proportion in the tails. The RMSE of estimated ES-implied correlations is less than the RMSE of estimated linear correlations almost at every probability level in the last three situations, where data in the tails are assumed to follow a different distribution. The estimated VaR-implied correlation is still untrustworthy around the center, leading to the mean and standard deviation of its RMSE very large at some probability levels.

Case 4 simulates a bivariate skewed T distribution. Besides non-standardized implied correlations, standardized implied correlations are also measured in this case for examining the difference between standardized and non-standardized correlations. The two-stage MC is employed to calculate the standardized implied correlations. See Beaulieu et al. (2014) about this technique. In the first stage, I use the Monte-Carlo to calculate the VaR and ES of the standard distribution of the individual asset and the portfolio. The sample size is 10^5 and replicate 10^4 times. In the second stage, I draw $m = 10^3$ times of the data from the bivariate skewed T distribution with sample size $T = 10^4$ and then calculate the standardized and non-standardized implied correlations.

The non-standardized correlations are expected to be very different from the linear correlation, but the standardized correlations should be close to the linear correlation since the effect from skewness is eliminated. Two figures are plotted in case 4. The left panel shows the subtle difference between the estimated standardized correlations and the estimated linear correlation. The right panel plots non-standardized correlations.

Panel (a) of figure 5 plots the average bias of the estimated linear correlation and the estimated standardized implied correlations during $m = 10^3$ repetitions. The estimated standardized implied correlation is closer to the actual linear correlation than the estimated linear correlation almost at every probability level. Panel (b) of figure 5 plots the estimated linear correlation, standardized implied correlations and non-standardized implied correlations. The estimated VaR-implied correlation is very unstable around $\alpha = 0.62$, at which probability level, the quantile is very close to the mean, leading to the numerator in the formula of the VaR-implied correlation being close to 0. The estimated non-standardized correlation is greater than the estimated linear correlation and the estimated standardized implied correlations when the probability level is high and less than the estimated linear correlation and the estimated standardized implied correlations when the probability level is small.

4 Empirical Analysis

In the empirical analysis, I investigate correlations between equity returns of G7 countries. Measuring their correlations can help investors to make an investment decision in the global financial market. The G7 countries are the United States (US), Canada (CA), the United Kingdom (UK), Italy (IT), Germany (DE), France (FR) and Japan (JP).

4.1 Data

This paper uses the equity returns of G7 countries from January 1st, 1973 to December 31st, 2015. To avoid time difference, I use weekly returns when calculating the dependence between G7 countries. The weekly frequency also avoids market microstructure at daily frequencies, yet provides sufficient observations in the tails. The data is acquired from Datastream and includes 2313 observations. Panel A of table 2 presents the summary statistics of returns for the whole period. The mean and standard deviation of returns are annualized by multiplying returns by 52. The average returns of all the countries, except Japan, are around 10%. Japan has an average return of only about 7%. Standard deviations of returns vary from 1.14 to 1.69. The row labelled "Skewness" reports the results of the D'Agostino test of skewness [D'Agostino (1970)]. It implies that the equity returns in all countries, except UK and Italy, are skewed to left at at least 5% significance level, while the equity returns of UK are significantly skewed to right, and the returns of Italy do not exhibit significant skewness. The row labelled "Kurtosis" reports the results of the Anscombe-Glynn test [Anscombe and Glynn (1983)] and implies that all equity returns have acute peaks and tend to be heavy-tailed. The results of the Anderson-Darling test [See Stephens (1986) and Thode (2002)] and the Shapiro-Wilk test [See Shapiro and Wilk (1965) and Royston (1982)] reported in rows labelled "AD test" and "Shapiro test" also indicate that the equity returns do not follow normal distribution.

To see how the 2008 financial crisis affects returns, I divide the whole sample period into two subperiods. The first period includes the first 30 years from January 1st, 1973 to December 31st, 2002. The second period includes the 2008 financial crisis and extends from January 1st, 2003 to December 31st, 2015. Panel B and panel C in table 2 report the summary statistics for these two subperiods. Except Germany and Japan, the average mean of all the return series decreases in the second period. While in the first subperiod, returns of most countries, except UK, Italy, and Japan, are skewed to left, the skewness of all countries in the second subperiod is negative, implying that large negative returns appeared. Results from the kurtosis, Anderson-Darling test and Shapiro-Wilk test in both subperiods suggest that equity returns of all countries are not normal.

4.2 Empirical results

In total, there are $C(7,2) = \frac{7\times 6}{2} = 21$ correlation coefficients between the seven countries. To save space, only correlations between US and other countries are reported. The weight vector used to compose the portfolio is $(0.5, 0.5)$.

Figure 6 plots the performance of the estimated linear correlation, VaR-implied correlation and ES-implied correlation in the whole sample period, where low probability levels correspond to bad market situations and high probability levels correspond to good market situations.

The figure implies that the estimated VaR-implied correlation is very unstable around the center and goes beyond 1 frequently. The estimated ES-implied correlation is higher than the estimated linear correlation in the left tail and lower than the estimated linear correlation in the right tail for all the countries, consistent with the empirical findings that dependence increases during market downturns and decreases during market upturns. Among all countries, the tail dependences between US and Canada and between US and Japan increase least. It is also noteworthy that although US exhibits a higher linear correlation with Canada than with other countries, the dependence between US and Canada does not increase as much as the dependence between US and other countries when the market declines. For example, the tail dependence between US and UK is even higher than the tail dependence between US and Canada, emphasizing the importance of estimating tail dependence.

Table 3 reports the H-statistics and AH-statistics measuring correlation asymmetries quantitatively. Panel A reports the results of the H-statistics. The H-statistics between US and other countries are positive for both VaR-implied and ES-implied correlations, implying that the linear correlation underestimates the dependence in the left tail and overestimates the dependence in the right tail. Row 1 and row 2 report the value and the significance of downside H statistics using the VaR-implied correlation and the ES-implied correlation, respectively. The results show that the VaR-implied correlation tends to produce higher correlation asymmetries than ES-implied correlation, but statistics of the VaR-implied correlation are less significant than statistics of the ES-implied correlation. In particular, the VaR-implied correlation indicates that relations between US and Italy, France, and Japan in the left tail increase significantly at 5% level, and the ES-implied correlation implies significant relations between US and UK, Italy, Germany, and France. Compared to rows 1 and 2, rows 3 and 4 suggest that the amounts and significance of correlations decreasing when the market rises are generally less than the amounts of correlations increasing when the market falls. Only 2 out of 6 upside H-statistics based on the VaR-implied correlation and 3 based on the ES-implied correlation are significant at 5% level.

Panel B reports the results of the AH-statistics. The AH-statistics have positive and negative signs, but the statistics with negative signs are all insignificant, implying that correlations are generally higher during market downturns and lower during market upturns than the linear correlation. The downside AH-statistics using the VaR-implied correlation indicate that relations between US and Canada and Japan are significant at 5% level. The downside AH-statistics using the ES-implied correlation imply that correlations between US and other countries during market downturns are significantly underestimated by the linear correlation at 10% level and higher. The upside AH-statistics using the VaR-implied correlation suggests a positive significant relation between US and Germany, and the upside AH-statistics using the ES-implied correlation imply that all correlations decrease significantly at 10% and higher.

I then estimate the linear correlation and implied correlations for two subperiods, one including the 2008 financial crisis and one not. Figure 7 plots the results. The black points, red points, and green points represent the results of the estimated linear correlation, VaR-implied correlation, and ES-implied correlation for the first subperiod from January 1973 to December 2002, respectively. The grey points, pink points, and blue points correspond to the estimated linear correlation, VaR-implied correlation, and ES-implied correlation for the second subperiod from January 2003 to December 2015. It is obvious that correlations increase in the second period. Since market situations get worse in the second period and the result from the whole period suggests that dependence increases when the market situation worsens, correlations in the second period are expected to be higher in the second period than in the first period. Notice that the correlation between US and Canada increases least among all correlations, consistent with results in Figure 6 and Table 3 that the dependence between US and Canada does not increase much when the market declines.

4.3 Effect of weights

This section tests the impact of portfolio weights on risk measure-implied correlations by two approaches. In the first approach, I use different values of weights and estimate the difference in the implied correlations. In the second approach, I compare the difference in the estimated correlations from exact identification and overidentification in order to check whether giving weight vectors more than the number of correlations could improve efficiency since Mittnik (2014) documented that assigning weights to other assets in the portfolio provides more information and reduces the problem that the VaR-implied correlation locates outside the $[-1, 1]$ interval frequently.

4.3.1 Result of changing the value of weights

Instead of using the weight vector (0.5, 0.5), this section uses the weight vector (0.2, 0.8), i.e., investing 20% in US and 80% in the other market. Figure 8 plots the difference in the correlations from using the weight vectors (0.5, 0.5) and (0.2, 0.8). The difference in the estimated VaR-implied correlations varies a little in the tails, but differs a lot around the center. The estimated ES-implied correlation does not show discernible difference across probability levels. Panel A of table 4 reports the statistics measuring the difference in correlations from using different weigths. Both H and AH from ES-implied correlations are smaller than the ones from VaR-implied correlations. All the test statistics are insignificant except the AH-statistic based on the VaR-implied correlation between US and Japan. Therefore, despite the fact that the ES-implied correlation requires a pre-specified vector of weights, the estimate is not affected by the value of the weights significantly, suggesting the possibility of applying the ES-implied correlation in asset allocation.

4.3.2 Difference in estimated correlations from exact identification and overidentification

Figure 9 reports the difference in estimated correlations from exact identification and overidentification. The weights for exact identification are obtained by assigning every two assets an equal weight (0.5, 0.5). The weights for overidentification are received by drawing k assets respectively and assigning them an equal weights $1/k$, where k is an integer from 2 to n, and n is the number of total assets. This finally generates $C(n, 2) + ... + C(n, n) = 2^n - n - 1$ vectors of weights. Figure 9 indicates that overall, the difference in estimated correlations from exact identification and overidentification is very small and is negligible compared to the values of estimated implied correlations. The figure shows that VaR-implied correlations are more affected by estimation methods than ES-implied correlations.

Panel B of table 4 reports the H and AH statistics measuring the difference in implied correlations from using exact identification and overidentification. All of the statistics are insignificant, implying no significant impact of estimation methods on implied correlations. The statistics from ES-implied correlations are still smaller than the statistics from VaR-implied correlations, implying that ES-implied correlations are less affected by using different estimation methods than VaR-implied correlations. The difference in estimated VaR-implied correlations from exact identification and overidentification is much closer to 0 than the difference in estimated VaR-implied correlations from using different weights. Overall, the estimated risk measure-implied correlations are autonomous from changing the value of weights and using different estimation methods.

4.4 Simple illustrations for potential applications

Using the ES-implied correlation, we have seen that asset correlations increase during market downturns. Thus, investors who diversify risk according to the linear correlation may underestimate the risk in the tails. Considering that investors care more about losses than gains in the real world, the ES-implied correlation has important applications in risk management and asset allocation. A direct application is using the ES-implied correlation to assess the risk of the portfolio. For example, the volatility of a portfolio composed of 50% US equity index and 50% Canada equity index is 1.22 using the linear correlation, but for investors who are concerned about volatility during market downturns, e.g., the volatility at 5%-quantile, they could use the ES-implied correlation instead and get a higher portfolio volatility of 1.26.

The ES-implied correlation also compensates the linear correlation in asset allocation. Since correlation increases when the market falls, the protection from diversifying investment also erodes. Thus for investors who care about risk diversification during market downturns, they could use the ES-implied correlation to construct portfolios. Figure 10 plots the classic efficient frontier and the efficient frontier using the ES-implied correlation at 5%-quantile in solid line and dashed line, respectively, as the portfolio moves from the US equity index to the Canada equity index. We can see that investors using the ES-implied correlation are more risk sensitive: they demand higher expected returns for one percentage increase in volatility than their mean-variance counterparts.

One concern of applying the ES-implied correlation in asset allocation is that the ES-implied correlation requires pre-knowledge of weights. Consider the efficient frontier generated by the ES-implied correlation using equal weights 50% and 50% at 5%-quantile, under which condition the ES-implied correlation is 0.7920. Given the annual risk-less rate of 5%, the tangency weights, which provide the highest Sharpe ratio, are 55.81% in US equity index and 44.19% in Canada equity index, corresponding to which the ES-implied correlation is 0.7924. The difference in the ES-implied correlation, 0.0004, is trivial compared to the ES-implied correlation. Then I use the newly updated tangency weights (55.81%, 44.19%) to calculate the ES-implied correlation, construct a new efficient frontier, and find a new set of tangency weights. I repeat this process until the difference in the ES-implied correlation is less than 10^{-7}. It appears that the ES-implied correlation converges after three iterations. The final tangency weight set is (55.28%,44.72%) and the corresponding ES-implied correlation is 0.7923. In this case, I choose a starting weight vector (50%, 50%) very close to the final optimal tangency weights. Even though I start from (20%, 80%), I still get tangency weights (55.28%,44.72%) after three iterations. Thus, not using the optimal weights to calculate the ES-implied correlation affects little of finding the optimal weights. Similar to section 4.3, I employ the H and AH statistics to test the impact on the ES-implied correlation when not using the optimal weights. Table 5 reports the difference in the ES-implied correlation from using the optimal tangency weights and using equal weights and finds no significant difference.

5 Conclusion

Tail-based dependence measures play a central role in risk management and asset allocation. It is well known that the linear correlation provides a poor indicator of co-movements of financial assets under extreme circumstances, particularly during market crashes. Other dependence measures, the exceedance correlation and the tail dependence coefficient, have conditioning biases and can not be used in measuring correlation asymmetry directly. An alternative measure, the VaR-implied correlation, does not have a conditioning bias; however, it has a number of disadvantages, including the fact that the VaR is not coherent, disregards the data beyond it, and the VaR-implied correlation does not work around the center of distributions.

A development of the VaR-implied correlation, the ES-implied correlation, using expected shortfall instead, can eliminate all the shortcomings of the VaR-implied correlation. Simulations indicate that the estimated ES-implied correlation is as accurate as the estimated linear correlation when the estimated linear correlation is appropriate, but is much more accurate than the estimated linear correlation when the estimated linear correlation is inappropriate with respect to average bias, standard deviation and root-mean-square errors. VaR-implied correlations are much less stable than ES-implied correlations, especially around the center of the distribution and when the sample size is small.

In the empirical analysis of international equity indices, VaR-implied correlations frequently violate the $[-1, 1]$ interval, especially around the center of return distribution. However, ES-implied correlations are much more steady. More importantly, ES-implied correlations clearly show that the linear correlation underestimates the correlation during market downturns and overestimates it during market upturns. Thus using the linear correlation may underestimate risk and cause large losses when the market declines.

A series of H statistics are developed for measuring and testing correlation asymmetries. The H statistics involving ES-implied correlations clearly demonstrate that correlations between US and other G7 countries are significantly underestimated by the linear correlation during market downturns.

The H-statistics can also be used to measure and examine the impact of the choice of weights on VaR-implied correlations and ES-implied correlations. The H-statistics imply that the implied correlations are overall independent of the choice of weights, which suggests the possibility of applying the risk measure-implied correlations in asset allocation.

Besides risk management and asset allocation, the ES-implied correlation has another two potential applications. First, the standardized ES-implied correlation can be used to test or find the distribution of returns. If the assumption of the distribution is correct, the standardized ES-implied correlation should be close to the linear correlation. Thus a statistic measuring the difference between the linear correlation and the standardized ES-implied correlation can evaluate the accuracy of the hypothesis of asset distributions. By varying the hypothesis, one can find the true distribution of returns.

Second, the ES-implied correlation can be applied to measure the dependence of distributions whose second moment does not exist, for example, the stable distribution. The linear correlation is no longer accessible in the situation. However, expected shortfall is accessible as long as the first moment exists. Thus if we obtain a relation between the ES of the portfolio and the ES of individual assets, we can have a formula to estimate the dependence between assets whose second moments do not exist.

References

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, *26*(7), 1487–1503.

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies*, *15*(4), 1137–1187.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, *63*(3), 443–494.

Anscombe, F. J., & Glynn, W. J. (1983). Distribution of the kurtosis statistic b2 for normal samples. *Biometrika*, *70*(1), 227–234.

Artavanis, N. T. (2014). On the estimation of systematic downside risk. *Working Paper*.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1997). Thinking coherently. *Risk*, *10*(18), 68–71.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

Beaulieu, M.-C., Dufour, J.-M., & Khalaf, L. (2014). Exact confidence sets and goodness-of-fit methods for stable distributions. *Journal of Econometrics*, *181*(1), 3–14.

Campbell, R., Forbes, C. S., Koedijk, K. G., & Kofman, P. (2008). Increasing correlations or just fat tails? *Journal of Empirical Finance*, *15*(2), 287–309.

Campbell, R., Koedijk, K., & Kofman, P. (2002). Increased correlation in bear markets. *Financial Analysts Journal*, *58*(1), 87–94.

Castellacci, G., & Siclari, M. J. (2003). The practice of Delta–Gamma VaR: Implementing the quadratic portfolio model. *European Journal of Operational Research*, *150*(3), 529–545.

Cotter, J., & Longin, F. (2011). Implied correlation from VaR. arXiv preprint arXiv:1103.5655.

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g1. *Biometrika*, 679–681.

Danielsson, J., & De Vries, C. G. (2000). Value-at-risk and extreme returns. *Annals of Economics and Statistics*, 239–270.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, *64*(2), 247–254.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, *74*(1), 33–43.

Dufour, J.-M. (2006). Monte carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics*, *133*(2), 443–477.

Dufour, J.-M., & Kurz-Kim, J.-R. (2010). Exact inference and optimal invariant estimation for the stability parameter of symmetric α-stable distributions. *Journal of Empirical Finance*, *17*(2), 180–194.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *The Annals of Mathematical Statistics*, 181–187.

Fernández, C., & Steel, M. F. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, *93*(441), 359–371.

Fernandez, V. (2010). Extreme value theory and value at risk. *Revista de Análisis Económico–Economic Analysis Review*, *18*(1), 57–85.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, *10*(4), 507–521.

Fortin, I., & Kuzmics, C. (2002). Tail-dependence in stock-return pairs. *International Journal of Intelligent Systems in Accounting, Finance & Management*, *11*(2), 89–107.

Garcia, R., & Tsafack, G. (2011). Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking & Finance*, *35*(8), 1954–1970.

Gayen, A. (1951). The frequency distribution of the product-moment correlation coefficient in random samples of any size drawn from non-normal universes. *Biometrika*, *38*(1-2), 219–247.

Gilks, W. R. (2005). *Markov chain monte carlo*. Wiley Online Library.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica: Journal of the econometric society*, 413–430.

Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, *57*(1), 97–109.

Inui, K., & Kijima, M. (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of banking & finance*, *29*(4), 853–864.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance*, *56*(2), 649–676.

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative risk management: concepts, techniques, and tools*. Princeton university press.

Mittnik, S. (2014). VaR–implied Tail–correlation matrices. *Economics Letters*, *122*, 69-73.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence*. *International Economic Review*, *47*(2), 527–556.

Royston, J. (1982). An extension of shapiro and wilk's w test for normality to large samples. *Applied Statistics*, 115–124.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Shorack, G. R., & Wellner, J. A. (2009). *Empirical processes with applications to statistics* (Vol. 59). SIAM.

Stephens, M. A. (1986). Tests based on edf statistics. *Goodness-of-fit Techniques*, *68*, 97–193.

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking & Finance*, *26*(7), 1519–1533.

Thode, H. C. (2002). *Testing for normality* (Vol. 164). CRC press.

von Rohr, P., Hoeschele, I., et al. (2002). Bayesian QTL mapping using skewed Student-t distributions. *Genetics Selection Evolution*, *34*(1), 1–22.

Yamai, Y., & Yoshiba, T. (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance*, *29*(4), 997–1015.

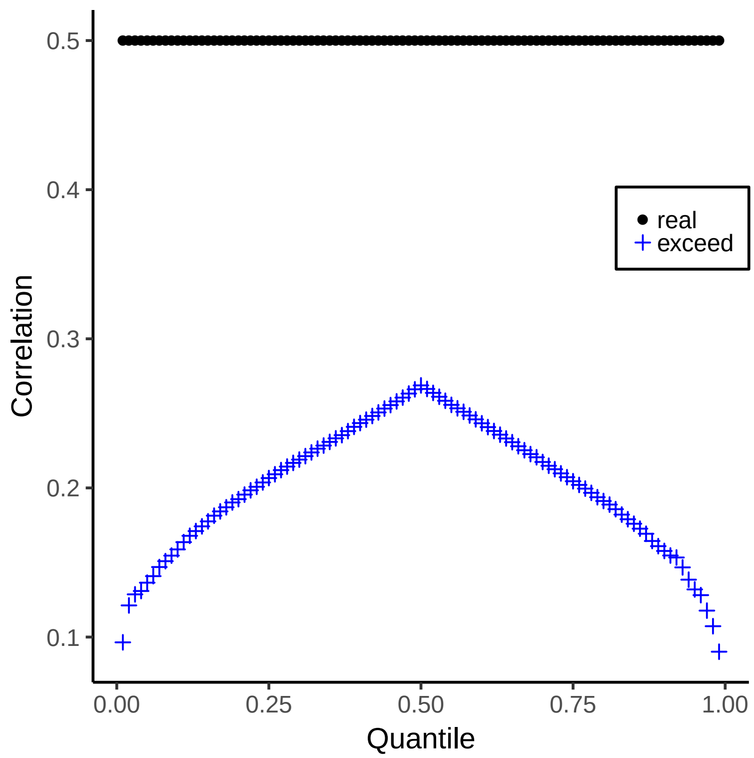
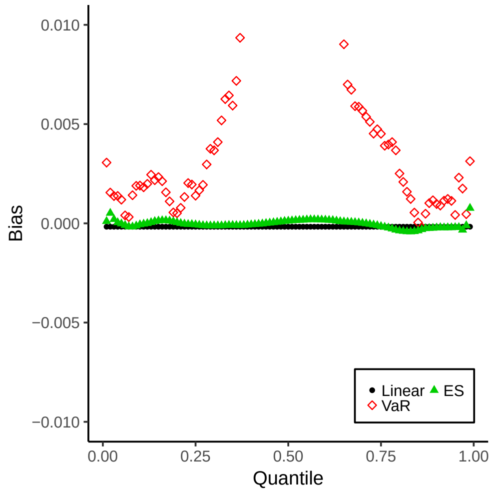


Figure 1: Exceedance correlation under normal assumption

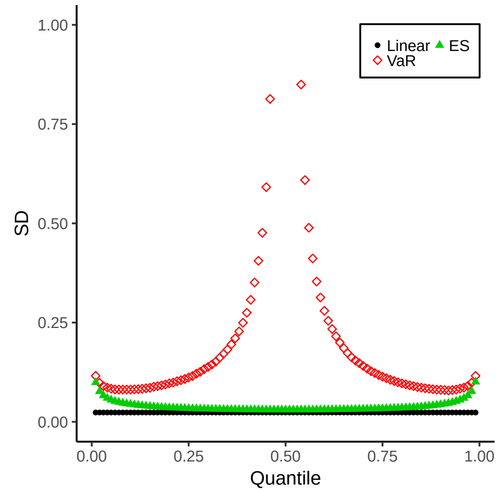
This figure plots the average of the exceedance correlation conditional on quantiles across the distribution of returns. Returns are drawn from a standard bivariate normal distribution with correlation of 0.5. The sample size is 10^3 and the exceedance correlation reported is the average over 10^4 repetitions. The black points plot the actual correlation of 0.5. The blue points plot exceedance correlations between X and Y given by

$$\rho(\alpha) = \begin{cases} Corr(X, Y|X \leq q_X(\alpha), Y \leq q_Y(\alpha)), \alpha < 0.5 \\ Corr(X, Y|X \geq q_X(\alpha), Y \geq q_Y(\alpha)), \alpha > 0.5 \end{cases}$$

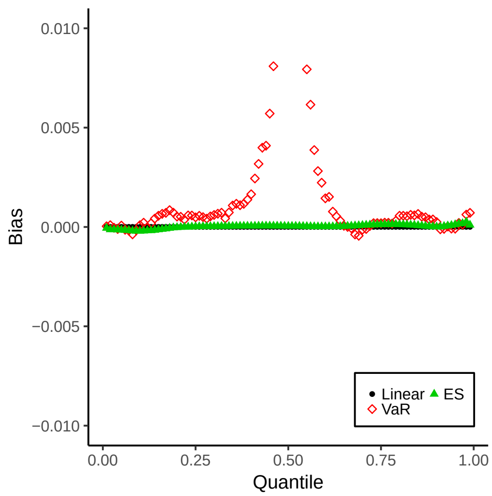
where α is the probability level.



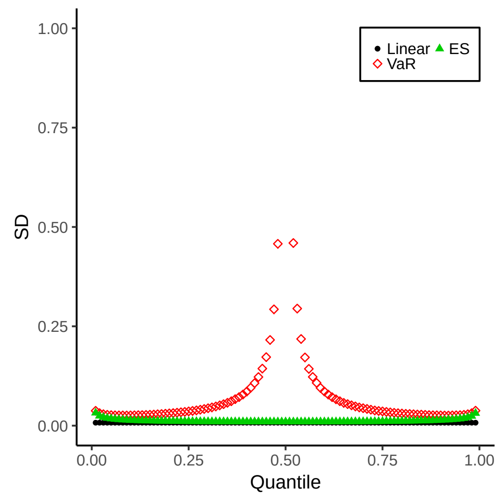
(a) Average bias, $T = 10^3$



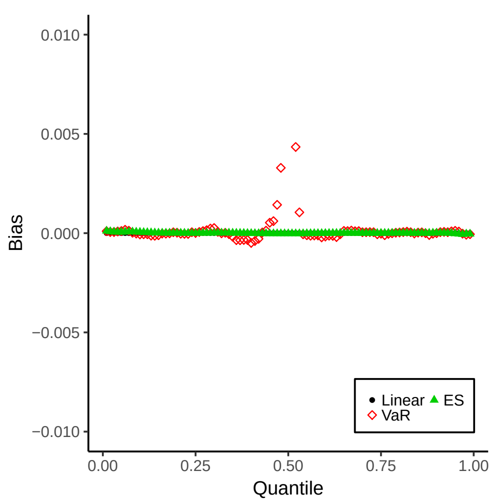
(b) Standard deviation, $T = 10^3$



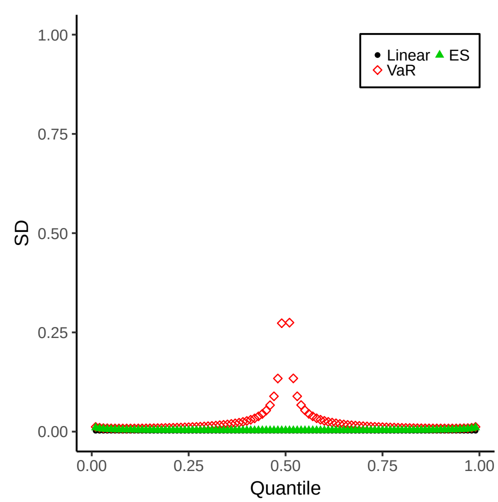
(c) Average bias, $T = 10^4$



(d) Standard deviation, $T = 10^4$

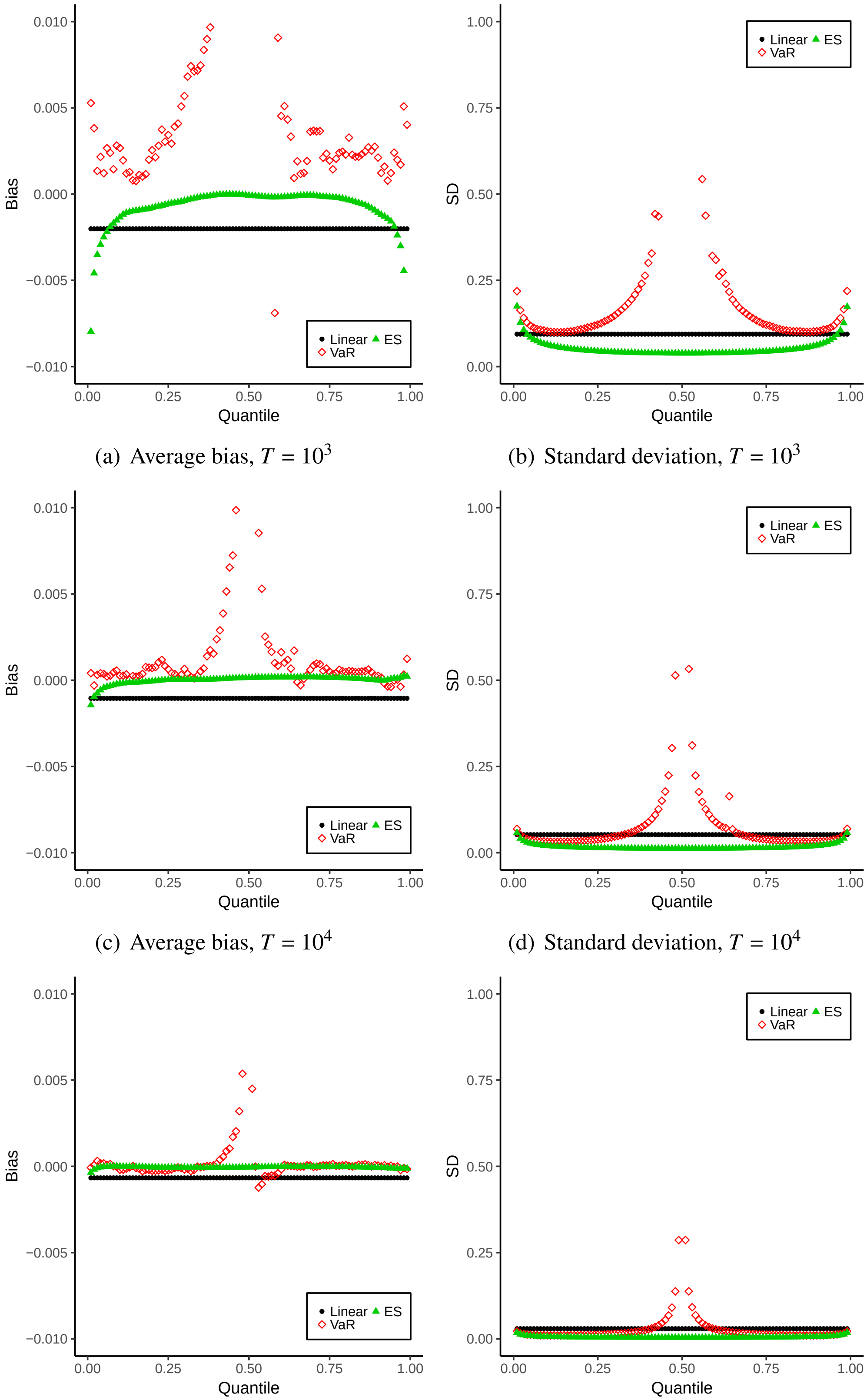


(e) Average bias, $T = 10^5$



(f) Standard deviation, $T = 10^5$

Figure 2: Average bias and standard deviation of estimated correlations over 10^4 repetitions in the multivariate normal distribution, $\rho = 0.5$, $T = 10^3, 10^4$ and 10^5.



(a) Average bias, $T = 10^3$
(b) Standard deviation, $T = 10^3$
(c) Average bias, $T = 10^4$
(d) Standard deviation, $T = 10^4$
(e) Average bias, $T = 10^5$
(f) Standard deviation, $T = 10^5$

Figure 3: Average bias and standard deviation of estimated correlations over 10^4 repetitions in the multivariate T distribution, $\rho = 0.5$, $T = 10^3, 10^4$ and 10^5.

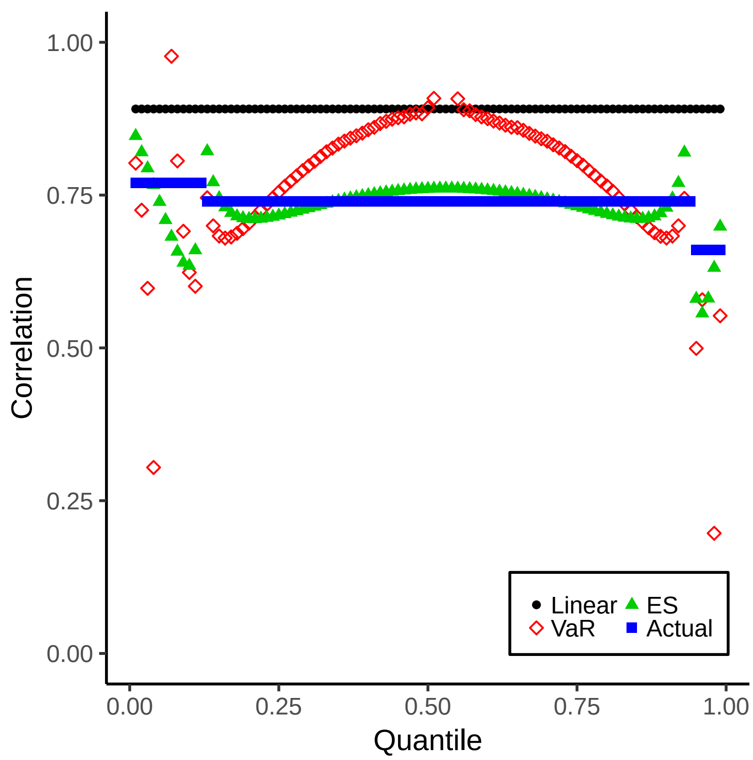
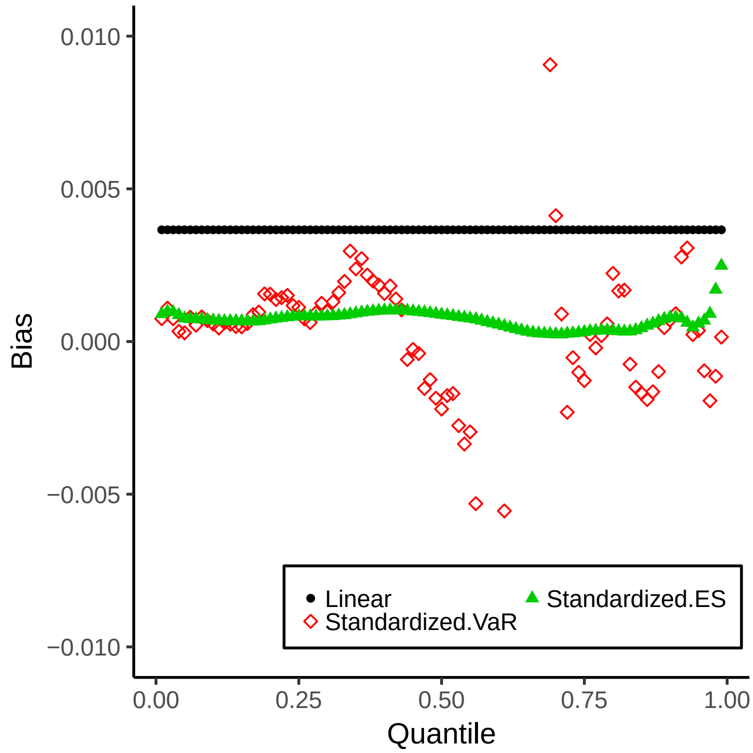


Figure 4: Estimated correlations in simulations of the non-constant model. The sample size is 10^4, and the process is repeated 10^3 times, $\rho = 0.77$ when $\alpha \leq 12\%$; $\rho = 0.74$, when $12\% < \alpha \leq 94\%$ and $\rho = 0.66$, when $\alpha \geq 94\%$.

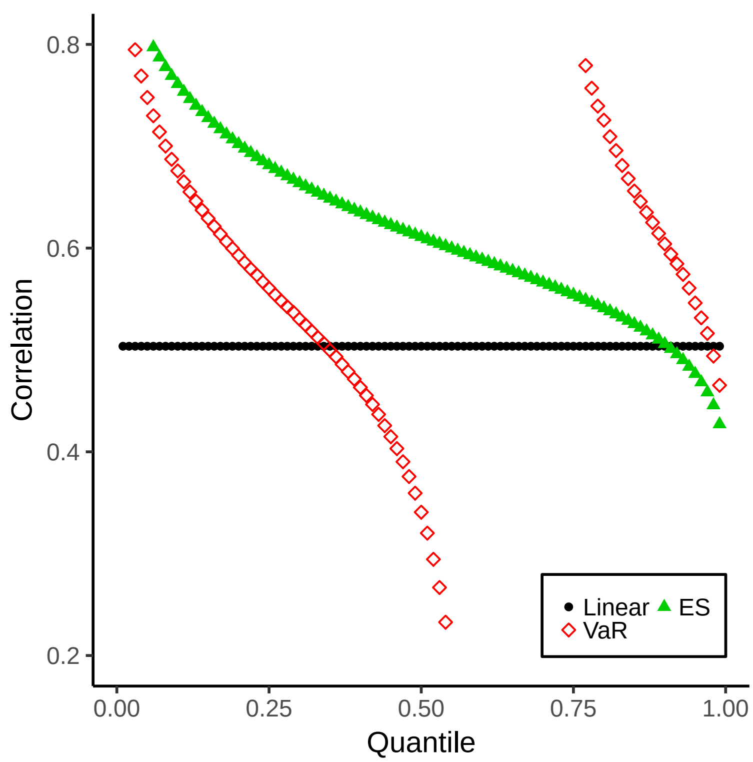
Table 1: Summary statistics of the RMSE in the simulation of case 3

This table presents the summary statistics of the RMSE of the estimated correlations in the simulation of case 3. $100p_1\%$ of data falls into $(-\infty,-1]\times(-\infty,-1]$ with the correlation 0.77, $100p_2\%$ of data falls into $(-1,1]\times(-1,1]$ with the correlation 0.74, and $100p_3\%$ of data falls into $(1,\infty)\times(1,\infty)$ with the correlation 0.66. Four sets of $p_1, p_2, \text{and} p_3$ are considered: $p_1 = 0, p_2 = 1, p_3 = 0$; $p_1 = 2\%, p_2 = 96\%, p_3 = 2\%$; $p_1 = 12\%, p_2 = 82\%, p_3 = 6\%$; and $p_1 = 12\%, p_2 = 76\%, p_3 = 12\%$. The quantiles of RMSE at probability level $\alpha = 1\%, 5\%, 10\%, 90\%, 95\%$ and 99% are also reported.

Statistics	Linear	VaR	ES	Statistics	Linear	VaR	ES
Situation 1: p_1=0, p_2=1, p_3=0				*Situation 2: p_1=2%, p_2=96%,p_3=2%*			
Mean	0.005	2.045	0.021	Mean	0.092	2.416	0.022
Std	0	19.100	0.013	Std	0.014	21.800	0.016
Min	0.005	0.022	0.008	Min	0.005	0.023	0.008
Max	0.005	190.200	0.094	Max	0.158	217	0.098
$\alpha = 1\%$	0.005	0.023	0.008	$\alpha = 1\%$	0.093	0.023	0.008
$\alpha = 5\%$	0.005	0.024	0.008	$\alpha = 5\%$	0.093	0.025	0.008
$\alpha = 10\%$	0.005	0.0273	0.009	$\alpha = 10\%$	0.093	0.030	0.009
$\alpha = 90\%$	0.005	0.225	0.032	$\alpha = 90\%$	0.093	0.283	0.033
$\alpha = 95\%$	0.005	0.353	0.033	$\alpha = 95\%$	0.093	0.551	0.042
$\alpha = 99\%$	0.005	4.926	0.093	$\alpha = 99\%$	0.093	12.300	0.093
Situation 3: p_1=12%, p_2=82%, p_3=6%				*Situation 4: p_1=12%, p_2=76%, p_3=12%*			
Mean	0.201	61.770	0.031	Mean	0.226	29.920	0.033
Std	0.034	59.49	0.029	Std	0.038	212.900	0.030
Min	0.119	0.021	0.009	Min	0.140	0.022	0.009
Max	0.274	5919	0.139	Max	0.295	2019	0.138
$\alpha = 1\%$	0.119	0.022	0.009	$\alpha = 1\%$	0.140	0.022	0.009
$\alpha = 5\%$	0.119	0.025	0.009	$\alpha = 5\%$	0.140	0.027	0.010
$\alpha = 10\%$	0.119	0.033	0.010	$\alpha = 10\%$	0.140	0.036	0.011
$\alpha = 90\%$	0.209	0.391	0.080	$\alpha = 90\%$	0.295	0.439	0.080
$\alpha = 95\%$	0.216	0.722	0.096	$\alpha = 95\%$	0.295	0.753	0.133
$\alpha = 99\%$	0.274	295.700	0.134	$\alpha = 99\%$	0.295	601.100	0.133



(a) Average bias



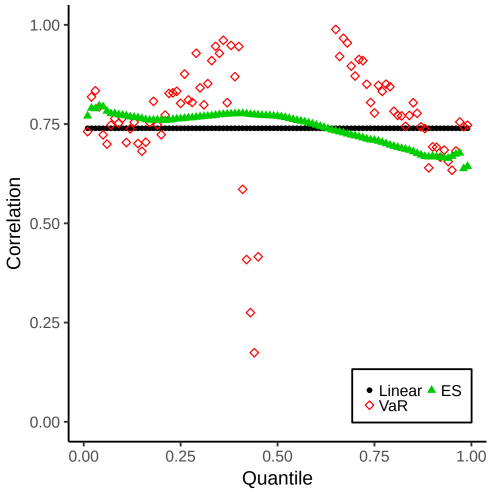
(b) Correlations

Figure 5: Estimated standardized and non-standardized correlations in the skewness T distribution, $T = 10^4$, $\rho = 0.5$.

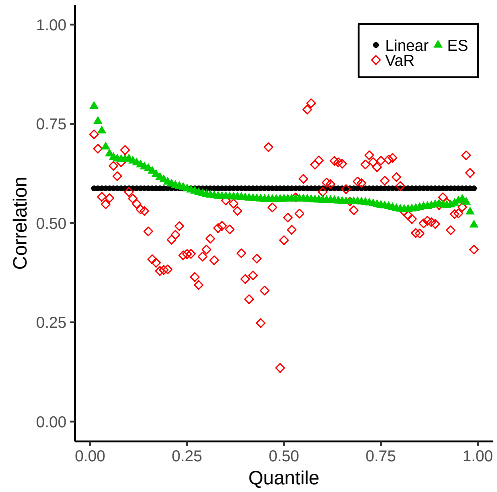
Table 2: Summary statistics of G7 countries

This table presents the summary statistics of annualized returns of G7 equity indices. Panel A reports statistics of returns in the whole period from 1/1/1973 to 31/12/2015. Panel B reports statistics of returns in the first subperiod from 1/1/1973 to 31/12/2002 and panel C reports statistics of the second period from 1/1/2003 to 31/12/2015. The superscripts *, ** and *** represent significance at 10%, 5%, and 1%, respectively.

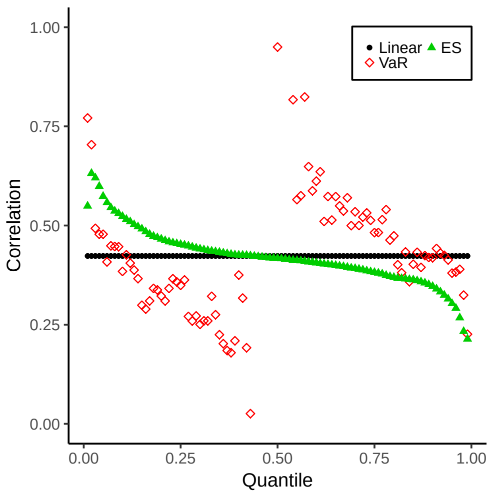
Statistics	US	CA	UK	IT	DE	FR	JP
Panel A: Summary Statistics in whole period							
Mean	0.113	0.106	0.128	0.126	0.096	0.132	0.068
SD	1.230	1.139	1.325	1.689	1.264	1.433	1.363
Min	-7.808	-8.425	-8.461	-8.693	-7.482	-9.368	-10.080
Max	7.628	7.111	12.546	10.388	6.511	7.727	8.666
Skewness	-0.295***	-0.602***	0.227***	-0.020	-0.440***	-0.371***	-0.226***
Kurtosis	7.315***	8.482***	11.141***	5.579***	6.102***	6.033***	7.441***
AD test	14.640***	18.476***	19.608***	12.121***	15.201***	12.225***	17.574***
Shapiro test	0.957***	0.945***	0.934***	0.974***	0.964***	0.969***	0.958***
Panel B: Summary Statistics in period I							
Mean	0.117	0.111	0.143	0.150	0.086	0.144	0.059
SD	1.223	1.102	1.365	1.762	1.217	1.449	1.258
Min	-7.808	-6.725	-8.461	-8.693	-7.482	-9.368	-6.005
Max	6.627	5.293	12.546	10.388	6.317	6.718	8.305
Skewness	-0.360***	-0.495***	0.366***	0.054	-0.497***	-0.423***	0.044
Kurtosis	6.776***	6.309***	11.997***	5.540***	6.395***	6.062***	5.926***
AD test	7.732***	7.544***	12.916***	7.685***	9.135***	7.264***	12.928***
Shapiro test	0.964***	0.968***	0.930***	0.975***	0.963***	0.970***	0.967***
Panel C: Summary Statistics in period II							
Mean	0.103	0.096	0.094	0.072	0.118	0.102	0.090
SD	1.245	1.222	1.226	1.509	1.367	1.397	1.580
Min	-7.519	-8.425	-5.409	-6.590	-5.993	-5.999	-10.080
Max	7.628	7.111	8.090	6.957	6.511	7.727	8.666
Skewness	-0.151	-0.773***	-0.239**	-0.333***	-0.353***	-0.241**	-0.546***
Kurtosis	8.475***	11.584***	7.587***	5.181***	5.496***	5.953***	8.185***
AD test	8.009***	13.009***	7.231***	5.486***	6.140***	5.600***	4.596***
Shapiro test	0.937***	0.892***	0.946***	0.969***	0.964***	0.962***	0.949***



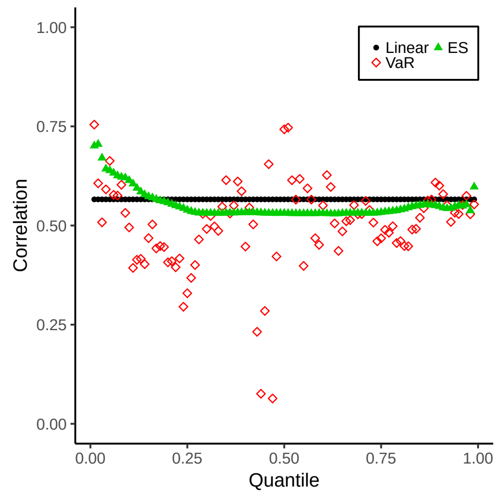
(a) US and CA



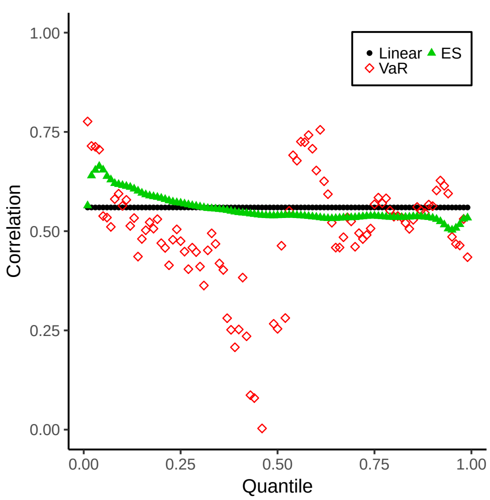
(b) US and UK



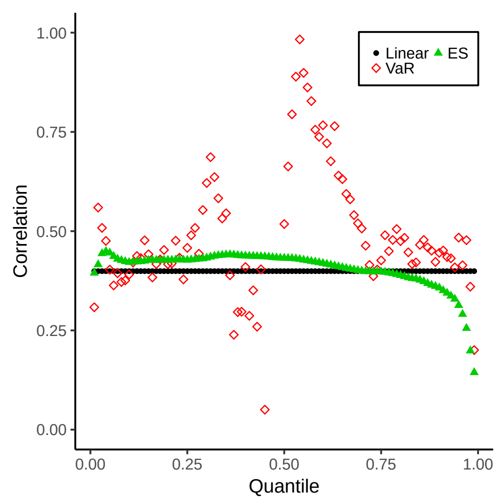
(c) US and IT



(d) US and DE



(e) US and FR



(f) US and JP

Figure 6: Estimated correlations between US and other G7 countries in the whole period.

Table 3: H statistics measuring correlation asymmetries

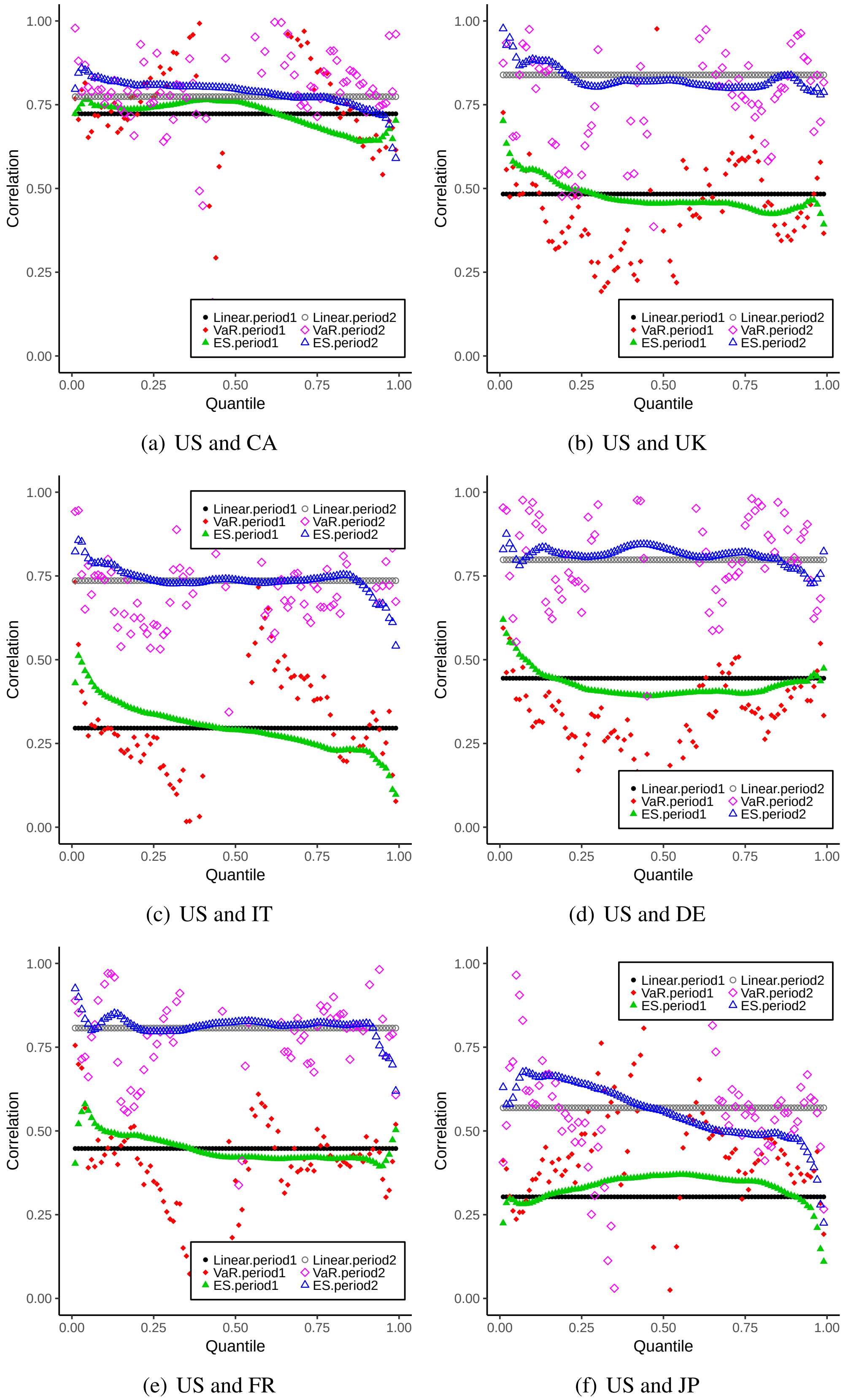
This table reports the H and AH statistics. Panel A reports downside and upside H statistics based on the VaR-implied and the ES-implied correlations, respectively, where downside and upside H statistics are the supremums of deviations of linear correlation from conditional correlations in the left tail and right tail, respectively. Panel B reports downside and upside AH statistics, which correspond to the average of deviations in the left tail and right tail, respectively. The superscripts *, ** and *** represent significance at 10%, 5%, and 1%, respectively.

	CA	UK	IT	DE	FR	JP
			Panel A: H statistics			
H^-_{VaR}	0.167	0.134	0.335***	0.186*	0.295***	0.219**
H^-_{ES}	0.054	0.212***	0.221***	0.143***	0.105**	0.048
H^+_{VaR}	0.111	0.157	0.181**	0.119	0.111	0.185**
H^+_{ES}	0.106**	0.092*	0.215***	0.036	0.056	0.247***
			Panel B: AH statistics			
AH^-_{VaR}	0.033**	-0.084	-0.036	-0.092	-0.029	0.045**
AH^-_{ES}	0.030***	0.050***	0.080***	0.022*	0.040***	0.027*
AH^+_{VaR}	-0.025	0.027	-0.012	0.048***	0.021	-0.043
AH^+_{ES}	0.057***	0.046***	0.080***	0.022*	0.030**	0.045***

Table 4: H statistics measuring impact of weights

This table reports H and AH statistics measuring the impact of weights. H is the supremum of absolute difference when changing values of weights or using different estimation methods across the left and right tail. AH is the average of the absolute difference across the tails. Panel A reports the difference in implied correlations from choosing different values for portfolio weights. Panel B reports the difference in implied correlations from using exact identification and overidientification. The superscripts *, ** and *** represent significance at 10%, 5%, and 1%, respectively.

	CA	UK	IT	DE	FR	JP
Panel A: Difference from different values of weights						
H_{VaR}	0.130	0.213	0.192	0.140	0.199	0.187
H_{ES}	0.026	0.080	0.056	0.061	0.037	0.077
AH_{VaR}	0.046	0.063	0.058	0.057	0.049	0.080**
AH_{ES}	0.012	0.009	0.010	0.009	0.007	0.020
Panel B: Difference from exact and overidentification						
H_{VaR}	0.098	0.063	0.100	0.064	0.090	0.083
H_{ES}	0.045	0.025	0.033	0.013	0.018	0.036
AH_{VaR}	0.018	0.023	0.026	0.019	0.021	0.031
AH_{ES}	0.007	0.004	0.004	0.004	0.002	0.005



(a) US and CA

(b) US and UK

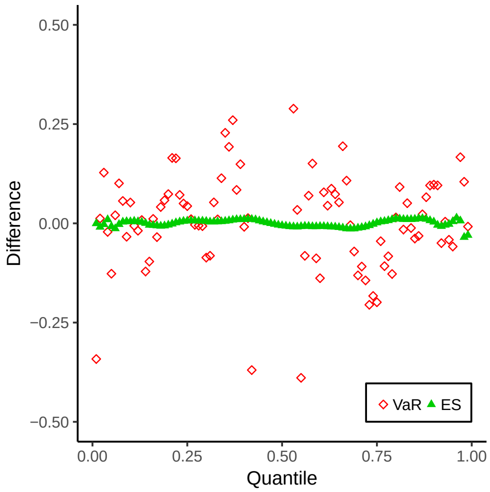
(c) US and IT

(d) US and DE

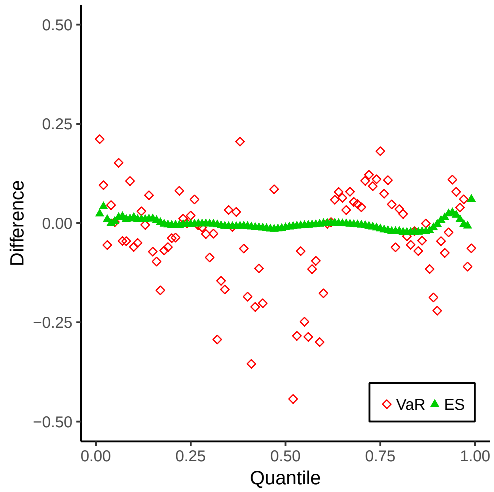
(e) US and FR

(f) US and JP

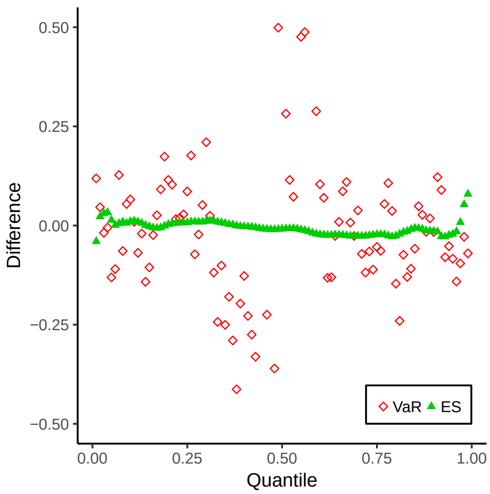
Figure 7: Estimated correlations between US and other G7 countries in the subperiods.



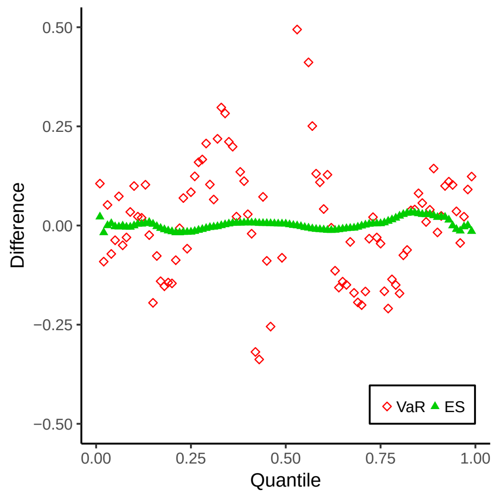
(a) US and CA



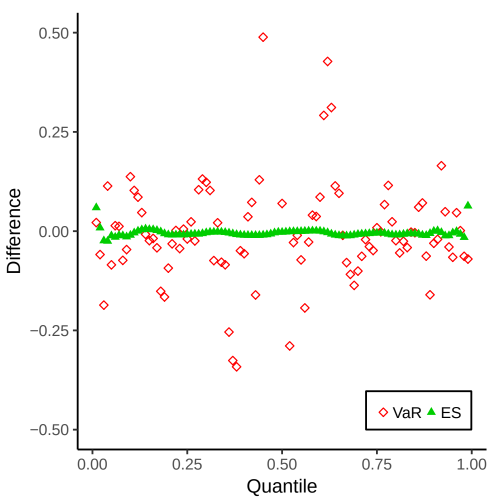
(b) US and UK



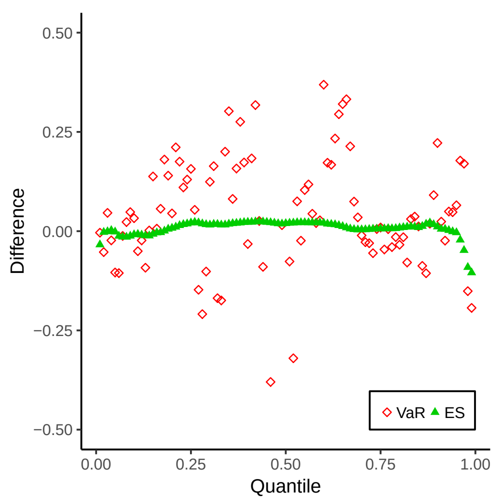
(c) US and IT



(d) US and DE

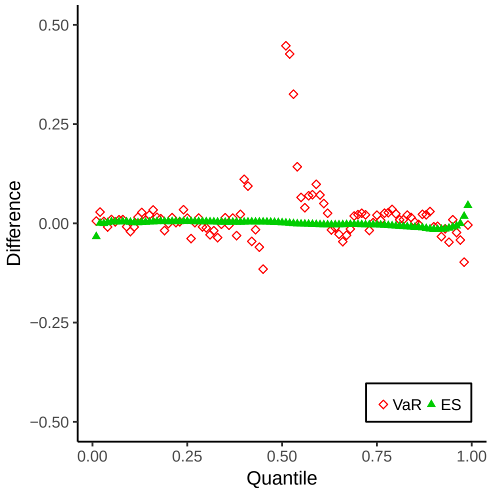


(e) US and FR

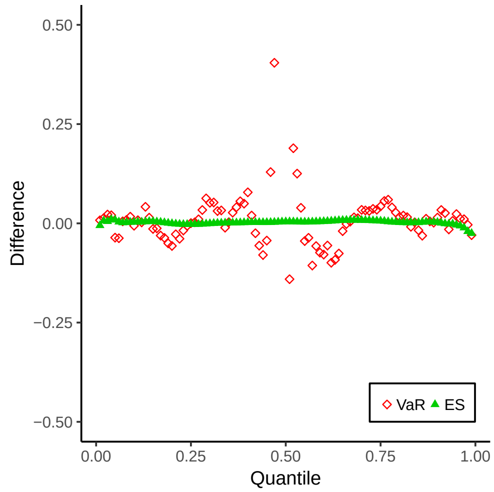


(f) US and JP

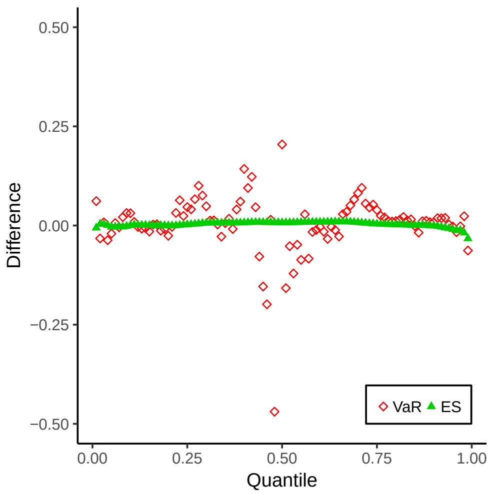
Figure 8: Difference in estimated correlations between US and other G7 countries from using different values of weights.



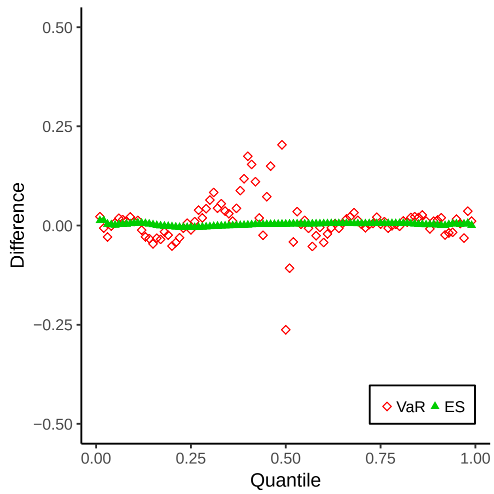
(a) US and CA



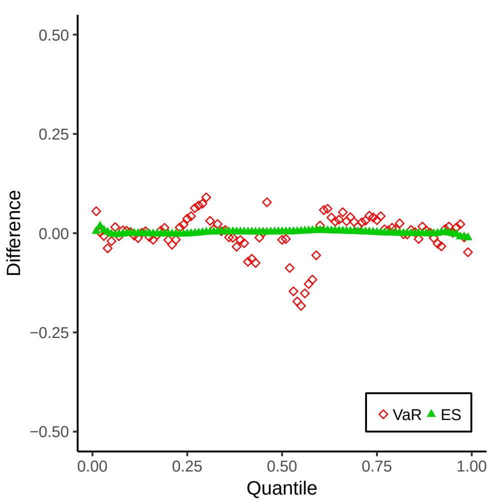
(b) US and UK



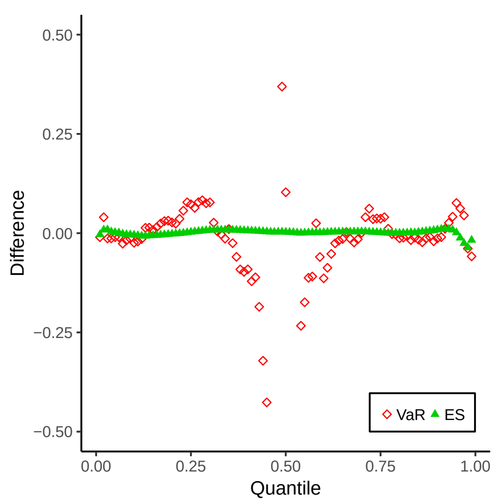
(c) US and IT



(d) US and DE



(e) US and FR



(f) US and JP

Figure 9: Difference in estimated correlations between US and other G7 countries from exact identification and overidentification.

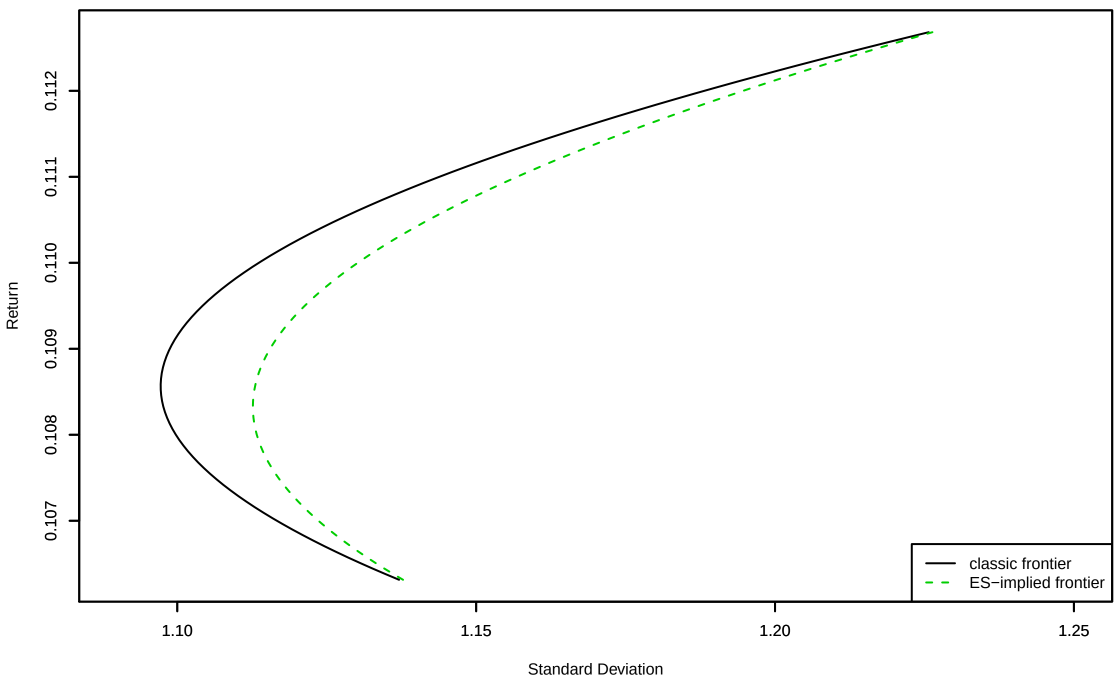


Figure 10: The classic mean-variance efficient frontier and the efficient frontier adjusted by the ES-implied correlation at 5%-quantile.

Table 5: H statistics measuring the impact of using optimal weights

This table reports H and AH statistics measuring the difference in implied correlations from using equal weights and using the optimal tangency weights in the mean-variance framework with variance constructed by implied correlations. The superscripts *, ** and *** represent significance at 10%, 5%, and 1%, respectively.

	CA	UK	IT	DE	FR	JP
H_{VaR}	0.045	0.108	0.253	0.346^{*}	0.101	0.482^{*}
H_{ES}	0.002	0.027	0.040	0.042	0.033	0.060
AH_{VaR}	0.009	0.031	0.083	0.104^{*}	0.032	0.176^{*}
AH_{ES}	0.001	0.008	0.013	0.014	0.006	0.018