Asymmetry in Stock Returns:

An Entropy Measure

Abstract

In this paper, we provide an entropy measure for asymmetric comovements between an asset return and the market return. This measure yields a model-free test for asymmetry in stock returns that has greater power than 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. In the cross-section of stock returns, we find a risk premium (discount) for stocks with high downside (upside) comovement with the market. Asymmetric comovement leaned toward the downside also earns a premium. The risk premia associated with downside comovement. Moreover, downside comovement premium is almost twice as large as the premium of downside beta. Our findings are consistent with theoretical implications of a representative agent model with disappointment aversion preferences.

Keywords: Asymmetric dependence, metric entropy, asset pricing, anomaly.

JEL Classification: C12, C15, C32, G11, G12, G17

1 Introduction

Asymmetric characteristics of asset returns, i.e. stocks co-move more strongly when market goes down than when market goes up, have been found by a number of prior studies. Ball and Kothari (1989); Schwert (1989); Conrad, Gultekin, and Kaul (1991); Cho and Engle (1999); Bekaert and Wu (2000); Ang and Chen (2002); Bae, Karolyi, and Stulz (2003); Ang, Chen, and Xing (2006), among others, documented asymmetries in covariances, correlations, volatilities, and betas of stock returns. Such asymmetric characteristics of stock returns are important both for portfolio selection and risk management, because effective hedging relies on the dependence between assets hedged and financial instruments used. If the dependence structure is varying with the state of the market, portfolio managers may need to worry about the effectiveness of their hedges when they are most needed.

Almost all these previous studies are model dependent. For example, Ang and Chen (2002) find correlation asymmetries among various portfolios under joint normality assumption, which leaves the room for possibility that the data features unexplained by the joint normal model may well be explained by some other symmetric model. Hong, Tu, and Zhou (2007) proposes the first and the only model-free test of asymmetry to date. This test, despite of its novelty, has two weaknesses. First, it detects only asymmetric correlations, and does not address asymmetry beyond the second moment. It is well known that the correlation coefficient is only a measure of linear dependence and cannot capture the full dependence structure for non-normal distributions, while several papers documented that realized stock returns are non-normally distributed (see, e.g., Embrechts, McNeil, and Straumann, 2002; Ang and Chen, 2002). Second, its finite-sample power seems low in empirical applications. For example, the test cannot detect any asymmetry in portfolios sorted by book-to-market ratio, and finds only one significant asymmetric portfolio among momentum sorted portfolios.

In this paper, we first propose an entropy measure to exactly measure the asymmetric dependence between individual stock return and the market return. Using the entropy measure, we propose a new model-free test for asymmetric dependence. The test statistic is a normalized metric entropy proposed by Granger, Maasoumi, and Racine (2004) that have been widely applied in econometrics (see, e.g., Maasoumi and Racine, 2002; Racine and Maasoumi, 2007). The entropy measure is defined based on the joint probability density functions, so it can summarize all the information of a given joint distribution and hence can capture general asymmetric dependence structure existed in all the moments. With Monte Carlo simulations, we find that the entropy-based test has good size and power properties. Using sorted portfolios based on size, book-to-market ratio and momentum, the entropy-based test detects statistically significant asymmetry in all kinds of the sorted portfolios. For example, in contrast to the Hong, Tu, and Zhou (2007) test, we find asymmetry in 2 portfolios at the 5% significance level, and in 7 portfolios at the 10% level, out of the 10 decile portfolios sorted by the book-to-market ratio.

What is the asset pricing implication of asymmetry in the cross-section of expected stock returns? It is actually a less studied question in the literature. Under classical Capital Asset Pricing Model (CAPM), it is sufficient to consider only linear correlations (captured by the CAPM beta) between individual stock returns and the market portfolio return. (see Sharpe, 1964; Lintner, 1965). However, more recent studies find supporting evidence that asymmetry features of the joint distribution of individual stock and market returns also determine the expected stock returns. For example, Harvey and Siddique (2000) show that the conditional coskewness plays an important role in explaining the cross-sectional returns. Ang, Chen, and Xing (2006) find that asymmetric risk premia are associated with downside and upside betas. They show that stocks with higher downside betas have on average higher returns, but have mixed evidence on whether higher upside betas are associated with lower returns. Since downside and upside betas are highly correlated with market betas (the correlations are above 0.75 as shown in Ang, Chen, and Xing (2006)), i.e. an increase in downside or upside betas are associated with an increase in the CAPM beta, it is difficult to distinguish the effects of downside or upside covariation from the overall covariation between the stock and the market returns. Alcock and Hatherley (2013) tries to overcome this problem by constructing a beta-invariant asymmetric dependence measure that is a modified J statistic proposed by Hong, Tu, and Zhou (2007). Although beta-invariant, their measure still does not capture full dependence structure since it is constructed based on exceedance correlations that can only capture conditional dependence to the second moment (linear dependence).

Some recent papers start to examine cross-sectional asset pricing implications of higher order dependence. Contemporary paper by Chabi-Yo, Ruenzi, and Weigert (2014), in a non-normal distribution framework, uses parametric copula-based tail dependence measure to explain the cross-sectional expected returns. This paper differs from theirs in that they focus on extreme lower tail dependence, or the crash sensitivities of stocks, while, muck like Ang, Chen, and Xing (2006), we focus on the downside and upside dependence when market returns are above or below the mean. Using the proposed entropy measure, we study the asset pricing implications of asymmetric dependence. Theoretically, Ang, Chen, and Xing (2006) shows that under a simple representative agent model with disappointment aversion (DA) utility (Gul, 1991), agents require a premium to hold stocks with strong covariation with the downside market, while are willing to hold stocks with high upside potential at a discount, all else being equal. Motivated from this insight, we expect stocks with stronger downside asymmetric dependence, i.e. the dependence with the downside market is stronger than with the upside market, to earn higher average returns, because those stocks are highly risky in the sense that they may incur large loss when the wealth level is low, meanwhile they do not have high upside potential when the market goes up. Furthermore, as pointed out by Ang, Chen, and Xing (2006), the DA utility is kinked at certainty equivalence wealth level, so the higher-order co-moments derived from Taylor expansion, like coskewness and cokurtosis, may not approximate the utility function well globally. This is a theoretical motivation why there may exist asymmetric effects of downside and upside dependence.

We construct proxies for downside and upside dependence using estimated probabilities that individual stock and market returns both fall below or above the sample means. Using Center for Research in Securities Prices (CRSP) data from 1962 to 2013, we find empirical evidence that stocks with high downside (upside) dependence earn a premium (discount). Both effects are statistically and economically significant after controlling for other known characteristics in cross-sectional Fama and MacBeth (1973) regressions. The findings support the theoretical implications of a representative agent model with DA utility. The value-weighted average return (Carhart (1997) four factor adjusted alpha) of the top quintile portfolio sorted based on downside asymmetric dependence outperforms the lowest quintile portfolio by 12.34% (12.89%) per annum. In Fama-Macbeth (1973) regressions, the premium of downside asymmetric dependence cannot be explained by known characteristics, such as CAPM beta, downside or upside betas, coskewness and cokurtosis, size, book-to-market ratio, past returns and maximum daily return within a month. The downside asymmetric dependence is time-varying and shows limited predictability using its own lag. Yet when using the lagged asymmetric dependence to form a trading strategy, the spread portfolio still earns an average equal-weighted annualized return of 4.5%. The premium is both economically and statistically significant.

The rest of the paper is organized as follows. Section 2 introduces the entropy-based test for asymmetric dependence. Section 3 examines the test size and power using Monte Carlo simulations. Section 4 applies the entropy test to investigate asymmetry in commonly used portfolios. Section 5 analyzes the impact of asymmetry on the cross-section stock returns. Section 6 concludes.

2 Test of Asymmetry Dependence

For easy understanding, in this section we first review the standard asymmetric correlation test, then extend the concept to more general asymmetric dependence, and finally provide our entropy-based test.

2.1 Asymmetric Correlation

In the finance literature, Ang and Chen (2002) and Hong, Tu, and Zhou (2007) provide asymmetry tests, but they test only asymmetric correlation instead of asymmetric dependence. To see why, Let $\tilde{x}$ and $\tilde{y}$ be pairs of standardized return series. Both of the tests rely on exceedance correlations, i.e. the conditional correlations evaluated when both individual stock and the market returns are below or above certain exceedance levels, to construct measures of asymmetry of the joint return distribution. The exceedance correlations at some given level c are defined as

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

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

The null hypothesis of interest is

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

Ang and Chen (2002) is the first to propose a formal statistical test for the asymmetric correlation hypothesis, whose test statistic is defined as

$$H = \left[\sum_{i=1}^{m} w(c_i)(\rho(c_i, \phi) - \hat{\rho}(c_i))^2 \right]^{1/2} \tag{3}$$

where $c_1, \ldots, c_m$ are m pre-selected exceedance levels, $w(c_1), ..., w(c_m)$ are weights, $\hat{\rho}(c_i)$ stands for sample realization of $\rho^+(c_i)$ or $\rho^-(c_i)$ and $\rho(c_i, \phi)$ is the population exceedance correlation implied by a given model with parameters ϕ. Their test addresses the interesting question whether the asymmetric correlations in the data can be explained by the given model. But there is a weakness that the data may still have asymmetric correlation if a given symmetric model, like the normality model they used, cannot explain it.

To overcome that weakness, Hong, Tu, and Zhou (2007) propose a model-free test

$$J_\rho = T(\hat{\rho}^+ - \hat{\rho}^-)' \hat{\Omega}^{-1} (\hat{\rho}^+ - \hat{\rho}^-) \tag{4}$$

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 chi-squared distribution, $J_\rho \xrightarrow{d} \chi^2_m$. The test answers the question whether the data have asymmetric at all. In other words, if the test rejects the null, it implies that no distributions of symmetric correlations can fit the data.

2.2 Asymmetric Dependence

It is well known that linear correlation coefficient is only a measure of linear dependence and cannot reflect dependence structure beyond the second moment. In general, zero correlation does not imply independence except for the joint normal case. Hence, tests based on linear correlations ignore possible higher order dependence entirely. On the other hand, it is also well known that financial time series usually display heavy tails and have non-standard higher order moments. Therefore, it is of interest to have a test for asymmetric dependence that involves all higher order moments.

Analogous to the exceedance correlations, we define exceedance densities by

$$f^+(c) = f(\tilde{x}, \tilde{y} | \tilde{x} > c,\ \tilde{y} > c \ or \ \tilde{x} < -c,\ \tilde{y} < -c), \tag{5}$$

$$f^-(c) = f(-\tilde{x}, -\tilde{y} | \tilde{x} > c,\ \tilde{y} > c \ or \ \tilde{x} < -c,\ \tilde{y} < -c), \tag{6}$$

where $f(\tilde{x}, \tilde{y})$ is the joint probability density function of standardized return series $\tilde{x}_{i,t}$ and $\tilde{y}_t$ in ranges of $\tilde{x} > c,\ \tilde{y} > c$ or $\tilde{x} < -c,\ \tilde{y} < -c$. Since $\tilde{x}_{i,t}$ and $\tilde{y}_t$ have zero means, $f(-\tilde{x}, -\tilde{y})$ denotes the joint probability density function of the rotated return series around the mean. 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 asymmetric dependence is

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

If this hypothesis is rejected, then the data must possess asymmetry as their density functions must be different at least in one of the two symmetric regions.

[Insert Figure 1 here]

As an example, Figure 1 illustrates a case of symmetric dependence and compares it with a case of asymmetric dependence. Subfigure (a) shows a scatter plot of 2000 data points that are generated by the Clayton copula that is known to have a stronger left tail dependence than right tail dependence. Subfigure (b) is a similar plot but is generated by using a bivariate normal distribution that has symmetric dependence at both tails. When we consider symmetric/asymmetric dependence, we examine the dependence structure over the shaded areas located in the first and third quadrants. In subfigure (a), it is clear that the data are more concentrated in the left tail than in the right tail, indicating stronger dependence (greater mass of the joint densities) in the left tail. Subfigure (b) has roughly equal joint densities in both tails, indicating symmetric dependence. The lines in both figures are fitted linear regression lines that indicate linear dependence. With visual inspection, we see that the linear dependence line does not differ very much in the two cases, but the actual dependence structures are quite different. Hence, this example also highlights the danger of failing to discover asymmetry when we focus only on linear dependence.

2.3 A Metric Entropy Measure

The important question is how to test the above null hypothesis. Intuitively, the joint distribution is symmetric if the distance between $f^+(c)$ and $f^-(c)$ is zero almost everywhere and is otherwise asymmetric if the distance is not zero on a set with positive measure. To do so, we have to use certain measures of distance between two probability density functions. Originated from physics and information theory as a measure of uncertainty, entropy has a long history of being used as a measure of divergence between distributions. It was first introduced by Shannon (1948), and later extended by Kullback, Leibler et al. (1951). Ullah (1996) provides an excellent survey of various entropy measures and their wide applications in econometrics. More recently, entropy has drawn more attention by financial economists and has been more and more applied in finance research. Some important applications of entropy include Sims (2003); Backus, Chernov, and Martin (2011); Hansen (2012). Among most recent notable examples, Cabrales, Gossner, and Serrano (2013) use Shannon's entropy (Shannon, 1948) to quantify the informativeness of a ruin-averse investor's beliefs on the state of nature; Backus, Chernov, and Zin (2014) use Kullback-Leibler relative entropy (Kullback, Leibler et al., 1951) to measure the differences between physical and risk-neutral probabilities and derive appropriate bounds for stochastic discount factors of popular asset pricing models. The entropy measure we use belongs to the same K-class entropy as the Kullback-Leibler divergence measure. First proposed by Granger, Maasoumi, and Racine (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 is the only metric entropy within its class. Besides being a metric, as shown by Granger, Maasoumi, and Racine (2004), this measure has many other desirable properties, and so we use its bivariate version in this paper.

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

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

which is clearly a function of c. The exceedance level c is chosen according to empirical interests. For example, when $c = 0$, $S_\rho(0)$ measures asymmetric dependence in the first quadrant (where both standardized returns are positive) and the third quadrant (where both standardized returns are negative). If the asymmetric decadence in the distribution tails are of interests, $S_\rho(c)$ can be measured at other exceedance levels, such as 0.5 and 1 standard deviations away from the mean. The S_ρ measure is well defined for both continuous and discrete data. Its value is between 0 and 1, and equals 0 if and only if the densities are the same or there is symmetric dependence. Econometrically, 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 commonly used logarithm transformation.

Consider now the case where we have multiple exceedance levels, $c_1, \ldots, c_m$. For example, while the singleton of 0 level is common, the set of the levels $\{0; 0.5; 1; 1.5\}$ is also commonly used in previous studies. In the multiple level case, we can apply (7) for each of the individual levels, and then aggregate using any arbitrary function. For simplicity, we use the simple average,

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

where $S_\rho(c_j)$ is computed from (7) for $j = 1, \ldots, m$. Then, our entropy is well defined for both cases of exceedance levels.

To carry out the entropy test in practice, we have to estimate first the density functions from data, and then compute the integral in (7) to obtain the statistic $\hat{S}_\rho$. Finally, we need to have a procedure to determine the P-value of the test. The task is unfortunately much more complex than the asymmetric correlation tests. These issues are addressed in the following two subsections.

2.4 Non-parametric estimation

Let us consider now how to estimate the densities in 7 given the data. Following Maasoumi and Racine (2002); Racine and Maasoumi (2007) among others, the unknown joint densities are estimated using non-parametric kernel estimators. In this paper, we use the popular smoothing non-parametric density estimation method due to Rosenblatt (1956) and Parzen (1962). In the univariate case, the "Parzen-Rosenblatt" kernel density estimator is defined as

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

where n is sample size of the data $\{X_i\}$, h is a smoothing parameter or bandwidth and

$k(\cdot)$ is a nonnegative bounded kernel function. In this paper, we extend it to the bivariate case by applying the "product kernel function," which is constructed as the product of univariate kernel functions. That is, our candidate density function of the joint data series is

$$\hat{f}(x, y) = \frac{1}{nh_1h_2} \sum_{i=1}^{n} k\left(\frac{x_i - x}{h_1}\right) \times k\left(\frac{y_i - y}{h_2}\right) \tag{9}$$

where n is the sample size, $k(\cdot)$ is a suitable univariate kernel function, h_1 and h_2 are bandwidths for each of the two components, and $\{(x_i, y_i)\}$ are the observed data pairs. It should be noted that n is equal to T, the length of the return series in the empirical applications.

The accuracy of the nonparametric kernel density estimator clearly relies on the selection of both the kernel function and the bandwidth. Since the choice of kernel function plays a much less important role than the selection of bandwidth and the return data are continuous, we choose the standard Gaussian kernel, $k(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, in our estimation below. On selecting the bandwidth, we use the well-known Kullback-Leibler cross-validation method. This procedure minimizes the Kullback-Leibler relative entropy between the actual density and the estimated one (see Li and Racine, 2007, for details). Numerically, the cross-validation solves the following maximization problem of the log-likelihood function,

$$\max_{h_1, h_2} \mathcal{L} = \sum_{i=1}^{n} \ln\left[\hat{f}_{-i}(x_i, y_i)\right], \tag{10}$$

where

$$\hat{f}_i(x_i, y_i) = \frac{1}{nh_1h_2} \sum_{j \neq i}^{n} k\left(\frac{x_j - x}{h_1}\right) \times k\left(\frac{y_j - y}{h_2}\right) \tag{11}$$

which is equal to $\hat{f}(x, y)$ without the i-th term. Based on efficient market hypothesis, portfolio returns can be seen as i.i.d. or weakly time dependent series, and under such assumptions the estimated density converges to the exact density. (see, e.g., Li and Racine, 2007, for detailed proofs).

With the method above, we can estimate the densities in (7), and then obtain the test statistic $\hat{S}_\rho$ 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 (). To conduct statistical inference on asymmetry, we need to know the sampling distribution of the test statistic $\hat{S}_\rho$ under the null hypothesis.

2.5 Distribution of test statistic

Asymptotic theory for the class of entropy measures with similar functional forms has been developed by Skaug and Tjøstheim (1993); Tjøstheim (1996) and Hong and White (2005). Asymptotic distribution of $\hat{S}_\rho$ derived under the null hypothesis does not depend on the bandwidth choice. This is partly because the bandwidth is a quantity that vanishes asymptotically. However, for a given finite sample size, the computed value of the test statistic depends crucially on the bandwidth choice (see also (Maasoumi and Racine, 2008)). This really raises concerns about using simple asymptotic distributions to conduct inference in empirical applications, since the results of such asymptotic-based tests tend to be highly sensitive to the bandwidth and there are many competing approaches for bandwidth selection. Therefore, following Racine and Maasoumi (2007) and others, rather than relying on asymptotic distribution for inference, we choose to use a bootstrap resampling approach to determine the empirical distribution of $\hat{S}_\rho$ (see Efron, 1982; Hall, 1992; Horowitz, 2001, for more discussions on bootstrap resampling approach).

To construct a sample under the null hypothesis in the bootstrap resampling, let

$$Z = \{(x_1, y_1), (x_2, y_2), ..., (x_T, y_T); (-x_1, -y_1), (-x_2, -y_2), ..., (-x_T, -y_T)\},$$

which is a vector obtained by stacking together the original data $(x_i,\ y_i)$ with the rotated data $(-x_i,\ -y_i)$. By bootstrapping samples from Z, we construct the empirical distribution of $\hat{S}_\rho$. We repeat the bootstrap draws B times for Z, we then obtain B samples of $\hat{S}_\rho$.

The choice of which bootstrap resampling procedure to use depends on the nature of the data. As stock return are known to be stationary and weakly dependent at most, the block bootstrap that takes such dependence structure into account is the natural choice (see Künsch (1989)). Politis and Romano (1994) shows that using overlapping blocks with lengths that are sampled randomly from the geometric distribution yields stationary bootstrap data series, while overlapping or non-overlapping blocks with fixed lengths may not ensure stationary. Hence, we use the stationary bootstrap in our paper. Moreover, we apply the data-driven and automatic method of Politis and White (2004); Patton, Politis, and White (2009) to select the optimal block length in our bootstrap. Econometrically, their method is beneficial as it minimizes the mean squared error of the estimator of the long-run variance of the time series.

In terms of choosing B, it is obviously true that the greater the B, the more accurate the bootstrapped distribution is. However, unlike the bootstrap in regressions, kernel estimation can be enormously time-consuming. In some similar problems, Davidson and MacKinnon (2000) suggests the use of $B = 399$ for simulations computing the P-value of a test at the 5% significance level. In this paper, although we find that a value of $B = 199$ yields similar results, following the suggestion of Davidson and MacKinnon (2000), we will use $B = 399$ throughout our bootstraps for each portfolio.

After having computed B replications of $\hat{S}_\rho^*$, the sampling distribution of $\hat{S}_\rho$ can be easily obtained. To find out the critical values for rejection at different confidence levels, we can we reorder the bootstrapped estimates from smallest to largest and denote by $\hat{S^*_{\rho,1}}, \hat{S^*_{\rho,2}}, \ldots, \hat{S^*_{\rho,B}}$, and then determining those percentiles from these ordered statistics. For example, to conduct the symmetry test at the 5% level, the null H_0 will be rejected if $\hat{S}_\rho > \widehat{S^*_{\rho,379}}$, where $\widehat{S^*_{\rho,379}}$ is the 95th percentile of the ordered bootstrapped estimates. Empirical p-values may also be obtained by counting the proportion of the ordered bootstrapped statistics that exceeds the estimated statistic from the original sample.

3 Size and Power of the Entropy Test

In this section, using copula-GARCH Monte Carlo simulations, we examine the size and power of the entropy texts and show that the entropy test has reliable sizes, and has higher power in finite samples than Hong, Tu, and Zhou (2007) test.

3.1 Modeling dependence with copulas

Since we are testing the joint distribution of two random variables, the simulation procedures involve in generating random samples from some joint distribution with certain dependence structure. Copulas are probably the most commonly used method to model the complete dependence structure between random variables (see Patton, 2004; Rodriguez, 2007; Okimoto, 2008, for some applications of copulas in finance). Sklar (1959) proves that all bivariate distribution functions $F(x_1, x_2)$ can be completely described by the univariate marginal distributions $F_1(x_1)$ and $F_2(x_2)$ and a copula function $C : [0,1]^2 \mapsto [0,1]$. Copula, a word chosen by Sklar, is a multivariate probability distribution function that describes such dependence structure between the two (or more) marginal distributions (see Nelsen, 1999, for a more detailed introduction to copulas).

Many copulas with different dependence structures have been developed and commonly applied in the literature. Some of those parametric copulas, such as Gaussian, Student's t and Frank copulas, are known to have symmetric tail dependence structure. Some copulas are constructed to have asymmetric tail dependence. For example, Clayton copula is known for strong left tail dependence, whereas Gumbel copula shows strong right tail dependence. As stock returns usually show stronger left tail dependence than right tail dependence with the market return (see Ang and Chen, 2002), Clayton copula seems to be a natural choice. However, it is not wise to completely rule out those copulas with symmetric dependence. Figure 2 gives the scatter plots of random samples generated by Gaussian, Clayton and mixed Gaussian-Clayton copulas, as well as the actual data plots of size 1 portfolio returns. It is clear that Clayton copula generated data with strong left tail dependence, as the plots are highly concentrated at the left tail, but the dependence seems to be much stronger than that is actually reflected in the scatter plot of size 1 portfolio. Even compared to the smallest size portfolio, which has shown to have the strongest asymmetric dependence in the following section and in Hong, Tu, and Zhou (2007), the generated data plots do not look much like the actual data plots. As shown in subfigure (C), the scatter plots generated by equal-weighted mixed Gaussian-Clayton copula look more similar to the actually data plots in subfigure (D). Therefore, we choose to use those mixed copulas as the data generating process in simulations. Such mixture copula models are also used in Hong, Tu, and Zhou (2007).

[Insert Figure 2 here]

A bivariate Gaussian copula is given by

$$C_{nor}(u,\, v;\, \rho) = \Phi_\rho(\Phi^{-1}(u),\, \Phi^{-1}(v)) \tag{12}$$

where $\rho \in (-1, 1)$ is correlation coefficient between the marginal distributions. Φ^{-1} is standard normal CDF inverse, and Φ_ρ is the standard bivariate normal distribution function with correlation ρ.

A bivariate Clayton copula is defined as

$$C_{clay}(u,\, v;\, \tau) = (u^{-\tau} + v^{-\tau} - 1)^{-\frac{1}{\tau}} \tag{13}$$

where $\tau > 0$ governs the dependence between the marginals. Higher τ indicates stronger dependence. Tawn (1988) proves that every convex combination of existing copula functions is still a copula. We can construct a mixture Gaussian-Clayton copula with pre-chosen weights by a convex combination of the two copulas.

The mixture Gaussian-Clayton copula used in this paper takes the following specification.

$$C_{mix}(u,\, v; \rho, \tau, \kappa) = \kappa C_{nor}(u,\, v;\, \rho) + (1-\kappa) C_{clay}(u,\, v;\, \tau) \tag{14}$$

where κ indicates the weight put on the bivariate normal (Gaussian) copula. The mixture copula shown in equation (14) nests both Gaussian and Clayton copulas as special cases. When $\kappa = 1$, the mixture copula reduces to Gaussian copula. When $\kappa = 0$, the mixture copula reduces to Clayton copula. In the following simulation, we take different κ values of 0, 0.25, 0.375, 0.5 and 1 to generate random samples with different levels of asymmetric dependence from the highest to lowest.

3.2 Simulation with Copula-GARCH Model

The joint distribution of generated random samples is governed by the mixture Gaussian-Clayton copula model. We still need to specify some model to best mimic the marginal distributions of asset returns. GARCH(1,1) process is a well-known parsimonious model for stock returns. Following Hong, Tu, and Zhou (2007), we model the marginal distributions of the return series with a GARCH(1,1) with no ARMA components. Basically, the return series is modeled to be equal to an expected return component plus a random error term that follows a GARCH(1,1) process. We first fit the copula-GARCH model to the data to estimate the related parameters using Maximum Likelihood (ML) approach. We then plug the ML estimates back in the model and use it as the data generating process (DGP) in simulation. To be conservative, instead of using portfolios that show clear asymmetric dependence, such as the smallest stock portfolio or momentum portfolios, etc., we use the 5^{th} smallest value-weighted size portfolio and the market return to estimate copula and GARCH parameters. Empirically, we do not find any evidence for asymmetric dependence for size 5 portfolio, hence using it to calibrate the parameters impose a harder challenge for the tests. It is interesting to see whether the tests have reasonable power under such parameter settings. Hong, Tu, and Zhou (2007) has done a similar practice in their simulation exercises.

Taking those ML estimates as the population parameters, we are able to simulate the data with the copula-GARCH model using the following detailed steps.

1. For a given κ, draw a bivariate uniform random sample of size T from the mixture

Gaussian-Clayton copula model;

2. Apply inverse standard normal CDF transformation to get a bivariate standard normal random sample with pre-specified 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 simulated joint return series;

4. The simulated data vectors will each follow a GARCH(1,1) process and the perceived dependence structure governed by the mixture copula model.

5. Repeat step 1 to 4 for 1,000 times to get 1,000 simulated random samples.

6. Repeat step 1 to 5 with different sample sizes T. Specifically, we consider $T = 240,\ 420$ and 600.

[Insert Table 1 about here]

The sample sizes are common choices in the literature. $T = 240$ stands for 20 years of monthly frequency data. $T = 420$ is the length of the subsample data period as used in Hong, Tu, and Zhou (2007). $T = 600$ stands for 50 years of monthly frequency data and is close to the full sample data length ($T = 588$) used in this paper. In simulation, we use one fixed bandwidth for each 1,000 random samples generated from the same DGP. The fixed bandwidth is set to be equal to the average of the 1,000 bandwidths computed for each of the 1,000 random samples via likelihood cross-validation. Similar practice is conducted for the optimal block length selection. The expected block length for each 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 (2009) algorithm. Averaging bandwidth and block length across random samples drawn from the same DGP could potentially reduce some sampling randomness and make the simulation results more stable.

Table 1 reports the empirical size and power for both tests when the nominal size is set at 5% based on 1,000 simulations. Powers are reported with different DGPs of different degrees of asymmetric dependence levels (from $\kappa = 1$ to $\kappa = 0$) and at various sample sizes. We report size and power of Hong, Tu, and Zhou (2007) test (HTZ test hereafter) computed based on both asymptotic distribution and stationary bootstrap with 399 replications. Based on the standard paired bootstrap procedure described in Cameron and Trivedi (2005) and following Horowitz (2001), we construct a pivotal (standardized) statistic when bootstrapping HTZ test statistic to achieve asymptotic refinement. We obtain the variance estimates of HTZ test statistic via sub-bootstrap, i.e. within each bootstrap replication, we bootstrap the replicated sample again to estimate the standard error based on a series of sub-bootstrapped statistics. Since we are estimating the variance rather than tail quantiles or critical values, a fairly small number of resamples is sufficient for consistent estimates. Following Racine (1997), we set the number of sub-bootstrap replications at one tenth of the original number bootstrap replications, i.e. $B_{sub} = 20$. But the bootstrap results of HTZ test does not yield better power than their asymptotic counterparts. However, the empirical sizes are much closer to the nominal values than those based on asymptotic theory. We also find that the powers are much better with the singleton test than the joint test. So we only report the simulation results of HTZ test under exceedance level of $c = 0$. The last column reports the power increase when inference of both tests is based on stationary bootstrap and the exceedance level is set at 0. Since the inference method is the same, we attribute this power increase to better information summarized by the entropy measure. The average power increase is computed as mean of differences among all the simulation scenarios considered in this paper.

We find a pattern that the average power increase is getting more significant as the nominal test size decreases, i.e. the entropy test gives better inference results when we want to report the testing results in a more accurate manner. At nominal size of 10%, the average power difference is only 0.03 or 4%. While at nominal size of 5%, the average power increase is 0.103 or 17.3% and when the nominal size is set at 1%, the the average power increase is 0.245 or a huge 84.6% increase. We can see that the role of information is very significant in making better statistical inference.

It indicates that the entropy test on average has higher power than HTZ test for different DGP that reflects various degree of asymmetric dependence. The difference in power varies with the dependence structure of the DGP. When the simulated data have very strong asymmetric dependence, the performance of both tests are close to each other. If the DGP is a bivariate Clayton copula ($\kappa = 0$), the difference in power is quite small (about 0.14 for $T = 240$) and the difference vanishes as sample size increase to $T = 600$. The power difference is most pronounced when the degree of asymmetric dependence is not very strong. When the DGP is a 37.5% mixed Gaussian-Clayton copula, the power of the entropy test is about 4 times higher than the power of HTZ test for smaller sample sizes ($T = 240$ or $T = 420$). The difference shrinks as the sample size increases to 600, but the power of the entropy test is still twice as large as the power of HTZ test. The improvement of power for both tests with larger sample size is expected, especially for HTZ test based on asymptotic distribution. We tried to make inference of HTZ test using stationary bootstrap, but the results are not as good as results derived from asymptotic inference, so we decide to report the size and power of their test based on asymptotic theory. When the underlying DGP is of symmetric dependence, i.e. the bivariate Gaussian case with $\kappa = 1$, the probability of rejection is the empirical size of the tests. The sizes of both tests are reported in the top left panel in Table 1.

[Insert Table 2 about here]

Table 2 reports the empirical size and power for both tests when the nominal size is set at 1% based on 1,000 simulations. The results reaffirm the conclusions drawn from Table 1. The entropy test shows higher power for all different DGPs. The power difference is more significant for data that show less strong asymmetric dependence.

4 Is Asymmetry Rare?

In this section, we apply our entropy measure to test whether there exists statistically significant asymmetry in common portfolios sorted by size, book-to-market ratio and momentum.

4.1 Data

Following existing studies on asymmetric correlation, we consider portfolios of stocks sorted by popular characteristics, size, book-to-market ratio, and momentum. As in Ang and Chen (2002), we use value-weighted returns of for both size and book-to-market decile portfolios, and use equal-weighted returns for decile momentum portfolios which are formed on prior 2 to 12 month return. Return on CRSP (Center for Research in Security Prices) value-weighted market index based on stocks listed in NYSE/AMEX/NASDAQ is used as a proxy for the market return. All returns are at the monthly frequency and are in excess of the risk-free rate which is taken as the one-month T-bill rate. The entire data are available from Kenneth French's site.[1] The sample period is from January 1965 to December 2013 (588 observations in total).

4.2 Test results

Panel A in Table 3 provides the results on the size portfolios. At the usual 5% level, the entropy test rejects symmetry for all size portfolios from the 1st to 6th smallest size portfolios. In contrast, the existing model-free test of Hong, Tu, and Zhou (2007) can only reject symmetry for the smallest size portfolio based on either the singleton exceedance level $\{0\}$ or the multiple exceedance levels $\{0, 0.5, 1, 1.5\}$. It is interesting to observe that the entropy test statistics decrease monotonically as the firm size increases with the only exception of the 9th decile portfolio. Similar patterns also hold for the asymmetric correlation test statistic. Intuitively, this should be true too. The larger the firm, the more it resembles the market, and hence the asymmetry relative to the market reduces.

[Insert Table 3 about here]

Note that while the P-values of the entropy test are computed based on 399 stationary

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

bootstraps (as discussed in Section 2), the P-values of the asymmetric correlation test are computed from the asymptotic Chi-square distribution with the degree of freedom 1 in the singleton test case and with the degree of freedom 4 for multiple exceedance level case, respectively. For the entropy test, because consistent nonparametric kernel estimation requires a fairly large sample size, we use only the singleton exceedance level. In contrast, estimating correlations do not require as large sample as in the density estimation case. Hence, it is not surprising that the asymmetric correlation test yields similar results with the singleton or with the multiple exceedance levels.

The empirical testing results from both tests for value-weighted book-to-market ratio portfolios are reported in Panel B of Table 3. The entropy test finds that the 9th highest book-to-market ratio portfolio shows significant asymmetric dependence with the market return at 1% level. The S_ρ measure shows a roughly increasing pattern when we go from low to high book-to-market ratio portfolios. In Ang and Chen (2002), they also find an increasing pattern of their H statistic when moving from growth (low book-to-market ratio) stocks to value (high book-to-market ratio) stocks.

Panel C of Table 3 gives the empirical testing results for equal-weighted momentum portfolios. Both tests find significant asymmetry for the return of the highest momentum portfolio (the highest past winner portfolio). The entropy test, in addition, find statistically significant asymmetric dependence in the return of the lowest momentum portfolio (the biggest past loser portfolio). This finding is consistent with Ang and Chen (2002), which shows that bivariate normal model is rejected when fitting to the past loser portfolio returns, i.e. the returns exhibit asymmetric correlations. However, HTZ test fails to detect such asymmetric correlation in the past loser portfolio. We also find vast majority of equal-weighted momentum portfolios shows significant asymmetric dependence at conventional significance levels, except for the one of decile 4. S_ρ increases when we go the either lower or higher ends and is the lowest in the middle deciles. The pattern is consistent with that of the J statistic in HTZ test, but again due to lower power in finite sample, their test fails to attain statistical significance. As we have shown in the simulation results in Section 3, such failure is due to low finite sample power of HTZ test.

5 Asset Pricing Implications of Asymmetry in the Cross-section

In this section, we introduce a downside asymmetric risk measure based on the asymmetric entropy test statistic, and examine its implications in the cross-section of expected stocks returns.

5.1 Downside Asymmetric Dependence

Note that $S_\rho(c)$ is a normalized metric that always takes values in-between 0 and 1, so it gives no direction of asymmetry dependence, i.e. it does not indicate whether the dependence is stronger in the downside or upside. In finance, investors are more concerned about the downside risk of an asset (see, e.g., Ang, Chen, and Xing, 2006). Therefore, we need a measure to distinguish the direction of asymmetric dependence.

Graphically, the degree of concentration of return pairs in a given region reflects the degree of dependence of the two variables in the local area. For example, if the points are more concentrated in the third quadrant than in the first quadrant, it indicates stronger dependence during the downside market. A proxy for the direction of asymmetric dependence can be constructed using joint probabilities of return pairs being in each region. The proxy, excessive downside probability (EDP), can be defined as the difference between a lower quadrant probability (LQP) and an upper quadrant probability (UQP). Specifically, LQP and UQP are given by

$$\mathrm{LQP}^c = Pr(\tilde{x} \leq -c,\ \tilde{y} \leq -c) = \int_{-\infty}^{c}\int_{-\infty}^{c} f(\tilde{x},\tilde{y})\, d\tilde{x} d\tilde{y}, \quad (15)$$

$$\mathrm{UQP}^c = Pr(\tilde{x} \geq c,\ \tilde{y} \geq c) = \int_{c}^{+\infty}\int_{c}^{+\infty} f(\tilde{x},\tilde{y})\, d\tilde{x} d\tilde{y}. \quad (16)$$

They measure probabilities of individual stock and market return pairs being both above or below the exceedance level c. When $c = 0$, higher LQP^0 (UQP^0) indicates higher tendency for the stock to co-move with the market below (above) the average levels, and hence is a good proxy for downside (upside) dependence with the market. The EDP is defined as

$$\begin{aligned} \text{EDP}^c &= \text{LQP}^c - \text{UQP}^c \\ &= \int_{-\infty}^{c} \int_{-\infty}^{c} [f(\tilde{x}, \tilde{y}) - f(-\tilde{x}, -\tilde{y})]\, d\tilde{x} d\tilde{y}. \end{aligned} \tag{17}$$

EDP is a function of exceedance level c. When c is taken to be 0, if $\text{EDP}^0 > 0$, the probability that the asset goes below the mean with the market is greater than the probability that it goes up above the mean with the market, indicating stronger downside dependence. When c equals other values, EDP^c indicates the dependence difference in farther tails. In empirical applications, EDP^c is obtained using kernel estimated cumulative distribution functions proposed by Li, Li, and Racine (2014).[2] Everything else equal, intuitively, most investors dislike the excessive downside probability defined above. From the viewpoint of utility theory, for example, investors with the disappointment aversion (DA) preference, which is introduced by Gul (1991) and excellently analyzed by Ang, Bekaert, and Liu (2005), weight outcomes below a certain reference point strictly more heavily than those above it if the DA coefficient is of a usual value less than 1. In other words, the greater the EDP^c, the more they require to be compensated for.

However, the degree of this asymmetric dependence is not fully reflected by EDP^c. For a fixed value of EDP^c, the greater the $S_\rho(c)$, the greater the asymmetry, and there should be a greater downside risk too. Hence, we define our downside asymmetric dependence risk measure (DownAsy) by

$$\text{DownAsy}^c = Sign(\text{EDP}^c) S_\rho(c), \tag{18}$$

where $Sign(x)$ is a sign function that takes the value of 1 if EDP^c is positive and equals -1

[2] Note that EDP^c can also be consistently estimated by empirical distribution functions. However, the empirical distribution function is a non-smooth step function that jumps up by $1/n$ at each of the n data points. The estimate is mechanically equal to 0 (1) at the sample minimum (maximum), while the true population support may not be bounded by the sample minimum and maximum. The problem is more prominent when the sample size is relatively small.

otherwise. Conditional on an asset has excessive downside probability, $Sign(\text{EDP}^c) = 1$, DownAsy^c indicates the degree of the downside asymmetric dependence risk. If there is no excessive downside probability, $Sign(\text{EDP}^c) = -1$, DownAsy^c indicates the degree of upside potential of the asset when the market is up. Intuitively, investors are likely to require greater risk premium for holding downside risk stocks than holding the upside potential stocks.

It is an empirical matter whether stocks with greater downside asymmetric dependence risk earn higher risk premia. It is interesting to exam the asset pricing implications of downside asymmetric dependence evaluated at the sample means ($c = 0$), as it closely mimics the way how Ang, Chen, and Xing (2006) define downside and upside betas, the conditional linear dependence with the market. The empirical analysis mainly emphasizes on the case of $c = 0$, so the results are directly comparable to Ang, Chen, and Xing (2006). The results of asymmetric dependence measures at farther tails are also reported as robustness checks.

5.2 Data and Empirical Research Design

Stock market data are from the CRSP that cover the sample period from January 1962 to December 2013. The data include all common stocks (with share codes of 10 or 11) listed on NYSE, AMEX and NASDAQ. In order to make the trading volume in NASDAQ comparable to NYSE and AMEX, volumes are adjusted based on the way proposed by Gao and Ritter (2010). Turnover ratio is calculated is calculated as the adjusted monthly trading volume divided by shares outstanding. Amihud (2002) ratio is also computed using the adjusted trading volumes. Following Acharya and Pedersen (2005), I also normalize the Amihud ratio to adjust for inflation and truncated it at 30 to eliminate the effect of outliers. The detailed steps are given in appendix.

The book value information comes from COMPUSTAT and is supplemented by the hand-collected book value data from Kenneth French's web site.[3] The book-to-market

[3]The data are available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

ratio is calculated by the book value of equity (assumed to be available six months after the fiscal year end) divided by current market capitalization. It is truncated at 0.5% percentile and 99.5% percentile to eliminate the effect of extreme values. Following the literature, I take natural logarithm of size, turnover ratio, and book to market ratio before controlling them as firm characteristics.

Following Jegadeesh and Titman (1993), I use returns over past six months to control for the momentum effect. The sample is restricted to stocks with beginning-of-month prices between \$5 and \$1,000 to eliminate stocks whose transaction cost is a huge part of their price and those that have very high prices. I construct β^-, β^+, coskewness and cokurtosis using the definitions given in Ang, Chen, and Xing (2006). Idiosyncratic volatility is calculated as the standard deviation of the CAPM residuals over 12-month horizon. Max is the maximum daily return in a month following the definition of Bali, Cakici, and Whitelaw (2011).

Factor pricing models focus on the contemporaneous risk return relationship. Classical CAPM indicates that stocks that have higher exposure to the market risk earn higher average returns over the same time period. The empirical research design of this paper closely follows Ang, Chen, and Xing (2006); Lewellen and Nagel (2006); Chabi-Yo, Ruenzi, and Weigert (2014) by investigating the contemporaneous relations between the realized risk exposure and realized average returns. This approach may raise some concern that the results may be driven by endogeneity. However, several papers documented that market risk exposures may be time-varying (see,e.g., Fama and French, 1992; Ang and Chen, 2007). In section 4, I also find evidence that the downside asymmetric dependence measure is time-varying, since the past DownAsy is not a good predictor of current DownAsy. Following the approach proposed by Ang, Chen, and Xing (2006), the dependence measures (LQP, UQP and DownAsy) are estimated using realized daily return data over overlapping 12-month periods. The estimates are updated monthly. Since the measures are estimated using non-parametric kernel methods that require sufficient data points for reliable estimates, I restrict the sample so that in each stock 12-month combination there are at least 100 daily observations. Furthermore, using 12-month horizon could better capture the time-varying feature of the dependence measures. Very long time intervals may lead to noisy estimates. Other risk measures (β, β^{-}, β^{+}, Ivol, Coskew, Cokurt) are estimated using the same way. As advocated by Ang, Chen, and Xing (2006); Lewellen and Nagel (2006), such estimation procedure provides greater statistical power with possible time-varying risk measures.

Except for estimating the risk measures, all the empirical asset pricing analyses are done using CRSP monthly frequency data. After applying the data filters, the number of firms in each month over the sample period ranges from 955 to 4364. In the empirical results to follow, all the dependence measures (LQP, UQP and DownAsy) are evaluated at the sample mean ($c = 0$), except for some cases that are specifically denoted.

Table 1 reports time series averages across months of the cross-sectional correlations of main variables, lower quadrant probability (LQP), upper quadrant probability (UQP), downside asymmetric dependence (DownAsy), CAPM beta (β), downside beta (β^{-}), upside beta (β^{+}), log of market capitalization (Size), log of book-to-market ratio (Bm), turnover ratio (Turn), normalized Amihud illiquidity measure (Illiq), past six-month return (Mom), idiosyncratic volatility (Ivol), coskewness (Coskew), cokurtosis (Cokurt) and the maximum daily return over the past one month (max), used in this study. At the beginning of each month t, all risk characteristics (LQP, UQP, DownAsy, β, β^{-} , β^{+}, Ivol, Coskew, Cokurt) are calculated using daily realized stock and market excess returns over the next 12-month period. Size, Bm, Turn, Illiq, Mom and Max are calculated using information available at the end of month $t-1$. All the variables are updated monthly. A detailed description of these variables is in the appendix.

[Insert Table 4 about here]

From Table 1, we can tell how the new dependence measures are linearly related to traditional variables that have explanatory powers on cross-sectional stock returns. The average correlation between LQP and UQP is relatively modest, at -0.08, which means that the tendency of stock to go up or down with the market may appear independently.

This finding also justifies the approach to separately estimate lower and upper quadrant probabilities, which allows for asymmetric dependence in the lower and upper quadrants. As expected, DownAsy has strong positive correlation (0.462) with LQP and negative correlation (-0.515) with UQP, because mechanically the sign of DownAsy coincides with the sign of (LQP$-$UQP). If a variable has similar correlations with LQP and UQP, DownAsy will show little correlation with the variable. Hence we see that DownAsy has almost no correlation with all the other variables. It is more interesting to focus on the correlations with LQP and UQP.

Both LQP and UQP are positively correlated with the CAPM β with correlation coefficients of 0.463 and 0.436 respectively. It is as expected because β captures the linear dependence between individual stock return and market return, while quadrant probabilities measure the general dependence that also captures linear dependence as one component. Stocks with higher β will have high probabilities to be above (below) its sample mean when the market is above (below) average.

On the other hand, β^- and β^+ both have very high positive correlations (around 0.8) with β due to construction. β^- and β^+ are also highly positively correlated with correlation coefficient equal to 0.528. This finding indicates that could β^- (β^+) are not clear measures of downside risk (upside potential). A higher β^- or β^+ is most likely associated with a higher CAPM β. It may explain why Ang, Chen, and Xing (2006) fail to find a negative premium for stocks with high β^+, which is implied by their theoretical representative agent model with disappointment aversion preferences.

Interestingly, size is positively correlated with both LQP and UQP with fairly large correlation coefficients of 0.299 and 0.473 respectively. It indicates that excess returns of larger stocks are more likely to be above (below) the sample mean when market is above (below) the average level. Note that UQP increases more strongly with size than LQP, which indicates that larger stocks have less degree of downside asymmetry than small stocks. It is also confirmed by the negative correlation between size and DownAsy. The finding is consistent with Ang and Chen (2002); Hong, Tu, and Zhou (2007), who find that small size portfolios show stronger asymmetric co-movements with the market using formal statistical tests.

LQP and UQP has little correlation with coskewness, but they have high positive correlations with the fourth co-moment, cokurtosis. The findings with coskewness seems odd, but upon scrutiny, it is not surprising. Just like skewness for univariate distribution, coskewness is more related with length of the tails in a joint distribution. LQP and UQP are measured at the sample mean, where the probability mass is more concentrated. Compared to the probability mass at the center, the probability difference at the tails are much less important. In unreported results, it is shown that LQP and UQP measured at 0.5 and 1 standard deviations away from the sample mean have much higher correlations with coskewness.[4] Cokurtosis measures the fatness of the tails in a given joint distribution. A fatter tail indicates higher probability in that quadrant. It is natural to see that both LQP and UQP are positively correlated with cokurtosis.

5.3 Portfolio Sorts

In this subsection, I study the impact of those dependence measures on the cross-section of average stock returns using simple univariate portfolio sorts.

5.3.1 Univariate Portfolio Sorts

At the beginning of each 12-month period at time t, I sort stocks into five quintile portfolios based on their realized LQP, UQP and DownAsy over the next 12 months. The portfolio returns are also computed as the average realized excess returns over the same 12-month period.

Table 2 shows the contemporaneous relationship between excess returns and LQP (Panel A), UQP (Panel B), and DownAsy (Panel C). Both equal-weighted and value-weighted excess returns and Carhart (1997) four factor adjusted alphas are reported. The row labeled "High - Low" gives the difference between the returns of portfolio 5 and portfolio

[4]The results are available upon request.

1, with corresponding statistical significance levels. Although I use a 12-month horizon, the quintile portfolios are updated at a monthly frequency. Using overlapping information to compute the returns/alphas is more efficient but the 12-month returns/alphas are autocorrelated by construction. To account for the autocorrelations, I report t-statistics of returns/alphas differences computed using Newey and West (1987) heteroskedasticity and autocorrelation consistent (HAC) standardized errors with 12 lags.[5] The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. For robustness checks, I also conduct cross-sectional regression analysis using nonoverlapping yearly periods in the later subsection.

[Insert Table 5 about here]

Panel A of Table 2 shows an increasing pattern between realized LQP and average annualized returns and Carhart alphas. For value-weighted returns/alphas, the increasing pattern is monotonic. In the second column, Quintile 1 (5) shows an average equal-weighted excess return of -5.89% (16.15%) per annum, and the spread in average excess returns is a 22.03% per annum, with a corresponding Newey-West t-statistic of 10.31. In the fourth column, Quintile 1 (5) shows an average value-weighted excess return of -9.16% (8.89%) per annum, and the value-weighted spread is a 18.05% per annum, statistically significant at the 1% level. Return spread is smaller when more weights are given to larger stocks, but the reduction is not much, at about 4% per annum. The results only consider the effect of one variable, while it is shown that LQP is correlated with other variables that also affect returns, such as CAPM β, size and book-to-market ratio. To account for the effects of market, size (SMB), book-to-market (HML), and momentum (UMD) factors, I calculate the equal-weighted and value-weighted alphas from Carhart (1997) four factor model for the quintile and spread portfolios. The results are listed in the third and fifth columns respectively, with equal-weighted (value-weighted) four factor alpha for the spread

[5] Although the theoretical number of lags is 11, to follow the practice of Ang, Chen, and Xing (2006), I use 12 lags. Adding one more lag is more conservative and leads to smaller t-statistics than using 11 lags.

portfolio is equal to 19.89% (16.45%). Both alphas are economically large and statistically significant at the 1% level.

Panel B of Table 2 shows a decreasing pattern between realized UQP and average annualized returns. For Carhart alphas and value-weighted returns, such decreasing pattern is monotonic. Quintile 1 (5) portfolios have an average equal-weighted excess return of 17.33% (-1.79%) per annum, and the spread in average excess returns is a -19.12% per annum. The return difference is statistically significant at the 1% level and cannot be explained by Carhart (1997) four factor model. The four factor equal-weighted alpha spread is -19.00% per annum. The return to the spread portfolio is much lower, at -8.22% per annum, when the returns are weighted by firm's market capitalization. The four factor value-weighted alpha spread is -9.87% per annum. Both value-weighted returns and alphas are still highly statistically significant, at the 1% level. However, the economic significance is much reduced compared to the equal-weighted case.

As pointed out by (Ang, Chen, and Xing, 2006), the reason is due to the fact that asymmetric market risk exposure is bigger among smaller stocks, so (Ang, Chen, and Xing, 2006) choose to focus on equal-weighted results. Previous studies on testing asymmetric correlations (Ang and Chen, 2002; Hong, Tu, and Zhou, 2007) on testing asymmetric dependence all find that smaller stocks tend to co-vary with the market more strongly during the downside market than during the upside market. Such downside asymmetry is statistically significant according to their testing results, but these studies do not find any statistically significant asymmetry in large size portfolios. The findings indicate that the joint return distribution is more symmetric for larger stocks, which further indicates that the correlations between LQP and UQP should be positive among large stocks.Indeed, among top 10 percentile biggest firms, I find that the time series average correlation between LQP and UQP is 0.31, much larger than its full sample counterpart, -0.08. Such large correlation leads to a reduction in the value-weighted return spread, due to the opposite effects of LQP and UQP on contemporaneous returns. For example, a very big firm with high UQP is sorted into quintile 5 portfolios, but it may also have high LQP. High UQP leads to a lower excess return, but high LQP also leads to a higher excess return, so the combined effect makes the stock to earn higher return than the other stocks in the quintile with similar UQP. Value-weighted results put very large weights on such biggest companies. Thus the average value-weighted return for quintile 5 portfolio is higher than the average equal-weighted return. Similarly, very big stock may have both low LQP and UQP, the combined effect makes the average value-weighted returns for lower quintile portfolios to be smaller than the average equal-weighted returns. This is exactly the case as shown in Panel B. A lower value-weighted return spread is also observed in Panel A, but empirically the reduction is not as large as in Panel B.

The patterns shown in Panel A and B are consistent with theoretical predictions of a representative agent model with DA utility (Ang, Chen, and Xing, 2006). The DA preferences allow agents to have greater weights on losses than gains. In equilibrium, a representative agent requires a premium to hold stocks with high downside risk, but is willing to hold stocks with high upside potential at a discount, holding other things equal. Empirically, I find that there is a positive premium for high LQP stocks and a negative premium (discount) for stocks with high UQP. The effect is stronger for LQP than for UQP (22.03% v.s. -19.12% for the equal-weighted case and 18.05% v.s. -8.22% for the value-weighted case).

Given the opposite effects of LQP and UQP on returns and the fact that LQP and UQP are only modestly correlated, we expect the downside asymmetric dependence measure (DownAsy) to be positively associated with returns, since higher DownAsy indicates stronger dependence with the downside market while limited upside potential. Agents dislike this kind of stocks and should require a risk premium to hold them. The risk premium is expected to be larger than that of LQP, because it also combines the effect of UQP. Panel C of Table 2 shows average returns for portfolios sorted by DownAsy. We can see a monotonically increasing pattern between realized DownAsy and average annualized returns as well as Carhart alphas. In the second column, Quintile 1 (5) shows an average equal-weighted excess return of -6.96% (21.21%) per annum, and the spread in average excess returns is a 28.17% per annum, which is statistically significant at the 1% level. The equal-weighted four factor alpha for the spread portfolio is 25.58% per annum. Both excess return and alpha are higher compared to those for the spread portfolio sorted by either LQP or UQP. The fourth and fifth columns show value-weighted results. The average value-weighted excess return (alpha) of the spread portfolio is 12.34% (12.89%) per annum, higher than the UQP return spread, but lower than the LQP spread. The reason is the same as for the UQP sorted portfolios. While the direction and statistical significance of the relationship between the dependence measures and returns hold for both an average stock (equal weighting) or an average dollar (value weighting), the magnitude is smaller with value-weighting. Follow Ang, Chen, and Xing (2006), I will mainly report equal-weighted results in the following analysis.

Since betas are wildly used in the literature as linear dependence measures with the market, for comparison purposes, I also sort stocks into quintile portfolios based on their contemporaneous realized β^- (Panel A), β^+ (Panel B) and $\beta^- - \beta^+$ (Panel C) over 12-month periods. The method and sample used is the same as in Table 2. With a longer sample period and with all stocks listed on NYSE/AMEX/NASDAQ, I have got similar findings as Ang, Chen, and Xing (2006). The results are reported in Table 3.

[Insert Table 6 about here]

Panel A of Table 3 shows a monotonically increasing pattern between realized β^- and average annualized returns/alphas. The average equal-weighted excess return of the spread portfolio in average is a 12.23% per annum, which is statistically significant at the 1% level. However, after accounting for Carhart (1997) four factors, the alpha spread is only 5.49% per annum. Although the downside beta does not exactly reflect the exposure to the market factor, the market, size, boot-to-market and momentum factors can still explain more than half of the excess return difference. Compared to the LQP sorted portfolios, the magnitude of downside beta premium is only about half, the alpha difference is even more prominent (5.49% vs. 19.89%). It is clear evidence that the non-linear downside dependence measure, LQP, can better capture the downside risk than the downside beta. While LQP is also estimated using information from the joint distribution of individual stock and market returns, there is no linear structure involved. It may explain why the market factor, combined with the other three factors, fail to explain much of the excess return to the LQP sorted spread portfolio.

Panel B of Table 3 shows an increasing pattern between realized β^+ and average annualized returns/alphas. As noted in Ang, Chen, and Xing (2006), the pattern is inconsistent with their model predictions and the reason is due to high correlation between β^+ and CAPM β. The results indicate that the upside beta is not a clean measure of upside potential. In comparison, the excess return (four factor alpha) to the spread portfolio sorted by UQP is negative and significant both economically and statistically.

Panel C of Table 3 shows an increasing pattern in average annualized returns/alphas with increasing realized $(\beta^- - \beta^+)$. This measure gauges the effect of downside linear dependence relative to upside linear dependence and can be considered as a linear downside asymmetric dependence measure.[6] Compared to the non-linear downside asymmetric dependence measure, the spread in equal-weighted returns is much smaller (9.13% v.s. 28.17% per annum), and a large portion can be explained by Carhart (1997) four factors. The findings suggest that the entropy-based downside asymmetric dependence measure better captures the asymmetry in market risk exposure.

5.3.2 Dependent Portfolio Sorts

The univariate return patterns could be driven by differences in other risk measures or firm characteristics known to affect contemporaneous returns. As shown in Table 4, LQP and UQP are correlated with some other variables, such as CAPM β, size and cokurtosis. To see a clearer picture of the composition of the other variables across the LQP and UQP sorted portfolios, Table 7 presents summary statistics of the related variables for the stocks sorted into decile portfolios by LQP (Panel A) and UQP (Panel B). Specifically, at

[6] Ang, Chen, and Xing (2006) report the portfolio sorting results based on $(\beta^+ - \beta^-)$. They find a decreasing pattern with a -7.81% equal-weighted excess return for the spread portfolio.

the beginning of each month t, I rank all stocks into decile portfolios based on realized LQP and UQP measures over the next 12 months. The table reports for each decile the time-series average across months of the cross-sectional mean values within each month of the same set of variables as appeared in Table 4.

[Insert Table 7 about here]

From Panel A of Table 4, we can see that there is enough dispersion in LQP across the deciles, with the smallest being 0.224 and largest being 0.370. As we move from the low LQP to the high LQP decile, all three betas increase monotonically. The pattern may raise some concern that the positive risk premium in Table 2 may be driven by higher linear dependence with the market or higher downside beta. We rule out this possibility using dependent portfolio sorts that control for the variations in β in the analysis to follow. As LQP increases across deciles, firm market capitalization (size) increases and illiquidity (Illiq) decreases, indicating that high LQP stocks tend to be larger and more liquid. This is good news for the univariate results reported in Panel A of Table 2, since previous studies have documented that larger (Banz, 1981) and more liquid (Amihud, 2002) stocks tend to earn a return discount, not the return premium observed in the data. The fact that high LQP portfolios contain larger and more liquid stocks but still earn higher average returns works to strength the effect of LQP. It is also observed that cokurtosis (cokurt) is increasing with LQP. Dittmar (2002) document that stocks with higher cokurtosis earn higher average returns. Therefore, the premium of LQP may be explained by the difference in cokurtosis across the deciles. It motivates me to do dependent portfolio sorts with LQP and cokurtosis. The book-to-market ratio (Bm) does not show a clear pattern, although the high LQP portfolios seem to have more growth stocks. There is no clear pattern found for other control variables, such as past six-month return (Mom), idiosyncratic volatility (Ivol), coskewness (Coskew), and maximum daily return in the previous month (Max).

Panel B of Table 4 shows some interesting patterns for decile portfolios sorted by UQP. As we move from the low UQP to the high UQP decile portfolio, the average across months of the mean UQP of stocks increases from 0.201 in the Decile 1 to 0.361 in Decile 10. Similar as the LQP case, all three betas increase monotonically with UQP. This finding works to strength the effect of UQP, because stocks with higher market β or β^{-} tend to have higher average returns instead of the lower average returns indicated by high UQP. Size increases with UQP, which works to weaken the effect of UQP as larger firms tend to earn lower average returns. Illiquidity also decreases with for the high UQP deciles, consistent with these portfolios containing larger stocks. It may also be confounding the effect of UQP, as more liquid stocks tend to have lower returns. The book-to-market ratio decreases with UQP, so it may also explain part of the return discount of high UQP stocks. Past six-month return and cokurtosis show increasing pattern as UQP increases, which strengths the effect of UQP since stocks with high momentum (Jegadeesh and Titman, 1993) and high cokurtosis (Dittmar, 2002) tend to earn higher average returns instead of the observed lower returns. Idiosyncratic volatility seems to decrease with UQP, but the variation in Ivol is not very large. Other variables, like coskewness and Max, do not show a clear pattern.

Those control variables show almost identical co-movement patterns with either LQP or UQP. Since I use excessive downside probability (EDP), defined as (LQP$-$UQP), to determine the sign of downside asymmetric dependence (DownAsy), the opposite patterns given by LQP and -UQP almost cancel out. There is no clear pattern for any of these control variables, when decile portfolios are formed based on realized DownAsy, which means that the return premium due to DownAsy should not be driven by other known characteristics that affect cross-sectional returns. Therefore, the summary statistics for decile portfolios sorted by DownAsy are not reported.[7]

Motivated by the patterns observed in Table 4, I conduct dependent portfolio sorts to explicitly control for the effects of the other stock characteristics that co-vary most with both LQP and UQP, i.e. the CAPM β, size, coskewness, and cokurtosis. I include coskewness in the double sorts mainly due to theoretical consideration, since coskewness is a moment-based measure of asymmetry. Although the linear beta exposure to market

[7]The results are available upon request.

and the size effect can be controlled by looking at the Carhart alphas in the univariate portfolio sorts, dependent portfolio sorts can account for some potential nonlinear impact of these control variables.

At each month, I first form quintile portfolios sorted on each of β, size, coskewness, and cokurtosis, then within each quintile, I further sort stocks into five portfolios based on their realized lower quadrant probability (LQP). The results are reported in Table 8. The row labeled "High - Low" reports the difference between the excess returns of portfolio 5 and portfolio 1 in each β, size, coskewness and cokurtosis first-sort quintile with corresponding statistical significance levels. The column labeled "Average" reports the average excess return of stocks in each second-sort quintile. Newey-West (1987) 12-lag adjusted t-statistics are reported in row labeled "t-stat".[8]

[Insert Table 8 about here]

Panel A of Table 8 reports equal-weighted portfolio excess returns of $\beta\times$ LQP portfolios. Within each β quintile, the return to the high LQP portfolio is larger than the return to the low LQP portfolio. The return spreads are both economically and statistically significant. They range from 24.59% per annum in β quintile 1 to 17.39% per annum in β quintile 5. The average difference in excess returns is 20.57% per annum, only slightly smaller than the return spread in the univariate sorting case. Therefore, market β, although correlated with LQP, can only account for a tiny part of the premium associated with high LQP.

Panel B of Table 8 repeats the same analysis as Panel A, with β being replaced by firm size. Within each size quintile, the equal-weighted return to the "How - Low" portfolio is highly significant both economically and statistically, ranging from 32.51% to 14.26% per annum. The return difference decreases as as we move to high size quintile. As mentioned above, the reason is due to higher correlation between LQP and UQP among larger stocks. It is difficult to purge the effect of UQP from LQP among large size quintile, which leads to a shrunk return spread. Due to the same argument, we also observe similar pattern, among

[8] For tables to fit in one page, I only report results for the quintile 5 and quintile 1 second-sort portfolios, and the "High - Low" portfolio. Detailed results are available upon request.

size $\times$ UQP portfolios in Panel B of Table 9. Despite the decreasing pattern, the average return difference among five size quintiles is still 24.95% per annum and highly significant statistically. The magnitude is even higher than the return spread in the univariate sorting case, which indicates that size cannot explain the risk premium associated with high LQP either.

Panel C of Table 8 reports equal-weighted portfolio excess returns of coskewness $\times$ LQP portfolios. Empirically, I find coskewness is not much correlated with either downside or upside dependence measures. The reason is that coskewness captures more of the asymmetry in the length of the tails, while LQP and UQP are measured at the sample mean and are less sensitive to the probability difference in the tails.[9] Given the low correlation, we do not expect coskewness to account for the risk premium due to LQP. Within each coskewness quintile, the return of the spread portfolio is large and statistically significant at 1% level, with an average spread of 22.09% per annum. Meanwhile, we can confirm that coskewness has negative impact on returns as documented in Harvey and Siddique (2000).

Panel D of Table 8 reports equal-weighted cokurtosis $\times$ LQP double-sorted portfolio excess returns. Within each cokurtosis quintile, the return of the spread portfolio is large and statistically significant at the 1% level, ranging from 29.57% to 10.77% per annum. Although we observe a decreasing return spread as cokurtosis increases, the average spread is 21.89% and is highly significant at the 1% level, indicating that cokurtosis cannot account for the return premium associated with LQP.

Table 9 repeats the same exercises as Table 8, only replacing LQP by UQP. In general, none of β, size, coskewness or cokurtosis can account for the discount for holding stocks with high UQP. Although if anything, firm market capitalization seems to reduce the negative excess return earned by the UQP spread portfolio. Panel B of Table 9 shows that the return spread within size quintile 1 is -19.26%. While within the highest size quintile, the spread is narrowed to -5.48% due to the reason I mentioned above. The average return spread across all size quintiles is -14.89%. The magnitude shrinks compared to -19.12%, the

[9]When measured at 0.5 and 1 standard deviations away from the sample mean, LQP anf UQP do show much higher correlations with coskewness.

return spread in the univariate sorting case. It is consistent with the finding that the value-weighted UQP return spread is much smaller than the equal-weighted UQP return spread in terms of absolute value. Therefore, size can explain a small part of the return discount due to high UQP, but the unexplained part remains to be quite large. β, coskewness and cokurtosis cannot account for any proportion of the return discount due to high UQP.

[Insert Table 9 about here]

Finally, in Table 10, I report the double-sorting results using the downside asymmetric dependence measure (DownAsy) with β (Panel A), size (Panel B), coskewness (Panel C), and cokurtosis (Panel D). By construction, DownAsy reflects the combined pattern of LQP and UQP. For example, the magnitude of the return spread decreases as we move from small to large size quintile for both LQP and UQP second-sort portfolios. We may expect similar pattern in the size $\times$ DownAsy portfolios. Panel B of Table 10 shows that the return spread within size quintile 1 is 30.86% and monotonically decreases to 12.08% in size quintile 5. It is consistent with the findings in the existing literature that small stocks are more exposed to asymmetric downside market risk. The average return difference 23.92%, although is slightly smaller than the return difference in univariate sorts, is still highly significant both economically and statistically. Patterns of return spreads for double-sorted portfolios with β, coskewness and cokurtosis are much similar to the LQP case, as shown in Table 8. None of the control variables can largely explain the risk premium earned for holding stocks with high downside asymmetric dependence.

[Insert Table 10 about here]

In summary, the results of dependent portfolio sorts provide strong evidence that that the risks associated with LQP, UQP and DownAsy are weakly related to size, but clearly are different from risks associated with CAPM β, size, coskewness and cokurtosis. Dependent sorts allow us to control for potential nonlinear impact, but only one other stock characteristic can be controlled for at one time. In the following subsection, I conduct a series of Fama and MacBeth (1973) cross-sectional regressions at the firm level, which allows us to examine the impact of the dependence measures while controlling for many other firm characteristics at the same time.

5.4 Fama-Macbeth Regressions

Following several prior studies (see, e.g., Brennan, Chordia, and Subrahmanyam, 1998; Ang, Chen, and Xing, 2006; Ang, Liu, and Schwarz, 2010; Chabi-Yo, Ruenzi, and Weigert, 2014) that test asset pricing models with individual stock data, I run Fama-MacBeth (1973) regressions at the individual stock level over the sample period from January 1962 to December 2013.[10] I regress stock excess returns on realized dependence measures with respect to the market risk (LQP, UQP, and DownAsy), realized betas (β, β^{-}, and β^{+}) and other firm characteristics using 12-month rolling periods. Since the regressions are run at monthly frequency with a 12-month horizon, I report t-statistics of the estimated coefficients computed using 12 Newey-West (1987) lags. For each month, the risk characteristics (LQP, UQP, DownAsy, β^{-}, β^{+}, Ivol, Coskew, Cokurt) are calculated contemporaneously over the same 12-month period as the excess returns. Log firm size, log book-to-market ratio, turnover ratio, normalized Amihud (2002)illiquidity ratio, past six-month return, and maximum are calculated at the beginning of each month t. All the independent variables are winsorized at the 0.5% and 99.5% levels to avoid some extreme observations driving the results. All the main findings hold no matter I choose to do winsorization or not. Table 11 report the regression results with various sets of control variables. For easier interpretation, the second to last column shows time series averages of cross-sectional mean and standard deviation of each independent variable. To test whether the stock characteristics are still significant after taking the effects of commonly used factors into account, I use 12-month Carhart (1997) four factor adjusted excess return as the dependent variable in regressions

[10]Estimates of risk loadings, such as the realized betas, from individual stock data are less precise than using portfolios as the test assets, which leads to well-known errors-in-variables (EIV) problem. However, Ang, Liu, and Schwarz (2010) argue that with individual stock data, the estimated factor loadings have greater dispersion that reduces the variance of the risk premium estimator and hence is statistically more efficient. Furthermore, Lo and MacKinlay (1990); Lewellen, Nagel, and Shanken (2010) also argue that the method used to form portfolios can lead to very distinct results in asset pricing tests, while using individual stocks as test assets can avoid this arbitrary element in portfolio grouping choice.

(8) and (9). The risk-adjusted returns are used by Brennan, Chordia, and Subrahmanyam (1998) to test factor based asset pricing models. This method avoids the errors-in-variables bias in estimating the risk premia of stock characteristics by putting the factor loadings on the left hand side as the dependent variable. The last column reports the change in 12-month Carhart (1997) four factor adjusted excess return given a one standard deviation increase in the respective independent variable based on regressions (8) and (9).

[Insert Table 11 about here]

Regression (1) and (2) only include LQP and UQP respectively as the explanatory variable to see the univariate effect. Both variables are highly significant economically and statistically with opposite impacts on returns. A one standard deviation increase in LQP is associated with $2.049 \times 0.042 = 8.6\%$ higher average excess returns per annum. UQP has shown a significantly negative impact and a one standard deviation increase in UQP leads to $1.471 \times 0.047 = 6.9\%$ lower average excess returns per annum. In regression (3), I include both LQP and UQP as the independent variables to see the joint effects. The positive (negative) coefficient of LQP (UQP) remains unchanged with even higher economic magnitude. Still we can see that UQP has a smaller impact on returns than LQP, indicating that investors show stronger aversion to the downside risk than preference for upside potential. The findings are consistent with the theoretical predictions of a representative agent model with DA preferences. Regression (4) includes only DownAsy as the explanatory variable. Consistent with the findings in univariate portfolio sorts, the impact of downside asymmetric dependence is negative and statistically significant at the 1% level. A one standard deviation increase in DownAsy is associated with $2.443 \times 0.046 = 11.2\%$ higher average excess returns per annum. The impact is economically more significant than the univariate effect of LQP.

In regression (5), I check the effects of β^- and β^+, the linear counterparts of LQP and UQP. With the sample used in this paper, I can confirm the results from Ang, Chen, and Xing (2006) that downside beta earns a risk premium (5.0% per annum), and upside beta earns a discount (-1.3% per annum), both impacts are statistically significant at the 1% level. The economic magnitude of downside beta premium is much higher compared to the upside beta discount. Regression (6) adds a full set of control variables along with LQP and UQP. The results show that the effects of LQP and UQP are still highly significant with similar magnitudes compared to regression (3). The estimated return premium for bearing one standard deviation downside dependence risk is 11.6% per annum. The impact of upside dependence (UQP) is slightly lower, but still earns -6.3% discount with a one standard deviation increase. In comparison, the effect of upside beta is no longer significant, consistent with the findings in Ang, Chen, and Xing (2006). It indicates that the upside beta may not be a good measure of upside risk, as the results are inconsistent with theoretical model prediction. The findings in regression (6) confirm many patterns that have been documented in the literature. For example, small size stocks and stocks with high book-to-market ratios have high average returns (Fama and French, 1993). Stocks with high past six-month returns earn high average returns during the next 12 months (Jegadeesh and Titman, 1993). Anomaly documented by Ang et al. (2006, 2009) is confirmed that high realized idiosyncratic volatility is associated with low average returns. Less liquid stocks tend to earn lower average returns (Amihud, 2002). Stocks with high coskewness earn low average returns (Harvey and Siddique, 2000) and stocks with positive cokurtosis have high returns (Dittmar, 2002).

In regression (7), I replace the upside and downside dependence measures (LQP and UQP) by the downside asymmetric dependence measure (DownAsy) and include the same set of controls as in regression (6). We can see that the effect of DownAsy is highly significant and the economic magnitude only slightly reduced compared to the univariate regression (4). Interestingly, I find that the effects of Ivol and coskewness are no longer significant statistically after including the entropy-based downside asymmetric dependence measure. Even in regression (6), the effects are not economically significant. The finding suggests that the anomalies due to volatility and coskewness may be explained by the nonlinear dependence with the market risk. However, cokurtosis is still highly significant in both regression (6) and (7).

Regression (8) and (9) use the same controls as in (6) and (7), but replace the dependent variable as Carhart (1997) four factor adjusted return to see whether the characteristics still have explanatory power after accounting for the effects of the four factors. It is clear that the effects of the nonlinear dependence measures are robust even using the risk-adjusted return as the dependent variable. The economic significance of each independent variable is reported in the last column mainly based on regression (9), except for LQP and UQP that are based on regression (8). Among all the explanatory variables, the downside dependence with the market (LQP) has the strongest impact, 9.08% higher adjusted return per annum given a one standard deviation increase. The downside beta, although still negative and significant statistically, has much lower impact (3.68% per annum) on adjusted return, which suggests that LQP is a more accurate measure of downside risk. Downside asymmetric dependence (DownAsy) is also positive and highly significant with 8.59% impact per annum. The upside dependence (UQP) has a significant negative impact of -5.14% per annum on the risk-adjusted return. The magnitude of the discount is much smaller than the risk premium associated with the downside dependence risk or the downside asymmetry risk. The evidence suggests that investors dislike stocks exhibiting strong dependence with the downside market, while prefer stocks with strong upside potential. The aversion to downside risk is stronger compared to the attraction to upside potential.

5.5 Robustness Checks

In this subsection, I run a series of Fama-Macbeth (1973) regressions using different weighting schemes, samples and measures of asymmetric dependence at other exceedance levels to check the robustness of the findings in Table 11. I use Carhart (1997) four factor adjusted excess return as the dependent variable with the full set of controls in these regressions. The results are reported in Table 12.

[Insert Table 12 about here]

Regression (1) and (2) report the value-weighted regression results with full set of controls. The weighting variable is firm's market capitalization at the beginning of each month. The regression coefficients now reflect the impacts for each dollar invested. Similar as the findings in the univariate portfolio sorts, the signs and statistical significance of LQP, UQP and DownAsy remain intact, but the economic magnitude of the impacts are reduced for LQP and DownAsy. Regression (3) and (4) report the regression results when the sample is restricted to NYSE stocks only. Since stocks listed on NYSE tend to be larger size stocks, the findings are similar to value-weighted results. Regression (5) and (6) report the regression results using non-overlapping yearly observations. Using non-overlapping periods are less efficient statistically, but do not cause the returns to be autocorrelated, so the standard t-statistics are reported. The findings are almost the same as the results using overlapping periods, with only small changes to some coefficients.

Regressions (7) to (10) test whether the impacts of those nonlinear dependence measures still hold when they are evaluated at other exceedance levels, such as 0.5 and 1 standard deviations away from the mean. The measures evaluated at farther tails capture the tendency of a stock to move drastically with large market movements and hence are proxies for joint tail risks. The findings are largely consistent with the previous findings when the measured are evaluated at the sample mean. The only exception is that the effect of UQP^1 is no longer statistically significant as shown in regression (9), indicating that investors's attraction to stocks with high upper tail dependence with the market is not robust. On the other hand, the aversion to downside risk is significant and robust at any exceedance level. Similar findings are also documented by Chabi-Yo, Ruenzi, and Weigert (2014).

5.6 Past Downside Asymmetric Dependence and Future Returns

The empirical results documented above demonstrate significant positive relationship between high downside (asymmetric) dependence with the market and the average stock returns over the same period. If the dependence characteristics are stable or predictable over time, then investors can exploit this cross-sectional return relationship and form investable trading strategies based on stocks' asymmetric exposure to the downside risk and upside risk. Since portfolios formed based on contemporaneous DownAsy gives the highest return spread as shown in Table 5. In this section, I examine the time-series persistence of downside asymmetric dependence and check if we can predict such asymmetric downside risk exposure in a future period using prior information.

5.6.1 Determinants of Downside Asymmetric Dependence

I explore the determinants of DownAsy using cross-sectional Fama-MacBeth (1973) regressions. Specifically, at each month, I regress realized downside asymmetric dependence (DownAsy) over the next 12-month period on a set of past risk measures and firm characteristics variables including the lagged DownAsy estimated over the previous 12-month period. At the beginning of each month t, the past risk measures (β^-, β^+, Ivol, Coskew, Cokurt) are estimated over the previous 12-month period ($t-12$ to $t-1$). Size, Bm, Turn, Illiq, Mom and Max are calculated at the beginning of the month t. The regression results are in Table 13.

[Insert Table 13 about here]

We can see that DownAsy is not persistent over time and hard to predict even with all these past variables. When using lagged DownAsy as the only predictor variable, the coefficient 0.058, although highly significant ($t = 6.6$), is far from 1. The corresponding R^2 is lower than 0.01, meaning most variations in current DownAsy cannot be explained by past DownAsy. Size effect is still clear in the predictive setting. Large market capitalization predicts low future downside asymmetric risk. The relationship between current book-to-market ratio and future DownAsy is positive and significant. High current cokurtosis predicts low DownAsy in the future. But the R^2 is only 0.056 even if we put all these past variables as predictors.

An alternative approach is to examine the average 12-month decile portfolio transition matrix, i.e. the average probability $p_{i,j}$ that a stock in decile i during the previous 12-month period will be in decile j in the next 12-month. In the unreported results, I find that stocks in decile 10 according to lagged 12-month DownAsy have 17.64% chance to be in the same decile over the next 12-month period and the chance of staying in the top 3 deciles is 40.69%. It indicates that stocks with high downside asymmetric risk exposure tend to have slightly higher chance to retain that characteristics over the next 12 months compared to the case when DownAsy is totally random.

The findings indicate that there is limited predictability in a stock's asymmetric exposure to the downside risk and support time-varying risk exposures as suggested in Lewellen and Nagel (2006).

5.6.2 Trading Strategy

Although the predictability is limited, yet it is still interesting to examine whether it is possible to generate abnormal return spreads based on past realized DownAsy. At the beginning of each month, I sort stocks into quintile (1-5) portfolios based on their realized DownAsy over the previous 12 months. Then, I examine equal-weighted average returns of these portfolios over the next 12-month period (Panel A) and over the next one month period (Panel B). The data used in this paper range from January 1962 to December 2013. As I use first 12-month data to estimate the first lagged DownAsy, so the first portfolios are formed in January 1963. Then I update those portfolios in a monthly frequency. The results are reported in Table 14 below.

[Insert Table 14 about here]

Panel A of Table 14 shows 12-month holding period returns for portfolios sorted based on lagged DownAsy. Newey-West (1987) standard errors with 12-lags are used to compute the t-statistics (in parentheses) to account for autocorrelations of the returns. In the second column, quintile 1 (5) shows an average equal-weighted excess return of 10.63% (13.86%). The spread in average excess returns is a 3.22% per annum, which is statistically significant at the 1% level. To purge any effect due to exposures to systematic risk factors, I regress the return series of each quintile portfolio and the spread portfolio on the market factor, Fama and French (1993) three factors, and Carhart (1997) four factors respectively. The alphas are reported in the third to fifth columns. CAPM alpha spread is at 2.97% per annum, showing that a small part of the premium can be explained by the market factor. After controlling for the size factor (SMB) and the book-to-market factor (HML), alpha increases to 3.47% per annum. Adding the momentum factor (UMD) reduces the alpha spread to only 1.23% per annum that is marginally significant at the 10% level. It indicates that the part of the return based on the trading strategy is due to exposure to the momentum factor.

We find that DownAsy is not persistent over time. During shorter holding period, DownAsy may change less than during the longer period, so in Panel B, I show 1-month holding period returns for portfolios sorted based on lagged DownAsy. Non-overlapping 1-month returns are usually considered to have no autocorrelations, so the standard t-statistics are reported in parentheses. As expected, the trading strategy of investing in high DownAsy stocks and shorting low DownAsy stocks yields an economically significant one month return of 0.37% per month, which amounts to a compounded return premium of 4.53% per annum. The return difference is also statistically significant at the 1% level. The CAPM alpha spread is at 0.34% per month (4.16% per annum). Adding Fama-French factors increase the alpha spread to 0.38% per month (4.66% per annum). When taking the momentum factor into account, the alpha spread decreases to 0.19% per month (2.30% per annum). Exposure to the momentum factor can explain part of the return spread, but still the four factor alpha spread is statistically significant and economically meaningful.

In summary, DownAsy has limited predictability based on past information. It is difficult to exploit the strong contemporaneous relation between downside asymmetry. Although the return to the spread portfolio formed on lagged DownAsy is smaller than the contemporaneous return spread, it is still economically and statistically significant. In comparison, Ang, Chen, and Xing (2006) find that a trading strategy based on past downside beta using all stocks does not yield an economically significant return spread. It also suggests that the nonlinear asymmetric dependence measures can better capture the downside risk than the downside beta does.

6 Conclusion

Asymmetric dependence in stock returns are important for both portfolio management and risk hedging. However, existing tests focus only on asymmetric correlation, a special case of asymmetric dependence because correlation coefficient is only a measure of linear dependence that ignores higher order dependency. In this paper, we propose an entropy measure, which is model-free and provides a direct test for general asymmetric dependence. Econometrically, our test extends the univariate asymmetry test proposed by Racine and Maasoumi (2007) to the bivariate case that is of interest in finance. With Monte Carlo simulations, we find that the entropy test has good size properties, and has greater power in finite samples than the existing asymmetric correlation test of Hong, Tu, and Zhou (2007). Empirically, based on the entropy test, we find statistically significant asymmetry for most common portfolios, such as those sorted on size, book to market and momentum. In contrast, existing studies only identify a few.

Furthermore, this paper examines whether a stock's nonlinear dependence with the downside and upside market have significant impact on the cross section of stock returns. Using dependence measures constructed with a metric entropy and estimated quadrant probabilities of the joint distribution of stock and the market returns, I find a risk premium (discount) for stocks that are more likely to covary with the market during market declines (rises). The asymmetry between the downside and upside dependence with the market is earns a risk premium as well. The risk premia associated with the downside dependence and downside asymmetric dependence are higher compared to the discount due to upside dependence. The findings suggest that investors' aversion to downside losses are stronger than their attraction to the upside gains, which are consistent with theoretical implications of a representative agent model with disappointment aversion (DA) utility.

Fama-Macbeth (1973) regressions show that the contemporaneous impacts of the dependence measures on cross-sectional returns cannot be explained by a set of well-known stock characteristics, such as the market beta (linear exposure to the market risk), downside or upside betas (asymmetric exposure to downside and upside market risk), size and book-to-market effects, illiquidity risk, momentum effect, coskewness, cokurtosis, and stock's lottery feature as captured by the maximum daily return within a month. The estimated cross-sectional excess return premium for bearing downside dependence risk is approximately 11.6% per annum, almost twice as large as the effect of the downside beta (6% as reported in Ang, Chen, and Xing (2006)). The downside premium, downside asymmetry premium and upside discount are robust across a battery of robustness checks. In addition, I also find that the downside dependence and downside asymmetric dependence measures have low cross-sectional correlations with coskewness, since coskewness captures more of the asymmetry in the lengths of the tails. In Fama-Macbeth (1973) regressions, adding the downside and upside dependence measures (or the downside asymmetric dependence measure) along with the downside and upside betas can empirically rule out the effect of coskewness. The finding suggests that exploring the nonlinear dependence with the market factor may help explain some CAPM anomalies.

The downside asymmetric dependence is not persistent over time and shows limited predictability. However, a trading strategy that forms portfolios based on past asymmetric dependence can still earn an average equal-weighted annualized return of 4.5%. Such a premium is both economically and statistically significant. However, a trading strategy based on downside beta fails to yield a economically meaningful return spread. All the findings suggest that there exist gain when going beyond traditional linear dependence measures. Nonlinear dependence measures may better capture the market risk than their linear counterparts. It is interesting to further explore in theoretical asset pricing models to explain why nonlinear dependence measures can lead to gains in empirical results. This topic will be left for future research.

A Appendix: Variable Definitions

Let us denote a stock i's demeaned daily excess return as $\tilde{r}_{i,d}$, and demeaned daily market excess return as $\tilde{r}_{m,d}$.

CAPM BETA: β is estimated at each month t over the next 12-month, using the following formula

$$\hat{\beta}_{i,t} = \frac{\sum_{d=1}^{d=D_t} \tilde{r}_{i,d}\tilde{r}_{m,d}}{\sum_{d=1}^{d=D_t} \tilde{r}_{m,d}^2}, \tag{A–1}$$

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

DOWNSIDE and UPSIDE BETAS: Denote the sample average of demeaned daily market excess return during a 12-month period starting from month t as $\hat{\mu}_{m,t}$. Further denote demeaned excess return and demeaned market excess return conditional on 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 in Ang, Chen, and Xing (2006),

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

COSKEWNESS: Following Harvey and Siddique (2000), coskewness of stock i over a 12-month period starting at month t is given by

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

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

COKURTOSIS: Cokurtosis of stock i over a 12-month period starting at month t is similarly defined as

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

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

IDIOSYNCRATIC VOLATILITY: Ivol of stock i at the beginning of each month t is defined as the standard deviation of the CAPM residual series over the next 12 months.

SIZE: Following the existing literature, firm size at each month t is measured using the natural logarithm of the market value of equity at the end of month $t-1$.

BOOK-TO-MARKET: Following Fama and French (1992), a firm's book-to-market ratio in month t is calculated using the market value of equity at the end of December of the last year and the book value of common equity plus balance-sheet deferred taxes for the firm's latest fiscal year ending in the prior calendar year.

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 the previous one month skipped, i.e. the cumulative return from month $t-7$ to month $t-2$.

TURNOVER: Turnover ratio is calculated monthly as the adjusted monthly trading volume divided by shares outstanding.

ILLIQUIDITY: Following Amihud (2002), the proxy for the stock illiquidity is from normalizing $L_{i,t} = |r_{i,t}|/dv_{i,t}$. It is the ratio of absolute change of price $r_{i,t}$ to the dollar trading volume $dv_{i,t}$ for stock i at day t. The monthly illiquidity ratios are the daily average of the illiquidity ratio for each stock. To get an accurate estimate of monthly Amihud ratio, we drop the months for stocks if the number of the monthly observations is smaller than 15. Following Acharya and Pedersen (2005), we also normalize the Amihud ratio to adjust for inflation and truncated it at 30 to eliminate the effect of outliers (the stocks with transaction cost larger than 30% of the price).

$$\text{ILLIQ}_{i,t} = \min\left(0.25 + 0.3L_{i,t} \times \frac{\text{capitalization of market portfolio}_{t-1}}{\text{capitalization of market portfolio}_{July1962}},\ 30\right) \quad \text{(A–5)}$$

MAXIMUM: Following Bali, Cakici, and Whitelaw (2011), Max of stock i at month t is defined as the maximum daily excess return within that month.

References

Acharya, V. V., and L. H. Pedersen. 2005. Asset pricing with liquidity risk. *Journal of Financial Economics* 77:375–410.

Alcock, J., and A. Hatherley. 2013. Characterising the asymmetric dependence premium. *Available at SSRN 2476803* .

Amihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5:31–56.

Ang, A., G. Bekaert, and J. Liu. 2005. Why stocks may disappoint. *Journal of Financial Economics* 76:471–508.

Ang, A., and J. Chen. 2002. Asymmetric correlations of equity portfolios. *Journal of Financial Economics* 63:443–94.

———. 2007. Capm over the long run: 1926–2001. *Journal of Empirical Finance* 14:1–40.

Ang, A., J. Chen, and Y. Xing. 2006. Downside risk. *Review of Financial Studies* 19:1191–239.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *The Journal of Finance* 61:259–99.

———. 2009. High idiosyncratic volatility and low returns: International and further us evidence. *Journal of Financial Economics* 91:1–23.

Ang, A., J. Liu, and K. Schwarz. 2010. Using stocks or portfolios in tests of factor models. *Working Paper, Columbia University and University of Pennsylvania* .

Backus, D., M. Chernov, and I. Martin. 2011. Disasters implied by equity index options. *The Journal of Finance* 66:1969–2012.

Backus, D., M. Chernov, and S. Zin. 2014. Sources of entropy in representative agent models. *The Journal of Finance* 69:51–99.

Bae, K.-H., G. A. Karolyi, and R. M. Stulz. 2003. A new approach to measuring financial contagion. *Review of Financial Studies* 16:717–63.

Bali, T. G., N. Cakici, and R. F. Whitelaw. 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99:427–46.

Ball, R., and S. Kothari. 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25:51–74.

Banz, R. W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics* 9:3–18.

Bekaert, G., and G. Wu. 2000. Asymmetric volatility and risk in equity markets. *Review of Financial Studies* 13:1–42.

Brennan, M. J., T. Chordia, and A. Subrahmanyam. 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics* 49:345–73.

Cabrales, A., O. Gossner, and R. Serrano. 2013. Entropy and the value of information for investors. *The American Economic Review* 103:360–77.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge university press.

Carhart, M. M. 1997. On persistence in mutual fund performance. *The Journal of finance* 52:57–82.

Chabi-Yo, F., S. Ruenzi, and F. Weigert. 2014. Crash sensitivity and the cross-section of expected stock returns. *Available at SSRN 2011746* .

Cho, Y.-H., and R. F. Engle. 1999. Time-varying betas and asymmetric effect of news: Empirical analysis of blue chip stocks. Working Paper, National Bureau of Economic Research.

Conrad, J., M. N. Gultekin, and G. Kaul. 1991. Asymmetric predictability of conditional variances. *Review of Financial Studies* 4:597–622.

Davidson, R., and J. G. MacKinnon. 2000. Bootstrap tests: How many bootstraps? *Econometric Reviews* 19:55–68.

Dittmar, R. F. 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance* 57:369–403.

Efron, B. 1982. *The jackknife, the bootstrap, and other resampling plans*, vol. 38. Society for industrial and applied mathematics.

Embrechts, P., A. McNeil, and D. Straumann. 2002. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond* 176–223.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *The Journal of Finance* 47:427–65.

———. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33:3–56.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy* 81:607–36.

Gao, X., and J. R. Ritter. 2010. The marketing of seasoned equity offerings. *Journal of Financial Economics* 97:33–52.

Granger, C., E. Maasoumi, and J. Racine. 2004. A dependence metric for possibly nonlinear processes. *Journal of Time Series Analysis* 25:649–69.

Gul, F. 1991. A theory of disappointment aversion. *Econometrica* 59:667–86.

Hall, P. 1992. The bootstrap and edgeworth expansion. *Springer series in statistics* .

Hansen, L. P. 2012. Dynamic valuation decomposition within stochastic economies. *Econometrica* 80:911–67.

Harvey, C. R., and A. Siddique. 2000. Conditional skewness in asset pricing tests. *The Journal of Finance* 55:1263–95.

Hong, Y., J. Tu, and G. Zhou. 2007. Asymmetries in stock returns: Statistical tests and economic evaluation. *Review of Financial Studies* 20:1547–81.

Hong, Y., and H. White. 2005. Asymptotic distribution theory for nonparametric entropy measures of serial dependence. *Econometrica* 73:837–901.

Horowitz, J. L. 2001. The bootstrap. *Handbook of econometrics* 5:3159–228.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance* 48:65–91.

Kullback, S., R. Leibler, et al. 1951. On information and sufficiency. *The Annals of Mathematical Statistics* 22:79–86.

Künsch, H. R. 1989. The jackknife and the bootstrap for general stationary observations. *Annals of statistics* 17:1217–41.

Lewellen, J., and S. Nagel. 2006. The conditional capm does not explain asset-pricing anomalies. *Journal of Financial Economics* 82:289–314.

Lewellen, J., S. Nagel, and J. Shanken. 2010. A skeptical appraisal of asset pricing tests. *Journal of Financial Economics* 96:175–94.

Li, C., H. Li, and J. S. Racine. 2014. Cross-validated mixed datatype bandwidth selection for nonparametric cumulative distribution/survivor functions. *McMaster University, Working paper* .

Li, Q., and J. S. Racine. 2007. *Nonparametric econometrics: Theory and practice*. Princeton University Press.

Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics* 47:13–37.

Lo, A. W., and A. C. MacKinlay. 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies* 3:431–67.

Maasoumi, E., and J. Racine. 2002. Entropy and predictability of stock market returns. *Journal of Econometrics* 107:291–312.

Maasoumi, E., and J. S. Racine. 2008. A robust entropy-based test of asymmetry for discrete and continuous processes. *Econometric Reviews* 28:246–61.

Nelsen, R. B. 1999. *An introduction to copulas.* Springer.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–08.

Okimoto, T. 2008. New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis* 43:787–815.

Parzen, E. 1962. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33:1065–76.

Patton, A., D. N. Politis, and H. White. 2009. Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. *Econometric Reviews* 28:372–5.

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.

Politis, D. N., and J. P. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89:1303–13.

Politis, D. N., and H. White. 2004. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* 23:53–70.

Racine, J. 1997. Consistent significance testing for nonparametric regression. *Journal of Business & Economic Statistics* 15:369–78.

Racine, J. S., and E. Maasoumi. 2007. A versatile and robust metric entropy test of time-reversibility, and other hypotheses. *Journal of Econometrics* 138:547–67.

Rodriguez, J. C. 2007. Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14:401–23.

Rosenblatt, M. 1956. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics* 27:832–7.

Schwert, G. W. 1989. Why does stock market volatility change over time? *The Journal of Finance* 44:1115–53.

Shannon, C. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27:379–423.

Sharpe, W. F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19:425–42.

Sims, C. A. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50:665–90.

Skaug, H. J., and D. Tjøstheim. 1993. A nonparametric test of serial independence based on the empirical distribution function. *Biometrika* 80:591–602.

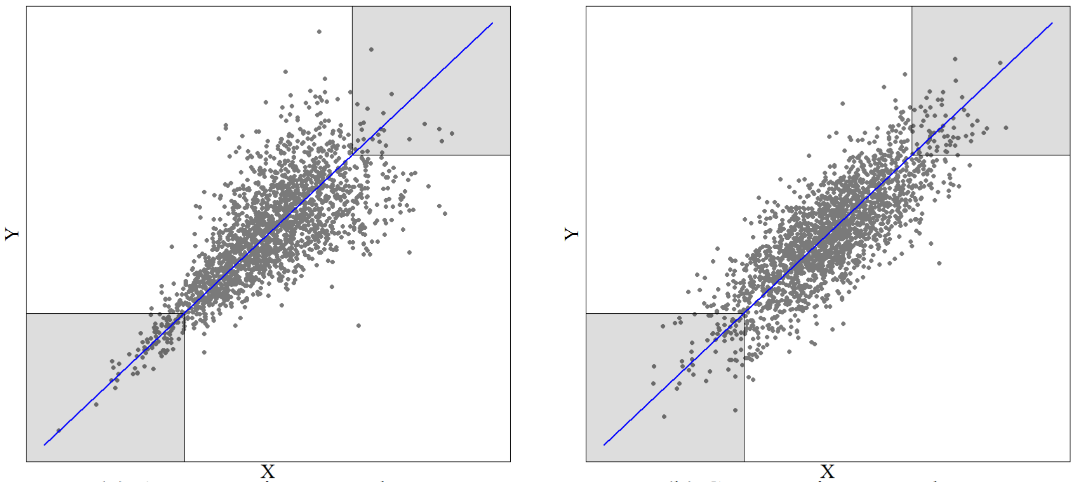
Sklar, M. 1959. *Fonctions de répartition à n dimensions et leurs marges.* Université Paris 8.

Tawn, J. A. 1988. Bivariate extreme value theory: models and estimation. *Biometrika* 75:397–415.

Tjøstheim, D. 1996. Measures of dependence and tests of independence. *Statistics: A Journal of Theoretical and Applied Statistics* 28:249–84.

Ullah, A. 1996. Entropy, divergence and distance measures with econometric applications. *Journal of Statistical Planning and Inference* 49:137–62.

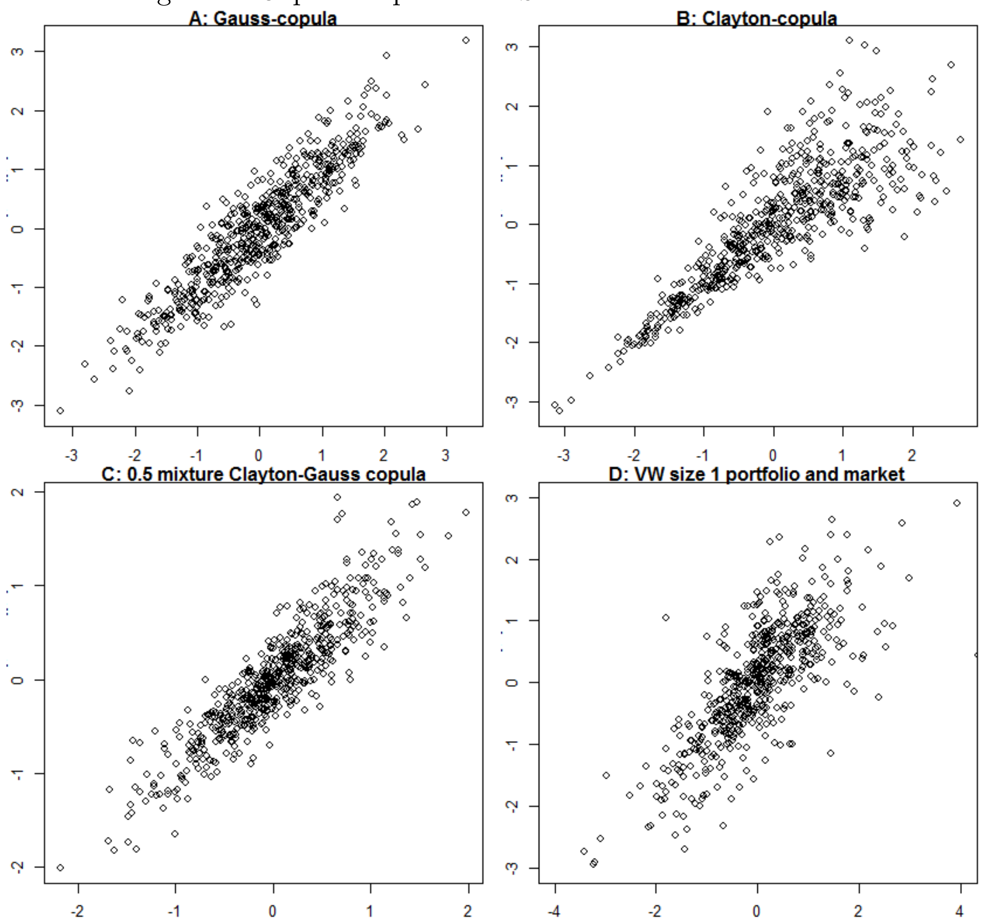
Figure 1: Illustration of Asymmetric Dependence



(a) Asymmetric Dependence (b) Symmetric Dependence

This figure shows two scatter plots from two data-generating processes with different dependence structures. Subfigure (a) shows a scatter plot of of 2000 points generated from a Clayton copula that is known to have stronger left tail dependence than right tail dependence. Subfigure (b) is a similar plot where the DGP is a bivariate normal distribution that has symmetric dependence at both tails. The blue lines in both subfigures are linear regression lines fitted to the data.

Figure 2: Copula Dependence Structures and Data Plots



This figure shows scatter plots of random samples generated by Gaussian copula (A), Clayton copula (B) and mixed Gaussian-Clayton copula with mixing weights of 0.5 each (C), as well as the actual data plots of the value-weighted returns of the smallest size portfolio and the market returns.

Table 1
Size and Power: Entropy-based test and HTZ test

	Entropy-Test	HTZ Test: Asymptotic Theory				HTZ Test: Bootstrap	Power Increase due to Information
	C={0}	C={0}	C={0, 0.5}	C={0, 0.5, 1}	C={0, 0.5, 1,1.5}	C={0}	C={0}
	κ=100% (size)						
T = 240	0.022	0.000	0.000	0.003	0.023	0.039	N/A
T = 420	0.033	0.000	0.000	0.001	0.008	0.045	N/A
T = 600	0.044	0.000	0.000	0.000	0.004	0.043	N/A
	κ=50%						
T = 240	0.094	0.005	0.021	0.046	0.092	0.233	-0.139
T = 420	0.223	0.003	0.022	0.034	0.068	0.323	-0.100
T = 600	0.405	0.014	0.050	0.060	0.088	0.469	-0.064
	κ=37.5%						
T = 240	0.312	0.086	0.086	0.104	0.167	0.423	-0.111
T = 420	0.729	0.215	0.142	0.140	0.176	0.582	0.147
T = 600	0.937	0.426	0.299	0.249	0.263	0.758	0.179
	κ=25%						
T = 240	0.748	0.478	0.325	0.299	0.380	0.549	0.199
T = 420	0.991	0.791	0.618	0.504	0.510	0.725	0.266
T = 600	1.000	0.969	0.867	0.763	0.723	0.854	0.146
	κ=0%						
T = 240	0.952	0.857	0.742	0.690	0.717	0.614	0.338
T = 420	1.000	0.983	0.937	0.895	0.880	0.766	0.234
T = 600	1.000	0.993	0.982	0.972	0.958	0.859	0.141
Avg. Power	0.699	0.485	0.424	0.396	0.419	0.596	0.103 (17.3%)

The nominal size of the tests is set at 5%. The table reports the probabilities of rejecting the null hypothesis of symmetric dependence based on 1,000 Monte Carlo simulations. Different values of κ governs the degree of left tail dependence of the underlying data generating process (DGP). When κ=100%, the DGP is a joint normal distribution and the respective rejecting probabilities are empirical sizes. In all other cases, the rejection probabilities are powers. The last column reports power increases when inferences of both tests are based on 399 stationary bootstraps and the exceedance level is set at {0}. The average power increase is computed as mean of differences among all the simulation cases considered in this paper.

Table 2
Size and Power: Entropy-based test and HTZ test

	Entropy-Test	HTZ Test: Asymptotic Theory				HTZ Test: Bootstrap	Power Increase due to Information
	C={0}	C={0}	C={0, 0.5}	C={0, 0.5, 1}	C={0, 0.5, 1,1.5}	C={0}	C={0}
	κ=100% (size)						
T = 240	0.006	0.000	0.000	0.001	0.009	0.006	N/A
T = 420	0.005	0.000	0.000	0.000	0.000	0.007	N/A
T = 600	0.008	0.000	0.000	0.000	0.000	0.007	N/A
	κ=50%						
T = 240	0.013	0.003	0.003	0.011	0.039	0.074	-0.061
T = 420	0.055	0.000	0.003	0.004	0.024	0.092	-0.037
T = 600	0.129	0.000	0.006	0.009	0.025	0.133	-0.004
	κ=37.5%						
T = 240	0.084	0.010	0.026	0.034	0.074	0.145	-0.061
T = 420	0.380	0.032	0.029	0.035	0.054	0.241	0.139
T = 600	0.733	0.084	0.060	0.069	0.103	0.367	0.366
	κ=25%						
T = 240	0.370	0.159	0.111	0.113	0.204	0.221	0.149
T = 420	0.915	0.415	0.262	0.218	0.259	0.325	0.590
T = 600	0.998	0.740	0.537	0.419	0.401	0.507	0.491
	κ=0%						
T = 240	0.745	0.559	0.466	0.446	0.506	0.273	0.472
T = 420	0.992	0.865	0.754	0.668	0.668	0.437	0.555
T = 600	1.000	0.960	0.922	0.877	0.857	0.659	0.341
Avg. Power	0.535	0.319	0.265	0.242	0.268	0.290	0.245 (84.6%)

The nominal size of the tests is set at 1%. The table reports the probabilities of rejecting the null hypothesis of symmetric dependence based on 1,000 Monte Carlo simulations. Different values of κ governs the degree of left tail dependence of the underlying data generating process (DGP). When κ=100%, the DGP is a joint normal distribution and the respective rejecting probabilities are empirical sizes. In all other cases, the rejection probabilities are powers. The last column reports power increases when inferences of both tests are based on 399 stationary bootstraps and the exceedance level is set at {0}. The average power increase is computed as mean of differences among all the simulation cases considered in this paper.

Table 3
Dependence Symmetry Tests: Size Portfolios

	Entropy-based Test		HTZ Test			
	C={0}		C={0}		C={0, 0.5, 1,1.5}	
Portfolios	Sρ×100	P-value	Test-stat	P-value	Test-stat	P-value
Size 1	2.027	0.010	4.212	0.040	9.715	0.046
Size 2	1.963	0.000	2.049	0.152	3.281	0.512
Size 3	1.868	0.020	0.937	0.333	1.108	0.893
Size 4	1.689	0.013	0.613	0.434	2.095	0.718
Size 5	1.690	0.030	0.431	0.512	5.015	0.286
Size 6	1.596	0.045	0.234	0.629	3.134	0.536
Size 7	1.477	0.065	0.092	0.761	0.849	0.932
Size 8	1.510	0.085	0.099	0.753	0.146	0.997
Size 9	1.695	0.075	0.005	0.945	0.030	1.000
Size 10	1.511	0.055	0.008	0.930	0.029	1.000

Dependence Symmetry Tests: Book-to-Market Portfolios

	Entropy-based Test		HTZ Test			
	C={0}		C={0}		C={0, 0.5, 1,1.5}	
Portfolios	Sρ×100	P-value	Test-stat	P-value	Test-stat	P-value
BE/ME 1	1.248	0.115	0.023	0.880	0.341	0.987
BE/ME 2	1.208	0.085	0.024	0.876	0.093	0.999
BE/ME 3	1.003	0.263	0.060	0.807	0.066	0.999
BE/ME 4	1.626	0.138	0.064	0.800	1.829	0.767
BE/ME 5	1.610	0.055	0.145	0.703	2.769	0.597
BE/ME 6	1.815	0.025	0.054	0.817	1.099	0.894
BE/ME 7	1.805	0.058	0.082	0.774	0.590	0.964
BE/ME 8	1.571	0.098	0.226	0.634	2.954	0.566
BE/ME 9	2.162	0.005	0.447	0.504	1.667	0.797
BE/ME 10	1.657	0.075	0.805	0.370	2.133	0.711

Dependence Symmetry Tests: Momentum Portfolios

	Entropy-based Test		HTZ Test			
	C={0}		C={0}		C={0, 0.5, 1,1.5}	
Portfolios	Sρ×100	P-value	Test-stat	P-value	Test-stat	P-value
L	3.597	0.003	5.191	0.023	6.369	0.173
2	2.572	0.003	2.354	0.125	5.022	0.285
3	2.237	0.018	1.452	0.228	5.407	0.248
4	1.784	0.050	1.008	0.315	5.018	0.285
5	2.155	0.005	0.915	0.339	4.468	0.346
6	1.981	0.003	0.775	0.379	0.921	0.922
7	2.385	0.000	0.717	0.397	3.915	0.418
8	1.959	0.000	1.029	0.311	2.591	0.628
9	2.298	0.000	1.850	0.174	3.507	0.477
W	2.338	0.000	3.329	0.068	13.141	0.011

The sample period is from January 1965 to December 2013. P-values of entropy test are computed based on 399 stationary bootstraps. P-values of Hong et al. (2007) test are computed based on asymptotic Chi-square distribution.

Table 4: Time series average of cross-sectional correlations

	LQP^0	UQP^0	$DownAsy^0$	β	β^-	β^+	Size	Bm	Turn	Illiq	Mom	Ivol	Coskew	Cokurt	Max
LQP^0	1														
UQP^0	-0.08	1													
$DownAsy^0$	0.462	-0.515	1												
β	0.463	0.436	-0.029	1											
β^-	0.3	0.273	-0.036	0.8	1										
β^+	0.356	0.34	-0.011	0.797	0.528	1									
Size	0.299	0.473	-0.079	0.222	0.097	0.233	1								
Bm	-0.071	-0.207	0.066	-0.259	-0.212	-0.202	-0.28	1							
Turn	0.164	0.2	-0.032	0.451	0.382	0.331	0.109	-0.201	1						
Illiq	-0.297	-0.405	0.048	-0.337	-0.233	-0.298	-0.637	0.256	-0.29	1					
Mom	0.009	0.051	-0.019	0.114	0.119	0.072	0.012	-0.237	0.153	-0.061	1				
Ivol	-0.091	-0.203	0.029	0.308	0.336	0.17	-0.525	-0.058	0.232	0.357	0.042	1			
Coskew	0.033	0.024	0.038	0.013	-0.343	0.388	0.038	-0.01	-0.025	0.01	-0.04	0.007	1		
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	
Max	0.01	-0.041	0.009	0.277	0.273	0.178	-0.259	-0.1	0.335	0.129	0.024	0.498	-0.001	-0.021	1

The table reports the averages across months of the cross-sectional correlations of main variables used in this study: lower quadrant probability (LQP), upper quadrant probability (UQP), downside asymmetric dependence (DownAsy), CAPM beta (β), downside beta (β^-), upside beta (β^+), log of market capitalization (Size), log of book-to-market ratio (Bm), turnover ratio (Turn), normalized Amihud illiquidity measure (Illiq), past six-month return (Mom), idiosyncratic volatility (Ivol), coskewness (Coskew), cokurtosis (Cokurt) and the maximum daily return over the past one month (max). At the beginning of each month t, all risk characteristics (LQP, UQP, DownAsy, β, β^- , β^+, Ivol, Coskew, Cokurt) are calculated using daily realized stock and market excess returns over the next 12-month period. Size, Bm, Turn, Illiq, Mom and Max are calculated using information available at the end of month $t-1$. A detailed description of these variables is summarized in the appendix. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013.

Table 5: Univariate Portfolio Sorts: Dependence Measures

Panel A: Lower Quadrant Probability

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	-5.89%	-11.92%***	-9.16%	-13.85%***
2	11.68%	-1.42%	1.37%	-5.99%***
3	17.23%	5.59%***	6.26%	0.19%
4	18.11%	8.39%***	7.84%	1.40%**
5 High	16.15%	7.97%***	8.89%	2.60%***
High - Low	22.03%***	19.89%***	18.05%***	16.45%***
t-stat	(10.31)	(8.56)	(8.11)	(7.72)

Panel B: Upper Quadrant Probability

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	17.33%	10.18%***	12.46%	8.02%***
2	18.74%	8.66%***	10.86%	4.72%***
3	15.61%	3.25%***	9.08%	2.13%***
4	8.61%	-3.11%***	6.30%	-0.42%
5 High	-1.79%	-8.82%***	4.24%	-1.85%***
High - Low	-19.12%***	-19.00%***	-8.22%***	-9.87%***
t-stat	(-12.55)	(-12.06)	(-4.09)	(-5.79)

Panel C: Downside Asymmetric Dependence

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	-6.96%	-12.80%***	-1.56%	-7.54%***
2	8.37%	-3.88%***	5.17%	-1.86%***
3	15.82%	4.08%***	8.38%	1.66%***
4	20.27%	9.89%***	10.61%	4.40%***
5 High	21.21%	12.78%***	10.78%	5.35%***
High - Low	28.17%***	25.58%***	12.34%***	12.89%***
t-stat	(21.45)	(16.54)	(9.46)	(8.84)

This table reports both equal-weighted and value-weighted average annualized returns and Carhart's (1997) four factor alphas of stock portfolios sorted by contemporaneous lower quadrant probability, upper quadrant probability and downside asymmetric dependence evaluated at the mean (exceedance $c = 0$). In each month, I rank stocks into quintile (1-5) portfolios based on the next 12 month realized measures and report the average excess returns over the same 12 months for each portfolio. The row labeled "High - Low" reports the difference between the returns of portfolio 5 and portfolio 1, with corresponding statistical significance levels. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Since the 12-month returns are computed using overlapping periods, the t-statistics computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 6: Univariate Portfolio Sorts: Beta Measures

Panel A: Downside Beta (β^-)

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	7.36%	0.33%	4.36%	-0.28%
2	8.54%	0.73%	5.02%	0.20%
3	10.08%	0.66%	6.45%	0.05%
4	12.94%	1.74%**	8.63%	0.33%
5 High	19.60%	5.82%***	11.24%	1.31%
High - Low	12.23%***	5.49%***	6.88%**	1.59%
t-stat	(4.08)	(2.62)	(2.40)	(0.73)

Panel B: Upside Beta (β^+)

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	10.41%	1.55%*	5.71%	-0.54%
2	10.53%	1.90%**	6.31%	0.18%
3	11.02%	2.03%***	6.46%	1.03%*
4	11.91%	1.47%**	6.38%	0.04%
5 High	13.87%	1.86%*	6.60%	-2.04%
High - Low	3.46%	0.30%	0.89%	-1.50%
t-stat	(1.58)	(0.19)	(0.38)	(-0.68)

Panel C: Downside Beta - Upside Beta ($\beta^- - \beta^+$)

Portfolio	EW Return	CAR-Alpha	VW Return	CAR-Alpha
1 Low	7.67%	-1.32%	3.16%	-2.15%**
2	9.61%	1.15%*	6.69%	1.12%**
3	11.06%	2.08%***	7.30%	1.36%***
4	13.12%	2.90%***	8.59%	0.77%
5 High	16.80%	4.11%***	9.49%	-0.17%
High - Low	9.13%***	5.43%***	6.33%***	1.99%
t-stat	(6.87)	(4.88)	(3.24)	(1.00)

This table reports both equal-weighted and value-weighted average annualized excess returns and Carhart's (1997) four factor alphas of stock portfolios sorted by contemporaneous β^-, β^+ and $\beta^- - \beta^+$. In each month, I rank stocks into quintile (1-5) portfolios based on the next 12 month realized beta measures and report the average excess returns over the same 12 months for each portfolio. The row labeled "High - Low" reports the difference between the returns of portfolio 5 and portfolio 1, with corresponding statistical significance levels. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Since the 12-month returns are computed using overlapping periods, the t-statistics computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 7: Summary Statistics

Panel A: Summary statistics for decile portfolios sorted by LQP

Decile	LQP^0	β	β^-	β^+	Size	Bm	Turn	Illiq	Mom	Ivol	Coskew	Cokurt	Max
1 Low	0.224	0.370	0.531	0.311	4.060	-0.310	0.028	13.366	0.065	0.026	-0.079	0.910	0.052
2	0.265	0.611	0.777	0.526	4.765	-0.498	0.052	8.532	0.094	0.026	-0.090	1.380	0.053
3	0.288	0.743	0.908	0.653	5.175	-0.583	0.066	5.940	0.111	0.025	-0.099	1.668	0.052
4	0.302	0.831	0.984	0.744	5.355	-0.601	0.071	4.805	0.113	0.024	-0.103	1.852	0.052
5	0.314	0.897	1.033	0.817	5.448	-0.600	0.073	4.304	0.111	0.024	-0.101	1.980	0.052
6	0.323	0.955	1.068	0.876	5.514	-0.587	0.075	4.064	0.106	0.024	-0.097	2.068	0.052
7	0.332	1.019	1.107	0.940	5.567	-0.578	0.077	3.895	0.102	0.024	-0.094	2.161	0.053
8	0.341	1.085	1.147	1.002	5.625	-0.569	0.080	3.755	0.098	0.024	-0.091	2.245	0.054
9	0.352	1.158	1.181	1.071	5.700	-0.547	0.082	3.650	0.092	0.024	-0.087	2.327	0.054
10 High	0.370	1.266	1.203	1.184	5.904	-0.531	0.080	3.343	0.084	0.024	-0.073	2.464	0.055

Panel B: Summary statistics for decile portfolios sorted by UQP

Decile	UQP^0	β	β^-	β^+	Size	Bm	Turn	Illiq	Mom	Ivol	Coskew	Cokurt	Max
1 Low	0.201	0.412	0.622	0.336	3.704	-0.158	0.026	15.098	0.049	0.029	-0.081	0.914	0.054
2	0.234	0.624	0.797	0.524	4.232	-0.325	0.045	10.706	0.075	0.028	-0.093	1.325	0.055
3	0.254	0.753	0.901	0.662	4.701	-0.437	0.059	7.369	0.090	0.026	-0.095	1.606	0.054
4	0.270	0.843	0.975	0.754	5.054	-0.511	0.069	5.409	0.098	0.025	-0.096	1.800	0.054
5	0.284	0.914	1.030	0.827	5.311	-0.561	0.073	4.386	0.104	0.024	-0.097	1.940	0.053
6	0.297	0.982	1.079	0.902	5.539	-0.605	0.078	3.569	0.110	0.024	-0.094	2.070	0.053
7	0.309	1.043	1.121	0.969	5.769	-0.650	0.082	3.007	0.113	0.023	-0.092	2.199	0.053
8	0.323	1.081	1.138	1.014	5.952	-0.674	0.083	2.744	0.113	0.022	-0.090	2.301	0.052
9	0.337	1.119	1.144	1.054	6.106	-0.690	0.083	2.678	0.109	0.022	-0.087	2.383	0.051
10 High	0.361	1.151	1.125	1.071	6.327	-0.692	0.077	2.722	0.103	0.021	-0.082	2.474	0.050

This table shows summary statistics of the main variables for decile portfolios formed every month based on realized lower quadrant probability (Panel A) and upper quadrant probability (Panel B). The table reports for each decile the average across the months in the sample of the cross-sectional mean values within each month of the variables–CAPM β, downside beta (β^-), upside beta (β^+), log of market capitalization (Size), log of book-to-market ratio (Bm), turnover ratio (Turn), normalized Amihud illiquidity measure (Illiq), past six-month return (Mom), idiosyncratic volatility (Ivol), coskewness (Coskew), cokurtosis (Cokurt) and the maximum daily return over the past one month (max). At the beginning of each month t, all risk characteristics (LQP, UQP, β, β^-, β^+, Ivol, Coskew, Cokurt) are calculated using daily realized excess returns over the next 12-month period. Size, Bm, Turn, Illiq, Mom and Max are calculated using information available at the end of month $t-1$. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013.

Table 8: Dependent Portfolio Sorts: Lower Quadrant Probability (LQP)

Panel A: Beta (β) and LQP

Portfolio	1 Low β	2	3	4	5 High β	Average
1 Low LQP	-9.21%	-7.66%	-6.72%	-4.55%	1.57%	-5.32%
5 High LQP	15.39%	13.97%	13.87%	14.10%	18.96%	15.26%
High - Low	24.59%***	21.63%***	20.59%***	18.65%***	17.39%***	20.57%***
t-stat	(20.45)	(13.34)	(11.20)	(9.17)	(7.56)	(13.40)

Panel B: Size and LQP

Portfolio	1 Low size	2	3	4	5 High size	Average
1 Low LQP	-6.27%	-9.23%	-8.44%	-6.24%	-2.81%	-6.60%
5 High LQP	26.23%	20.83%	18.03%	15.23%	11.45%	18.36%
High - Low	32.51%***	30.07%***	26.47%***	21.47%***	14.26%***	24.95%***
t-stat	(17.92)	(14.05)	(11.87)	(10.07)	(7.26)	(14.39)

Panel C: Coskewness and LQP

Portfolio	1 Low coskew	2	3	4	5 High coskew	Average
1 Low LQP	-1.75%	-3.68%	-6.36%	-7.46%	-7.51%	-5.35%
5 High LQP	20.89%	19.59%	17.03%	14.58%	11.59%	16.73%
High - Low	22.64%***	23.27%***	23.39%***	22.04%***	19.10%***	22.09%***
t-stat	(9.88)	(11.24)	(11.55)	(11.09)	(9.03)	(11.44)

Panel D: Cokurtosis and LQP

Portfolio	1 Low cokurt	2	3	4	5 High cokurt	Average
1 Low LQP	-10.16%	-9.08%	-6.00%	-2.31%	4.79%	-4.55%
5 High LQP	19.36%	18.04%	17.28%	16.44%	15.56%	17.33%
High - Low	29.57%***	27.12%***	23.28%***	18.75%***	10.77%***	21.89%***
t-stat	(23.96)	(18.16)	(12.31)	(8.82)	(5.00)	(14.24)

This table reports equal-weighted average annualized excess returns of portfolios double-sorted by realized lower quadrant probability (LQP) and realized CAPM β (Panel A), firm market capitalization (Panel B), realized coskewness (Panel C) and realized cokurtosis (Panel D), respectively. LQP is evaluated at the sample mean. For each month, I compute LQP, β, coskewness and cokurtosis using daily realized stock and market excess returns over the next 12 months. Size is computed at the beginning of each month using information at the end of previous month. First, I form quintile portfolios sorted on β, size, coskewness and cokurtosis respectively. Then, I rank stocks within each first-sort quintile into additional quintiles based on LQP. The row labeled "High - Low" reports the difference between the returns of portfolio 5 and portfolio 1 in each β, size, coskewness and cokurtosis first-sort quintile with corresponding statistical significance levels. The column labeled "Average" reports the average return of stocks in each second-sort quintile. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 9: Dependent Portfolio Sorts: Upper Quadrant Probability (UQP)

Panel A: Beta (β) and UQP

Portfolio	1 Low β	2	3	4	5 High β	Average
1 Low UQP	16.26%	19.60%	20.44%	22.55%	26.54%	21.08%
5 High UQP	-7.82%	-7.16%	-5.59%	-2.60%	4.72%	-3.69%
High - Low	-24.37%***	-26.76%***	-26.03%***	-25.16%***	-21.82%***	-24.80%***
t-stat	(-18.61)	(-20.05)	(-19.63)	(-14.02)	(-10.04)	(-17.71)

Panel B: Size and UQP

Portfolio	1 Low size	2	3	4	5 High size	Average
1 Low UQP	19.37%	16.04%	14.73%	12.06%	9.27%	14.29%
5 High UQP	0.11%	-3.66%	-2.65%	-0.58%	3.79%	-0.60%
High - Low	-19.26%***	-19.70%***	-17.38%***	-12.63%***	-5.48%***	-14.89%***
t-stat	(-7.83)	(-9.12)	(-8.54)	(-6.69)	(-3.11)	(-8.57)

Panel C: Coskewness and UQP

Portfolio	1 Low coskew	2	3	4	5 High coskew	Average
1 Low UQP	19.44%	19.11%	17.44%	16.14%	14.58%	17.34%
5 High UQP	3.71%	0.55%	-2.13%	-4.38%	-6.59%	-1.77%
High - Low	-15.72%***	-18.56%***	-19.57%***	-20.52%***	-21.17%***	-19.11%***
t-stat	(-9.12)	(-11.36)	(-11.95)	(-13.42)	(-13.26)	(-12.79)

Panel D: Cokurtosis and UQP

Portfolio	1 Low cokurt	2	3	4	5 High cokurt	Average
1 Low UQP	17.55%	18.51%	18.48%	18.55%	18.40%	18.30%
5 High UQP	-10.86%	-11.32%	-7.95%	-2.89%	6.23%	-5.36%
High - Low	-28.77%***	-29.84%***	-26.43%***	-21.43%***	-12.17%***	-23.67%***
t-stat	(-23.40)	(-18.91)	(-15.57)	(-14.16)	(-8.81)	(-18.53)

This table reports equal-weighted average annualized excess returns of portfolios double-sorted by realized upper quadrant probability (UQP) and realized CAPM β (Panel A), firm market capitalization (Panel B), realized coskewness (Panel C) and realized cokurtosis (Panel D), respectively. UQP is evaluated at the sample mean. For each month, I compute LQP, β, coskewness and cokurtosis using daily realized stock and market excess returns over the next 12 months. Size is computed at the beginning of each month using information at the end of previous month. First, I form quintile portfolios sorted on β, size, coskewness and cokurtosis respectively. Then, I rank stocks within each first-sort quintile into additional quintiles based on LQP. The row labeled "High - Low" reports the difference between the returns of portfolio 5 and portfolio 1 in each β, size, coskewness and cokurtosis first-sort quintile with corresponding statistical significance levels. The column labeled "Average" reports the average return of stocks in each second-sort quintile. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 10: Dependent Portfolio Sorts: Downside Asymmetric Dependence (DownAsy)

Panel A: Beta (β) and DownAsy

Portfolio	1 Low β	2	3	4	5 High β	Average
1 Low DownAsy	-8.55%	-7.77%	-7.46%	-6.28%	-0.28%	-6.07%
5 High DownAsy	16.90%	20.31%	21.77%	24.74%	28.68%	22.48%
High - Low	25.45%***	28.08%***	29.22%***	31.01%***	28.96%***	28.55%***
t-stat	(24.86)	(26.14)	(21.46)	(18.83)	(14.89)	(23.35)

Panel B: Size and DownAsy

Portfolio	1 Low size	2	3	4	5 High size	Average
1 Low DownAsy	-4.74%	-8.18%	-7.24%	-4.92%	-1.13%	-5.24%
5 High DownAsy	26.12%	22.47%	19.21%	14.63%	10.95%	18.68%
High - Low	30.86%***	30.65%***	26.46%***	19.54%***	12.08%***	23.92%***
t-stat	(22.50)	(17.87)	(15.04)	(12.84)	(11.33)	(18.85)

Panel C: Coskewness and DownAsy

Portfolio	1 Low coskew	2	3	4	5 High coskew	Average
1 Low DownAsy	-1.97%	-5.25%	-7.58%	-8.83%	-9.49%	-6.62%
5 High DownAsy	24.89%	24.07%	21.63%	19.31%	16.17%	21.21%
High - Low	26.86%***	29.32%***	29.20%***	28.14%***	25.65%***	27.83%***
t-stat	(16.23)	(21.83)	(21.48)	(21.59)	(19.87)	(22.36)

Panel D: Cokurtosis and DownAsy

Portfolio	1 Low cokurt	2	3	4	5 High cokurt	Average
1 Low DownAsy	-10.28%	-10.61%	-7.89%	-4.00%	3.69%	-5.82%
5 High DownAsy	19.65%	21.83%	22.20%	21.94%	20.46%	21.22%
High - Low	29.93%***	32.44%***	30.09%***	25.94%***	16.77%***	27.03%***
t-stat	(30.12)	(24.50)	(18.90)	(14.53)	(10.44)	(22.06)

This table reports equal-weighted average annualized excess returns of portfolios double-sorted by realized downside asymmetric dependence (DownAsy) and realized CAPM β (Panel A), firm market capitalization (Panel B), realized coskewness (Panel C) and realized cokurtosis (Panel D), respectively. DownAsy is evaluated at the sample mean. For each month, I compute LQP, β, coskewness and cokurtosis using daily realized stock and market excess returns over the next 12 months. Size is computed at the beginning of each month using information at the end of previous month. First, I form quintile portfolios sorted on β, size, coskewness and cokurtosis respectively. Then, I rank stocks within each first-sort quintile into additional quintiles based on LQP. The row labeled "High - Low" reports the difference between the returns of portfolio 5 and portfolio 1 in each β, size, coskewness and cokurtosis first-sort quintile with corresponding statistical significance levels. The column labeled "Average" reports the average return of stocks in each second-sort quintile. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 11: Firm-Level Cross-Sectional Return Regressions

	(1) Return	(2) Return	(3) Return	(4) Return	(5) Return	(6) Return	(7) Return	(8) Car R_{adj}	(9) Car R_{adj}	Mean (Std Dev)	Econ Sig
LQP^0	2.049***		2.738***			2.755***		2.162***		0.311	9.08%
	(11.36)		(12.45)			(15.36)		(15.73)		(0.042)	
UQP^0		-1.471***	-1.706***			-1.339***		-1.093***		0.287	-5.14%
		(-12.68)	(-6.84)			(-4.27)		(-4.19)		(0.047)	
$DownAsy^0$				2.443***			2.330***		1.867***	0.013	8.59%
				(11.25)			(11.10)		(11.19)	(0.046)	
β^-					0.071***	0.071***	0.070***	0.052***	0.052***	0.993	3.68%
					(5.14)	(5.95)	(5.60)	(4.98)	(4.64)	(0.708)	
β^+					-0.017***	-0.009	-0.008	-0.006	-0.005	0.811	-0.39%
					(-2.65)	(-1.23)	(-0.99)	(-0.82)	(-0.63)	(0.783)	
Size						-0.031***	-0.025***	-0.022***	-0.017***	5.265	-2.88%
						(-7.02)	(-6.33)	(-6.31)	(-5.51)	(1.693)	
Bm						0.013**	0.015***	-0.004	-0.002	-0.530	-0.16%
						(2.29)	(2.59)	(-0.88)	(-0.49)	(0.775)	
Turn						-0.470***	-0.423***	-0.383***	-0.349***	0.071	-2.62%
						(-6.65)	(-6.49)	(-5.00)	(-4.65)	(0.075)	
Illiq						0.003***	0.003***	0.002***	0.002***	5.838	1.72%
						(5.68)	(6.46)	(5.07)	(5.85)	(8.591)	
Mom						0.057***	0.056***	0.005	0.003	0.096	0.08%
						(4.01)	(3.83)	(0.34)	(0.26)	(0.277)	
Ivol						-1.869**	-0.845	-1.116*	-0.314	0.024	-0.31%
						(-2.17)	(-0.91)	(-1.73)	(-0.46)	(0.010)	
Coskew						-0.046*	-0.039	-0.037	-0.031	-0.091	-0.51%
						(-1.72)	(-1.36)	(-1.37)	(-1.20)	(0.163)	
Cokurt						0.029***	0.061***	0.024***	0.047***	1.902	5.33%
						(4.00)	(7.49)	(3.25)	(6.10)	(1.133)	
Max						-0.045	-0.032	-0.326***	-0.315***	0.054	-1.13%
						(-0.65)	(-0.43)	(-5.29)	(-4.87)	(0.036)	
Constant	-0.517***	0.539***	-0.211***	0.096***	0.055***	-0.191***	0.130***	-0.212***	0.027		
	(-9.52)	(17.38)	(-3.48)	(4.83)	(2.98)	(-3.36)	(4.63)	(-4.76)	(1.17)		
Obs	1,307,423	1,307,423	1,307,423	1,307,423	1,307,423	1,307,423	1,307,423	1,307,423	1,307,423		
R^2	0.082	0.065	0.131	0.067	0.061	0.248	0.202	0.146	0.116		

This table gives the results of multivariate Fama-MacBeth (1973) regressions. 12-month firm-level excess returns over the risk-free rate are regressed on LQP, UQP, downside asymmetric dependence (DownAsy) all evaluated at the sample mean and a set of other explanatory variables that have been shown to affect cross-sectional stock returns by previous studies. At the beginning of each month t, all risk characteristics (LQP, UQP, DownAsy, β^-, β^+, Ivol, Coskew, Cokurt) are calculated using daily realized stock excess returns and market returns over the following 12-month period. The dependent return variables are computed contemporaneously over the same period. Size, Bm, Turn, Illiq, Mom and Max are calculated at the beginning of each month using information available at the end of month $t-1$. The second to last column displays the time series averages of cross-sectional mean and standard deviation of each independent variable. The last column reports the change in 12-month Carhart (1997) four factor adjusted excess returns for a one standard deviation increase in the respective independent variable based on regressions (8) and (9). The dependent variable is the 12-month excess return in model (1-7) and Carhart (1997) four factor adjusted return in model (8-9). The adjusted returns are calculated following the method suggested by Brennan, Chordia, and Subrahmanyam (1998). The sample includes all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 12: Firm-Level Cross-Sectional Return Regressions: Robustness Checks

	Value-weighted		NYSE only		Nonoverlapping		Other Exceedance levels			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LQP^0	1.555***		1.918***		2.323***					
	(13.77)		(16.78)		(14.19)					
UQP^0	-1.198***		-1.331***		-1.184***					
	(-5.85)		(-5.10)		(-4.29)					
$DownAsy^0$		1.157***		1.705***		1.935***				
		(10.31)		(10.72)		(10.37)				
β^-	0.052***	0.049***	0.063***	0.061***	0.050***	0.049***	0.006	0.032**	0.031***	0.062***
	(4.13)	(3.34)	(5.92)	(5.21)	(3.61)	(3.39)	(0.47)	(2.50)	(2.65)	(4.16)
β^+	-0.002	0.003	-0.017**	-0.014*	-0.004	-0.002	0.026***	0.018**	0.000	0.018**
	(-0.16)	(0.27)	(-2.07)	(-1.78)	(-0.48)	(-0.17)	(3.48)	(2.52)	(0.01)	(2.04)
Size	-0.021***	-0.017***	-0.020***	-0.016***	-0.026***	-0.022***	-0.023***	-0.019***	-0.025***	-0.025***
	(-6.35)	(-5.28)	(-6.50)	(-5.49)	(-7.14)	(-6.30)	(-6.61)	(-6.26)	(-7.07)	(-6.33)
Bm	-0.020***	-0.020***	-0.009**	-0.007	-0.001	0.000	-0.002	-0.000	0.001	-0.004
	(-3.63)	(-3.60)	(-2.09)	(-1.55)	(-0.26)	(0.08)	(-0.37)	(-0.10)	(0.14)	(-0.73)
Turn	-0.204**	-0.173*	-0.314***	-0.295***	-0.599***	-0.552***	-0.408***	-0.365***	-0.401***	-0.311***
	(-2.12)	(-1.85)	(-3.90)	(-3.75)	(-4.87)	(-4.72)	(-5.31)	(-4.76)	(-5.09)	(-4.00)
Illiq	0.004***	0.005***	0.003***	0.004***	0.001**	0.002***	0.002***	0.002***	0.002***	0.005***
	(3.51)	(4.49)	(3.81)	(5.27)	(2.40)	(2.98)	(5.11)	(5.66)	(4.70)	(4.77)
Mom	0.024	0.019	0.000	-0.002	-0.023	-0.024	0.004	0.002	-0.000	0.002
	(1.56)	(1.19)	(0.02)	(-0.10)	(-0.99)	(-1.07)	(0.33)	(0.14)	(-0.00)	(0.11)
Ivol	-6.892***	-6.707***	-3.160***	-2.249***	-0.093	0.634	-0.192	-0.895	0.300	-2.616**
	(-7.81)	(-7.24)	(-3.86)	(-2.72)	(-0.10)	(0.68)	(-0.27)	(-1.19)	(0.40)	(-2.36)
Coskew	-0.070*	-0.054	-0.027	-0.017	-0.022	-0.013	-0.080***	-0.043	0.016	-0.002
	(-1.78)	(-1.45)	(-0.74)	(-0.53)	(-0.59)	(-0.32)	(-2.88)	(-1.54)	(0.62)	(-0.06)
Cokurt	0.007	0.014	0.020**	0.033***	0.032***	0.056***	0.034***	0.043***	0.027***	0.031***
	(0.62)	(1.35)	(2.48)	(3.96)	(3.29)	(6.09)	(4.06)	(5.32)	(3.10)	(3.44)
Max	-0.082	-0.034	-0.251***	-0.236***	-0.536***	-0.520***	-0.341***	-0.298***	-0.343***	-0.311***
	(-0.94)	(-0.37)	(-3.83)	(-3.47)	(-6.12)	(-6.03)	(-5.21)	(-4.39)	(-5.21)	(-4.03)
$LQP^{0.5}$							2.715***			
							(16.68)			
$UQP^{0.5}$							-1.766***			
							(-9.31)			
$DownAsy^{0.5}$								0.432***		
								(11.59)		
LQP^1									2.459***	
									(10.03)	
UQP^1									0.064	
									(0.22)	
$DownAsy^1$										0.090***
										(5.02)
Constant	0.083	0.141***	-0.044	0.063**	-0.215***	0.049*	-0.018	0.058**	0.020	0.113***
	(1.37)	(4.48)	(-0.82)	(2.36)	(-3.79)	(1.76)	(-0.70)	(2.28)	(0.78)	(3.14)
Obs	1,307,423	1,307,423	702,879	702,879	108,675	108,675	1,307,423	1,301,597	1,307,423	823,063
R^2	0.203	0.171	0.162	0.127	0.151	0.120	0.096	0.087	0.087	0.105

This table reports the results of a battery of multivariate Fama-MacBeth (1973) regressions under different specifications for robustness checks. Carhart (1997) four factor adjusted returns is used as the dependent variable in all the 10 regressions. Model 1 and 2 report the value-weighted regression results with full set of controls and the weighting variable is firm's market capitalization. Model 3 and 4 report the regression results with the same specification but the sample is restricted to NYSE stocks only. Model 5 and 6 report the regression results using non-overlapping yearly observations. Model 7 to 10 report the regression results with the same set of controls, but lower quadrant probability (LQP), upper quadrant probability (UQP) and downside asymmetric dependence (DownAsy) are evaluated at other exceedance levels, i.e. 0.5 and 1 standard deviations away from the mean. The sample includes all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1962 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses, except for model 5 and 6 where standard t-statistics are reported. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 13: Determinants of Downside Asymmetric Dependence

	Past Variables											
DownAsy^0	β^-	β^+	Size	Bm	Turn	Illiq	Mom	Ivol	Coskew	Cokurt	Max	R^2
0.058***												0.009
(6.60)												
	-0.003***											0.009
	(-3.82)											
		-0.003***										0.008
		(-4.04)										
			-0.002***									0.017
			(-5.30)									
				0.005***								0.013
				(5.72)								
					-0.05***							0.006
					(-3.68)							
						0.000***						0.012
						(3.21)						
							-0.003**					0.004
							(-2.50)					
								0.066				0.013
								(0.96)				
									0.002			0.004
									(1.24)			
										-0.004***		0.013
										(-4.34)		
											-0.014	0.006
											(-1.08)	
0.046***	0.000	-0.001***	-0.001***	0.003***	-0.027***	-0.000**	-0.003***	0.052	0.011***	-0.002**	-0.019***	0.056
(5.52)	(0.31)	(-2.83)	(-4.09)	(5.56)	(-3.77)	(-2.50)	(-3.43)	(0.78)	(3.43)	(-2.09)	(-5.10)	

The table reports the results of Fama-MacBeth (1973) regressions of realized downside asymmetric dependence (DownAsy) over a 12-month period on a set of past firm characteristics and risk measure variables including the 12-month lagged DownAsy. At the beginning of each month t, the past risk measures (β^-, β^+, Ivol, Coskew, Cokurt) are estimated over the previous 12 months ($t-12$ to $t-1$) that does not overlap with the current 12-month period when the dependent variable is evaluated. Size, Bm, Turn, Illiq, Mom and Max are calculated at the beginning of the month t. The sample includes all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1963 to December 2013, with the last 12-month period starting in January 2013. Newey-West (1987) 12-lag adjusted t-statistics are reported in parentheses. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.

Table 14: Trading Strategy Based on Past Downside Asymmetric Dependence (DownAsy)

Panel A: Portfolios Returns with 12-month Holding Period

Portfolio	Return	CAPM-Alpha	FF-Alpha	CAR-Alpha
1 Low	10.63%	4.89%***	0.51%	1.38%**
2	9.66%	3.57%***	0.48%	1.02%*
3	10.28%	4.16%***	1.05%*	1.17%**
4	11.81%	5.47%***	2.15%***	1.47%***
5 High	13.86%	7.86%***	3.98%***	2.61%***
High - Low	3.22%***	2.97%***	3.47%***	1.23%*
t-stat	(3.73)	(3.35)	(3.52)	(1.78)

Panel B: Portfolios Returns with 1-month Holding Period

Portfolio	Return	CAPM-Alpha	FF-Alpha	CAR-Alpha
1 Low	0.65%	0.15%	-0.12%*	0.02%
2	0.75%	0.17%	-0.03%	0.06%
3	0.85%	0.27%***	0.07%*	0.12%***
4	0.94%	0.36%***	0.16%***	0.16%***
5 High	1.03%	0.49%***	0.25%***	0.21%***
High - Low	0.37%***	0.34%***	0.38%***	0.19%**
t-stat	(4.39)	(4.04)	(4.61)	(2.47)

This table reports equal-weighted average returns and alphas of stock portfolios sorted by past DownAsy evaluated at the sample mean. In each month, I rank stocks into quintile (1-5) portfolios based on the past 12-month realized DownAsy. I report the average excess returns/alphas over the next 12 months for each portfolio in Panel A and the average excess returns/alphas over the next 1 month in Panel B. The row labeled "High - Low" reports the difference between the returns/alphas of portfolio 5 and portfolio 1, with corresponding statistical significance levels. The sample covers all U.S. common stocks traded on the NYSE/AMEX/NASDAQ, and the sample period is from January 1963 to December 2013, with the last 12-month period starting in January 2013. The t-statistics computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags are reported in parentheses for Panel A and standard t-statistics in parentheses for Panel B. *, ** and *** indicate significance levels at 0.1, 0.05 and 0.01 respectively.