JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 53, No. 4, Aug. 2018, pp. 1479–1507

doi:10.1017/S0022109018000340

Asymmetry in Stock Comovements: An Entropy Approach

Lei Jiang, Ke Wu, and Guofu Zhou*

Abstract

We provide an entropy approach for measuring the asymmetric comovement between the return on a single asset and the market return. This approach yields a model-free test for stock return asymmetry, generalizing the correlation-based test proposed by Hong, Tu, and Zhou (2007). Based on this test, we find that asymmetry is much more pervasive than previously thought. Moreover, our approach also provides an entropy-based measure of downside asymmetric comovement. In the cross section of stock returns, we find an asymmetry premium: Higher downside asymmetric comovement with the market indicates higher expected returns.

I. Introduction

A number of studies have found asymmetries in asset returns with respect to the market. Ball and Kothari (1989), Schwert (1989), Conrad, Gultekin, and Kaul (1991), Kroner and Ng (1998), Cho and Engle (1999), Bekaert and Wu (2000), Ang and Chen (2002), Bae, Karolyi, and Stulz (2003), and Ang, Chen, and Xing (2006), among others, document asymmetries in the covariances, correlations,

*Jiang, jianglei@sem.tsinghua.edu.cn, Tsinghua University School of Economics and Management; Wu (corresponding author), ke.wu@ruc.edu.cn, Renmin University of China Hanqing Advanced Institute of Economics and Finance; and Zhou, zhou@wustl.edu, Washington University in St. Louis Olin School of Business. We are grateful to Vikas Agarwal, Doron Avramov, Jonathan Brogaard, Jeffrey Busse, Hui Chen, Tarun Chordia, Lauren Cohen, Philip Dybvig, Andrey Ermolov, Ping He, Zhiguo He, David Jacho-Chavez, Ying Jiang, Raymond Kan, Omesh Kini, Junghoon Lee, Tingjun Liu, Esfandiar Maasoumi, Michael Powers, Jun Tu, Aurelio Vasquez, Baozhong Yang, and seminar participants at Cornerstone Research, Case Western Reserve University, Emory University, Georgia State University, Harbin Institute of Technology, ITAM, Peking University, Washington University in St. Louis, the 2014 and 2015 Tsinghua Finance Workshops, the 2015 China Finance Review International Conference, the 2015 Five-Star Workshop in Finance, the 2015 China International Conference in Finance, and the 2017 Midwest Finance Association Annual Meeting for numerous helpful comments and especially to Jennifer Conrad (the editor) and Riccardo Colacito (the referee) for their many insightful and detailed comments that have substantially improved the article. Jiang gratefully acknowledges financial support from the AXA Research Fund, the Initiative Research Program supported by Tsinghua University (20151080398), the National Natural Science Foundation of China (71572091), and the Tsinghua National Laboratory for Information Science and Technology. The authors thank Jing Ding and Jinyu Liu for excellent research assistance.

and betas of stock returns. Recently, Lettau, Maggiori, and Weber (2014) have also found asymmetric betas in currency portfolios. Such asymmetries in asset returns are important for both portfolio selection and risk management because diversification and hedging strategies rely critically on the ways in which asset returns are dependent on each other. Whereas earlier tests of symmetry in the stock market are model dependent, Hong, Tu, and Zhou (HTZ) (2007) propose a model-free test. This test, despite its great potential, has some weaknesses. First, it does not address asymmetry beyond the second moment and hence cannot capture the full dependence structure for nonnormal distributions; this is a concern because realized stock returns are well known to be nonnormally distributed (e.g., Zhou (1993), Ang and Chen (2002)). Second, the test has low power and fails to detect asymmetry in many portfolios. For example, no asymmetry is found in book-to-market portfolios, which appears inconsistent with the results of Ang and Chen (2002).

In this article, we propose an entropy measure to quantify asymmetry in the joint distribution of an individual stock return and the market return and then extend it to measure their downside asymmetric comovement. Based on the entropy asymmetry measure, we provide a new model-free test to determine whether or not asymmetric comovement exists in the stock market. Statistically, the test statistic is a bivariate normalized version of the entropy measure originally proposed by Granger, Maasoumi, and Racine (2004), which has been widely applied in econometrics (see, e.g., Maasoumi and Racine (2002), Racine and Maasoumi (2007)). The entropy measure, based on the joint probability density functions, summarizes all the information in the joint distribution, thus capturing the general asymmetric dependence structure that exists in all moments beyond the second.

With Monte Carlo simulations, we find that the entropy test has correct empirical size and is generally more powerful than the HTZ (2007) test. Using portfolios sorted by size, book-to-market ratio, or prior (2- to 12-month) returns, we find that the entropy test detects statistically significant asymmetry in all of them. For example, in contrast to the HTZ test, which fails to reject symmetry for all deciles of book-to-market portfolios, the entropy test rejects symmetry in 7 out of 10 portfolios at the 10% significance level using a single exceedance level.[1]

What are the asset-pricing implications of asymmetric comovements on the cross section of stock returns? To address this question, we extend the entropy asymmetry measure to a measure of downside asymmetric comovement. Ang, Bekaert, and Liu (2005) show that downside asymmetry risk matters to investors with disappointment aversion (DA) preferences, and Dahlquist, Farago, and Tédongap (2017) further develop portfolio strategies under asymmetry. Because investors dislike stocks with downside asymmetric comovement with the market, we expect a risk premium for such stocks. Whether or not the size of this premium is of economic significance is an empirical question that we investigate here.

[1]Using multiple exceedance levels, symmetry is rejected in 4 out of the 10 portfolios at the 10% level.

By sorting stocks into quintile portfolios by their prior downside asymmetry, we find that the return difference between the highest quintile and lowest quintile can be as high as 4.50% per annum, which is both economically and statistically significant.[2] Because this difference in returns is a compensation for downside asymmetry risk, it is not surprising that it cannot be explained by standard factor models such as the 4-factor model of Carhart (1997). Moreover, the risk premium is time-varying, rising during periods of low investor sentiment, low market volatility, and high market liquidity.

We also compare our entropy-based downside asymmetry measure with downside beta (Ang et al. (2006)) and coskewness (Harvey and Siddique (2000)). The downside beta captures a stock's systematic risk exposure to the downside market, and coskewness describes the comovement of an asset return with the square of the market return. Differences in returns between stocks sorted by either of these two measures are much smaller and can be fully explained by Carhart's (1997) 4-factor model. Motivated by the definition of the HTZ (2007) test statistic, we construct a downside asymmetry risk measure based on the exceedance correlations. The return spread between stocks sorted by this measure earns a risk premium of 3.24% per annum, but the Carhart 4-factor alpha is still not significant. In contrast, the entropy-based measure yields an approximately 40% higher risk premium and is highly significant after the 4-factor adjustment. Because existing asymmetry measures are all based on the second and third moments only, the results show that our asymmetry measure, based on the entire joint density function, is indeed relevant in practice.

Backus, Boyarchenko, and Chernov (2018) and Chabi-Yo and Colacito (2016), motivated from different applications, propose two alternative co-entropy measures that capture nonlinear codependence between two random variables beyond the Pearson correlation. Interestingly, when we apply their measures to our sample, we also find statistically significant risk premiums greater than that from the downside beta. These results suggest that downside asymmetry in nonlinear codependence beyond the second moment is important and commands a substantial and statistically significant premium in the cross section of stocks, robust to various entropy-based measures.

There are several recent studies emphasizing tail risk and cross-sectional asset-pricing implications. For example, Kelly and Jiang (2014) propose a left-tail-risk factor and find that the higher exposure to the tail risk, the higher the expected returns. Chabi-Yo, Ruenzi, and Weigert (2018) examine copula-based left-tail dependence and find that it has explanatory power on the cross section of stock returns. In contrast, we focus on the difference between the left and right tails rather than their absolute magnitude. In other words, no matter how fat the left tail of the joint distribution of a stock return and the market return may be, we do not consider the distribution asymmetrical as long as it has an equally fat right tail. In a related work, Alcock and Hatherley (2017) construct a beta-invariant asymmetry measure based on the HTZ (2007) test statistic and document a negative premium for upside asymmetry. However, their asymmetry

[2]To avoid the look-ahead bias, we sort portfolios based on predetermined variables only. If the contemporaneous downside asymmetry measure were used, the return spread would more than double.

measure is still based on the second moments and does not capture asymmetry existing in higher-order moments. Our article also differs from studies of the asset-pricing effect of asymmetry in individual stock distributions or asymmetry in the distributions of macroeconomic variables. For example, Conrad, Dittmar, and Ghysels (2013) and Amaya, Christoffersen, Jacobs, and Vasquez (2015) find evidence that stocks with right-skewed return distributions have lower cross-sectional average returns. Colacito, Ghysels, Meng, and Siwasarit (2016) introduce skewness into consumption-based asset-pricing models and provide empirical evidence that the skewness of analysts' forecasts for gross domestic product (GDP) growth predicts equity returns. In another recent paper, Jiang, Wu, Zhou, and Zhu (2018) study general univariate asymmetry and its asset pricing implications. Unlike these studies, we focus on the asymmetry in the comovements of individual stock returns with the market return.

The remainder of the article is organized as follows: Section II introduces the entropy test for joint distribution asymmetry. Section III examines the test's size and power using Monte Carlo simulations. Section IV applies the entropy test to investigate asymmetry in common portfolios. Section V analyzes the impact of asymmetry on the cross section of stock returns. Section VI concludes.

II. Tests of Asymmetry

In this section, we first briefly review standard symmetry tests based on exceedance correlations, then extend the concept to more general asymmetry, and finally provide our entropy test.

A. Asymmetric Correlation

Ang and Chen (2002) and HTZ (2007) provide tests of asymmetry for a stock's comovement with the market, but both of these tests address only exceedance correlation asymmetry.

To see why, let x and y be pairs of stock return series, and let $\tilde{x}$ and $\tilde{y}$ be their standardizations. Both tests rely on exceedance correlations, that is, the conditional correlations calculated when both individual stock and market returns are below or above certain exceedance levels. The exceedance correlations at some given level c are defined as

$$\rho^+(c) = \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} > c, \tilde{y} > c), \tag{1}$$

$$\rho^-(c) = \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} < -c, \tilde{y} < -c), \quad \forall c \geq 0. \tag{2}$$

Clearly, both $\rho^+(c)$ and $\rho^-(c)$ measure conditional correlations between two return series conditioning on whether both series are above or below a certain exceedance level c. The null hypothesis of interest is

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

If the null is true, the exceedance correlations are symmetric.

Ang and Chen's (2002) test addresses whether the asymmetric correlations observed in the data can be explained by a given model. Therefore, the test is model dependent, and its results rely on the choice of model. One weakness associated with such a test is that rejection of the symmetry hypothesis could be a result of rejection of the given model.

To overcome this weakness, HTZ (2007) propose a model-free test. Their test statistic is defined as

$$J_\rho \quad = \quad T(\hat{\rho}^+ - \hat{\rho}^-)'\hat{\Omega}^{-1}(\hat{\rho}^+ - \hat{\rho}^-), \tag{3}$$

where T is the sample size, $\hat{\rho}^+$ and $\hat{\rho}^-$ are $m \times 1$ vectors of sample exceedance correlations, and $\hat{\Omega}$ is a consistent estimator of the covariance matrix of $\sqrt{T}(\hat{\rho}^+ - \hat{\rho}^-)$. Under the null of symmetric correlations and certain regularity conditions, the test has a simple asymptotic χ^2 distribution, $J_\rho \xrightarrow{d} \chi^2_m$. The test answers the question of whether or not asymmetric correlations exist in the data at all. In other words, if the test rejects the null, no distributions with symmetric correlations can fit the data well.

B. Asymmetric Comovement

The methodologies presented in the previous subsection use correlations to test for asymmetric comovement in bivariate stock returns. However, it is well known that a correlation coefficient measures only linear dependence and cannot reflect a dependence structure beyond the second moment. In general, a zero correlation does not imply independence if the joint distribution is nonnormally distributed. Hence, tests based on correlations ignore possible higher-order codependence. Because research has established that financial and macroeconomic time series are usually leptokurtic and have heavy tails (e.g., Zhou (1993), Cont (2001), Ang and Chen (2002), and Bekaert and Engstrom (2017)),[3] studying asymmetric comovement in asset returns requires a test that involves all higher-order moments.

Because the joint density function uniquely defines a joint distribution, directly testing for asymmetry in the joint probability density function certainly involves all higher-order moments. We start with a univariate test of distributional asymmetry proposed by Racine and Maasoumi (2007) and extend it to the case of joint distributions.

Analogous to the exceedance correlations, we define exceedance densities as

$$f^+(c,\tilde{x},\tilde{y}) \quad = \quad f(\tilde{x},\tilde{y}|\tilde{x} > c, \tilde{y} > c), \tag{4}$$

$$f^-(c,\tilde{x},\tilde{y}) \quad = \quad f(-\tilde{x},-\tilde{y}|\tilde{x} > c, \tilde{y} > c), \quad \forall c \geq 0. \tag{5}$$

where $f^+(c,\tilde{x},\tilde{y})$ is the joint probability density function of standardized returns in the ranges of $\tilde{x} > c$ and $\tilde{y} > c$.[4] Restrictions to these ranges, where both returns are above or below a certain exceedance level, follow from the work of Ang and Chen (2002) and HTZ (2007) to capture the comovements in the lower and upper tails of the joint return distribution. $f^-(c,\tilde{x},\tilde{y})$ denotes the joint density function of the rotated return series around the mean in the ranges of $\tilde{x} > c$

[3]Beyond returns, Bekaert, Engstrom, and Ermolov (2015) provide an asymmetric model of volatility.

[4]We may consider $\tilde{x}$ as the return to a single stock and $\tilde{y}$ as the market return.

and $\tilde{y} > c$.[5] If the joint distribution is truly symmetric, then the two densities should be the same almost everywhere. Intuitively, the distance between the two density functions reflects the degree of asymmetry of the joint return distribution. Hence, our null hypothesis for testing symmetric comovement is

$$(6) \qquad H_0: \quad f^{+}(c,\tilde{x},\tilde{y}) \quad = \quad f^{-}(c,\tilde{x},\tilde{y}), \quad \text{for all } (c,\tilde{x},\tilde{y}) \in \mathbb{R}_{+} \times \mathbb{R}^2.$$

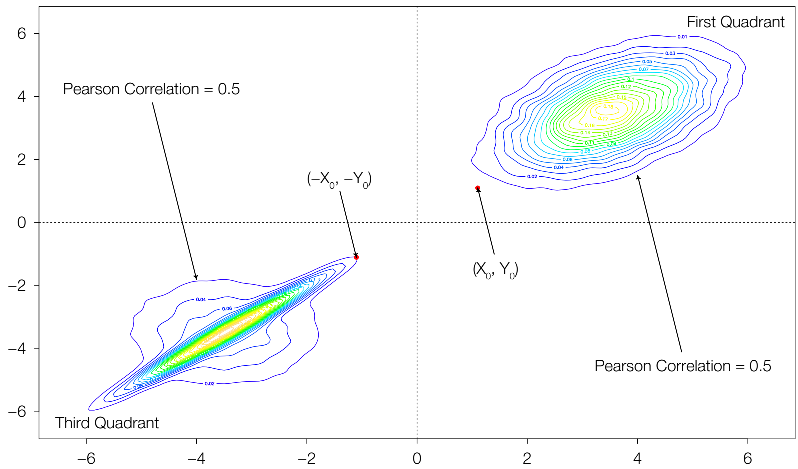
If this null hypothesis is rejected, then the data must possess asymmetry because the density function after a certain rotation differs from the original density function.

To explain the intuition behind this proposed entropy test, we give an illustrative example. Consider two random variables X and Y, which may be regarded as the standardized returns on a stock and on the market. Let their joint distribution be characterized by the contour plots of Figure 1. In the first quadrant, X and Y are drawn from a joint normal distribution with means of 3.5, with unit variances and a Pearson correlation of 0.5, whereas in the third quadrant, they are drawn independently from two joint normal distributions with equal probability of 0.5. Both joint normal distributions have means of -3.5 and unit variances, but one of them has 0 correlation, and the other one has a correlation coefficient equal to 1. Then, by construction, we know that the Pearson correlation coefficient is exactly the same, 0.5, in both quadrants. However, it is clear from Figure 1 that the codependence between X and Y looks quite different across the quadrants. This is

FIGURE 1

Illustrative Example: Symmetric Correlation but Asymmetric Comovement

Figure 1 shows contour plots of the joint distributions of two random variables X and Y. In the first quadrant, the joint distribution is a bivariate normal distribution with correlation 0.5. In the third quadrant, the distribution is a mixture of two bivariate normal distributions where the mixing parameter is set to 0.5. All normal distributions employed here have unit variances. The means of the normal distributions in the first quadrant are set to 3.5, whereas those in the third quadrant are set to -3.5. The points (X_0, Y_0) and $(-X_0, -Y_0)$ are symmetric about the origin.



[5] $f^{-}(c,\tilde{x},\tilde{y})$ can be equivalently defined as $f(\tilde{x},\tilde{y}|\tilde{x} < -c, \tilde{y} < -c)$, the joint density function of the original return series in the ranges of $\tilde{x} < -c$ and $\tilde{y} < -c$. We choose the definition in equation (5) so that the entropy measure of asymmetry as in equation (7) can be more concisely defined.

because Pearson's correlation can only pick up the linear codependence between two random variables and thus becomes an inaccurate measure of codependence in the presence of nonlinearities.

Consider the point (X_0, Y_0) in the first quadrant and its associated reflection point around the means $(-X_0, -Y_0)$ in the third quadrant. Clearly, the density associated with (X_0, Y_0) is equal to 0, whereas the density associated with $(-X_0, -Y_0)$ is not. Had the distribution been symmetric, then the two points would have been characterized by the same probability mass. Hence, differences in densities between any point (x, y) and its symmetric counterpart $(-x, -y)$ provide a measure of asymmetry for the joint distribution of X and Y. This simple figure illustrates the motivation for using equation (6) to test for symmetry.

C. Entropy Measure of Asymmetry

The important question is how to test the null hypothesis of symmetry provided in equation (6). Intuitively, the joint distribution is symmetric if the distance between $f^+(c)$ and $f^-(c)$ is 0 almost everywhere. Otherwise, it is asymmetric. To formalize this null hypothesis, we rely on measures of distance between the two probability density functions.

Originating in physics and information theory as a measure of uncertainty, entropy has a long history as a measure of divergence between distributions. It was first introduced by Shannon (1948) and later extended by Kullback and Leibler (1951).[6] More recently, entropy has drawn attention from financial economists. Recent examples include Cabrales, Gossner, and Serrano (2013), who use Shannon's (1948) entropy to quantify the informativeness of ruin-averse investors' beliefs on the state of nature. Backus, Chernov, and Zin (2014) use Kullback and Leibler's (1951) relative entropy to measure differences between physical and risk-neutral probabilities and derive appropriate bounds for the stochastic discount factors of consumption-based asset-pricing models. In more recent work, Backus et al. (2018) propose the use of co-entropy as a new measure of codependence between random variables; this is a bivariate variant of the relative entropy concept proposed by Backus et al. (2014). Chabi-Yo and Colacito (2016) propose an alternative co-entropy measure to evaluate the performance of international finance models. In Section V, we present a comparison of our methodology and these co-entropy measures proposed in the literature.

The entropy measure we use belongs to the same K-class entropy as the Kullback–Leibler relative entropy. First proposed by Granger et al. (2004), the measure is a special case of K-class entropy, with $K = 1/2$, which is a normalization of the Hellinger distance measure and the only metric entropy within its class. As shown by Granger et al. (2004), this metric measure has many other desirable properties.

Consider first, for simplicity, the case in which we have only one exceedance level c. The entropy measure of asymmetry is defined as

$$S_\rho(c) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(f^+(c,\tilde{x},\tilde{y})^{1/2} - f^-(c,\tilde{x},\tilde{y})^{1/2})^2\, d\tilde{x}\, d\tilde{y}, \tag{7}$$

[6] Ullah (1996) provides an excellent survey of various entropy measures and their wide applications in econometrics.

which is a function of the exceedance level c. In practice, c is chosen according to empirical interests. For example, when $c=0$, $S_\rho(0)$ measures the asymmetric comovement in the first quadrant (where both standardized returns are positive) and third quadrant (where both standardized returns are negative). If the asymmetric comovement in the tails of the distribution is of interest, $S_\rho(c)$ can be measured at other exceedance levels, such as 0.5 or 1 standard deviation away from the mean.

The entropy measure $S_\rho(c)$ is well defined for both continuous and discrete data. It takes values between 0 and 1 and equals 0 if and only if the densities are equal, which indicates symmetric comovement. Mathematically, it is a true measure of distance because it satisfies the triangular inequality. Moreover, the measure is invariant under continuous and strictly increasing transformations, such as the common logarithm transformation.

Consider now the case in which we have multiple exceedance levels, $c_1,\ldots,c_m$, and want to determine whether jointly symmetric comovement exists at all exceedance levels. For example, whereas the singleton set of $c=\{0\}$ is usually of interest in the literature, the set of levels $c=\{0,0.5,1,1.5\}$ is also common in previous studies, such as those of Ang and Chen (2002) and HTZ (2007). For the multilevel case, we can also apply the statistic in equation (7) to each of the individual levels and then aggregate the results. Because $S_\rho(c)$ is a metric and is always nonnegative, we could simply take the arithmetic average as an aggregation:

$$S_\rho = \frac{S_\rho(c_1)+\cdots+S_\rho(c_m)}{m}, \tag{8}$$

where $S_\rho(c_j)$ is computed from equation (7) for $j=1,\ldots,m$. Thus, the entropy test statistic is well defined for either the single test case with one exceedance level or the joint test case with multiple exceedance levels.

To carry out the entropy test in practice, we first need to estimate the joint density function of the data and then compute the integral in equation (7) to obtain the statistic $\hat{S}_\rho(c)$. Finally, we need to determine the distribution of the test statistic under the null hypothesis for computing the p-values of the test. This task is, unfortunately, much more complex than that of the asymmetric correlation test. The challenges inherent in the calculations are addressed in the following two subsections.

D. Nonparametric Estimation

In this subsection we provide a methodology to estimate the density functions in equation (7). Following Racine and Maasoumi (2007), among others, we use a nonparametric kernel smoothing method to consistently estimate unknown joint densities. Specifically, we use the popular Parzen (1962)–Rosenblatt (1956) kernel density estimator. For the univariate case, the Parzen–Rosenblatt kernel density estimator is defined as

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{X_i-x}{h}\right), \tag{9}$$

where n is the sample size of the data $\{X_i\}$, h is a smoothing parameter (commonly referred to as the bandwidth), and $k(\cdot)$ is a nonnegative bounded kernel function.

Because we are dealing with bivariate density functions, in estimation we employ a commonly used product kernel function constructed as the product of univariate kernel functions. Our candidate kernel density estimator of the data is given by

$$(10) \qquad \hat{f}(x,y) = \frac{1}{nh_1h_2}\sum_{i=1}^{n} k\left(\frac{x_i - x}{h_1}\right) \times k\left(\frac{y_i - y}{h_2}\right),$$

where n is the sample size, $k(\cdot)$ is a suitable univariate kernel function, h_1 and h_2 are the respective bandwidths for each of the two components, and $\{(x_i, y_i)\}$ is an observed data pair. In the empirical applications that follow, n is set equal to T, the length of the return series.

The accuracy of the nonparametric kernel density estimator relies on the selection of both the kernel function and the bandwidth. Because the return data are continuous, we use a standard Gaussian kernel, $k(z)=(1/(\sqrt{2\pi}))e^{-z^2/2}$, as the univariate kernel function in the estimation. The choice of kernel function is less important than the bandwidth selection. On selecting the bandwidth, we use the likelihood cross-validation method proposed by Duin (1976). This procedure solves the following maximization problem of the log-likelihood function;

$$(11) \qquad \max_{h_1,h_2} \mathcal{L} = \sum_{i=1}^{n} \ln[\hat{f}_{-i}(x_i, y_i)],$$

where

$$(12) \qquad \hat{f}_{-i}(x_i,y_i) = \frac{1}{nh_1h_2}\sum_{j\neq i}^{n} k\left(\frac{x_j - x}{h_1}\right) \times k\left(\frac{y_j - y}{h_2}\right),$$

which is equal to $\hat{f}(x,y)$ without the ith term. Econometrically, this data-driven bandwidth-selection method yields density estimates that are close to the actual density in terms of the Kullback–Leibler relative entropy (for details, see Li and Racine (2007)).

Using the method described previously, the density functions in equation (7) can be consistently estimated. The test statistic $\hat{S}_\rho(c)$ can then be obtained by computing the integral using a standard numerical procedure in the case of a single exceedance level. In the presence of multiple exceedance levels, the test statistic is computed from equation (8).

E. Distribution of the Test Statistic

To make statistical inferences for the entropy test of asymmetry, we need to know the sampling distribution of the test statistic under the null hypothesis of symmetry. There are two potential methods for determining the sampling distribution: asymptotic theory and the bootstrap resampling method.

The asymptotic theory for the class of entropies with similar functional forms was developed by Skaug and Tjøstheim (1993), Tjøstheim (1996), and Hong and White (2005). Under certain regularity conditions, $\hat{S}_\rho(c)$ follows an asymptotically normal distribution, and the distribution derived under the null hypothesis does not depend on the choice of bandwidth. This happens in part because the bandwidth is a quantity that vanishes in the limit. However, for a given finite sample size, the computed value of the test statistic critically depends on the bandwidth selection (see Maasoumi and Racine (2008)). This raises strong concerns about using the asymptotic distribution to conduct statistical inference in empirical applications. Therefore, following the suggestions of Hong and White (2005) and many others, rather than rely on an asymptotic distribution, we use the bootstrap resampling approach to determine the empirical null distribution of $\hat{S}_{\rho}(c)$ (for further discussions on the bootstrap resampling approach, see Efron (1982), Hall (1992), and Horowitz (2001)).[7]

III. Size and Power of the Entropy Test

In this section, we use Monte Carlo simulations based on copula–generalized autoregressive conditional heteroskedasticity (GARCH) models to examine the size and power of the entropy test and show that it has reliable size and fairly high power in finite samples.

A. Modeling Dependence with Copulas

Because we are testing the joint distribution of two random variables, the simulation requires generating random samples from a joint distribution with a certain dependence structure. The copula is a widely used method to model the complete dependence structure between random variables.

As documented in the literature (see, e.g., Ang and Chen (2002), Ang et al. (2006)), stock returns usually show stronger downside comovement than the upside comovement with the market return. A Clayton copula that features stronger left-tail dependence is a natural candidate for the data-generating process (DGP). It has the functional form given by

$$C_{\text{clay}}(u, v;\, \tau) \quad = \quad (u^{-\tau} + v^{-\tau} - 1)^{-1/\tau},$$

where $\tau > 0$ governs the dependence between the marginal distributions. A higher τ indicates stronger left-tail dependence.

A bivariate Gaussian copula features symmetric dependence and is defined as

$$C_{\text{nor}}(u, v;\, \rho) \quad = \quad \Phi_{\rho}(\Phi^{-1}(u), \Phi^{-1}(v)),$$

where ρ is the correlation coefficient between the marginal distributions, Φ^{-1} is the inverse of the standard normal cumulative distribution function, and Φ_{ρ} is the standard bivariate normal distribution function with correlation ρ. To best fit our return data, we maximize flexibility by using a mixed Gaussian–Clayton copula as the DGP,[8] which is defined as

$$(13) \qquad C_{\text{mix}}(u, v;\, \rho, \tau, \kappa) \quad = \quad \kappa C_{\text{nor}}(u, v;\, \rho) + (1-\kappa) C_{\text{clay}}(u, v;\, \tau),$$

where κ indicates the weight on the bivariate normal (Gaussian) copula. The mixed copula nests both the Gaussian and Clayton copulas as special cases:

[7]See Section IA.I of the Internet Appendix (available at http://apps.olin.wustl.edu/faculty/zhou/Appendix_Co-Asymmetry.pdf) for a discussion of the choice of bootstrap resampling methods suitable for the entropy test of asymmetry and detailed procedures for implementing the resampling method.

[8]Tawn (1988) proves that every convex combination of existing copula functions is still a copula.

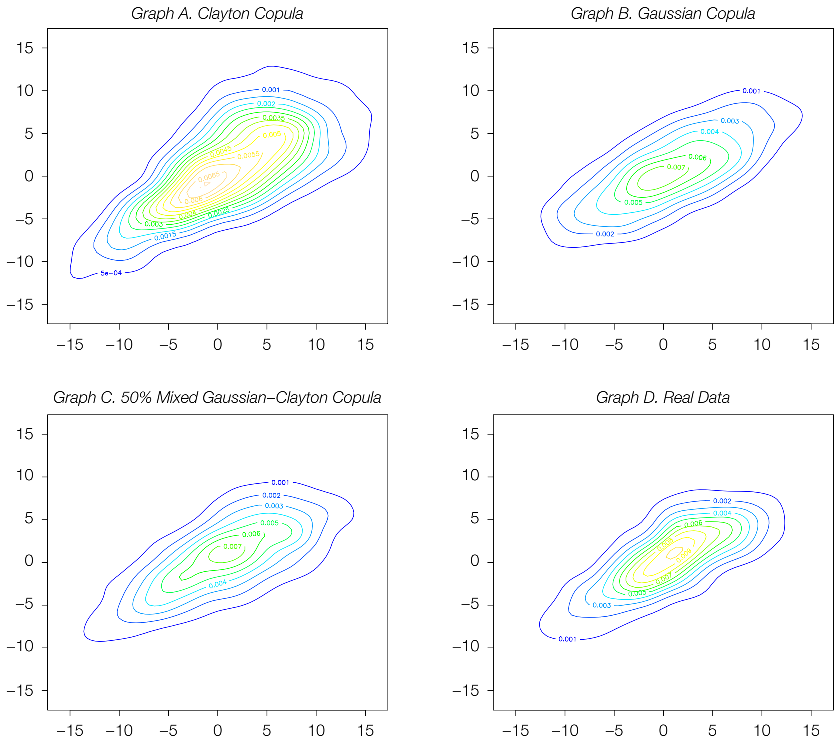
When $\kappa = 1$, the mixed copula reduces to a Gaussian copula; when $\kappa = 0$, the mixed copula reduces to a Clayton copula. In the simulations that follow, we use κ values of 0, 0.25, 0.375, 0.5, 0.75, and 1 to generate random samples with different levels of asymmetric comovement.

Figure 2 shows the contour plots of random samples simulated using copula–GARCH models, which are described in detail in the following section. The model parameters are fitted to the value-weighted monthly returns of the smallest size portfolio and the market portfolio using the maximum likelihood method. The samples are generated by assuming Clayton, Gaussian, and equal-weighted mixed Gaussian–Clayton copula dependence structures, respectively. The contour plots of simulated data (Graphs A–C) are plotted along with the actual return data (Graph D). It is clear that the Clayton copula generates data with a stronger lower-tail dependence because the plots are highly concentrated in the lower left tail. The smallest size portfolio is known to have the strongest asymmetric correlation, but the data plots generated by the pure Clayton copula seem to have stronger

FIGURE 2

Contour Plots of Copula Dependence Structures

Figure 2 shows contour plots of random samples simulated using copula–GARCH models. The value-weighted returns of the smallest size portfolio and the market portfolio are used as the base asset to fit parameters for the data-generating process. Graph A shows a Clayton copula, Graph B a Gaussian copula, and Graph C a mixed Gaussian–Clayton copula with mixing weights of 0.5 each. Graph D shows the actual data of the value-weighted returns of the smallest size portfolio and the market. The returns are in percentage points, with the horizontal axis denoting the size portfolio returns and the vertical axis denoting the market returns.



lower-tail dependence than the actual data. As shown in Graph C, the scatter plots generated by the equal-weighted mixed Gaussian–Clayton copula look most similar to the actual data plots in Graph D, which justifies our decision to use mixed copulas as the DGP in the simulations.

B. Simulation with Copula–GARCH Models

To ensure that the random samples drawn from the mixed Gaussian–Clayton copula model are as close to the actual data as possible, we must specify a model that best mimics the marginal distribution of the asset returns. We model an asset return as the sum of an expected return component and a residual following a GARCH(1,1) process, a parsimonious way to model conditional heteroscedasticity. In the implementation, we first fit the data to the copula–GARCH model using the maximum likelihood method. We then plug the maximum likelihood estimates back into the model and use them as the true parameters to generate data to compute the simulated p-values. To be conservative, instead of using portfolios that clearly show asymmetric comovement, such as the smallest size or momentum portfolios, we use the value-weighted returns of the fourth-largest size portfolio[9] and the market portfolio to estimate the copula and GARCH parameters. The maximum likelihood estimates of the Clayton copula parameter τ and Gaussian copula parameter ρ are 5.768 and 0.951, respectively. The estimated GARCH parameters are reported in Table IA.1 of the Internet Appendix. The results show that all the GARCH parameters are statistically significant at the 5% level, indicating that the GARCH(1,1) is indeed a good description of the return data. It is of interest to see whether the entropy test has reasonable power under the parameter settings.[10]

In summary, we simulate the data as follows:

(1) For a given κ, draw a bivariate uniform random sample of size T from the mixed Gaussian–Clayton copula model.

(2) Apply an inverse standard normal cumulative distribution function transformation to obtain a bivariate standard normal random sample with a prespecified dependence structure.

(3) Feed each series of the joint normal random sample into the univariate GARCH(1,1) process as the innovation terms to generate a simulated joint return series.[11]

[9]Based on the entropy test results in Section IV, the fourth-largest size portfolio is the smallest size portfolio that cannot reject the null of symmetry at the 5% significance level.

[10]We also check the robustness of the test power using the threshold GARCH(1,1) (TGARCH) model as the underlying DGP for the marginal distributions. The estimated TGARCH parameters are reported in Table IA.2 of the Internet Appendix. The test results, reported in Tables IA.5, IA.6, and IA.7 in the Internet Appendix, are qualitatively similar. The asymmetry of the joint distribution is mainly governed by the copula function. Therefore, the leverage effect in the conditional heteroscedasticity of the marginal distributions has little effect on the joint asymmetry.

[11]This ensures that the simulated data vectors each follow a GARCH(1,1) process and that the desired dependence structure is governed by the mixed copula model.

(4) Repeat steps 1–3 1,000 times.

(5) Repeat steps 1–4 with different sample sizes $T = \{240, 420, 600, 840\}$.

Our choice of sample sizes represents typical data lengths encountered in empirical studies: $T = 240$ represents 20 years of monthly frequency data, $T = 420$ is the length of the subsample data period used by HTZ (2007), $T = 600$ represents 50 years of monthly frequency data and is close to the full sample data length ($T = 588$) used in this article, and $T = 840$ represents 70 years of monthly frequency data. In the simulation, we use one fixed bandwidth for every 1,000 random samples generated from the same DGP. The fixed bandwidth is set to the average of the 1,000 bandwidths computed for each of the 1,000 random samples via the likelihood cross-validation. The selection of optimal block length is similar in that the expected block length for every 1,000 random samples generated from the same DGP is fixed to be the average of the 1,000 optimal block lengths computed using Patton, Politis, and White's (2009) algorithm. Averaging the bandwidths and block lengths across random samples drawn from the same DGP can potentially reduce sampling randomness and make the simulation results more stable.[12]

Table 1 reports the empirical size and power for both tests based on 1,000 Monte Carlo simulations in which the nominal size is set to 5%. Powers are reported for different DGPs with different asymmetric comovement levels (from $\kappa = 1$ to $\kappa = 0$) and at various sample sizes. Following HTZ (2007), we focus on exceedance levels $\{0\}$ and $\{0, 0.5, 1, 1.5\}$. For comparison, we also report the size and power of the HTZ (2007) test. When the underlying DGP is symmetrically distributed with $\kappa = 1$, the true probability of rejection is the size of the tests. The empirical sizes of both tests are reported in Panel A. The empirical size of the entropy test at $\{c = 0\}$ is fairly close to 5% and approaches the true nominal size as the sample size increases. When the sample size is 840, the empirical size is 0.055, which is very close to 5%, the nominal size. The joint entropy test at $c = \{0, 0.5, 1, 1.5\}$ has an empirical size of 0.047 with a sample size of 840, which is also close to the nominal size. We find that the HTZ test shows a tendency toward underrejection at both the single and joint exceedance levels, with empirical sizes too close to 0 in many cases. Similar bootstrap procedures can improve the HTZ test so that its size is as reliable as that of the entropy test, but its power, on average, remains comparatively low.

We next evaluate the empirical powers of both tests. We find that, in most cases, the single test has higher power than the joint test does; this is the case for both the HTZ (2007) and entropy tests. The results indicate that the entropy test, in most cases, has higher average rejection rates than the HTZ test at both the single and joint exceedance levels. The difference in power varies with the dependence structure of the DGP. When the simulated data are strongly asymmetric, both tests perform similarly. If the DGP is a bivariate Clayton copula ($\kappa = 0$) with an exceedance level of $c = 0$, the difference in power is not large (0.211 for $T = 240$) and vanishes as the sample size increases to

[12] We also use the optimal bandwidth and block length for each simulated sample without averaging and obtain similar results.

TABLE 1

Size and Power: Entropy Test and HTZ Test

Table 1 reports the rejection rates for the null hypothesis of symmetric comovement based on 1,000 Monte Carlo simulations, where the nominal size of the tests is set to 5%. When $\kappa=100\%$, the data-generating process (DGP) is a joint normal distribution, and the rejection rates are the empirical sizes. In all other cases, the rejection rates reflect empirical power. The Clayton copula parameter $\tau=5.768$, and the normal copula parameter $\rho=0.951$. The specification for the marginal distribution is a standard GARCH(1,1): $r_{i,t}=\mu_i+\varepsilon_{i,t}$, where $\varepsilon_{i,t}$ is normally distributed with a time-varying variance $\sigma_{i,t}^2=\omega_i+\alpha_i\varepsilon_{i,t-1}^2+\beta_i\sigma_{i,t-1}^2$.

	Entropy Test		HTZ Test	
	$c=\{0\}$	$c=\{0,0.5,1,1.5\}$	$c=\{0\}$	$c=\{0,0.5,1,1.5\}$
Panel A. $\kappa=100\%$ (SIZE)				
$T=240$	0.032	0.025	0.000	0.005
$T=420$	0.031	0.037	0.000	0.000
$T=600$	0.044	0.049	0.000	0.000
$T=840$	0.055	0.058	0.000	0.000
Panel B. $\kappa=75\%$				
$T=240$	0.065	0.058	0.000	0.022
$T=420$	0.120	0.112	0.001	0.008
$T=600$	0.233	0.186	0.002	0.003
$T=840$	0.430	0.314	0.004	0.006
Panel C. $\kappa=50\%$				
$T=240$	0.278	0.251	0.059	0.099
$T=420$	0.679	0.555	0.152	0.113
$T=600$	0.912	0.803	0.283	0.129
$T=840$	0.991	0.955	0.482	0.188
Panel D. $\kappa=37.5\%$				
$T=240$	0.513	0.460	0.169	0.185
$T=420$	0.897	0.826	0.369	0.219
$T=600$	0.984	0.951	0.584	0.288
$T=840$	0.999	0.995	0.778	0.434
Panel E. $\kappa=25\%$				
$T=240$	0.733	0.695	0.382	0.314
$T=420$	0.978	0.958	0.688	0.450
$T=600$	1.000	0.997	0.855	0.580
$T=840$	1.000	1.000	0.962	0.760
Panel F. $\kappa=0\%$				
$T=240$	0.970	0.957	0.759	0.657
$T=420$	1.000	0.999	0.953	0.838
$T=600$	1.000	1.000	0.987	0.928
$T=840$	1.000	1.000	0.997	0.978

$T=840$. However, the power difference becomes most pronounced when the degree of asymmetry is not very strong. For example, when the DGP is a 50% mixed Gaussian–Clayton copula, the power of the entropy test is about twice as high as the power of the HTZ test, even when the sample size is 840. In general, both tests increase in power as the sample size increases.

Our results are qualitatively similar to those in Table 1 if the nominal size is set at 1% or 10% (see Tables IA.3 and IA.4 of the Internet Appendix for details). This fact reaffirms that the proposed entropy test generally has the correct size and higher power than the HTZ (2007) test. There are two possible reasons for this. First, the entropy measure utilizes more information from the joint distribution, giving it an information advantage over the HTZ test, which uses only the second moment. Second, we rely on the bootstrap resampling method for inference, so the null distribution of the test statistics is potentially more accurate than the asymptotic distribution.

IV. Is Asymmetry Rare?

In this section, we apply our entropy measure to test whether statistically significant asymmetry exists in common portfolios sorted by size, book-to-market ratio, and momentum. Previous studies have rarely found asymmetry; however, as we show in the following discussion, the entropy test shows very different results.

A. Test Portfolios

Following prior studies on asymmetric correlation, we consider portfolios of stocks sorted by market capitalization, book-to-market ratio, and momentum. We use value-weighted returns for all portfolios. The return on the Center for Research in Security Prices (CRSP) value-weighted market index, based on stocks listed on the New York Stock Exchange (NYSE)/American Stock Exchange (AMEX)/National Association of Securities Dealers Automated Quotations (NASDAQ), is used as a proxy for the market return. All returns are calculated at a monthly frequency and in excess of the risk-free rate, which is defined as the 1-month T-bill rate. All of these data are available on Kenneth French's Web site.[13] The sample period is from Jan. 1965 to Dec. 2013 (588 observations in total).[14]

B. Test Results

Table 2 shows the asymmetry test results based on both the HTZ (2007) test and the entropy test with the exceedance levels of $\{0\}$ and $\{0, 0.5, 1, 1.5\}$. Panel A provides the test results for the size portfolios. At the usual 5% level, the entropy test at the single exceedance level $c = \{0\}$ rejects symmetry for the six smallest size portfolios. When testing at the multiple exceedance levels $c = \{0, 0.5, 1, 1.5\}$, the entropy test rejects symmetry for the same six size portfolios at the 10% level. In contrast, the HTZ test can only reject symmetry for the smallest size portfolio, regardless of whether we use single or multiple exceedance levels and whether a 5% or 10% significance level is used. It is interesting to note that the entropy test statistics generally decrease as firm size increases; a similar pattern also holds for the HTZ test statistics. This finding is consistent with the work of Ang and Chen (2002).

For comparison, the last two columns in Panel A of Table 2 also report conventional moment-based statistics that describe asymmetry: the skewness and coskewness. The skewness does not show any pattern across size portfolios, whereas the coskewness is generally larger in absolute value for smaller stocks. Clearly, these moment-based measures alone do not provide enough information to draw any conclusions regarding the relationship between asymmetry in stock returns and firm size. In contrast, our entropy test provides evidence of strong asymmetry for the six smallest size portfolios.

Panel B of Table 2 reports the test results for book-to-market portfolios. Unlike the size portfolios, there are no apparent patterns for the entropy statistic

[13]We are grateful to Kenneth French for making the data available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[14]We choose the starting date to be consistent with HTZ (2007). The results are similar when using Jan. 1963 as the starting date. We also report the test results using exactly the same sample period as in HTZ (2007) (Jan. 1965 to Dec. 1999) in Table IA.8 of the Internet Appendix.

TABLE 2
Testing for Asymmetry

Table 2 reports the HTZ (2007) test statistic (J_ρ) and the entropy test statistic along with their associated p-values. The last two columns report skewness and coskewness. The test assets are the value-weighted size, book-to-market, and momentum portfolios. The data are monthly, and the sample period is from Jan. 1965 to Dec. 2013.

	Entropy Test				HTZ Test					
	$c=\{0\}$		$c=\{0, 0.5, 1, 1.5\}$		$c=\{0\}$		$c=\{0, 0.5, 1, 1.5\}$			
Portfolios	$S_\rho \times 100$	p-Value	$S_\rho \times 100$	p-Value	J_ρ	p-Value	J_ρ	p-Value	Skewness	Coskewness
Panel A. Size										
Size 1	2.027	0.010	1.094	0.013	4.212	0.040	9.715	0.046	−0.167	−0.568
Size 2	1.963	0.000	1.191	0.010	2.049	0.152	3.281	0.512	−0.262	−0.543
Size 3	1.868	0.020	1.178	0.028	0.937	0.333	1.108	0.893	−0.448	−0.550
Size 4	1.689	0.013	1.040	0.073	0.613	0.434	2.095	0.718	−0.510	−0.555
Size 5	1.690	0.030	1.060	0.060	0.431	0.512	5.015	0.286	−0.524	−0.556
Size 6	1.596	0.045	1.081	0.073	0.234	0.629	3.134	0.536	−0.540	−0.543
Size 7	1.477	0.065	1.020	0.125	0.092	0.761	0.849	0.932	−0.497	−0.521
Size 8	1.510	0.085	1.046	0.125	0.099	0.753	0.146	0.997	−0.470	−0.514
Size 9	1.695	0.075	1.098	0.160	0.005	0.945	0.030	1.000	−0.450	−0.485
Size 10	1.511	0.055	1.025	0.133	0.008	0.930	0.029	1.000	−0.347	−0.461
Panel B. Book-to-Market										
BM 1	1.248	0.115	0.817	0.226	0.023	0.880	0.341	0.987	−0.230	−0.412
BM 2	1.208	0.085	0.871	0.168	0.024	0.876	0.093	0.999	−0.458	−0.493
BM 3	1.003	0.263	0.670	0.414	0.060	0.807	0.066	0.999	−0.527	−0.515
BM 4	1.626	0.138	1.065	0.120	0.064	0.800	1.829	0.767	−0.474	−0.504
BM 5	1.610	0.055	0.959	0.180	0.145	0.703	2.769	0.597	−0.427	−0.497
BM 6	1.815	0.025	1.243	0.073	0.054	0.817	1.099	0.894	−0.432	−0.487
BM 7	1.805	0.058	1.211	0.060	0.082	0.774	0.590	0.964	−0.137	−0.375
BM 8	1.571	0.098	1.084	0.135	0.226	0.634	2.954	0.566	−0.464	−0.459
BM 9	2.162	0.005	1.438	0.015	0.447	0.504	1.667	0.797	−0.325	−0.473
BM 10	1.657	0.075	1.155	0.083	0.805	0.370	2.133	0.711	0.069	−0.439
Panel C. Momentum										
L	2.758	0.008	1.791	0.003	1.188	0.276	7.831	0.098	0.663	−0.324
2	1.331	0.170	0.811	0.361	0.309	0.578	0.980	0.913	0.224	−0.326
3	1.345	0.080	0.910	0.128	0.070	0.791	0.907	0.924	0.317	−0.293
4	1.974	0.093	1.334	0.153	0.070	0.791	1.009	0.908	−0.127	−0.392
5	1.498	0.053	0.987	0.095	0.069	0.793	3.081	0.544	−0.278	−0.453
6	2.059	0.013	1.497	0.010	0.242	0.623	1.924	0.750	−0.408	−0.531
7	1.614	0.010	1.125	0.040	0.048	0.826	0.274	0.991	−0.482	−0.500
8	1.487	0.068	1.056	0.100	0.167	0.683	0.388	0.983	−0.360	−0.471
9	2.297	0.000	1.446	0.005	0.275	0.600	3.467	0.483	−0.590	−0.558
W	2.221	0.003	1.373	0.005	0.749	0.387	5.384	0.250	−0.431	−0.490

across stocks with different book-to-market ratios. The same holds true for the HTZ (2007) test statistic, skewness, and coskewness. At the 10% level, the entropy test at $c=\{0\}$ rejects 7 out of 10 book-to-market portfolios, and the joint entropy test rejects 4 portfolios, whereas the HTZ tests reject none.

Panel C of Table 2 reports the test results for the momentum portfolios. The single entropy test finds statistically significant asymmetry in the returns of both the past loser and winner portfolios at the 1% level, and the joint entropy test gives similar testing results. Significant asymmetry in the past loser portfolio is consistent with the results of Ang and Chen (2002), who show that the bivariate normal model is rejected when fitted to the past loser portfolio returns; in other words, the returns of the past loser portfolios exhibit asymmetric correlations. In contrast, the HTZ (2007) test detects hardly any significant asymmetric correlation among the momentum portfolios at the 5% level. Only the past loser portfolio shows marginal significance at the 10% level, with a p-value of 0.098. We find that the entropy test statistic S_ρ is largest at the lower and higher ends and smaller in the middle deciles. The J_ρ statistic in the HTZ test follows a similar pattern. Finally, no obvious pattern of asymmetry can be inferred from skewness or coskewness.

V. Asset Pricing and Asymmetric Comovement

What are the asset-pricing implications of the asymmetric comovement of individual stocks with the market return that we found in many portfolios? In this section, we introduce a downside asymmetric comovement measure based on the entropy test statistic and examine its effect on the cross-sectional variations of expected returns.

A. Downside Asymmetric Comovement

We have shown that the entropy test statistic $S_\rho(c)$ defined in equation (7) can successfully capture the degree of asymmetric comovement in the stock market because the entropy test demonstrates good finite sample power. $S_\rho(c)$ is a normalized metric that is always nonnegative and therefore does not indicate whether stock comovement with the market is stronger in the downside or upside. However, investors are generally more concerned about downside risk and require higher premiums for holding assets with higher downside risk. Similar to the uncertainty measure of Segal, Shaliastovich, and Yaron (2015), we also need a measure that distinguishes the direction of asymmetry.

We define a lower-quadrant probability (LQP) and an upper-quadrant probability (UQP) as follows:

$$\text{(14)} \qquad \text{LQP} = \Pr(\tilde{x} \leq -c, \tilde{y} \leq -c) = \int_{-\infty}^{-c}\int_{-\infty}^{-c} f(\tilde{x},\tilde{y})\,d\tilde{x}\,d\tilde{y},$$

$$\text{(15)} \qquad \text{UQP} = \Pr(\tilde{x} \geq c, \tilde{y} \geq c) = \int_{c}^{+\infty}\int_{c}^{+\infty} f(\tilde{x},\tilde{y})\,d\tilde{x}\,d\tilde{y}.$$

These variables measure the concentration of individual stock and market return pairs being either above or below the exceedance level c. When $c=0$, a higher LQP (UQP) indicates a higher tendency for the stock to comove with the market below (above) average levels, thus capturing the downside (upside) comovement of individual stock returns with the market. We define our downside asymmetric comovement (DOWN_ASY) as

$$\text{(16)} \qquad \text{DOWN\_ASY} = \text{sign}(\text{LQP} - \text{UQP})S_\rho(c),$$

where $\text{sign}(\cdot)$ is a sign function that takes the value of 1 if the difference between LQP and UQP is nonnegative, and -1 otherwise. Higher DOWN_ASY indicates greater downside risk relative to upside potential.

To see the intuition behind DOWN_ASY, consider the case where the stochastic discount factor (SDF) of an asset-pricing model is only a function of the market return that is given by $g(R_{m,t+1})$, of which the capital asset pricing model (CAPM) is a special case. Let p_t be the price of an asset today, and let x_{t+1} be its future payoff in the next period. Then p_t is the discounted expected payoff,

$$p_t = E_t[g(R_{m,t+1})x_{t+1}].$$

We assume that the SDF is positive to eliminate arbitrage (Hansen and Jagannathan (1991)). It is then clear that $g(R_{m,t+1})x_{t+1}$ is negative in the third quadrant. Holding the properties of the asset returns constant outside the exceedance regions, then the more density mass in the third quadrant, the lower the current price. If the mass of the density in the third quadrant is greater than the mass in the first quadrant, then the asset return is asymmetric to the downside. In this case, DOWN_ASY lowers the current price, implying in general a risk premium for the downside asymmetry risk. For example, for a class of utility functions, Gul (1991) and Ang et al. (2005) rationalize that investors weigh payoffs below a certain reference point more heavily than those above it. This suggests that the downside asymmetry risk premium from their models can be higher than that from a typical model with constant relative risk aversion (CRRA) utility.

B. Downside Asymmetric Comovement and Firm Characteristics

We construct the comovement measures LQP, UQP, and DOWN_ASY, as well as downside beta β^-, upside beta β^+, coskewness, and cokurtosis, based on CRSP data from Jan. 1962 to Dec. 2013. The Appendix provides detailed formulas for these variables. The estimation is carried out using realized daily return data from overlapping 12-month periods. We drop any year with fewer than 100 daily return observations for each stock. The estimates are then updated monthly.

Table 3 reports the time-series averages of the cross-sectional correlations of the lower-quadrant probability, the upper-quadrant probability, downside asymmetric comovement (DOWN_ASY), the CAPM beta (β), downside beta (β^-), upside beta (β^+), the natural log of market capitalization (SIZE), the natural log of the book-to-market ratio (BM), the turnover ratio (TURN), the normalized Amihud illiquidity measure (ILLIQ), the past-6-month return (MOM), idiosyncratic volatility (IVOL), coskewness (COSKEW), cokurtosis (COKURT), and the maximum daily return over the past month (MAX). It is interesting to note that the correlation between LQP and UQP is fairly modest (−0.080), which means that the tendencies of the stock to comove with the market in the upside or downside are almost linearly independent. This finding justifies our approach of separately estimating the lower and upper quadrant probabilities.

Both LQP and UQP are positively correlated with the CAPM beta with correlation coefficients of 0.463 and 0.436, respectively. β captures the linear dependence between an individual stock return and the market return, and quadrant probabilities measure the tendency of the returns to occur above (below) the sample averages. Stocks with a higher β will have a higher probability of being above (below) their sample means when the market is above (below) average. LQP and UQP have little correlation with coskewness. However, they have high positive correlations with the cokurtosis because cokurtosis measures the fatness of the tails in a given joint distribution, and a fatter tail indicates a higher probability in that quadrant. DOWN_ASY has little correlation with firm characteristics such as size and book-to-market ratio, which indicates that DOWN_ASY is distinct from these well-known measures and provides additional new information.

In Table 4 we sort stocks into deciles based on their realized DOWN_ASY values each month. We report the time-series averages, which range from −0.065 to 0.107, in the second column. None of the other firm characteristics

TABLE 3
Correlations

Table 3 reports the the time-series averages of the cross-sectional correlations of various risk measures and firm characteristics: lower-quadrant probability (LQP), upper-quadrant probability (UQP), downside asymmetric comovement (DOWN_ASY), CAPM beta (β), downside beta (β^-), upside beta (β^+), natural log of market capitalization (SIZE), natural log of book-to-market ratio (BM), turnover ratio (TURN), normalized Amihud illiquidity measure (ILLIQ), past-6-month return (MOM), idiosyncratic volatility (IVOL), coskewness (COSKEW), cokurtosis (COKURT), and maximum daily return over the past 1 month (MAX). The sample covers all U.S. common stocks listed on the New York Stock Exchange (NYSE)/American Stock Exchange (AMEX)/National Association of Securities Dealers Automated Quotations (NASDAQ), and the sample period is from Jan. 1962 to Dec. 2013.

	LQP	UQP	DOWN_ASY	β	β^-	β^+	SIZE	BM	TURN	ILLIQ	MOM	IVOL	COSKEW	COKURT	MAX
LQP	1.000														
UQP	−0.080	1.000													
DOWN_ASY	0.462	−0.515	1.000												
β	0.463	0.436	−0.029	1.000											
β^-	0.300	0.273	−0.036	0.800	1.000										
β^+	0.356	0.340	−0.011	0.797	0.528	1.000									
SIZE	0.299	0.473	−0.079	0.222	0.097	0.233	1.000								
BM	−0.071	−0.207	0.066	−0.259	−0.212	−0.202	−0.280	1.000							
TURN	0.164	0.200	−0.032	0.451	0.382	0.331	0.109	−0.201	1.000						
ILLIQ	−0.297	−0.405	0.048	−0.337	−0.233	−0.298	−0.637	0.256	−0.290	1.000					
MOM	0.009	0.051	−0.019	0.114	0.119	0.072	0.012	−0.237	0.153	−0.061	1.000				
IVOL	−0.091	−0.203	0.029	0.308	0.336	0.170	−0.525	−0.058	0.232	0.357	0.042	1.000			
COSKEW	0.033	0.024	0.038	0.013	−0.343	0.388	0.038	−0.010	−0.025	0.010	−0.040	0.007	1.000		
COKURT	0.453	0.497	−0.059	0.687	0.577	0.647	0.523	−0.203	0.239	−0.472	0.075	−0.234	0.013	1.000	
MAX	0.010	−0.041	0.009	0.277	0.273	0.178	−0.259	−0.100	0.335	0.129	0.024	0.498	−0.001	−0.021	1.000

TABLE 4

Summary Statistics of Asymmetry Portfolios

Table 4 reports various risk measures and firm characteristics of deciles sorted based on their realized DOWN_ASY each month: downside asymmetric comovement (DOWN_ASY), CAPM beta (β), downside beta (β^-), upside beta (β^+), natural log of market capitalization (SIZE), natural log of book-to-market ratio (BM), turnover ratio (TURN), normalized Amihud illiquidity measure (ILLIQ), past-6-month return (MOM), idiosyncratic volatility (IVOL), coskewness (COSKEW), cokurtosis (COKURT), and maximum daily return over the past 1 month (MAX). The sample covers all U.S. common stocks listed on the New York Stock Exchange (NYSE)/American Stock Exchange (AMEX)/National Association of Securities Dealers Automated Quotations (NASDAQ), and the sample period is from Jan. 1962 to Dec. 2013.

Decile	DOWN_ASY	β	β^-	β^+	$\beta^- - \beta^+$	SIZE	BM	TURN	ILLIQ	MOM	IVOL	COSKEW	COKURT	MAX
1 (Low)	−0.065	0.655	0.779	0.536	0.243	4.791	−0.437	0.057	8.839	0.076	0.025	−0.084	1.270	0.053
2	−0.023	0.876	0.982	0.782	0.200	5.553	−0.620	0.071	4.704	0.102	0.023	−0.093	1.848	0.051
3	−0.009	0.984	1.076	0.921	0.155	5.851	−0.650	0.074	3.426	0.114	0.023	−0.091	2.284	0.050
4	0.001	1.006	1.097	0.952	0.145	5.830	−0.637	0.075	3.352	0.113	0.023	−0.091	2.368	0.050
5	0.009	0.996	1.088	0.941	0.147	5.706	−0.618	0.075	3.633	0.111	0.023	−0.098	2.284	0.051
6	0.015	0.988	1.082	0.927	0.155	5.557	−0.595	0.074	4.115	0.107	0.024	−0.101	2.227	0.053
7	0.022	0.965	1.061	0.895	0.165	5.342	−0.558	0.072	4.902	0.101	0.025	−0.098	2.083	0.054
8	0.031	0.928	1.033	0.843	0.191	5.051	−0.510	0.070	6.051	0.097	0.026	−0.092	1.851	0.056
9	0.047	0.849	0.958	0.747	0.211	4.731	−0.428	0.064	7.667	0.083	0.027	−0.087	1.573	0.057
10 (High)	0.107	0.677	0.775	0.570	0.205	4.557	−0.298	0.052	9.466	0.063	0.026	−0.075	1.245	0.053

displays a similar monotonic pattern. This result reaffirms the findings in Table 3 that DOWN_ASY is not similar to other firm characteristics. Therefore, if DOWN_ASY explains the cross section of asset returns, then its impact on asset pricing is different from the effects of these known firm characteristics documented in the literature.

C. Comparison with Other Asymmetry Measures

We sort stocks into quintiles based on their realized asymmetry measures estimated over the previous 12 months. We then compute the equal-weighted average returns on these portfolios. Because we need the first 12 months of data to estimate the measures, the quintile portfolio returns are calculated from Jan. 1963 to Dec. 2013.

Panel A of Table 5 reports the univariate portfolio sorting results based on DOWN_ASY and other conventional asymmetry measures. The average returns of the DOWN_ASY portfolios increase from 0.65% to 1.03% per month. The trading strategy of investing in high-DOWN_ASY stocks and shorting low-DOWN_ASY stocks therefore yields an economically significant 1-month return of 0.38%, which amounts to a return premium of 4.5% per annum. The return difference is also statistically significant at the 1% level. To attribute this excess return to exposure to previously identified systematic risk factors, we regress the return series of each quintile portfolio and the spread portfolio on Carhart's (1997) 4 factors.

TABLE 5

Portfolios Sorted by Asymmetry Measures

Table 5 reports the equal-weighted average returns and alphas, in percentage points, for stock portfolios sorted by different realized asymmetry measures. DOWN_ASY is the downside asymmetry measure, DOWN_CORR is the downside asymmetric correlation, and β^{-} is the downside beta. BBC and CC co-entropies are based on Backus et al. (2018) and Chabi-Yo and Colacito (2016), respectively. The row labeled "High − Low" reports the difference between the returns/alphas of portfolios 5 and 1 and the associated t-statistics (reported in parentheses). The sample covers all U.S. common stocks listed on the New York Stock Exchange (NYSE)/American Stock Exchange (AMEX)/National Association of Securities Dealers Automated Quotations (NASDAQ) for the period from Jan. 1963 to Dec. 2013. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Downside Asymmetry and Conventional Asymmetry Measures

	DOWN_ASY		DOWN_CORR		β^{-}		COSKEW	
Portfolio	Return	Carhart Alpha	Return	Carhart Alpha	Return	Carhart Alpha	Return	Carhart Alpha
1 (Low)	0.652	0.045	0.691	0.107	0.772	0.181	0.956	0.126
2	0.746	0.077	0.775	0.113	0.837	0.174	0.948	0.190
3	0.849	0.138	0.891	0.167	0.886	0.157	0.863	0.155
4	0.943	0.177	0.902	0.135	0.909	0.133	0.774	0.092
5 (High)	1.027	0.225	0.958	0.142	0.817	0.022	0.680	0.103
High − Low	0.375***	0.180**	0.267***	0.035	0.045	−0.159	−0.276***	−0.023
	(4.39)	(2.36)	(3.48)	(0.59)	(0.21)	(−1.32)	(−2.70)	(−0.25)

Panel B. Alternative Entropy Measures

	BBC Co-Entropy		CC Co-Entropy	
Portfolio	Return	Carhart Alpha	Return	Carhart Alpha
1 (Low)	0.679	0.094	0.682	0.097
2	0.776	0.100	0.778	0.102
3	0.884	0.127	0.884	0.129
4	0.912	0.134	0.908	0.126
5 (High)	0.949	0.175	0.948	0.176
High − Low	0.269***	0.081	0.266***	0.079
	(4.34)	(1.54)	(4.23)	(1.49)

The abnormal return of the spread portfolio after adjusting for the Carhart 4 factors is 0.18% per month (2.2% per annum). It is statistically significant at the 5% level and economically meaningful.

For comparison, we examine the asset-pricing implications of other conventional asymmetry measures. Motivated by the HTZ test statistic, we first consider an asymmetric correlation measure based on the exceedance correlations. Because the test statistic is always positive and does not indicate whether an excessive exceedance correlation exists in the downside or in the upside, we may simply define the downside asymmetric correlation, denoted DOWN_CORR, as the difference between the downside and upside exceedance correlations,

$$\text{DOWN\_CORR} = \rho^{-}(c) - \rho^{+}(c), \tag{17}$$

where c is the exceedance level. DOWN_CORR takes positive values if the exceedance correlation in the downside tail is greater than that in the upside tail, and it takes negative or 0 values otherwise.

We report the results for the quintile portfolios sorted by DOWN_CORR ($c=0$). Similar to DOWN_ASY, the greater the downside asymmetric correlation, the greater the expected return. However, the return difference is lower, 0.27% per month or 3.24% per annum, which is approximately 29% less than the spread return of DOWN_ASY portfolios. Moreover, this return difference can be fully explained by the Carhart 4 factors. Similarly, although there are average return differences between stocks in the highest and lowest downside beta (Ang et al. (2006)) and coskewness (Harvey and Siddique (2000)) quintiles, these differences can be fully explained by the Carhart 4 factors.

Finally, we compare DOWN_ASY with other downside asymmetry measures based on the co-entropies proposed by Backus et al. (2018) and Chabi-Yo and Colacito (2016). This is an important robustness check. If there is a genuine asymmetry risk premium, these measures should detect that premium as well. To apply these alternative co-entropy measures to our data, we need to first compute exceedance co-entropies, which are simply the co-entropies evaluated in the exceedance regions. Specifically, based on the co-entropy of Backus et al. (2018), we compute the exceedance co-entropies as

$$C^{+}_{\text{BBC}}(c,x,y) = \left.\frac{L(xy)-L(x)-L(y)}{L(x)+L(y)}\right|(x>c, y>c), \tag{18}$$

$$C^{-}_{\text{BBC}}(c,x,y) = \left.\frac{L(xy)-L(x)-L(y)}{L(x)+L(y)}\right|(x<-c,\ y<-c); \tag{19}$$

and, based on the definition of the co-entropy given by Chabi-Yo and Colacito (2016), we compute another set of exceedance co-entropies as

$$C^{+}_{\text{CC}}(c,x,y) = \left.\frac{L(x)+L(y)-L(x/y)}{L(x)+L(y)}\right|(x>c, y>c), \tag{20}$$

$$C^{-}_{\text{CC}}(c,x,y) = \left.\frac{L(x)+L(y)-L(x/y)}{L(x)+L(y)}\right|(x<-c, y<-c), \tag{21}$$

where $L(x)=\ln E(x)-E(\ln x)$. Then, $C^{-}_{\text{BBC}}(c,x,y)-C^{+}_{\text{BBC}}(c,x,y)$ and $C^{-}_{\text{CC}}(c,x,y)-C^{+}_{\text{CC}}(c,x,y)$ can both serve as alternative downside asymmetry measures. These two measures are computationally much easier to implement than

DOWN_ASY because they are defined based on expectations that can be consistently estimated by sample means.

Panel B of Table 5 reports the portfolio returns when stocks are sorted by the downside asymmetry measures based on the co-entropies of Backus et al. (2018) and Chabi-Yo and Colacito (2016), respectively, which are referred to as BBC and CC in the table. Despite apparent differences in these co-entropy measures, both the BBC and CC measures are equally successful in measuring asymmetry. The return spreads between the highest-asymmetry and lowest-asymmetry portfolios are approximately 0.27% per month (3.24% per annum) in both cases. They are similar in magnitude as the return spread generated by DOWN_CORR, but the t-statistic is higher (4.36 vs. 3.47), meaning that the spread is more precisely estimated. Moreover, the Carhart 4-factor alphas are very close to statistically significant at the 10% level (p-value = 0.116), whereas the alpha in the DOWN_CORR case is much smaller and highly insignificant. This indicates, as expected theoretically, that the co-entropy-based asymmetry measures capture the stocks' asymmetric comovement better than the correlation-based measure, although these spreads are still lower than that of the portfolios formed using DOWN_ASY. This is also unsurprising because DOWN_ASY utilizes all the information available in the entire joint return distribution, whereas neither of the co-entropy measures does. However, if we take the estimation error into account, the 95% confidence interval of the return spread for portfolios sorted by the BBC measure is approximately [0.148, 0.391]. Statistically, the asymmetry risk premiums of DOWN_ASY and the co-entropy-based downside asymmetry measures could possibly be equally high. This suggests that the co-entropy measures may be good alternatives for DOWN_ASY in this context. Overall, our finding of a substantial downside risk premium is robust because all the entropy-based measures yield statistically significant evidence.

D. Further Analysis

We next examine how asymmetry relates to investor sentiment, market volatility, and market liquidity. In addition, we determine whether our asymmetry measures remain statistically significant after controlling for downside asymmetric correlation, downside beta, or coskewness via a double-sorting procedure.

The average return differences between stocks with high and low DOWN_ASY change over time, with positive realizations in approximately 62% of the time periods. Columns 1–4 of Table 6 report the results of regressing the return difference on market volatility, investor sentiment (Baker and Wurgler (2006)), and market liquidity (Pástor and Stambaugh (2003)).[15] The asymmetry risk premium is strong during periods of low investor sentiment, low market volatility, and high market liquidity, although the relation to investor sentiment is not statistically significant. We also estimate DOWN_ASY for each quarter, then sort all stocks into quintiles according to their average quarterly asymmetry during the previous 12 months. Columns 5–8 report the regression results with the return difference as the dependent variable, using the average

[15]They are available at http://people.stern.nyu.edu/jwurgler/ and http://faculty.chicagobooth.edu/lubos.pastor/research/.

TABLE 6

Determinants of Time-Varying Premium of DOWN_ASY

Table 6 reports various regression results where the dependent variables are the asymmetry premium (realized or moving average (MA) smoothed), and the independent variables include market volatility, market liquidity, and investor sentiment. MKT_VOL is the monthly variance of the daily value-weighted market return. Aggregate stock market liquidity (LIQ) is provided by Pástor and Stambaugh (2003). The sample period is from Jan. 1963 to Dec. 2013. BW_SENT denotes the sentiment index of Baker and Wurgler (2006), available only since July 1965. t-statistics are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Premium Based on Realized DOWN_ASY				Premium Based on MA Smoothed DOWN_ASY			
	1	2	3	4	5	6	7	8
MKT_VOL	−0.124***			−0.087*	−0.118***			−0.086**
	(−2.83)			(−1.78)	(−3.05)			(−1.99)
LIQ		4.062***		2.963*		3.269***		2.316*
		(3.01)		(1.93)		(2.72)		(1.71)
BW_SENT			−0.036	−0.043			−0.166**	−0.170**
			(−0.39)	(−0.47)			(−2.04)	(−2.10)
Constant	0.494***	0.497***	0.381***	0.561***	0.419***	0.404***	0.301***	0.459***
	(5.22)	(5.29)	(4.26)	(5.42)	(4.99)	(4.83)	(3.81)	(5.03)
No. of obs.	612	612	582	582	612	612	582	582
R^2	0.013	0.015	0.000	0.020	0.015	0.012	0.007	0.026

quarterly asymmetry to sort portfolios. We can see that the findings are similar, except that the negative relation to investor sentiment becomes statistically significant.

To see the incremental effect of DOWN_ASY on stock returns after controlling for other asymmetry measures, we first sort stocks into quintiles by a given asymmetry measure, such as DOWN_CORR, $\beta-$, or coskewness. Within each quintile, we further divide stocks by DOWN_ASY. The results are reported in Table 7.

TABLE 7

Sequentially Double-Sorted Portfolios

Table 7 reports the equal-weighted average returns, in percentage points, of portfolios double-sorted by a realized asymmetry measure first and then by DOWN_ASY: DOWN_CORR (Panel A), β^- (Panel B), and coskewness (Panel C), respectively. DOWN_ASY is defined in equation (16). DOWN_CORR is defined in equation (17). β^- and coskewness are defined in equations (A-2) and (A-3). The row labeled "High − Low" reports the difference between the returns of portfolios 5 and 1 in each first-sort quintile and the corresponding t-statistics (reported in parentheses). The column labeled "Average" reports the average return of stocks in each second-sort quintile. The sample covers all U.S. common stocks listed on the New York Stock Exchange (NYSE)/American Stock Exchange (AMEX)/National Association of Securities Dealers Automated Quotations (NASDAQ), and the sample period is from Jan. 1963 to Dec. 2013. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. DOWN_CORR and DOWN_ASY

Portfolio	Low DOWN_CORR 1	2	3	4	High DOWN_CORR 5	Average
1 (Low DOWN_ASY)	0.604	0.616	0.657	0.641	0.753	0.654
2	0.522	0.606	0.833	0.828	0.983	0.754
3	0.706	0.786	0.862	0.946	0.963	0.853
4	0.780	0.886	1.044	1.007	1.056	0.955
5 (High DOWN_ASY)	0.852	0.982	1.058	1.089	1.039	1.004
High − Low	0.247***	0.366***	0.401***	0.448***	0.285***	0.349***
	(2.69)	(3.99)	(3.99)	(4.01)	(2.67)	(4.31)

(continued on next page)

TABLE 7 (continued)
Sequentially Double-Sorted Portfolios

Panel B. β^- and DOWN_ASY

Portfolio	Low β^- 1	2	3	4	High β^- 5	Average
1 (Low DOWN_ASY)	0.657	0.670	0.728	0.733	0.607	0.679
2	0.673	0.762	0.764	0.809	0.699	0.741
3	0.806	0.850	0.883	0.909	0.849	0.859
4	0.884	0.927	0.947	0.970	0.885	0.923
5 (High DOWN_ASY)	0.837	0.975	1.107	1.120	1.040	1.016
High – Low	0.180**	0.305***	0.379***	0.388***	0.433***	0.337***
	(2.22)	(3.66)	(4.26)	(3.67)	(3.26)	(4.54)

Panel C. COSKEW and DOWN_ASY

Portfolio	Low COSKEW 1	2	3	4	High COSKEW 5	Average
1 (Low DOWN_ASY)	0.686	0.758	0.617	0.580	0.558	0.639
2	0.929	0.831	0.769	0.657	0.509	0.739
3	0.961	0.958	0.865	0.782	0.719	0.857
4	1.085	1.107	1.065	0.879	0.803	0.988
5 (High DOWN_ASY)	1.116	1.087	0.991	0.973	0.806	0.995
High – Low	0.431***	0.329***	0.374***	0.393***	0.249***	0.355***
	(3.74)	(3.19)	(3.61)	(4.20)	(2.65)	(4.35)

Panel A of Table 7 reports the results from sorting first by DOWN_CORR and then by DOWN_ASY. The return difference remains sizable after controlling for downside asymmetric correlations, ranging from 0.247% to 0.448% across the five downside correlation groups. All are statistically significant at the 5% level and economically significant. The results indicate that the downside asymmetric correlation cannot rule out the importance of DOWN_ASY. Panel B reports the results based on β^-. Again, after controlling for the effect of the downside beta, economically and statistically significant return differences persist. Similar results can be seen in Panel C for coskewness.

Overall, we confirm that our entropy-based downside asymmetry measure contains information that is not captured by other asymmetry measures. Therefore, it offers a new perspective on asymmetry for practical use, complementing existing measures.

VI. Conclusion

Asymmetric comovement in stock returns is important for both portfolio management and risk hedging. However, existing measures of asymmetry use only the second moments, a special case of joint asymmetry. In particular, HTZ's (2007) asymmetric correlation test only measures linear dependence, ignoring higher-order dependence. Therefore, it has low power and can only detect a few cases of asymmetry for commonly used stock portfolios. In this article, we propose a model-free entropy measure that provides a direct test for asymmetry in the joint distribution. Econometrically, our test extends the univariate asymmetry test proposed by Racine and Maasoumi (2007) to the bivariate case, which has important applications in finance. Using Monte Carlo simulations, we find that the entropy test has reliable size and good power. Empirically, based on the entropy test, we find statistically significant asymmetry for many common portfolios, such as size, book-to-market, and momentum portfolios.

Our article also examines whether stocks' downside asymmetric comovement with the market has pricing implications in the cross section of stock returns. Using a modified entropy measure based on the quadrant probabilities of the joint distribution of stock and market returns, we find a significant return difference between stocks with high and low downside asymmetry. The findings are consistent with the theoretical implications of a representative-agent model with disappointment-aversion utility. A trading strategy that buys stocks with high downside asymmetry and sells stocks with low downside asymmetry can earn a return of 4.5% per year. Such a premium is both economically and statistically significant and cannot be explained by previously identified risk factors, such as the excess return on the market, small minus big (SMB), high minus low (HML), or momentum factors. Further analysis indicates that the premium also cannot be explained by other moment-based measures, such as the downside beta and coskewness.

Appendix. Variable Definitions

In this Appendix, we provide detailed definitions of various variables used in the article. A stock i's demeaned daily excess return is denoted by $\tilde{r}_{i,d}$, and the demeaned daily market excess return is denoted by $\tilde{r}_{m,d}$.

β (CAPM beta): β is estimated for each month t using past-12-month observations, based on the standard formula

$$\text{(A-1)} \qquad \hat{\beta}_{i,t} = \frac{\sum_{d=1}^{d=D_t} \tilde{r}_{i,d}\tilde{r}_{m,d}}{\sum_{d=1}^{d=D_t} \tilde{r}^2_{m,d}},$$

where D_t is the number of trading days in the past-12-month period ending in month t.

β^- and β^+ (downside beta and upside beta): We denote the sample average of the daily market excess return during a 12-month period ending in month t, $\hat{\mu}_{m,t}$. We further denote the demeaned excess return and demeaned market excess return conditional on the market excess return being below (above) $\hat{\mu}_{m,t}$ as $\tilde{r}^-_{i,d}$ ($\tilde{r}^+_{i,d}$) and $\tilde{r}^-_{m,d}$ ($\tilde{r}^+_{m,d}$), respectively. Following the definitions of Ang et al. (2006), we compute downside and upside betas as

$$\text{(A-2)} \qquad \hat{\beta}^-_{i,t} = \frac{\sum_{r_{m,d} < \hat{\mu}_{m,t}} \tilde{r}^-_{i,d}\tilde{r}^-_{m,d}}{\sum_{r_{m,d} < \hat{\mu}_{m,t}} \tilde{r}^{-2}_{m,d}}, \quad \text{and} \quad \hat{\beta}^+_{i,t} = \frac{\sum_{r_{m,d} > \hat{\mu}_{m,t}} \tilde{r}^+_{i,d}\tilde{r}^+_{m,d}}{\sum_{r_{m,d} > \hat{\mu}_{m,t}} \tilde{r}^{+2}_{m,d}}.$$

COSKEW (coskewness): Following Harvey and Siddique (2000), the coskewness of stock i over a 12-month period ending in month t is given by

(A-3) $$\widehat{\text{COSKEW}}_{i,t} = \frac{\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{i,d}\tilde{r}_{m,d}^2}{\sqrt{\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{i,d}^2}\left(\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{m,d}^2\right)},$$

where D_t is the number of trading days in a 12-month period ending in month t.

COKURT (cokurtosis): The cokurtosis of stock i over a 12-month period ending in month t is defined as

(A-4) $$\widehat{\text{COKURT}}_{i,t} = \frac{\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{i,d}\tilde{r}_{m,d}^3}{\sqrt{\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{i,d}^2}\left(\frac{1}{T}\sum_{d=1}^{d=D_t}\tilde{r}_{m,d}^2\right)^{3/2}},$$

where D_t is the number of trading days in a 12-month period ending in month t.

IVOL (idiosyncratic volatility): The measure IVOL of stock i at the end of each month t is defined as the standard deviation of the CAPM residual series over the past 12 months.

SIZE: Firm size for each month t is measured using the natural logarithm of the market value of equity at the end of month t.

BM (book-to-market ratio): Following Fama and French (1992), a firm's book-to-market ratio is calculated using its market value of equity and book value of equity. We assume the book value is available 6 months after the reporting date. We take book values from Compustat supplemented by book values from Ken French's Web site. The BM at month t is defined as the natural logarithm of the book-to-market ratio at the end of month t.

MOM (momentum): Following Jegadeesh and Titman (1993), the momentum effect of each stock in month t is measured by the cumulative return over the previous 6 months, with 1 month skipped, that is, the cumulative return from month $t-6$ to month $t-1$.

TURN (turnover): The turnover ratio is calculated monthly as the monthly trading volume divided by month-end shares outstanding.

ILLIQ (illiquidity): We first calculate the ratio of absolute price change to dollar trading volume for each stock each day. Then we take the monthly average of the ratio if the number of observations is greater than 15 in that month. Following Acharya and Pedersen (2005), we normalize the Amihud ratio and truncate it at 30.

MAX (maximum): Following Bali, Cakici, and Whitelaw (2011), MAX for stock i in month t is defined as the maximum daily excess return that month.

References

Acharya, V. V., and L. H. Pedersen. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics*, 77 (2005), 375–410.

Alcock, J., and A. Hatherley. "Characterizing the Asymmetric Dependence Premium." *Review of Finance*, 21 (2017), 1701–1737.

Amaya, D.; P. Christoffersen; K. Jacobs; and A. Vasquez. "Does Realized Skewness Predict the Cross-Section of Equity Returns?" *Journal of Financial Economics*, 118 (2015), 135–167.

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

Ang, A., and J. Chen. "Asymmetric Correlations of Equity Portfolios." *Journal of Financial Economics*, 63 (2002), 443–494.

Ang, A.; J. Chen; and Y. Xing. "Downside Risk." *Review of Financial Studies*, 19 (2006), 1191–1239.

Backus, D.; N. Boyarchenko; and M. Chernov. "Term Structures of Asset Prices and Returns." *Journal of Financial Economics*, 129 (2018), 1–23.

Backus, D.; M. Chernov; and S. Zin. "Sources of Entropy in Representative Agent Models." *Journal of Finance*, 69 (2014), 51–99.

Bae, K.-H.; G. A. Karolyi; and R. M. Stulz. "A New Approach to Measuring Financial Contagion." *Review of Financial Studies*, 16 (2003), 717–763.

Baker, M., and J. Wurgler. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance*, 61 (2006), 1645–1680.

Bali, T. G.; N. Cakici; and R. F. Whitelaw. "Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns." *Journal of Financial Economics*, 99 (2011), 427–446.

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

Bekaert, G., and E. Engstrom. "Asset Return Dynamics under Habits and Bad-Environment Good-Environment Fundamentals." *Journal of Political Economy*, 125 (2017), 713–760.

Bekaert, G.; E. Engstrom; and A. Ermolov. "Bad Environments, Good Environments: A Non-Gaussian Asymmetric Volatility Model." *Journal of Econometrics*, 186 (2015), 258–275.

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

Cabrales, A.; O. Gossner; and R. Serrano. "Entropy and the Value of Information for Investors." *American Economic Review*, 103 (2013), 360–377.

Carhart, M. M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, 52 (1997), 57–82.

Chabi-Yo, F., and R. Colacito. "The Term Structures of Co-Entropy in International Financial Markets." Available at https://ssrn.com/abstract=2341772 (2016).

Chabi-Yo, F.; S. Ruenzi; and F. Weigert. "Crash Sensitivity and the Cross Section of Expected Stock Returns." *Journal of Financial and Quantitative Analysis*, 53 (2018), 1059–1100.

Cho, Y.-H., and R. F. Engle. "Time-Varying Betas and Asymmetric Effect of News: Empirical Analysis of Blue Chip Stocks." Technical Report, National Bureau of Economic Research (1999).

Colacito, R.; E. Ghysels; J. Meng; and W. Siwasarit. "Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory." *Review of Financial Studies*, 29 (2016), 2069–2109.

Conrad, J.; R. F. Dittmar; and E. Ghysels. "Ex Ante Skewness and Expected Stock Returns." *Journal of Finance*, 68 (2013), 85–124.

Conrad, J.; M. N. Gultekin; and G. Kaul. "Asymmetric Predictability of Conditional Variances." *Review of Financial Studies*, 4 (1991), 597–622.

Cont, R. "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues." *Quantitative Finance*, 1 (2001), 223–236.

Dahlquist, M.; A. Farago; and R. Tédongap. "Asymmetries and Portfolio Choice." *Review of Financial Studies*, 30 (2017), 667–702.

Duin, R. P. "On the Choice of Smoothing Parameters for Parzen Estimators of Probability Density Functions." *IEEE Transactions on Computers*, C-25 (1976).

Efron, B. *The Jackknife, the Bootstrap, and Other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics 38. Philadelphia, PA: Society for Industrial and Applied Mathematics (1982).

Fama, E. F., and K. R. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), 427–465.

Granger, C.; E. Maasoumi; and J. Racine. "A Dependence Metric for Possibly Nonlinear Processes." *Journal of Time Series Analysis*, 25 (2004), 649–669.

Gul, F. "A Theory of Disappointment Aversion." *Econometrica*, 59 (1991), 667–686.

Hall, P. *The Bootstrap and Edgeworth Expansion (Springer Series in Statistics)*. New York, NY: Springer (1992).

Hansen, L. P., and R. Jagannathan. "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy*, 99 (1991), 225–262.

Harvey, C. R., and A. Siddique. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance*, 55 (2000), 1263–1295.

Hong, Y.; J. Tu; and G. Zhou. "Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation." *Review of Financial Studies*, 20 (2007), 1547–1581.

Hong, Y., and H. White. “Asymptotic Distribution Theory for Nonparametric Entropy Measures of Serial Dependence.” *Econometrica*, 73 (2005), 837–901.

Horowitz, J. L. “The Bootstrap.” *Handbook of Econometrics*, 5 (2001), 3159–3228.

Jegadeesh, N., and S. Titman. “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency.” *Journal of Finance*, 48 (1993), 65–91.

Jiang, L.; K. Wu; G. Zhou; and Y. Zhu. “Stock Return Asymmetry: Beyond Skewness.” Working Paper, Washington University in St. Louis (2018).

Kelly, B., and H. Jiang. “Tail Risk and Asset Prices.” *Review of Financial Studies*, 27 (2014), 2841–2871.

Kroner, K., and V. K. Ng. “Modeling Asymmetric Comovements of Asset Returns.” *Review of Financial Studies*, 11 (1998), 817–844.

Kullback, S., and R. Leibler. “On Information and Sufficiency.” *Annals of Mathematical Statistics*, 22 (1951), 79–86.

Lettau, M.; M. Maggiori; and M. Weber. “Conditional Risk Premia in Currency Markets and Other Asset Classes.” *Journal of Financial Economics*, 114 (2014), 197–225.

Li, Q., and J. S. Racine. *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ: Princeton University Press (2007).

Maasoumi, E., and J. Racine. “Entropy and Predictability of Stock Market Returns.” *Journal of Econometrics*, 107 (2002), 291–312.

Maasoumi, E., and J. S. Racine. “A Robust Entropy-Based Test of Asymmetry for Discrete and Continuous Processes.” *Econometric Reviews*, 28 (2008), 246–261.

Parzen, E. “On Estimation of a Probability Density Function and Mode.” *Annals of Mathematical Statistics*, 33 (1962), 1065–1076.

Pástor, L., and R. F. Stambaugh. “Liquidity Risk and Expected Stock Returns.” *Journal of Political Economy*, 111 (2003), 642–685.

Patton, A.; D. N. Politis; and H. White. “Correction to ‘Automatic Block-Length Selection for the Dependent Bootstrap’ by D. Politis and H. White.” *Econometric Reviews*, 28 (2009), 372–375.

Racine, J. S., and E. Maasoumi. “A Versatile and Robust Metric Entropy Test of Time-Reversibility, and Other Hypotheses.” *Journal of Econometrics*, 138 (2007), 547–567.

Rosenblatt, M. “Remarks on Some Nonparametric Estimates of a Density Function.” *Annals of Mathematical Statistics*, 27 (1956), 832–837.

Schwert, G. W. “Why Does Stock Market Volatility Change over Time?” *Journal of Finance*, 44 (1989), 1115–1153.

Segal, G.; I. Shaliastovich; and A. Yaron. “Good and Bad Uncertainty: Macroeconomic and Financial Market Implications.” *Journal of Financial Economics*, 117 (2015), 369–397.

Shannon, C. “A Mathematical Theory of Communication.” *Bell System Technical Journal*, 27 (1948), 379–423.

Skaug, H. J., and D. Tjøstheim. “A Nonparametric Test of Serial Independence Based on the Empirical Distribution Function.” *Biometrika*, 80 (1993), 591–602.

Tawn, J. A. “Bivariate Extreme Value Theory: Models and Estimation.” *Biometrika*, 75 (1988), 397–415.

Tjøstheim, D. “Measures of Dependence and Tests of Independence.” *Statistics: A Journal of Theoretical and Applied Statistics*, 28 (1996), 249–284.

Ullah, A. “Entropy, Divergence and Distance Measures with Econometric Applications.” *Journal of Statistical Planning and Inference*, 49 (1996), 137–162.

Zhou, G. “Asset-Pricing Tests under Alternative Distributions.” *Journal of Finance*, 48 (1993), 1927–1942.