Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation

Yongmiao Hong
Cornell University and Xiamen University

Jun Tu
Singapore Management University

Guofu Zhou
Washington University and CCFR

We provide a model-free test for asymmetric correlations in which stocks move more often with the market when the market goes down than when it goes up, and also provide such tests for asymmetric betas and covariances. When stocks are sorted by size, book-to-market, and momentum, we find strong evidence of asymmetries for both size and momentum portfolios, but no evidence for book-to-market portfolios. Moreover, we evaluate the economic significance of incorporating asymmetries into investment decisions, and find that they can be of substantial economic importance for an investor with a disappointment aversion (DA) preference as described by Ang, Bekaert, and Liu (2005). (*JEL* C12, C15, C32, G12)

A number of recent studies discuss the asymmetric characteristics of asset returns. Ball and Kothari (1989), Schwert (1989), Conrad, Gultekin and Kaul (1991), Cho and Engle (2000), and Bekaert and Wu (2000), among others, document asymmetries in covariances, volatilities, and betas of stock returns. Harvey and Siddique (2000) analyze asymmetries in higher moments. Of particular interest to this article are the asymmetric correlations of stock returns with the market indices, which were studied by Karolyi and Stulz (1996), Ang and Bekaert (2000), Longin and Solnik (2001), Ang and Chen (2002), and Bae, Karolyi, and Stulz (2003). In particular, Longin and Solnik (2001) find that international markets have

We are grateful to Yacine Aït-Sahalia (the Editor), Andrew Ang, Joseph Chen (the ACFEA discussant), Robert Connolly, Alex David, Heber Farnsworth, Michael Faulkender, Raymond Kan, Hong Liu, Tom Miller, Ľuboš Pástor, Andrew Patton (the AFA discussant), and Bruno Solnik, seminar participants at Tsinghua University and Washington University in St. Louis, the 14th Annual Conference on Financial Economics and Accounting and 2004 American Finance Association Meetings for many helpful comments, and especially to an anonymous referee for numerous insightful and detailed comments that improved the article drastically. We also thank Lynnea Brumbaugh-Walter for many helpful editorial comments. Hong and Tu acknowledge financial support for this project from NSF Grant SES-0111769 and the Changjiang Visiting Scholarship of Chinese Ministry of Education, and from Singapore Management University Research Grant C207/MSS4B023, respectively. Zhou is grateful to China Center for Financial Research (CCFR) of Tsinghua University for visits in April 2005 and May 2006 where part of the article was written. Address correspondence to Guofu Zhou, Olin School of Business, Washington University, St. Louis, MO 63130, or e-mail: zhou@wustl.edu.

© The Author 2007. Published by Oxford University Press on behalf of The Society for Financial Studies. All rights reserved. For Permissions, please email: journals.permissions@oxfordjournals.org.
doi:10.1093/rfs/hhl037 Advance Access publication September 23, 2006

greater correlations with the US market when it is going down than when it is going up, and Ang and Chen (2002) find strong asymmetric correlations between stock portfolios and the US market. The study of asymmetric correlations is important for two reasons. First, hedging relies crucially on the correlations between the assets hedged and the financial instruments used. The presence of asymmetric correlations can cause problems in hedging effectiveness. Second, though standard investment theory advises portfolio diversification, the value of this advice might be questionable if all stocks tend to fall as the market falls.

However, assessing asymmetric correlations requires care. Stambaugh (1995), Boyer, Gibson, and Loretan (1999), and Forbes and Rigobon (2002) find that a correlation computed conditional on some variables being high or low is a biased estimator of the unconditional correlation. Therefore, even if one obtains from the real data a conditional correlation that is much higher than the unconditional sample correlation, it is not sufficient to claim the existence of asymmetric correlations. A formal statistical test must then account for both sample variations and the bias induced by conditioning. Ang and Chen (2002) seem the first to propose such a test.[1] Given a statistical model for the data, their test compares the sample conditional correlations with those implied by the model. If there is a large difference, then the observed asymmetric correlations cannot be explained by the model. However, despite its novelty, Ang and Chen's test answers only the question whether the asymmetry can be explained by a given model.

The first contribution of this article is to propose a test to answer the question whether the data are asymmetric at all. The test has three appealing features. First, it is model free. One can use it without having to specify a statistical model for the data. In other words, if symmetry is rejected by our test, then the data cannot be modeled by *any* symmetry distributions (under standard regularity conditions). Second, unlike many asymmetry studies that impose the normality assumption on the data, our test allows for general distributional assumptions, such as the GARCH process. Third, the test statistic is easy to implement, and its asymptotic distribution follows a standard chi-square distribution under the null hypothesis of symmetry. Therefore, our proposed test can be straightforwardly applied elsewhere.

While asymmetric correlations seem obviously important from a management perspective in hedging risk exposures (e.g., Jaeger (2003, p. 216)), betas are closely related to asset pricing theories, and useful in understanding the riskiness of the associated stocks. Ball and Kothari

[1] Their test tends to over-reject based on their normal approximation to its distribution. The problem, however, disappears completely by using the asymptotic distribution provided in Appendix A of this article.

(1989), Conrad, Gultekin and Kaul (1991), Cho and Engle (2000), and Bekaert and Wu (2000), among others, document asymmetries in betas of stock returns, but provide no formal statistical tests. The second contribution of this article is to fill the gap by providing a model-free test of beta symmetry. In addition, we develop such a test for symmetric covariances. This is of interest because covariances are usually the direct parameter inputs for optimal portfolio choice, while betas are primarily useful in understanding assets' systematic risks associated with the market/factors in general.

Since the presence of statistically significant asymmetry may not necessarily be economically important (and vice versa), the third contribution of this article is to provide a Bayesian framework for modeling asymmetry and for assessing its economic importance from an investment perspective. A mixture model of normal and Clayton copulas is proposed for the data. We develop algorithms for drawing samples from both Bayesian posterior and predictive distributions. To assess the economic value of asymmetry, we consider the portfolio choice problem of an investor who is uncertain about whether asymmetry in stock returns exists. In the spirit of Kandel and Stambaugh (1996) and Pástor and Stambaugh (2000), we ask what utility gains an investor can achieve if he switches from a belief in symmetric returns to a belief in asymmetric ones. On the basis of the DA preference of Ang, Bekaert, and Liu (2005), we compute the utility gains when the investor's felicity function is of the power utility form and when his coefficient of DA is between 0.55 and 0.25. We find that he can achieve over 2% annual certainty-equivalent gains when he switches from a belief in symmetric stock returns to a belief in asymmetric ones.

The remainder of the article is organized as follows. Section 1 provides statistical tests for various symmetries. Section 2 applies the tests to stock portfolios grouped by size, book-to-market, and momentum, respectively, to assess their asymmetries, and introduces the copula model to capture the detected asymmetries. Section 3 discusses Bayesian portfolio decisions that incorporate asymmetries. Section 4 concludes the article.

1. Symmetry Tests

In this section, we provide three model-free symmetry tests. The first is on symmetry in correlations, and the other two are on symmetry in betas and covariances, respectively. While their intuition and asymptotic distributions are discussed in this section, their small sample properties are studied later because we need the data to calibrate parameters for simulations.

1.1 Testing correlation symmetry

Let $\{R_{1t}, R_{2t}\}$ be the returns on two portfolios in period t. Following Longin and Solnik (2001) and Ang and Chen (2002), we consider the exceedance correlation between the two series. A correlation at an exceedance level c is defined as the correlation between the two variables when both of them exceed c standard deviations away from their means, respectively,

$$\rho^+(c) = \text{corr}(R_{1t}, R_{2t} | R_{1t} > c, R_{2t} > c), \tag{1}$$

$$\rho^-(c) = \text{corr}(R_{1t}, R_{2t} | R_{1t} < -c, R_{2t} < -c), \tag{2}$$

where, following Ang and Chen (2002) and many others in the asymmetry literature, the returns are standardized to have zero mean and unit variance so that the mean and variance do not appear explicitly in the right-hand side of the definition, making easy both the computation and statistical analysis. The null hypothesis of symmetric correlation is

$$H_0: \quad \rho^+(c) = \rho^-(c), \quad \text{for all } c \geq 0. \tag{3}$$

That is, we are interested in testing whether the correlation between the positive returns of the two portfolios is the same as that between their negative returns. If the null hypothesis is rejected, there must exist asymmetric correlations. The alternative hypothesis is

$$H_A: \quad \rho^+(c) \neq \rho^-(c), \quad \text{for some } c \geq 0. \tag{4}$$

Longin and Solnik (2001) use extreme value theory to test whether $\rho^+(c)$ or $\rho^-(c)$ is zero as c becomes extremely large. In contrast, Ang and Chen (2002) provide a more direct test of the symmetry hypothesis. For a set of random samples, $\{R_{1t}, R_{2t}\}_{t=1}^T$, of size T, the exceedance correlations can be estimated by their sample analogues,

$$\hat{\rho}^+(c) = \hat{\text{corr}}(R_{1t}, R_{2t} | R_{1t} > c, R_{2t} > c), \tag{5}$$

$$\hat{\rho}^-(c) = \hat{\text{corr}}(R_{1t}, R_{2t} | R_{1t} < -c, R_{2t} < -c). \tag{6}$$

That is, $\hat{\rho}^+(c)$ and $\hat{\rho}^-(c)$ are the standard sample correlations computed based on only those data that satisfy the tail restrictions. On the basis of these sample estimates, Ang and Chen (2002) propose an H statistic for testing correlation symmetry:

$$H = \left[\sum_{i=1}^{m} w(c_i)(\rho(c_i, \phi) - \hat{\rho}(c_i))^2 \right]^{1/2}, \tag{7}$$

where $c_1, c_2, \ldots, c_m$ are m chosen exceedance levels, $w(c_i)$ is the weight (all weights sum to one), $\hat{\rho}(c_i)$ can be either $\hat{\rho}^+(c_i)$ or $\hat{\rho}^-(c_i)$, and $\rho(c_i, \phi)$ is the population exceedance correlation computed from a given model with

parameter ϕ. If H is large, this implies that the given model cannot explain the observed sample exceedance correlations. Hence, Ang and Chen (2002) test is useful in answering whether the empirical exceedance correlations can be explained by a given model.

In contrast, here we are interested in the question whether the data are asymmetric at all. This requires a model-free test. Intuitively, if the null is true, the following $m \times 1$ difference vector

$$\hat{\rho}^+ - \hat{\rho}^- = \left[\hat{\rho}^+(c_1) - \hat{\rho}^-(c_1), \ldots, \hat{\rho}^+(c_m) - \hat{\rho}^-(c_m)\right]' \tag{8}$$

must be close to zero. It can be shown (Appendix A) that, under the null of symmetry, this vector after being scaled by $\sqrt{T}$ has an asymptotic normal distribution with mean zero and a positive definite variance-covariance matrix Ω for all possible true distributions of the data satisfying some regularity conditions.

To construct a feasible test statistic, we need to estimate Ω. Let T_c^+ be the number of observations in which both R_{1t} and R_{2t} are larger than c simultaneously. Then the sample means and variances of the two conditional series are easily computed,

$$\hat{\mu}_1^+(c) = \frac{1}{T_c^+}\sum_{t=1}^{T} R_{1t}1(R_{1t} > c, R_{2t} > c),$$

$$\hat{\mu}_2^+(c) = \frac{1}{T_c^+}\sum_{t=1}^{T} R_{2t}1(R_{2t} > c, R_{2t} > c),$$

$$\hat{\sigma}_1^+(c)^2 = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T}[R_{1t} - \hat{\mu}_1^+(c)]^2 1(R_{1t} > c, R_{2t} > c),$$

$$\hat{\sigma}_2^+(c)^2 = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T}[R_{2t} - \hat{\mu}_2^+(c)]^2 1(R_{1t} > c, R_{2t} > c),$$

where $1(\cdot)$ is the indicator function. As a result, we can express the sample conditional correlation $\hat{\rho}^+(c)$ as

$$\hat{\rho}^+(c) = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T} \hat{X}_{1t}^+(c)\hat{X}_{2t}^+(c)1(R_{1t} > c, R_{2t} > c), \tag{9}$$

where

$$\hat{X}_{1t}^+(c) = \frac{R_{1t} - \hat{\mu}_1^+(c)}{\hat{\sigma}_1^+(c)}, \qquad \hat{X}_{2t}^+(c) = \frac{R_{2t} - \hat{\mu}_2^+(c)}{\hat{\sigma}_2^+(c)}.$$

Clearly, we can have a similar expression for $\hat{\rho}^-(c)$.

Then, under general conditions, a consistent estimator of Ω is given by the following almost surely positive definite matrix,

$$\hat{\Omega} = \sum_{l=1-T}^{T-1} k(l/p)\hat{\gamma}_l, \tag{10}$$

where $\hat{\gamma}_l$ is an $N \times N$ matrix with (i, j)-th element

$$\hat{\gamma}_l(c_i, c_j) = \frac{1}{T} \sum_{t=|l|+1}^{T} \hat{\xi}_t(c_i)\hat{\xi}_{t-|l|}(c_j), \tag{11}$$

and

$$\hat{\xi}_t(c) = \frac{T}{T_c^+}[\hat{X}_{1t}^+(c)\hat{X}_{2t}^+(c) - \hat{\rho}^+(c)]1(R_{1t} > c, R_{2t} > c) - \frac{T}{T_c^-}[\hat{X}_{1t}^-(c)\hat{X}_{2t}^-(c) - \hat{\rho}^-(c)]1(R_{1t} < -c, R_{2t} < -c); \tag{12}$$

and $k(\cdot)$ is a kernel function that assigns a suitable weight to each lag of order l, and p is the smoothing parameter or lag truncation order (when $k(\cdot)$ has bounded support). In this article, we will use the Bartlett kernel,

$$k(z) = (1 - |z|)1(|z| < 1), \tag{13}$$

which is popular and is used by Newey and West (1994) and others. With these preparations, we are ready to define a statistic for testing the null hypothesis as

$$J_\rho = T(\hat{\rho}^+ - \hat{\rho}^-)'\hat{\Omega}^{-1}(\hat{\rho}^+ - \hat{\rho}^-). \tag{14}$$

On the basis of our earlier discussions, J_ρ summarizes the deviations from the null at various values of the cs.

However, the value of p has to be provided to compute the test statistic. There are two ways to choose p. The first is to take p as a nonstochastic known number, especially in the case where one wants to impose some lag structure on the data. Another choice is to allow it to be determined by the data with either the Andrews (1991) or the Newey and West (1994) procedure. Let $\hat{J}_\rho$ be the same J_ρ statistic except using $\hat{p}$, the data-driven p.

The following theorem provides the theoretical basis for making statistical inference based on J_ρ and $\hat{J}_\rho$:

Theorem 1: *Under the null hypothesis H_0 and under certain regularity conditions given in Appendix A,*

$$J_\rho \to^d \chi_m^2, \tag{15}$$

and

$$\hat{J}_\rho \to^d \chi^2_m, \tag{16}$$

as the sample size T approaches infinity.

Theorem 1 (proofs of all theorems are given in Appendix A) says that the correlation symmetry test has a simple asymptotic chi-square distribution with degrees of freedom m. So, the P-value of the test is straightforward to compute, making its applications easy to carry out.

As can be seen from the regularity conditions given in the proof, the test is completely model free. It is also robust to volatility clustering which is a well-known stylized fact for many financial time series. We have also explicitly justified the use of a data-driven bandwidth, say $\hat{p}$, and show that $\hat{p}$ has no impact on the asymptotic distribution of the test provided that $\hat{p}$ converges to p at a sufficiently fast rate. For simplicity, we will use $p = 3$ in what follows because the time-consuming, data-driven bandwidth does not make much difference in our simulation experiments. In addition, since the kernel estimator has a known small sample bias, we will, following Den Haan and Levin (1997, p. 310), replace T/T_c^+ of Equation (12) by $(T - T_c^+)/T_c^+$ and do the same for the T/T_c^- term, to make the test have better finite sample properties. It should also be noted that the asymptotic theory does not provide any guidance for the choice of the exceedance levels except that they are required to be distinct numbers. Intuitively, more levels or larger ones are likely to increase power, but they may lead to imprecise estimation of Ω to yield poor small sample properties. For this reason, like Ang and Chen (2002), we focus on using $C = \{0\}$ and $C = \{0, 0.5, 1, 1.5\}$ which seem to have reasonable finite sample performance, as shown in our simulations later.

Econometrically, our test is similar to constructing a Wald test in the Hansen (1982) generalized method of moments (GMM) framework. However, unlike the standard GMM, the sample moments are conditional ones and hence stronger regulation conditions are needed to ensure the convergence of the sample analogues to their suitable asymptotic limits. In particular, we need to impose

Assumption A.1: For the prespecified exceedance levels $c = (c_1, c_2, \ldots, c_m)' \in \mathbb{R}^m$, the variance-covariance matrix Ω, with (i, j)-th element $\Omega_{ij} \equiv \sum_{l=-\infty}^{\infty} \text{cov}[\xi_t(c_i), \xi_{t-l}(c_j)]$, is finite and nonsingular, where

$$\xi_t(c) = \frac{X_{1t}^+(c)X_{2t}^+(c) - \rho^+(c)}{\Pr(R_{1t} > c, R_{2t} > c)} - \frac{X_{1t}^-(c)X_{2t}^-(c) - \rho^-(c)}{\Pr(R_{1t} < -c, R_{2t} < -c)},$$

with $X_{kt}^{+}(c) = [R_{kt} - E(R_{kt}|R_{1t} > c, R_{2t} > c)]/[\text{var}(R_{kt}|R_{1t} > c, R_{2t} > c)]^{1/2}$ and $X_{kt}^{-}(c) = [R_{kt} - E(R_{kt}|R_{1t} < -c, R_{2t} < -c)]/[\text{var}(R_{kt}|R_{1t} < -c, R_{2t} < -c)]^{1/2}$ for $k = 1$ and 2.

This assumption prevents degeneracy of our test statistics. A necessary but not sufficient condition is, for any of the chosen level c, $\Pr(R_{1t} > c, R_{2t} > c)$ and $\Pr(R_{1t} < -c, R_{2t} < -c)$ must be bounded away from zero. This should be obviously true empirically as one has to be able to estimate these conditional probabilities based on the data. In this article, we consider the use of a finite number of the exceedance levels only, which yields the simple χ^2 distribution of the tests, but makes the finite sample properties depend on the choice of the levels. One way to overcome this problem is to use all possible numbers. However, the resulting distribution will then be too complex to compute in many applications; we will therefore leave it to studies elsewhere.

1.2 Testing beta and covariance asymmetries

As mentioned earlier, betas are useful for understanding the riskiness of the associated stocks, and hence it is of interest to test their symmetry. To do so, we first define betas conditional on the market's up- and down-moves. Analogous to the conditional correlations, we can define the conditional betas at any exceedance level c as

$$\beta^{+}(c) = \frac{\text{cov}(R_{1t}, R_{2t}|R_{1t} > c, R_{2t} > c)}{\text{var}(R_{2t}|R_{1t} > c, R_{2t} > c)} = \frac{\sigma_1^{+}(c)}{\sigma_2^{+}(c)}\rho^{+}(c), \tag{17}$$

$$\beta^{-}(c) = \frac{\text{cov}(R_{1t}, R_{2t}|R_{1t} < -c, R_{2t} < -c)}{\text{var}(R_{2t}|R_{1t} < -c, R_{2t} < -c)} = \frac{\sigma_1^{-}(c)}{\sigma_2^{-}(c)}\rho^{-}(c), \tag{18}$$

where

$$\sigma_1^{+}(c)^2 = \text{var}(R_{1t}|R_{1t} > c, R_{2t} > c), \tag{19}$$

$$\sigma_2^{+}(c)^2 = \text{var}(R_{2t}|R_{1t} > c, R_{2t} > c), \tag{20}$$

and $\sigma_1^{-}(c)$ and $\sigma_2^{-}(c)$ are defined similarly. In particular, when $c = 0$, $\beta^{+}(c)$ and $\beta^{-}(c)$ are the upside and downside betas defined by Ang and Chen (2002). Clearly, even if $c \neq 0$, they can still be interpreted as the upside and downside betas except that they are now examined at a non-zero exceedance level. If we interpret R_{2t} as the return on the market, then $\sigma_1^{+}(c)/\sigma_2^{+}(c)$ is the ratio of upside asset standard deviation (asset risk) to the upside market standard deviation (market risk), and so the upside beta is the product of this ratio and the conditional correlation between the asset and the market. Because the ratio can be different in upside and downside markets, the betas can be asymmetric even if there are no asymmetries in correlations. Hence, our earlier test for symmetry in correlations cannot be used for testing symmetry in betas.

To obtain a test for symmetry in betas, we, as in the correlation case, evaluate the sample differences of the upside and downside betas,

$$\sqrt{T}(\hat{\beta}^{+} - \hat{\beta}^{-}) = \sqrt{T}\left[\hat{\beta}^{+}(c_1) - \hat{\beta}^{-}(c_1), \ldots, \hat{\beta}^{+}(c_m) - \hat{\beta}^{-}(c_m)\right]', \quad (21)$$

where $c_1, c_2, \ldots, c_m$ are a set of m chosen exceedance levels. Now, the symmetry hypothesis of interest is

$$H_0: \quad \beta^{+}(c) = \beta^{-}(c), \quad \text{for all } c \geq 0. \quad (22)$$

Under the null and some regularity conditions, similar to the earlier correlation case, we can show that $\sqrt{T}(\hat{\beta}^{+} - \hat{\beta}^{-})$ has an asymptotic normal distribution with mean zero and a positive definite variance-covariance matrix Ψ which can be consistently estimated by

$$\hat{\Psi} = \sum_{l=1-T}^{T-1} k(l/p)\hat{g}_l, \quad (23)$$

where $k(\cdot)$ is the kernel function as in Equation (13), p is the bandwidth, and $\hat{g}_l$ is an $m \times m$ matrix with (i, j)-th element

$$\hat{g}_l(c_i, c_j) = \frac{1}{T}\sum_{t=|l|+1}^{T} \hat{\eta}_t(c_i)\hat{\eta}_{t-|l|}(c_j), \quad (24)$$

where

$$\hat{\eta}_t(c) = \frac{T}{T_c^{+}}\left[\frac{\hat{\sigma}_1^{+}(c)}{\hat{\sigma}_2^{+}(c)}\hat{X}_{1t}^{+}(c)\hat{X}_{2t}^{+}(c) - \hat{\beta}^{+}(c)\right]1(R_{1t} > c, R_{2t} > c)$$

$$- \frac{T}{T_c^{-}}\left[\frac{\hat{\sigma}_1^{-}(c)}{\hat{\sigma}_2^{-}(c)}\hat{X}_{1t}^{-}(c)\hat{X}_{2t}^{-}(c) - \hat{\beta}^{-}(c)\right]1(R_{1t} < -c)1(R_{2t} < -c). \quad (25)$$

Then the test for beta symmetry can be constructed as

$$\mathcal{J}_\beta = T(\hat{\beta}^{+} - \hat{\beta}^{-})'\hat{\Psi}^{-1}(\hat{\beta}^{+} - \hat{\beta}^{-}), \quad (26)$$

where the bandwidth p is assumed as a fixed constant. As we did in the correlation case, we denote $\hat{\mathcal{J}}_\beta$ as the same statistic except using a stochastic value of p estimated from the data.

Because of its importance in portfolio selections, consider now how to test symmetry in covariances,

$$H_0: \quad \sigma_{12}^{+}(c) = \sigma_{12}^{-}(c), \quad \text{for all } c \geq 0, \quad (27)$$

where

$$\sigma_{12}^{+}(c) = \text{cov}(R_{1t}, R_{2t} | R_{1t} > c, R_{2t} > c) = \sigma_1^{+}(c)\sigma_2^{+}(c)\rho^{+}(c), \quad (28)$$

$$\sigma_{12}^{-}(c) = \text{cov}(R_{1t}, R_{2t} | R_{1t} < -c, R_{2t} < -c) = \sigma_1^{-}(c)\sigma_2^{-}(c)\rho^{-}(c). \quad (29)$$

As with the beta symmetry test, we can construct a test for covariance symmetry as

$$\mathcal{J}_{\sigma_{12}} = T(\hat{\sigma}_{12}^{+} - \hat{\sigma}_{12}^{-})' \hat{\Phi}^{-1} (\hat{\sigma}_{12}^{+} - \hat{\sigma}_{12}^{-}), \quad (30)$$

where

$$\hat{\sigma}_{12}^{+} - \hat{\sigma}_{12}^{-} = \left[\hat{\sigma}_{12}^{+}(c_1) - \hat{\sigma}_{12}^{-}(c_1), ..., \hat{\sigma}_{12}^{+}(c_m) - \hat{\sigma}_{12}^{-}(c_m)\right]', \quad (31)$$

$$\hat{\Phi} = \sum_{l=1-T}^{T-1} k(l/p)\hat{h}_l, \quad (32)$$

and $k(\cdot)$ is the kernel function as in Equation (13), $\hat{h}_l$ is an $m \times m$ matrix with (i, j)-th element

$$\hat{h}_l(c_i, c_j) = \frac{1}{T} \sum_{t=|l|+1}^{T} \hat{\phi}_t(c_i)\hat{\phi}_{t-|l|}(c_j), \quad (33)$$

where

$$\hat{\phi}_t(c) = \frac{T}{T_c^{+}} \left[\hat{\sigma}_1^{+}(c)\hat{\sigma}_2^{+}(c)\hat{X}_{1t}^{+}(c)\hat{X}_{2t}^{+}(c) - \hat{\sigma}_{12}^{+}(c)\right] 1(R_{1t} > c, R_{2t} > c)$$

$$- \frac{T}{T_c^{-}} \left[\hat{\sigma}_1^{-}(c)\hat{\sigma}_2^{-}(c)\hat{X}_{1t}^{-}(c)\hat{X}_{2t}^{-}(c) - \hat{\sigma}_{12}^{-}(c)\right] 1(R_{1t} < -c)1(R_{2t} < -c). \quad (34)$$

The bandwidth p has a meaning analogous to what it had before, and $\hat{J}_{\sigma_{12}}$ is defined in the same way as $\hat{\mathcal{J}}_{\beta}$.

For the asymptotic distributions of the above two symmetry tests, we have

Theorem 2: *Under the null hypotheses, Equations (22) and (27), and under certain regularity conditions, we have*

$$J_{\beta} \to^{d} \chi_m^2, \quad (35)$$

and

$$\mathcal{J}_{\sigma_{12}} \to^{d} \chi_m^2, \quad (36)$$

respectively, as the sample size goes to infinity. Moreover, both $\hat{J}_{\beta}$ and $\hat{\mathcal{J}}_{\sigma_{12}}$ have the same asymptotic χ_m^2 distributions.

Again, the tests are model free. It is unnecessary to choose a parametric model to fit the data to answer the question whether betas or covariances are symmetric. Once the null is rejected, the data cannot be modeled by using any regular symmetric distributions, and so we can legitimately claim that there are asymmetric betas or covariances.

2. Is There Asymmetry?

In this section, we apply first the proposed tests to examine the asymmetries of stock portfolios grouped by size, book-to-market and momentum, respectively, then explore the use of Clayton copula to capture the asymmetries, and finally provide a simulations study on the size and power of the proposed tests.

2.1 The data

While the tests can be easily carried out for any data set, we focus here on three of them. The first is monthly returns on the ten size portfolios of the Center for Research in Security Prices (CRSP), and the monthly market returns, taken here as the returns on the value-weighted market index based on stocks in NYSE/AMEX/NASDAQ, also available from the CRSP. Following Ang and Chen (2002), all risky returns below are in excess of the risk-free rate, which is approximated by the one-month Treasury bill rate available from French's homepage.[2] The other two data sets are book-to-market and momentum decile portfolios that have been getting increasingly popular. The former is again available from French's web, and the latter from Liu, Warner, and Zhang (2005) who provide an interesting recent study of the economic forces. The sample period is from January 1965 to December 1999 (420 observations) for all three data sets.

2.2 Statistical tests

Panel A of Table 1 provides the results for testing correlation symmetry for the size portfolios. The assets are in the first column. They range from the smallest (size 1) to the largest (size 10). The second column reports the P-values (in percentage points) of the correlation symmetry test, J_ρ, based on the singleton exceedance level $c = 0$. The P-values are less than 5% for the first four portfolios, and are greater than 5% for the rest. The fourth column reports the P-values of the same test but with a set of four exceedance levels, 0, 0.5, 1, and 1.5. The results are consistent with the singleton exceedance level case, and the rejections are often stronger because of the smaller P-values.[3] Overall, we find statistically

[2] We are grateful to Ken French for making it available at www.mba.tuck.dartmouth.edu/pages/faculty/ken.french.

[3] Theoretically, there is no reason for preferring one choice to the other. Simulations later also show that both perform almost equally well although the singleton choice has slightly higher power.

Table 1
Correlation, beta and covariance symmetry tests: size portfolios

	$C = \{0\}$		$C = \{0, 0.5, 1, 1.5\}$						
Portfolio	P-value	$c = 0$	P-value	$c_1 = 0$	$c_2 = 0.5$	$c_3 = 1.0$	$c_4 = 1.5$	Skewness	Kurtosis
				Panel A: Correlation					
Size 1	0.34	−0.523	0.00	−0.523	−0.387	−0.298	−0.426	0.87	7.12
Size 2	0.79	−0.440	0.06	−0.440	−0.349	−0.327	−0.147	0.25	5.89
Size 3	0.95	−0.425	0.02	−0.425	−0.402	−0.223	−0.424	0.01	5.99
Size 4	1.72	−0.409	0.00	−0.409	−0.335	−0.182	−0.389	−0.03	6.60
Size 5	5.18	−0.337	7.17	−0.337	−0.372	−0.445	−0.420	−0.31	6.43
Size 6	8.98	−0.279	30.85	−0.279	−0.365	−0.373	−0.368	−0.35	6.18
Size 7	20.56	−0.209	62.53	−0.209	−0.285	−0.312	−0.258	−0.53	6.42
Size 8	38.20	−0.145	77.09	−0.145	−0.198	−0.333	−0.711	−0.58	6.13
Size 9	61.85	−0.083	92.88	−0.083	−0.158	−0.200	−0.491	−0.59	6.28
Size 10	96.63	−0.006	100.00	−0.006	−0.014	−0.018	−0.051	−0.37	5.22
				Panel B: Beta					
Size 1	13.09	−0.282	1.12	−0.282	0.032	0.446	0.749	0.87	7.12
Size 2	10.02	−0.287	3.65	−0.287	−0.093	0.124	0.919	0.25	5.89
Size 3	4.26	−0.344	0.20	−0.344	−0.264	0.240	0.188	0.01	5.99
Size 4	7.51	−0.315	0.05	−0.315	−0.176	0.332	0.419	−0.03	6.60
Size 5	10.05	−0.292	27.85	−0.292	−0.330	−0.380	0.180	−0.31	6.43
Size 6	12.97	−0.257	48.97	−0.257	−0.345	−0.282	0.061	−0.35	6.18
Size 7	16.12	−0.238	55.40	−0.238	−0.334	−0.288	0.216	−0.53	6.42
Size 8	29.42	−0.178	59.49	−0.178	−0.230	−0.349	−0.725	−0.58	6.13
Size 9	46.96	−0.123	87.40	−0.123	−0.191	−0.262	−0.517	−0.59	6.28
Size 10	66.84	0.063	95.38	0.063	0.065	0.076	0.147	−0.37	5.22
				Panel C: Covariance					
Size 1	6.87	−0.241	6.11	−0.241	−0.192	−0.140	−0.338	0.87	7.12
Size 2	4.24	−0.271	12.81	−0.271	−0.271	−0.320	−0.367	0.25	5.89
Size 3	3.67	−0.285	9.26	−0.285	−0.319	−0.312	−0.566	0.01	5.99
Size 4	4.98	−0.280	4.90	−0.280	−0.283	−0.254	−0.590	−0.03	6.60
Size 5	5.35	−0.281	25.75	−0.281	−0.334	−0.406	−0.696	−0.31	6.43
Size 6	6.19	−0.260	27.43	−0.260	−0.341	−0.369	−0.611	−0.35	6.18
Size 7	6.33	−0.266	25.05	−0.266	−0.324	−0.384	−0.606	−0.53	6.42
Size 8	8.47	−0.247	42.87	−0.247	−0.301	−0.454	−0.836	−0.58	6.13
Size 9	12.44	−0.220	50.06	−0.220	−0.295	−0.405	−0.754	−0.59	6.28
Size 10	23.17	−0.154	69.27	−0.154	−0.211	−0.253	−0.432	−0.37	5.22

Panels A through C of the table report, respectively, the results of the correlation, beta and covariance symmetry tests between the market excess return and the excess return on one of the CRSP ten size portfolios. The data are monthly from January, 1965 to December, 1999 ($T = 420$ observations). Two sets of exceedance levels are used to compute the tests. The first is the singleton of $C = \{0\}$ and the second is $C = \{0, 0.5, 1, 1.5\}$. The P-values of the tests are in percentage points. Columns under $c = 0, 0.5$, etc., are the differences in sample conditional correlations at the corresponding exceedance level c.

significant evidence of asymmetry for the four smallest portfolios, but no such evidence for the rest.

It is interesting to observe, from columns 5 through 8, that the sample differences in the conditional correlation $\hat{\rho}^+ - \hat{\rho}^-$ are all negative at all four exceedance levels. This means that the sample downside correlations are greater than the upside ones. For example, $\hat{\rho}^-(0) - \hat{\rho}^+(0)$ for size 2 is as large as 44%! However, this does not mean that there is necessarily a genuine difference in the population parameters, because there are always differences in the sample estimates simply due to sample variations.

Nevertheless, the correlation symmetry test confirms that the correlation between size 2 and the market is indeed asymmetric. However, despite the seemingly large differences for sizes 5 and beyond, the test does not reject the symmetry hypothesis for them.

There are, in addition, two notable facts. First, the *P*-values tend to get greater as the firm size increases. This means that large firms tend to have symmetric up and down movements with the market. Second, the test statistic appears positively related to skewness. For example, size 1 has the smallest *P*-values and, at the same time, has the largest skewness. An intuitive explanation is that smaller firms usually drop more than others when the market goes down. Hence, their greater positively skewed returns are simply a reward for their higher downside risk.

For beta asymmetry, Panel B of Table 1 provides the results. Column 4 shows that the symmetry hypothesis is rejected for sizes 1 to 4 portfolios under the set of four exceedance levels. However, the beta symmetry test based on the singleton exceedance level rejects fewer. Overall, it is interesting that, as is likely to happen, an asset that is rejected by the correlation symmetry test is also rejected by the beta symmetry test.

Consider now asymmetry in covariances. Panel C of Table 1 reports the results that reject symmetry for sizes 2 to 4 at the usual 5% level, but do so for size 1 only at a significance level of 6.11%. In comparison with sizes 2 to 4, since size 1 has in general the greatest differences in down- and up-correlations, it is of interest to know why it does not show the greatest covariance asymmetry, especially given the fact that, if the up- and down-variances were equal for both size 1 and the market, the correlation and covariance asymmetries must be equivalent to each other. However, the up- and down-variances are different here as can easily be seen in the singleton exceedance case in which, based on Equation (28) and Equation (29), we have

$$\sigma_{12}^{+} = \sigma_{1}^{+}\sigma_{2}^{+}\rho^{+} = 0.835 \times 0.604 \times \rho^{+} = 0.504\rho^{+}, \tag{37}$$

$$\sigma_{12}^{-} = \sigma_{1}^{-}\sigma_{2}^{-}\rho^{-} = 0.625 \times 0.762 \times \rho^{-} = 0.476\rho^{-}. \tag{38}$$

Since $\rho^{+} < \rho^{-}$, the larger upside standard deviation of size 1, σ_{1}^{+}, helps to inflate σ_{12}^{+} substantially to narrow its difference with σ_{12}^{-}, and hence the *P*-value here is larger than that of the correlation symmetry test. Because such inflation is relatively greater for size 1 than for sizes 2 to 4, the *P*-value associated with size 1 is also larger than those associated with the latter. Intuitively, small firms are riskier and more sensitive to the market's down turn as evidenced by the asymmetric correlation. But they may still be fairly valued by investors to have relatively symmetric covariances. Indeed, while the size premium has been decreasing since the 1980s, the asymmetric correlation is persistent over time on the basis of diagnostic results not reported.

In the interest of comparison, we also conduct asymmetry studies on the book-to-market and momentum portfolios. First, Table 2 shows that there is no evidence of asymmetry for the book-to-market portfolios. Economically, a high BE/ME firm has high market leverage relative to book leverage, resulting in the so-called distress effect. While high BE/ME portfolios have negative values of sample measures of asymmetries, they are far smaller in absolute value than those in the size portfolio case, which explains why there are no statistical rejections of the symmetry hypothesis

Table 2
Correlation, beta and covariance symmetry tests: book-to-market portfolios

	$C = \{0\}$		$C = \{0, 0.5, 1, 1.5\}$						
Portfolio	P-value	$c = 0$	P-value	$c_1 = 0$	$c_2 = 0.5$	$c_3 = 1.0$	$c_4 = 1.5$	Skewness	Kurtosis
				Panel A: Correlation					
BE/ME 1	78.16	−0.041	97.86	−0.041	−0.055	−0.111	−0.274	−0.15	4.48
BE/ME 2	78.27	−0.043	99.81	−0.043	−0.085	−0.140	−0.119	−0.42	5.04
BE/ME 3	69.96	−0.062	93.62	−0.062	−0.056	−0.063	−0.129	−0.57	5.72
BE/ME 4	54.25	−0.099	62.44	−0.099	−0.091	−0.193	−0.637	−0.40	5.24
BE/ME 5	47.43	−0.119	64.00	−0.119	−0.191	−0.135	−0.437	−0.45	6.27
BE/ME 6	63.49	−0.082	91.04	−0.082	−0.116	−0.214	−0.641	−0.44	5.96
BE/ME 7	58.94	−0.083	74.42	−0.083	−0.077	−0.080	−0.363	0.08	5.21
BE/ME 8	37.11	−0.149	42.11	−0.149	−0.162	−0.249	−0.647	−0.03	5.52
BE/ME 9	27.41	−0.175	27.79	−0.175	−0.189	−0.062	−0.239	−0.14	5.30
BE/ME 10	23.94	−0.205	9.89	−0.205	−0.165	−0.264	−0.168	0.09	6.93
				Panel B: Beta					
BE/ME 1	54.53	0.087	95.97	0.087	0.095	0.084	−0.012	−0.15	4.48
BE/ME 2	94.16	−0.011	84.98	−0.011	−0.080	−0.202	−0.017	−0.42	5.04
BE/ME 3	63.47	−0.076	99.01	−0.076	−0.125	−0.199	−0.278	−0.57	5.72
BE/ME 4	71.76	−0.059	89.57	−0.059	−0.068	−0.188	−0.599	−0.40	5.24
BE/ME 5	64.39	−0.080	89.67	−0.080	−0.133	−0.124	−0.367	−0.45	6.27
BE/ME 6	76.67	−0.052	96.63	−0.052	−0.070	−0.166	−0.527	−0.44	5.96
BE/ME 7	55.51	0.094	95.90	0.094	0.203	0.372	0.348	0.08	5.21
BE/ME 8	89.93	−0.021	43.17	−0.021	0.097	0.121	−0.228	−0.03	5.52
BE/ME 9	67.50	−0.067	34.98	−0.067	−0.032	0.432	0.672	−0.14	5.30
BE/ME 10	76.97	−0.054	60.74	−0.054	0.048	0.067	1.098	0.09	6.93
				Panel C: Covariance					
BE/ME 1	26.15	−0.130	83.54	−0.130	−0.177	−0.215	−0.338	−0.15	4.48
BE/ME 2	18.04	−0.172	67.27	−0.172	−0.253	−0.348	−0.442	−0.42	5.04
BE/ME 3	13.99	−0.204	59.47	−0.204	−0.280	−0.381	−0.648	−0.57	5.72
BE/ME 4	17.21	−0.184	69.68	−0.184	−0.227	−0.389	−0.725	−0.40	5.24
BE/ME 5	17.75	−0.198	42.00	−0.198	−0.269	−0.330	−0.732	−0.45	6.27
BE/ME 6	21.59	−0.179	76.91	−0.179	−0.234	−0.391	−0.836	−0.44	5.96
BE/ME 7	35.74	−0.112	66.02	−0.112	−0.118	−0.084	−0.221	0.08	5.21
BE/ME 8	21.12	−0.161	49.23	−0.161	−0.173	−0.220	−0.492	−0.03	5.52
BE/ME 9	16.28	−0.180	47.89	−0.180	−0.197	−0.207	−0.404	−0.14	5.30
BE/ME 10	19.36	−0.181	19.20	−0.181	−0.153	−0.185	−0.428	0.09	6.93

Panels A through C of the table report, respectively, the results of the correlation, beta and covariance symmetry tests between the market excess return and the excess return on one of the book-to-market (BE/ME) decile portfolios. The data are monthly from January, 1965 to December, 1999 ($T = 420$ observations). Two sets of exceedance levels are used to compute the tests. The first is the singleton of $C = \{0\}$ and the second is $C = \{0, 0.5, 1, 1.5\}$. The P-values of the tests are in percentage points. Columns under $c = 0, 0.5$, etc., are the differences in sample conditional correlations at the corresponding exceedance level c.

here. Intuitively, the market risk may not be of primary concern when a firm is in distress, and hence its returns are fairly symmetric relative to the market because they are primarily driven by other risks. Second, Table 3 provides the results on the momentum portfolios. Because current losers (winners) are more likely to be future losers (winners), one may expect that the losers (winners) go down (up) more often with the market when the market is down. Surprisingly, however, there is no correlation asymmetry whatsoever. Therefore, it must be the case that when they are down, the losers must be down more in magnitude than when they

Table 3
Correlation, beta and covariance symmetry tests: momentum portfolios

	$C = \{0\}$		$C = \{0, 0.5, 1, 1.5\}$						
Portfolio	P-value	$c = 0$	P-value	$c_1 = 0$	$c_2 = 0.5$	$c_3 = 1.0$	$c_4 = 1.5$	Skewness	Kurtosis
				Panel A: Correlation					
L	23.73	−0.169	58.75	−0.169	−0.222	−0.312	−0.567	0.21	5.52
2	29.44	−0.150	60.97	−0.150	−0.186	−0.240	−0.467	−0.00	5.91
3	31.83	−0.153	82.97	−0.153	−0.236	−0.269	−0.540	−0.15	6.19
4	34.81	−0.152	86.48	−0.152	−0.242	−0.266	−0.539	−0.29	6.47
5	36.20	−0.150	86.21	−0.150	−0.246	−0.286	−0.557	−0.45	6.91
6	37.83	−0.149	82.87	−0.149	−0.239	−0.302	−0.631	−0.66	6.89
7	37.64	−0.148	80.58	−0.148	−0.237	−0.306	−0.638	−0.81	6.89
8	35.25	−0.154	67.02	−0.154	−0.214	−0.320	−0.692	−0.90	6.98
9	21.01	−0.196	30.85	−0.196	−0.219	−0.280	−0.681	−0.89	6.31
W	8.35	−0.258	14.85	−0.258	−0.273	−0.282	−1.084	−0.73	5.21
				Panel B: Beta					
L	77.21	0.044	98.98	0.044	0.043	0.027	−0.237	0.21	5.52
2	99.71	−0.001	98.61	−0.001	0.021	0.042	−0.150	−0.00	5.91
3	79.03	−0.044	99.21	−0.044	−0.084	−0.071	−0.256	−0.15	6.19
4	59.66	−0.092	95.55	−0.092	−0.132	−0.141	−0.329	−0.29	6.47
5	40.16	−0.148	79.09	−0.148	−0.210	−0.204	−0.446	−0.45	6.91
6	23.17	−0.215	45.13	−0.215	−0.291	−0.350	−0.671	−0.66	6.89
7	13.47	−0.267	26.87	−0.267	−0.363	−0.453	−0.756	−0.81	6.89
8	7.68	−0.312	8.99	−0.312	−0.406	−0.519	−0.861	−0.90	6.98
9	3.03	−0.356	1.12	−0.356	−0.410	−0.491	−0.841	−0.89	6.31
W	1.99	−0.357	3.15	−0.357	−0.418	−0.427	−1.080	−0.73	5.21
				Panel C: Covariance					
L	26.50	−0.129	79.84	−0.129	−0.160	−0.224	−0.474	0.21	5.52
2	17.58	−0.164	60.54	−0.164	−0.191	−0.229	−0.440	−0.00	5.91
3	14.65	−0.189	61.94	−0.189	−0.242	−0.275	−0.522	−0.15	6.19
4	14.49	−0.204	61.09	−0.204	−0.279	−0.312	−0.593	−0.29	6.47
5	11.81	−0.227	53.76	−0.227	−0.306	−0.382	−0.667	−0.45	6.91
6	8.64	−0.257	44.67	−0.257	−0.343	−0.452	−0.742	−0.66	6.89
7	6.60	−0.278	37.94	−0.278	−0.381	−0.517	−0.799	−0.81	6.89
8	5.05	−0.297	31.18	−0.297	−0.385	−0.535	−0.827	−0.90	6.98
9	3.10	−0.308	16.80	−0.308	−0.371	−0.521	−0.794	−0.89	6.31
W	2.33	−0.296	14.04	−0.296	−0.342	−0.459	−0.857	−0.73	5.21

Panels A through C of the table report, respectively, the results of the correlation, beta and covariance symmetry tests between the market excess return and the excess return on one of the momentum decile portfolios. The data are monthly from January, 1965 to December, 1999 ($T = 420$ observations). Two sets of exceedance levels are used to compute the tests. The first is the singleton of $C = \{0\}$ and the second is $C = \{0, 0.5, 1, 1.5\}$. The P-values of the tests are in percentage points. Columns under $c = 0, 0.5$, etc., are the differences in sample conditional correlations at the corresponding exceedance level c.

are up. The opposite must also be true for the winners. Interestingly though, there are apparent beta and covariance asymmetries, especially among the top two winner portfolios. Econometrically, these are driven by the up- and down-variances of the associated assets like the size portfolio case but in the opposite direction. For example, in the case of top winners under the singleton exceedance, σ_1^+/σ_2^+ and σ_1^-/σ_2^-, are 0.912 and 1.055, exacerbating the correlation difference substantially to yield an asymmetric beta statistic. Economically, Schwert (2003) and others find that the momentum premium remains an asset pricing anomaly that cannot be explained by the market. An asymmetric covariance is clearly consistent with these studies because a symmetric covariance risk might be explained by the market factor model while an asymmetric covariance risk certainly cannot.

2.3 Diagnostics and modeling

In the previous subsection, we have found evidence of asymmetries, with the strongest one shown between the market and size 1. In this subsection, we explore the intuition of this strongest asymmetric correlation, and illustrate why it cannot be modeled by a normal distribution, but can be captured by a mixture Clayton copula to be introduced below.

The top-left graph of Figure 1 is an 'empirical' contour plot of the standardized monthly returns on the market and size 1 portfolios. Visually, it is apparent, along the 45 degree line, that there are more observations near the lower portion than near the upper one. This clearly suggests asymmetry. In contrast, a theoretical contour plot based on the normal distribution (with the same sample unconditional correlation) shows a much more even distribution. This is expected. Because the normal distribution is symmetric, the down- and up-side comovements between the two assets must be the same. What we learn here is that a simple plot of the data reveals the shortcomings of the normal distribution in modeling asymmetric comovements.

While the plots are informative, they are not precise. To formally quantify the deviations from symmetry, we, following Hu (2004), use a contingency table approach. We divide the range of returns into $K = 6$ cells. That balances the tradeoff of having enough observations in each cell versus having enough cells for testing contingent dependence. Let $A_{i,j}$s be the numbers of observed frequencies in cell (i, j)s, which are reported in the upper-left panel of Table 4. The asymmetry shows up quantitatively in terms of these frequencies. For instance, the lower-left corner, cell (6, 1), has a value of 40, while the upper-right corner, cell (1, 6), has a value of 21, telling us that out of 420 observations there are almost twice as many occurrences of both asset returns in their respective lowest percentile as those in their top percentile.

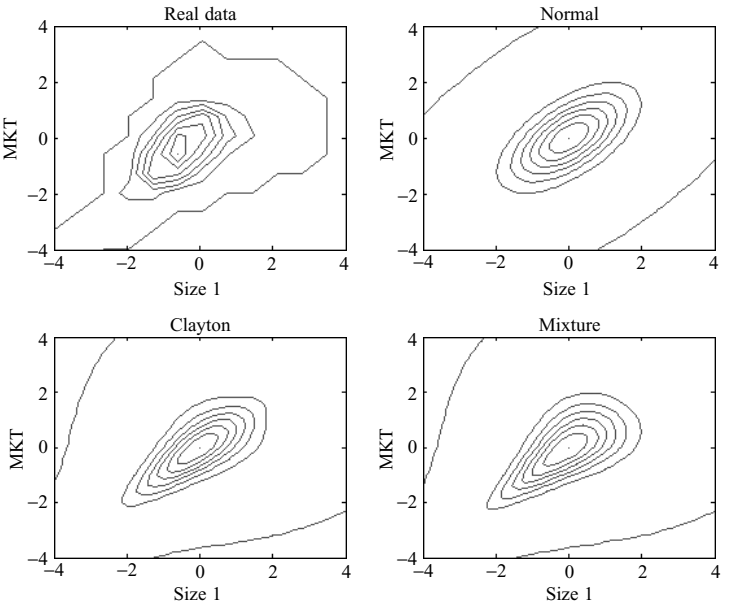

Figure 1
The figure reports four contour graphs. The first, on the top-left side, is the "empirical" contour graph of the observed standardized excess returns on the market and the smallest size portfolio, and those on the top-right, bottom-left and bottom-right are the theoretical contour graphs based on a fitted normal, Clayton, and mixture copula models, respectively.

Let $B_{i,j}$s be the numbers of predicted frequencies in cell (i, j)s based on data from the normal distribution, computed as the product of the exact probabilities of the data's falling into the cells and the sample size.[4] The upper-right panel of Table 4 provides the results. Both the lower-left and upper-right cells have identical values because of the symmetry of the normal distribution. In comparison, there are large differences across the cells in the actual data as shown by $A_{i,j}$s. Are the differences due to chance? The Pearson chi-squared test of the joint equality of the $A_{i,j}$s and $B_{i,j}$s has a P-value of 4.13%. This suggests statistically that the normal distribution is not capable of describing the asymmetry of the data, a result consistent with the symmetry tests.

What distribution might capture the asymmetry? One of the simplest asymmetric distributions is the Clayton copula. Conceptually, a copula is a multivariate distribution that combines two (or more) almost arbitrarily given marginal distributions into a single joint distribution, to which Nelsen (1999) provides an excellent introduction. Longin and Solnik (2001) seems

[4] Under normality, these probabilities can be evaluated by using the formulas provided by Ang and Chen (2002).

Table 4
Goodness of fit

Panel A: Observed frequencies						Panel B: Normal copula					
2	1	9	18	19	21	0.82	2.92	5.91	10.25	17.14	32.98
6	7	9	14	17	17	2.92	7.21	10.97	14.43	17.33	17.14
4	14	17	8	11	16	5.91	10.97	13.61	14.83	14.43	10.25
4	9	16	14	16	11	10.25	14.43	14.83	13.61	10.97	5.91
14	22	15	11	5	3	17.14	17.33	14.43	10.97	7.21	2.92
40	17	4	5	2	2	32.98	17.14	10.25	5.91	2.92	0.82
								P-value = 4.13 (%)			
Panel C: Clayton copula						Panel D: Mixture copula					
1.28	4.81	9.02	14.26	18.17	22.46	1.06	4.09	7.16	11.06	17.80	28.83
1.77	6.74	11.57	14.70	16.95	18.27	1.69	5.75	10.41	15.67	19.07	17.41
3.46	9.82	13.40	14.47	14.90	13.95	2.73	9.34	14.18	16.66	15.59	11.50
6.47	14.32	15.16	13.62	11.27	9.16	6.18	15.29	16.79	14.38	10.20	7.16
14.34	20.19	14.31	9.52	6.72	4.91	15.00	20.90	15.27	9.21	5.73	3.90
42.68	14.13	6.53	3.44	1.99	1.24	43.34	14.64	6.19	3.02	1.61	1.20
		P-value = 96.16 (%)						P-value = 30.25 (%)			

Dividing the range of returns into six cells, the table reports the frequencies of the real data (excess returns on the market and size 1 portfolios) and those frequencies implied by the three fitted models: the normal distribution, the Clayton copula and the mixture Clayton copula. The P-values under the frequencies are from the Pearson chi-squared test of the null of no differences between the given model implied frequencies and the observed ones.

to have made the first major application of the copula approach in finance. Patton (2004) shows that Clayton copulas, among others, do capture the asymmetries of many financial time series.

The copula idea is appealing in empirical studies. A multivariate distribution is usually specified to fit the data, but it often fails to capture some salient features of the univariate time series, i.e., the marginal distributions may not provide good descriptions for the individual data series. The copula solves exactly this problem. One can model the univariate series first, and then use a copula to assemble the univariate distributions into a coherent multivariate one. For example, let Φ be the standard univariate normal distribution function and Φ^{-1} be its inverse. If two data series are well modeled individually by univariate normal distributions, we can assemble them into a multivariate distribution with correlation $\rho \in [-1, 1]$ by using a copula,

$$C_{nor}(u, v; \rho) = \mathbf{\Phi}_\rho(\Phi^{-1}(u), \Phi^{-1}(v)), \tag{39}$$

where $\mathbf{\Phi}_\rho$ is the standard bivariate distribution function with correlation ρ. Since $C_{nor}(u, v; \rho)$ produces a bivariate normal distribution with normal marginals, it is referred to as "the normal copula". A bivariate Clayton copula is defined as

$$C_{clay}(u, v; \tau) = \left(u^{-\tau} + v^{-\tau} - 1\right)^{-\frac{1}{\tau}}, \tag{40}$$

where $\tau > 0$ is the parameter. For any given inverse marginal distributions of u and v, such as $\Phi^{-1}(u)$ and $\Phi^{-1}(v)$, the Clayton copula can be used to generate a bivariate distribution.

Since the normal distribution, though it does not capture the asymmetry of the data, is widely used in both theoretical and empirical studies, it might be extreme to rule it out completely. So, in what follows, we use a mixture model that mixes the normal with a Clayton copula. In the bivariate case, the density function is

$$f_{mix}(u, v; \rho, \tau, \kappa) = \kappa f_{nor}(u, v; \rho) + (1 - \kappa) f_{clay}(u, v; \tau), \qquad (41)$$

where κ is the mixture parameter, f_{nor} is the density of the normal copula, and f_{clay} is the density of the Clayton copula. The latter two densities are easily obtained as the partial derivatives of $C_{nor}(u, v; \rho)$ and $C_{clay}(u, v; \tau)$ with respect to u and v (see Appendix B). It is clear that the mixture model nests the normal distribution as a special case. Since the GARCH(1,1) process is a well-known parsimonious model for stock returns, we will in the rest of the article use it exclusively to model the univariate distributions of the asset returns. Then, their joint distribution is determined by Equation (41), the mixture Clayton model.

The maximum likelihood (ML) method is the standard approach for estimating the model. Rather than maximizing the ML function directly, we use the EM algorithm[5] of Redner and Walker (1984) to ensure fast convergence of the numerical solution to the optimum of the objective function. Panel A of Table 5 reports the estimation results for the bivariate series of the market and size 1 portfolios. The first case is to use a bivariate normal distribution to fit the data. The estimated correlation is 0.609 with a standard error of 0.037. However, the log likelihood value is 95.538, implying that the associated likelihood ratio test (LRT) of this model versus the mixture one has a P-value of virtually 0%. So the normal distribution is rejected soundly by the data. Similarly, the pure Clayton copula model is rejected too because the LRT has an almost zero P-value. Both rejections are also confirmed by the estimation result on κ. The ML estimate of κ is 0.275 with a standard error of 0.089, which is significantly different from either one or zero. The τ parameter, interestingly, is not much different in the pure Clayton and mixture copula models. It is also interesting to observe that the correlation in the mixture model is greater than that in the normality case, suggesting at least in this example that removing asymmetric data increases the correlation of the rest of the sample.

In the interest of comparison, we also estimate the model using the Bayesian approach. Details of this approach are provided in Appendix B.

[5] See McLachlan and Krishnan (1996) for an introduction and extensive applications of the EM algorithm.

Table 5
Model estimation and comparison

Panel A: Classical framework

	Normal		Clayton		Mixture	
	est.	std	est.	std	est.	std
ρ	0.609	0.037			0.855	0.022
τ			1.344	0.175	1.266	0.171
κ					0.275	0.089
likelihood (log)	95.538		111.666		183.947	
LRT	0.000		0.000			

Panel B: Bayesian framework

	Normal		Clayton		Mixture	
	est.	std	est.	std	est.	std
ρ	0.606	0.031			0.921	0.097
τ			1.349	0.118	1.351	0.203
κ					0.243	0.113

The table reports both the maximum likelihood (ML) and Bayesian estimates of the parameters in three fitted models for the excess returns on the market and size 1 portfolios: the normal distribution, the Clayton copula and the mixture Clayton copula. The first two models are nested in the third whose density is

$$f_{mix}(u, v; \rho, \tau, \kappa) = \kappa f_{nor}(u, v; \rho) + (1 - \kappa) f_{clay}(u, v; \tau).$$

The table also reports the log of likelihood values at the ML estimates and the P-values of the LRT of the first two models against the last one, respectively.

Panel B of Table 5 reports the results. Both the point estimates of the parameters and their standard errors are very similar between the ML and Bayesian approaches. Although the Bayesian approach only confirms the ML estimates here, it is a more flexible method. When we model all assets simultaneously in high dimensions, the ML is not feasible owing to difficulties in numerical maximization. But the Bayesian approach can still be used to obtain both parameter estimates and the predictive density of the data.

To see how well both the pure Clayton and the mixture models explain the asymmetry, the lower part of Figure 1 plots the theoretical contour graphs based on the two models, respectively, in the same way as we did for the normal. Now the graphs resemble more closely the asymmetric pattern of the real data than the normality case. To assess further quantitatively, the two panels of the lower part of Table 4 provide the frequencies that are well approximated by using 100,000 simulated data sets from the two models, respectively. Clearly, the pure Clayton model is more asymmetric and has a P-value of 96.16% in matching the observed frequencies. Interestingly though, the mixture model can also explain well the observed frequencies with a P-value of 30.25%. In addition, it is a much better model than the pure Clayton in explaining the overall features of the data, as suggested by the earlier LRT result.

2.4 Size and power

As the mixture model seems to explain the data well, it serves as a good distribution to draw data from, in order to assess the power of the proposed symmetry tests. For size of the tests, the data can be drawn straightforwardly from the normal distribution which is the standard benchmark.[6] While any of the ten size portfolios can be used in conjunction with the market to calibrate the parameters, we would like to choose a more sensible one, though the results are largely similar. In the earlier modeling case, we used size 1 for illustration because it is the most difficult to model. Now, size 5 is the first one whose symmetry the tests fail to reject, and hence it is of interest to use it to calibrate the parameters to see whether the tests have reasonable power in simulations. Hence, all parameters below (in this subsection) are calibrated by using the market and size 5 returns unless otherwise specified explicitly.

The nominal size of the tests is set at 5% on the basis of their asymptotic distributions. Columns 2 to 5 of Table 6 report the empirical size of the Ang and Chen test (based on the asymptotic distribution in Appendix A) and the proposed correlation, beta, and covariance symmetry tests. With varying sample sizes and exceedance levels, we see that the rejection rates do not change much. While some of them are close to 1%, by and large, the results are not much different from 5%, and all the tests seem fairly reliable under the null.

To assess power, κ is allowed to be less than 100%. The lower it is, the more it deviates from normality to have more asymmetry. When the data is not much different from normality with $\kappa = 75\%$, the rejection rates are very low when $T = 240$, and are mostly under 50% even when $T = 840$. Unlike the real data case, more exceedance levels do not necessarily yield more rejections, which seems to be due to relatively fewer samples available in the higher exceedance levels. However, the power increases dramatically when κ decreases to 50%, corresponding roughly to the calibrated posterior mean of 51.4% (in the mixture model for the pair of size 5 and the market).[7] For example, when $T = 420$, the four tests have rejection rates of 86%, 90%, 86%, and 49%, respectively, at the singleton exceedance level. Although not reported, the power is much greater when $\kappa = 25\%$. For example, when $\kappa = 25\%$ and $T = 420$, the rejection rates are 95%, 99%, 99%, and 57% in the singleton case. Notice that the Ang and Chen test imposes normality and it should in general have greater power than the other tests. What we find interesting here is that, within the class of the mixture Clayton copula distributions, the powers of all the tests are

[6] The bootstrap method is difficult to apply here for power studies, though it can be used to examine the size to obtain similar results (not reported here).

[7] The posterior mean of κ for the other nine pairs are 24.3%, 27.5%, 40.2%, 38.8%, 52.7%, 68.2%, 77.4%, 84.0% and 75.2%, respectively.

Table 6
Size and power

	$\kappa = 100\%$				$\kappa = 75\%$				$\kappa = 50\%$			
C	AC	Corr	Beta	Cov	AC	Corr	Beta	Cov	AC	Corr	Beta	Cov
T = 240												
{0}	2.80	2.46	6.41	4.93	16.00	22.03	26.70	9.37	49.30	68.03	66.24	27.73
{0, 0.5}	2.80	2.69	6.17	3.29	11.30	18.55	23.62	5.96	31.30	60.71	60.33	18.74
{0, 0.5, 1}	4.10	2.07	4.70	2.08	9.80	14.75	19.39	3.79	24.50	53.78	54.53	12.95
{0, 0.5, 1, 1.5}	7.30	1.47	3.41	1.54	12.80	11.23	15.48	2.33	21.40	46.39	48.19	9.24
T = 420												
{0}	2.50	2.02	5.09	4.59	30.60	37.12	39.27	15.12	86.00	89.94	86.76	48.61
{0, 0.5}	3.40	2.03	5.25	3.31	24.50	31.12	34.75	9.96	74.60	85.96	83.21	37.36
{0, 0.5, 1}	3.40	1.40	4.18	2.15	20.40	26.36	30.60	6.99	60.10	82.28	79.61	30.02
{0, 0.5, 1, 1.5}	4.20	1.00	2.74	1.46	20.30	20.47	24.92	4.64	53.10	76.70	74.84	23.52
T = 600												
{0}	3.20	1.88	4.69	4.21	43.90	47.69	48.74	19.36	95.90	97.22	95.05	63.21
{0, 0.5}	3.70	1.85	5.00	2.94	32.00	40.89	44.02	12.77	90.30	95.54	92.99	51.53
{0, 0.5, 1}	3.50	1.36	3.85	1.99	23.20	35.61	39.67	9.63	73.20	94.03	91.56	45.23
{0, 0.5, 1, 1.5}	4.20	0.88	2.69	1.27	20.80	28.97	33.34	6.76	62.80	91.44	88.71	36.69
T = 840												
{0}	2.50	1.33	4.92	4.29	54.50	61.90	59.73	25.37	99.20	99.50	98.74	77.52
{0, 0.5}	2.50	1.46	4.98	2.74	41.00	54.77	55.08	17.58	96.50	99.14	98.04	67.27
{0, 0.5, 1}	2.70	0.83	3.76	2.02	28.20	49.14	50.99	13.79	85.50	98.71	97.39	61.80
{0, 0.5, 1, 1.5}	3.60	0.57	2.83	1.31	22.20	42.14	44.40	9.94	74.70	97.86	96.20	53.90

The table reports the size and power of Ang and Chen's test, the correlation, beta and covariance symmetry tests, denoted by AC, Corr, Beta and Cov, respectively. The nominal size of the tests is set at 5% based on their asymptotic distributions. The results are based on 10,000 simulations drawn from the mixture copula model

$$f_{mix}(u, v; \rho, \tau, \kappa) = \kappa f_{nor}(u, v; \rho) + (1 - \kappa) f_{clay}(u, v; \tau),$$

where model parameters except κ are calibrated using excess returns on the market and size 5 portfolios. Under the null of no asymmetry, $\kappa = 100\%$ and the data-generating process is the normal distribution. $\kappa = 75\%$ and 25% represent two different degrees of asymmetries.

comparable. In summary, all the tests have good power against the null when $\kappa = 50\%$ or lower, but not so when $\kappa = 75\%$ or greater.

Since the tests are based on the conditional correlations, etc., it is of interest to know their population counterparts under the alternative. Put differently, we want to ask what degrees of asymmetry in correlation, beta, and covariance a given κ can generate. Because analytical formulas are unavailable, we estimate them by drawing data from the calibrated model with the varying specification of κ. As the sample size or the number of draws increases, the estimates should converge to the true parameters. Table 7 provides the results. In comparison with the real data case, there are two nice patterns. First, all the estimated population asymmetry measures are negative. Second, except for some lesser degree in covariance, their magnitudes resemble well the real data estimates when $\kappa = 50\%$ and 25% (corresponding closely to the κ's for sizes 5 and 1), respectively. Both results suggest that κ is indeed the key parameter of the copula model that drives asymmetries in correlation, beta, and covariance.

Table 7
Implied asymmetry measures of the calibrated model

C	$\kappa = 75\%$			$\kappa = 50\%$			$\kappa = 25\%$		
Correlation	ρ^+	ρ^-	$\rho^+ - \rho^-$	ρ^+	ρ^-	$\rho^+ - \rho^-$	ρ^+	ρ^-	$\rho^+ - \rho^-$
T = 240									
{0}	0.542	0.678	−0.136	0.455	0.729	−0.274	0.369	0.780	−0.411
{0.5}	0.460	0.617	−0.157	0.366	0.684	−0.318	0.268	0.748	−0.480
{1}	0.388	0.541	−0.153	0.309	0.619	−0.311	0.219	0.690	−0.471
{1.5}	0.296	0.436	−0.140	0.234	0.521	−0.286	0.153	0.588	−0.435
T = 840									
{0}	0.536	0.681	−0.145	0.441	0.732	−0.291	0.347	0.783	−0.437
{0.5}	0.455	0.621	−0.167	0.348	0.688	−0.340	0.234	0.753	−0.518
{1}	0.398	0.550	−0.153	0.305	0.628	−0.323	0.193	0.699	−0.507
{1.5}	0.344	0.466	−0.122	0.279	0.546	−0.266	0.190	0.614	−0.424
Beta	β^+	β^-	$\beta^+ - \beta^-$	β^+	β^-	$\beta^+ - \beta^-$	β^+	β^-	$\beta^+ - \beta^-$
T = 240									
{0}	0.549	0.687	−0.138	0.462	0.738	−0.277	0.375	0.790	−0.415
{0.5}	0.474	0.633	−0.159	0.378	0.702	−0.324	0.277	0.768	−0.491
{1}	0.415	0.572	−0.157	0.331	0.654	−0.323	0.237	0.728	−0.491
{1.5}	0.348	0.495	−0.147	0.274	0.583	−0.309	0.188	0.654	−0.466
T = 840									
{0}	0.541	0.688	−0.146	0.446	0.740	−0.294	0.350	0.792	−0.441
{0.5}	0.465	0.635	−0.170	0.356	0.704	−0.348	0.240	0.770	−0.530
{1}	0.417	0.575	−0.158	0.321	0.657	−0.336	0.203	0.731	−0.527
{1.5}	0.379	0.507	−0.129	0.309	0.594	−0.285	0.214	0.667	−0.454
Covariance	σ_{12}^+	σ_{12}^-	$\sigma_{12}^+ - \sigma_{12}^-$	σ_{12}^+	σ_{12}^-	$\sigma_{12}^+ - \sigma_{12}^-$	σ_{12}^+	σ_{12}^-	$\sigma_{12}^+ - \sigma_{12}^-$
T = 240									
{0}	0.208	0.257	−0.049	0.175	0.273	−0.098	0.142	0.289	−0.147
{0.5}	0.142	0.186	−0.044	0.113	0.202	−0.089	0.082	0.216	−0.134
{1}	0.102	0.136	−0.034	0.080	0.150	−0.069	0.056	0.160	−0.105
{1.5}	0.075	0.099	−0.024	0.059	0.111	−0.051	0.040	0.117	−0.077
T = 840									
{0}	0.205	0.259	−0.054	0.168	0.277	−0.109	0.132	0.294	−0.162
{0.5}	0.140	0.189	−0.049	0.106	0.206	−0.099	0.071	0.222	−0.151
{1}	0.103	0.139	−0.036	0.078	0.154	−0.076	0.049	0.167	−0.118
{1.5}	0.081	0.104	−0.023	0.065	0.117	−0.051	0.045	0.126	−0.081

The table reports the estimated values of the implied parameters, ρ^+, ρ^-, and $\rho^+ - \rho^-$, β^+, β^-, and $\beta^+ - \beta^-$ and σ_{12}^+, σ_{12}^-, and $\sigma_{12}^+ - \sigma_{12}^-$ at different exceedance levels for $T = 240$ and 840, respectively, which are based on 10,000 data sets drawn from the calibrated mixture copula model of Table 6.

3. Asset Allocation Perspective

Statistical tests in Section 2 show evidence of asymmetric correlations, betas, and covariances. The question we ask in this section is how important these asymmetries are from an investor's portfolio decision point of view.

Consider an investment universe consisting of cash plus n risky assets. Let R_t denote an n-vector with i-th element $r_{i,t}$, the return on the i-th risky asset at time t in excess of the return, $R_{f,t}$, on a riskless asset. Then the excess return of a portfolio with weight w_i on the i-th risky asset is $R_{p,t} = \sum_{i=1}^{n} w_i r_{i,t}$. Under the standard expected utility framework, the investor chooses portfolio weights $w = (w_1, \ldots, w_n)'$ to maximize the

expected utility,

$$\max_{w} E\left[U(W)\right], \tag{42}$$

where W is the next period wealth

$$W_{t+1} = 1 + R_{f,t+1} + \sum_{i=1}^{n} w_i r_{i,t+1}, \tag{43}$$

with the initial wealth set to be equal to one. The popular choices for $U(W)$ are the quadratic and CRRA utility functions, but the former utility does not capture the impact of higher moments and the latter one is still a locally mean-variance preference.

Following Ang, Bekaert and Liu (2005) who build their insights on Gul (1991), we use the DA preference in our assessment of the economic importance of asymmetries. The utility μ_W is implicitly defined by the following equation,

$$U(\mu_W) = \frac{1}{K}\left(\int_{-\infty}^{\mu_W} U(W)dF(W) + A\int_{\mu_W}^{\infty} U(W)dF(W)\right), \tag{44}$$

where $U(\cdot)$ is the felicity function chosen here as the power utility form, i.e., $U(W) = W^{(1-\gamma)}/(1-\gamma)$; A is the coefficient of DA; $F(\cdot)$ is the cumulative distribution of wealth; μ_W is the certainty-equivalent wealth (the certain level of wealth that generates the same utility as the portfolio allocation determining W) and K is a non-random scalar given by

$$K = Pr(W \leq \mu_W) + APr(W > \mu_W). \tag{45}$$

It is seen that μ_W also serves as the reference point in both determining K and the bracketed term in Equation (44).[8] This reference point is irrelevant only when $A = 1$ and in this case the DA preference reduces to the power utility. Since A is usually set to be less than 1, the outcomes below the reference point are weighted more heavily than those above it. For example, if A is equal to 0.5, the outcomes below μ_W are weighted as twice as important as the others. Intuitively, asymmetries make downside moves of a portfolio more likely, and so they should be more important to DA investors. Hence, the DA preference is of particular usefulness in analyzing asymmetries.[9]

[8] The subscript W in μ_W is purely a notation referring to the certainty-equivalent wealth and should not be confused with the W elsewhere in which it is the terminal wealth and is a random variable.

[9] Although beyond the scope of this article, the kinked utility function, used by Benartzi and Thaler (1995) and Aït-Sahalia and Brandt (2001), seems another important class of preference for analyzing asymmetries.

The DA preference is usually implemented on the basis of a point estimator of the model parameters for asset return dynamics since the true parameters are unknown in practice. This plug-in approach ignores the estimation risk as the parameter estimates are subject to random sampling errors. Another desirable feature of our approach is that we use a Bayesian decision framework to compute the utility that accounts for the estimation risk.[10] Let R denote the data available at time T. In the Bayesian framework, all information, sample variation and parameter uncertainty about future stock returns is summarized by $p(R_{T+1}|R)$, the predictive density of the returns conditional on the available data. When the data are normally distributed, the predictive density is analytically available from Zellner (1971) and more generally from Pástor and Stambaugh (2000). However, when the data are nonnormal, such as t-distributed, it can be determined only numerically, as shown by Tu and Zhou (2004). In the present case of a mixture copula model with asymmetries, the predictive density is more complex. We relegate the details of its computation to Appendix B.

Under the DA preference, the Bayesian investor's optimization problem is

$$\max_{w} U(\mu_W), \tag{46}$$

where the certainty-equivalent wealth is defined by Equation (44) and W is defined by Equation (43). The first-order condition is

$$\int_{W_{T+1}\le\mu_W} (W_{T+1}^{-\gamma} r_{i,T+1}) p(R_{T+1}|R) dR_{T+1} + A\int_{W_{T+1}>\mu_W} (W_{T+1}^{-\gamma} r_{i,T+1}) p(R_{T+1}|R) dR_{T+1} = 0, \tag{47}$$

for $i = 1, 2, \ldots, n$, where $W_{T+1} = 1 + R_{f,T+1} + \sum_{i=1}^{n} w_i r_{i,T+1}$ is the predictive wealth at $T+1$ when time T wealth W_T is set to \$1. In contrast with the classic framework of Ang, Bekaert and Liu (2005), the equation is identical except that we use the predictive distribution of the wealth, whereas they use an assumed true distribution.[11] Hence, other than the technical difficulty of determining $p(R_{T+1}|R)$, the optimization problem can be solved by using their approach with simple modifications to accommodate multiple assets.[12]

[10] See Kan and Zhou (2006) and references therein for recent studies on estimation risk.

[11] The Bayesian posterior mean estimates of the parameters may be used in the same way as the point estimates in the classic framework to evaluate the utility gains. The resulting allocation is, however, riskier than using Bayesian predictive density because risky assets are riskier now with estimation risk.

[12] An appendix on the details for this and Equation (48) is available upon request.

We are now ready to assess the economic value in portfolio choices when one switches from a belief in symmetric returns to a belief in asymmetric ones. Under the belief in symmetric returns, we assume that the investor obtains his optimal portfolio w^{Nor} under the benchmark normal data-generating process by solving the earlier optimization problem. Under the belief in asymmetric returns, we assume that the investor obtains w^{Asy} on the basis of the true data-generating process, the n-dimensional mixture Clayton copula, in solving the same problem. Let μ_W^{Nor} and μ_W^{Asy} be the associated certainty-equivalent wealth levels, respectively. Because they also serve as different reference points in the utilities, their difference does not readily capture the utility gain of switching from the symmetric belief to the asymmetric one. To truly capture the gain, we fix μ_W^{Nor} as the reference point and ask how much certainty-equivalent return, CE, the investor is willing to give up to maintain the same benchmark level of μ_W^{Nor} when he switches from w^{Nor} to w^{Asy}. Formally, taking the mixture Clayton copula as the true data-generating process, we solve CE in the following problem,

$$U(\mu_W^{Nor}) = \frac{1}{K}\left(\int_{W_{T+1}^* < \mu_W^{Nor}} \frac{(W_{T+1}^*)^{(1-\gamma)}}{1-\gamma} p(R_{T+1}^{Asy}|R) dR_{T+1}^{Asy} \right.$$

$$\left. + A \int_{W_{T+1}^* > \mu_W^{Nor}} \frac{(W_{T+1}^*)^{(1-\gamma)}}{1-\gamma} p(R_{T+1}^{Asy}|R) dR_{T+1}^{Asy}\right), \qquad (48)$$

where

$$W_{T+1}^* = 1 + R_{f,T+1} + \sum_{i=1}^{n} w_i^{Asy} r_{i,T+1}^{Asy} - CE$$

is the terminal wealth at $T+1$ generated by the optimal portfolio w^{Asy} after deducting an amount of CE, and $p(R_{T+1}^{Asy}|R)$ is the predictive density of the returns under the mixture copula model. CE can be interpreted as the "perceived" certainty-equivalent gain to the investor who switches his belief from symmetric returns into asymmetric ones. The idea of the CE approach can be traced back at least to Kandel and Stambaugh (1996), yielding many applications such as Fleming, Kirby, and Ostdiek (2001). The issue is how big this value can be. Imagine that there exists an investor who does not know how to incorporate asymmetry into his investment decision. If the gain is over 2% annually, he would be willing to pay a fund manager a 1% fee (reasonably high in the fund industry) to manage the money for him to yield a 1% extra gain. So, not surprisingly, values over a couple of percentage points per year are usually deemed economically significant.

In our applications, there are jointly 10 monthly excess returns on the size portfolios as well as monthly excess returns on the market, and so the dimensionality is $n = 11$. Table 8 provides the utility gains. When

Table 8
Utility gains

A	γ			
	2	4	6	8
0.55	1.49	2.05	2.59	3.13
0.45	3.28	3.82	5.03	5.63
0.35	5.49	6.11	6.66	7.21
0.25	8.44	9.16	9.87	10.67

The table reports the utility gains (measured as certainty-equivalent returns in percentage points and annualized) of switching from a belief in symmetric stock returns to a belief in asymmetric ones, where the beliefs are modeled by using the normal and mixture copula distributions, respectively, for making investment decisions. The investment opportunity set consists of the ten CRSP size portfolios, the market portfolio and a risk-free asset. The investor is assumed to have a DA preference of Ang, Bekaert and Liu (2005) with power felicity function. A is the coefficient of disappointment aversion and γ is the curvature parameter.

the DA $A = 0.55$, the annual gains increase from 1.49% to 3.13% as the curvature parameter γ varies from 2 to 8. The monotone relation seems to be due to the fact that as the investor becomes more risk-averse, the loss aversion becomes more important. As a result, the gains from symmetry to asymmetry are greater. When A goes down, the disappointment is valued more by the investor, and hence the gains increase. For example, when $A = 0.25$, the asymmetry matters substantially more than before,[13] and the gain is as high as 8.44% when $\gamma = 2$. It has an even more impressive value of 10.67% when $\gamma = 8$. Overall, the gains reported in Table 8 are clearly economically important. Although not reported in the table, as A goes up, the gains become smaller. When $A = 0.85$ or higher, most of the gains are under 1%. Hence, that our results do not claim asymmetry makes a big difference to all investors, though it does matter to investors with suitable DA parameters.

Related to utility gains, there are two interesting questions on the optimal portfolio weights. First, to what extent do asymmetries affect these portfolio weights? To address this question, we hold all other calibrated parameters constant while allowing κ to vary from 75% to 50% and to 25%. Recall that κ summarizes all three asymmetries, and so these values of κ reflect some asymmetry, more asymmetry, and severe asymmetry, respectively. We find that, other things being equal, the more the asymmetry, the less the holdings of the asymmetric assets. This is apparent with the portfolio allocation on size 1. Consider, for example, the case when $A = 0.55$ and $\gamma = 2$. The allocations are 20.8%, 15.7% and 9.2%, respectively, for the three varying κ values. Similar patterns are also found with alternative asset universes that contain fewer size portfolios (results are available upon request). The finding makes obvious economic

[13] Because of the non-participation result of Ang, Bekaert and Liu (2005), the investor will not invest in the stock market if A is small enough under either of the data-generating processes. In this case, there will be no utility gains.

sense. As asymmetry increases, the risk of loss increases too. To reduce the risk, the holdings of the risky assets must be reduced.

The second question is how beliefs about asymmetry and symmetry affect the allocations. We find results similar to the above but for a different reason. The holding on size 1, for example, reduces from 21.8% to 20.4% when switching beliefs from symmetry to asymmetry and when $A = 0.55$. The reduction becomes much greater, from 19.5% to 3.6%, when $A = 0.25$. The reason for the reduced holding is that size 1 appears less risky to those who believe symmetries because higher moments reflecting asymmetries are not incorporated into their objective function. In contrast, with belief of asymmetry, the higher moments matter and hence size 1 becomes riskier. To minimize the risk, its holding must be reduced. For those investors who are more disappointment-averse, the holding is reduced even more, as in the $A = 0.25$ case. In summary, an increase in asymmetries makes an otherwise identical portfolio riskier and hence the allocation to risky assets should be reduced. A belief that accounts for asymmetries versus one that does not can lead to the same results.

4. Conclusions

Many recent studies examine asymmetric characteristics of asset returns in both domestic and international markets. Of particular interest are asymmetric correlations in which stock returns tend to have higher correlations with the market when it goes down than when it goes up. Ang and Chen (2002) seem to be the first to provide a novel test for the null hypothesis of symmetric correlations, but their test is model dependent, testing the joint hypothesis of both symmetry and validity of a given model so that a rejection of symmetry may be solely due to a rejection of the model. In this article, we address the question of whether the data are symmetric at all by proposing a test that is completely model free. A rejection of the symmetry hypothesis by our test tells us that *any* symmetric model (under some standard regularity conditions) cannot explain the data. In addition, our test has a simple asymptotic chi-square distribution and can be adapted easily for testing beta and covariance symmetries.

Applying our tests to the CRSP 10 size portfolios, we find that asymmetric correlations, betas, and covariances are significant only for the first four smallest size portfolios, despite the fact that sample estimates all indicate asymmetries. We also apply our tests to both book-to-market and momentum decile portfolios. While there is no evidence of asymmetries for the book-to-market portfolios, we do find that the top two winner portfolios have strong asymmetric betas and covariances. Besides addressing the statistical significance of asymmetries, we propose a Bayesian framework, which accounts for both parameter and model

uncertainties, to model them as well as to assess their economic value. We find that incorporating assets' asymmetric characteristics can add substantial economic value in portfolio decisions. Finally, the methodology proposed in this article seems useful not only in testing asymmetric correlations, betas, and covariances but also in studying almost any asymmetric properties of the data.

Appendix A: Proofs

Proof of Theorem 1: In the proof below, we first spell out clearly what the regularity conditions are in addition to Assumption A.1 stated earlier, and then provide the rigorous proof. Throughout this proof, we use C to denote a generic bounded constant that may differ from place to place.

Assumption A.2: (i) The return series of the two portfolio returns, $\{R_{1t}, R_{2t}\}$, is a bivariate fourth order stationary process with $E(|R_{1t}|^{4\nu} + E|R_{2t}|^{4\nu}) \leq C$ for some $\nu > 1$; (ii) $\{R_{1t}, R_{2t}\}$ is an α-mixing process with α-mixing coefficient satisfying $\sum_{j=-\infty}^{\infty} j^2 \alpha(j)^{\frac{\nu}{\nu-1}} < \infty$.

Assumption A.3: The kernel function $k : \mathbb{R} \to [-1, 1]$ is symmetric about zero and is continuous at all points except a finite number of them on $\mathbb{R}$, with $k(0) = 1$ and $\int_{-\infty}^{\infty} |k(z)| dz < \infty$.

Assumption A.4: The bandwidth $p = p(T) \to \infty$, $p/T \to 0$ as the sample size $T \to \infty$.

Assumption A.5: (i) For some $b > 1$, $|k(z)| \leq C|z|^{-b}$ as $z \to \infty$; (ii) $|k(z_1) - k(z_2)| \leq C|z_1 - z_2|$ for any z_1, z_2 in $\mathbb{R}$.

Assumption A.6: $\hat{p}$ is a data-dependent bandwidth such that $\hat{p}/p = 1 + O_P(p^{1+b}/T^{\kappa(1+b)})$ for any $0 < \kappa < \frac{1}{2}$ and some nonstochastic bandwidth p satisfying $p = p(T) \to \infty$, $p/T^{\kappa} \to 0$.

Assumption A.2 allows for the existence of volatility clustering, which is a well-known stylized fact for most financial time series. The mixing condition is commonly used for a nonlinear time series analysis, as is the case with our test, because we only consider the cross-correlation in the tail distributions of the returns $\{R_{1t}, R_{2t}\}$. This condition characterizes temporal dependence in return series and rules out long memory processes. However, it is well known that returns of portfolios have weak serial correlations. Therefore, the mixing condition is quite reasonable in the present context.

Assumptions A.3 and A.4 are standard conditions on the kernel function $k(\cdot)$ and bandwidth p. These conditions are sufficient when we use nonstochastic bandwidths. Assumptions A.5 imposes some extra conditions on the kernel function, which is needed when we use data-dependent bandwidth $\hat{p}$. Many commonly used kernels, such as the Bartlett, Parzen, and quadratic-spectral kernels, are included. However, Assumption A.5 rules out the truncated and Daniell kernels. For various kernels, see, for example, Priestley (1981, p. 442) for a detailed discussion. Assumption A.6 imposes a rate condition on the data-driven bandwidth $\hat{p}$, which ensures that using $\hat{p}$ rather than p has no impact on the limit distribution of our test statistic. Commonly used data-driven bandwidths are the Andrews (1991) parametric plug-in method or the Newey and West (1994) nonparametric plug-in method. Note that the condition on p in Assumption A.6 is more restrictive than Assumption A.4, but it still allows for optimal bandwidths for most commonly used kernels. All these ensure that our test is completely model free. Right prior to the proof, we re-state Theorem 1 in the following technically clearer way.

Theorem 1: Suppose Assumptions A.1–A.4 hold. Then, under H_0, we have (i)

$$\mathcal{J}_\rho = (\hat{\rho}^+ - \hat{\rho}^-)' \hat{\Omega}^{-1} (\hat{\rho}^+ - \hat{\rho}^-) \to^d \chi^2_m \tag{A1}$$

as $T \to \infty$; and (ii), if, in addition, Assumptions A.5 and A.6 hold, $\hat{\mathcal{J}}_\rho - \mathcal{J}_\rho \to^p 0$, and

$$\hat{\mathcal{J}}_\rho \to^d \chi^2_m. \tag{A2}$$

Proof: (i) We first use the Cramer–Wold device to show $\sqrt{T}(\hat{\rho}^+ - \hat{\rho}^-) \to^d N(0, \Omega)$. Put $\hat{\xi}_t = \Sigma_{j=1}^m \lambda_j \hat{\xi}_t(c_j)$ and $\xi_t = \Sigma_{j=1}^m \lambda_j \xi_t(c_j)$, where $\hat{\xi}_t(c)$ and $\xi_t(c)$ are defined in (12) and Assumption A.1 respectively, and $\lambda = (\lambda_1, \ldots, \lambda_m)'$ is an $m \times 1$ vector such that $\lambda'\lambda = 1$. We then have $\lambda'(\hat{\rho}^+ - \hat{\rho}^-) = \Sigma_{j=1}^m \lambda_j[\hat{\rho}^+(c_j) - \hat{\rho}^-(c_j)] = T^{-1}\Sigma_{t=1}^T \hat{\xi}_t$. By tedious but straightforward algebra, this reduces to $\lambda'(\hat{\rho}^+ - \hat{\rho}^-) = T^{-1}\Sigma_{t=1}^T \xi_t + o_P(T^{-1/2})$. In other words, the replacement of the sample means, sample variances, and sample proportions with their population counterparts has no impact on the asymptotic distribution of $\sqrt{T}\lambda'(\hat{\rho}^+ - \hat{\rho}^-)$.

Given Assumption A.2, $\{R_{1t}, R_{2t}\}$ is an α-mixing process, as is ξ_t, which is an instantaneous function of (R_{1t}, R_{2t}). Under $H_0 : \rho^+(c) = \rho^-(c)$ for all c, we have $E(\xi_t) = 0$ because $E[\xi_t(c_j)] = 0$. In addition, given Assumptions A.1 and A.2, we have

$$\begin{aligned} V &= \lim_{T\to\infty} \operatorname{var}\left[T^{-1/2}\sum_{t=1}^T \xi_t\right] = \sum_{j=-\infty}^{\infty} \operatorname{cov}(\xi_t, \xi_{t-j}) \\ &= \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j \sum_{l=-\infty}^{\infty} \operatorname{cov}[\xi_t(c_i), \xi_{t-l}(c_j)] \\ &= \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j \Omega_{ij} \\ &= \lambda'\Omega\lambda. \end{aligned} \tag{A3}$$

Note that $0 < V < \infty$ for all λ such that $\lambda'\lambda = 1$, because Ω is positive definite. Thus, using the central limit theorem for mixing processes (e.g., White 1984, Theorem 5.19), we have

$$\sqrt{T}(\hat{\rho}^+ - \hat{\rho}^-)/\sqrt{V} \to^d N(0, 1). \tag{A4}$$

It follows from the Cramer–Wold device that $\sqrt{T}(\hat{\rho}^+ - \hat{\rho}^-) \to^d N(0, \Omega)$, and hence

$$T(\hat{\rho}^+ - \hat{\rho}^-)'\Omega^{-1}(\hat{\rho}^+ - \hat{\rho}^-) \to^d \chi^2_m. \tag{A5}$$

Next, we show $\hat{\Omega} \to^p \Omega$. Write $\hat{\Omega} - \Omega = [\hat{\Omega} - E\hat{\Omega}] + [E\hat{\Omega} - \Omega]$. By the Andrews (1991) Lemma 1, Assumption A.2 implies that Assumption A of Andrews (1991) holds. It follows from the Andrews (1991) Proposition 1(a) that $\operatorname{var}(\hat{\Omega}) = E[(\hat{\Omega} - E\hat{\Omega})(\hat{\Omega} - E\hat{\Omega})'] = O(p/T)$. Therefore we have $\hat{\Omega} - \Omega = O_P(p^{1/2}/T^{1/2})$ by Chebyshev's inequality. In addition, because Assumption A.2(ii) implies $\Sigma_{j=-\infty}^{\infty}\Omega(j) \le C$, and because of Assumption A.4 and dominated convergence, we have

$$E\hat{\Omega} - \Omega = \Sigma_{j=1-T}^{T-1}[(1 - |j|/T)k(j/p) - 1]\Omega(j) + \Sigma_{|j|>T}\Omega(j) \to 0 \tag{A6}$$

as $T \to \infty$. Consequently, $\hat{\Omega} \to^p \Omega$. By Slutsky's theorem, we then obtain

$$J = T(\hat{\rho}^+ - \hat{\rho}^-)'\hat{\Omega}^{-1}(\hat{\rho}^+ - \hat{\rho}^-) \to^d \chi^2_m. \tag{A7}$$

(ii) Let $\hat{\Omega}^*$ and $\hat{\Omega}$ be the kernel estimators for Ω using the bandwidth $\hat{p}$ and p respectively. It suffices to show $\hat{\Omega}^* - \hat{\Omega} \to^p 0$ and then we can apply Slutsky's theorem. By the definition

of $\hat{\Omega}$, we have for the (i, j)-th element,

$$\begin{aligned}\hat{\Omega}^*_{ij} - \hat{\Omega}_{ij} &= \sum_{l=1-T}^{T-1} \big[k(l/\hat{p}) - k(l/p)\big]\hat{\gamma}_l(c_i, c_j) \\ &= \sum_{|l|\le q} \big[k(l/\hat{p}) - k(l/p)\big]\hat{\gamma}_l(c_i, c_j) + \sum_{q<|l|<T} \big[k(l/\hat{p}) - k(l/p)\big]\hat{\gamma}_l(c_i, c_j) \\ &= \hat{A}_1(i, j) + \hat{A}_2(i, j), \text{ say,} \end{aligned} \tag{A8}$$

where $q = T^\kappa$ for κ as in Assumption A.6.

We now consider the first term $\hat{A}_1$. Using Assumption A.5(ii) and the triangle inequality, we have

$$\begin{aligned}|\hat{A}_1(i, j)| &\le \sum_{|l|\le q} C|(l/\hat{p}) - (l/p)| \cdot |\hat{\gamma}_l(c_i, c_j)| \\ &\le C|\hat{p}^{-1} - p^{-1}|q \sum_{|l|\le q} |\hat{\gamma}_l(c_i, c_j) - \gamma_l(c_i, c_j)| + C|\hat{p}^{-1} - p^{-1}|q \sum_{|l|\le q} |\gamma_l(c_i, c_j)| \\ &= |\hat{p}^{-1} - p^{-1}|O_P(q/T^{1/2} + q) \\ &= O(q|\hat{p}^{-1} - p^{-1}|), \end{aligned} \tag{A9}$$

where we have made use of the facts that $\Sigma_{l=-\infty}^{\infty}|\gamma_l(c_i, c_j)| \le C$ and $\sup_{0<l<T} E[\hat{\gamma}_l(c_i, c_j) - \gamma_l(c_i, c_j)]^2 = O(T^{-1})$, which follows by the Hannan (1970) Equation (3.3) and Assumption A.2 (recall that this assumption ensures that the fourth order cumulant condition holds).

For the second term $\hat{A}_2(i, j)$, using Assumption A.5(i), we have

$$\begin{aligned}|\hat{A}_2(i, j)| &\le \sum_{q<|l|<T} C(|l/\hat{p}|^{-b} + |l/p|^{-b})|\hat{\gamma}_l(c_i, c_j)| \\ &\le C(\hat{p}^b + p^b)q^{1-b}q^{-1} \sum_{q<|l|<T} (l/q)^{-b}|\hat{\gamma}_l(c_i, c_j) - \gamma_l(c_i, c_j)| \\ &\quad + C(\hat{p}^b + p^b)q^{-b} \sum_{q<|l|<T} |\gamma_l(c_i, c_j)| \\ &= C(\hat{p}^b + p^b)q^{-b}[O_P(q/T^{1/2}) + o_P(1)], \end{aligned} \tag{A10}$$

where again we have used the facts that $\Sigma_{l=-\infty}^{\infty}|\gamma_l(c_i, c_j)| \le C$ and $\sup_{0<l<T} E[\hat{\gamma}_l(c_i, c_j) - \gamma_l(c_i, c_j)]^2 = O(T^{-1})$.

Combining (A1)–(A3), $q = o(T^{1/2})$ and $\hat{p}/p = 1 + O_P(p^{1+b}/q^{1+b})$ as implied by Assumption A.6, we have $\hat{\Omega}^* - \hat{\Omega} = o_P(1)$. Q.E.D.

Proof of Theorem 2: The proof is similar to that of Theorem 1 and is thus omitted. Q.E.D.

Derivation for the Asymptotic Distribution of the Ang and Chen Test:
Consider a matrix expression for their test:

$$H^2 = \sum_{i=1}^{m} w(c_i)(\rho(c_i, \phi) - \hat{\rho}(c_i))^2 = (\hat{\rho} - \rho)'W(\hat{\rho} - \rho), \tag{A11}$$

where W is a diagonal matrix formed by the weights, and $\hat{\rho}$ and ρ are defined accordingly. Let V be the asymptotic covariance matrix of $\hat{\rho}$ as computed for our tests. Then, asymptotically, $Z = V^{-1/2}(\hat{\rho} - \rho) \sim N(0, I)$. Let λ be a diagonal matrix of the eigenvalues of the matrix

$U = V^{1/2} W V^{1/2} = C\lambda C$, where C is an orthogonal matrix. Then, asymptotically,

$$H^2 = (\hat{\rho} - \rho)' W (\hat{\rho} - \rho) = Z' U Z = \sum_{i=1}^{m} \lambda_i \chi_i^2, \tag{A12}$$

where $\chi_1^2, \ldots, \chi_m^2$ are independent chi-squared random variables with degrees of freedom 1. Based on the above, the asymptotic distribution of H^2, and hence H, can be easily determined by simulating chi-squared random variables of the right-hand side. Q.E.D.

Appendix B: Bayesian Inference in the Mixture Copula Model

The data-generating process for each asset is the standard GARCH(1,1) process: $r_{i,t} = \mu_i + \varepsilon_{it}$ with ε_{it} normally distributed with a time-varying variance $\sigma_{it}^2 = a_i + b_i \sigma_{it-1}^2 + c_i \varepsilon_{it-1}^2$. Let $u_{it} = \Phi(x_{it})$ with $x_{it} = (r_{i,t} - \mu_i)/\sigma_{it}$. Then, the joint distribution of the us is given by an n-dimensional form of Equation (39) with

$$f_{nor}(u_{1t}, u_{2t}, \ldots, u_{nt}; \Sigma) = \frac{1}{|\Sigma|^{\frac{1}{2}}} exp\left(-\frac{1}{2}\zeta_t^T(\Sigma^{-1} - I_n)\zeta_t\right), \tag{B1}$$

$$f_{clay}(u_{1t}, u_{2t}, \ldots, u_{nt}; \tau) = \left[\prod_{i=1}^{n}(1 + (i-1)\tau)\right]\left(\sum_{i=1}^{n} u_{it}^{-\tau} - n + 1\right)^{-\frac{1}{\tau}-n}\left(\prod_{i=1}^{n} u_{it}\right)^{-\tau-1}, \tag{B2}$$

where $\zeta_t = (\Phi^{-1}(u_{1t}), \ldots, \Phi^{-1}(u_{nt}))'$, Σ is the correlation coefficient matrix and I_n is the identity matrix of order n.

In the Bayesian framework, τ is viewed as a random variable. We model it discretely by assuming that it takes values from set $S_\tau = \{0.1, 0.2, 0.3, 0.4, \ldots, 9.9, 10\}$. The diffuse prior on τ can be written as

$$p_0(\tau) = \frac{1}{|S_\tau|}, \tag{B3}$$

where $|S_\tau| = 100$, the number of elements in set S_τ. Then, we can use an almost diffuse prior for the model parameters,

$$p_0(\kappa, \tau, \Sigma) \propto p_0(\kappa) p_0(\tau) p_0(\Sigma), \tag{B4}$$

where

$$p_0(\kappa) \sim Beta(1, 1), \qquad p_0(\Sigma) \sim W(\nu_\Sigma, I_n), \tag{B5}$$

and $\nu_\Sigma = 14$ is the prior degree of freedom in the Wishart distribution.

To make Markov chain Monte Carlo (MCMC) posterior draws, we augment the data with independent and identically distributed (iid) samples $\{w_t\}_{t=1}^T$ from Binomial $(1, \kappa)$. Note that both the prior and the likelihood function conditional on the augmented data can be factored into two independent components on Σ and τ; it is hence feasible to draw a sample from the joint posterior distribution of w and $\theta = \{\kappa, \tau, \Sigma\}$. Ignoring w, the θ should be a sample from its marginal posterior. Starting with a κ from $p_0(\kappa) \sim Beta(1, 1)$ and iid $\{w_t\}_{t=1}^T$ from Binomial $(1, \kappa)$, the following steps implement the idea:

1. Divide the data $U_T = \{u^1, u^2, \ldots, u^T\}$, where $u^t = (u_{1t}, u_{2t}, \ldots, u_{nt})$, $t = 1, 2, \ldots, T$, into two groups, u_{nor} or u_{clay} according to whether $w_t = 1$ or 0;

2. Let T_{nor} denote the number of observations in u_{nor}. Then,

$$\Sigma^{-1} \mid u_{nor} \propto W\left(T_{nor} + \nu_{\Sigma},\ \left(I_n + T_{nor}\widehat{\Sigma}_{nor}\right)^{-1}\right), \tag{B6}$$

where $\widehat{\Sigma}_{nor} = \frac{1}{T_{nor}} \sum_{t=1}^{T} (x^t)' \times x^t 1(u^t \in u_{nor})$ and $x^t = (x_{1t}, x_{2t}, \ldots, x_{nt})$;

3. Draw τ from the posterior,

$$p(\tau|u_{clay}) \propto p_0(\tau) \prod_{u^t \in u_{clay}} f_{clay}(u_{1t}, u_{2t}, \ldots, u_{nt}; \tau); \tag{B7}$$

4. Draw κ from Beta $\left(T_{nor} + 1, T_{clay} + 1\right)$, where T_{clay} denotes the number of observations in u_{clay} and $T_{clay} = T - T_{nor}$;
5. Draw w_t from Binomial $(1, \kappa_t)$ for $t = 1, 2, \ldots, T$, where

$$\kappa_t = \frac{\kappa f_{nor}(u_{1t}, u_{2t}, \ldots, u_{nt}; \Sigma)}{\kappa f_{nor}(u_{1t}, u_{2t}, \ldots, u_{nt}; \Sigma) + (1-\kappa) f_{clay}(u_{1t}, u_{2t}, \ldots, u_{nt}; \tau)};$$

6. Repeat steps (1)–(5).

Let M be the number of total iterations in the above loop. Disregarding the first L ones for the burning period, the remaining $Q = M - L$ draws will be the posterior draws. Now, for each such draw of the parameters, say, Σ^q, κ^q and τ^q, we obtain u^{T+1} from the mixture copula. This way provides us Q draws from the predicative distribution. The following steps implement the idea:

1. Draw $u^{nor} = \{u_i^{nor}\}_{i=1}^n$ from the normal copula with correlation coefficient matrix Σ^q;
2. Draw $u^{clay} = \{u_i^{clay}\}_{i=1}^n$ from the Clayton copula with parameter τ^q;
3. Generate a draw from the mixture copula with mixture parameter κ^q as follows:
   (a) Simulate a uniform random variant, $d \sim U(0, 1)$;
   (b) Set $u_i^{mix} = u_i^{nor} 1(d < \kappa^q) + u_i^{clay} 1(d > \kappa^q)$, $i = 1, \ldots, n$, then $u^{mix} = (u_1^{mix}, u_2^{mix}, \ldots, u_n^{mix})$ is one draw from the mixture copula with parameter, $(\Sigma^q,\ \kappa^q,\ \tau^q)$, which is also a draw from the predicative distribution of $p(u^{T+1}|U_T)$.

Based on the above Q draws from the predicative distribution of $p(u^{T+1}|U_T)$, we can have a sample of size Q from the predictive distribution of $p(R^{T+1}|R)$.

References

Aï-Sahalia, Y., and M. Brandt. 2001. Variable Selection for Portfolio Choice. *Journal of Finance* 54:1297–351.

Andrews, D. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Estimation. *Econometrica* 59:817–58.

Ang, A., and G. Bekaert. 2000. International Asset Allocation with Time-varying Correlations. *Review of Financial Studies* 15:1137–87.

Ang, A., and J. Chen. 2002. Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics* 63:443–94.

Ang, A., G. Bekaert, and J. Liu. 2005. Why Stocks May Disappoint. *Journal of Financial Economics* 76:471–508.

Bae, K. H., G. A. Karolyi, and R. M. Stulz. 2003. A New Approach to Measuring Financial Market Contagion. *Review of Financial Studies* 16:717–64.

Ball, R., and S. P. Kothari. 1989. Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns. *Journal of Financial Economics* 25:51–74.

Bekaert, G., and G. Wu. 2000. Asymmetric Volatility and Risk in Equity Markets. *Review of Financial Studies* 13:1–42.

Benartzi, S., and R. Thaler. 1995. Myopic Loss Aversion and the Equity Premium Puzzle. *Quarterly Journal of Economics* 110:73–792.

Boyer, B. H., M. S. Gibson, and M. Loretan. 1999. Pitfalls in Tests for Changes in Correlations, *International Finance Discussion Paper*, Vol. 597, Board of Governors of the Federal Reserve System, Washington, DC.

Cho, Y. H., and R. F. Engle. 2000. Time-Varying Betas and Asymmetric Effects of News: Empirical Analysis of Blue Chip Stocks, Working Paper, National Bureau of Economic Research, Cambridge, MA.

Conrad, J., M. Gultekin, and G. Kaul. 1991. Asymmetric Predictability of Conditional Variances. *Review of Financial Studies* 4:597–622.

Den Haan, J., and A. Levin. 1997. A Practitioner's Guide to Robust Covariance Matrix Estimation, in G. S. Maddala, and C. R. Rao (eds), *Handbook of Statistics*. New York: Elsevier Science.

Fleming, J., C. Kirby, and B. Ostdiek. 2001. The Economic Value of Volatility Timing. *Journal of Finance* 56:329–52.

Forbes, K., and R. Rigobon. 2002. No Contagion, Only Interdependence: Measuring Stock Market Co-movements. *Journal of Finance* 57:2223–61.

Gul, F. 1991. A Theory of Disappointment Aversion. *Econometrica* 59:667–86.

Hannan, E. 1970. *Multiple Time Series*. New York: Wiley.

Hansen, L. P. 1982. Large Sample Properties of the Generalized Method of Moments Estimators. *Econometrica* 50:1029–54.

Harvey, C. R., and A. Siddique. 2000. Conditional Skewness in Asset Pricing Tests. *Journal of Finance* 55:1263–95.

Hu, L. 2004. Dependence Patterns across Financial Markets: a Mixed Copula Approach, Working Paper, Ohio State University.

Jaeger, R. 2003. *All About Hedge Funds*. New York: McGraw-Hill.

Kan, R., and G. Zhou. 2006. Optimal Portfolio Choice with Parameter Uncertainty, Working Paper, Washington University, St. Louis.

Kandel, S., and R. F. Stambaugh. 1996. On the Predictability of Stock Returns: An Asset-allocation Perspective. *Journal of Finance* 51:385–424.

Karolyi, A., and R. Stulz. 1996. Why Do Markets Move Together? An Investigation of US-Japan Stock Return Comovements. *Journal of Finance* 51:951–86.

Liu, L., J. B. Warner, and L. Zhang. 2005. Momentum Profits and Macroeconomic Risk, Working Paper, HKUST, University of Rochester and NBER.

Longin, F., and B. Solnik. 2001. Extreme Correlation of International Equity Markets. *Journal of Finance* 56:649–76.

McLachlan, G., and T. Krishnan. 1996. *The EM Algorithm and Extensions*. New York: Wiley.

Nelsen, R. B. 1999. *An Introduction to Copulas*. New York: Springer-Verlag.

Newey, W., and K. West. 1994. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* 61:631–53.

Pástor, L., and R. F. Stambaugh. 2000. Comparing Asset Pricing Models: An Investment Perspective. *Journal of Financial Economics Economics* 56:335–81.

Patton, A. J. 2004. On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation. *Journal of Financial Econometrics* 2:130–68.

Priestley, M. 1981. *Time Series and Spectral Analysis*. London: Academic Press.

Redner, R., and H. Walker. 1984. Mixture Densities, Maximum Likelihood and The EM Algorithm. *SIAM Review* 26:195–239.

Schwert, G. W. 1989. Why Does Stock Market Volatility Change over Time? *Journal of Finance* 44:1115–53.

Schwert, G. W. 2003. Anomalies and Market Efficiency, in G. Constantinides, M. Harris, and R. Stulz (eds), *Handbook of the Economics of Finance*. Amsterdam: North-Holland.

Stambaugh, R. F. 1995. Unpublished Discussion of Karolyi and Stulz (1996), *National Bureau of Economic Research*, Universities Research Conference on Risk Management, May 1995, Cambridge, MA.

Tu, J., and G., Zhou. 2004. Data-generating Process Uncertainty: What Difference Does It Make in Portfolio Decisions? *Journal of Financial Economics* 72:385–421.

White, H. 1984. *Asymptotic Theory for Econometricians*. San Diego: Academic Press.

Zellner, A. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.