Asymmetric Correlations of Equity Portfolios*

Andrew Ang†
Columbia University

Joseph Chen‡
Stanford University

First Version: 9 May, 2000
This Version: 22 January, 2001

Working Paper, Comments Welcome

JEL Classification: C12, C15, C32, G12
Keywords: stock return asymmetries, correlation, dispersion, model bias, GARCH, Jump model, regime-switching

*Any errors or omissions are the responsibility of the authors. The authors wish to thank Larry Glosten, Charlie Himmelberg, Harrison Hong and seminar participants at Columbia University and Ohio State University. We are especially grateful for suggestions from Geert Bekaert and Bob Hodrick. We also thank an anonymous referee whose comments and suggestions greatly improved the paper.

†Columbia Business School, 805 Uris Hall, 3022 Broadway, New York, NY 10027; ph: (212) 854-9154; fax: (212) 662-8474; e-mail: aa610@columbia.edu; WWW: http://www.columbia.edu/∼aa610.

‡Graduate School of Business, Stanford, CA 94305-5015; ph: (650) 723-4877; fax: (650) 725-7462; e-mail: chen.joseph@gsb.stanford.edu; WWW: http://www.stanford.edu/∼chenjs.

Abstract

Correlations between stocks and the aggregate market are much greater for downside moves, especially for extreme downside moves, than for upside moves. Ignoring this asymmetry leads to overestimating the benefits of diversification in portfolio allocation problems in falling markets. We develop a new statistic for measuring, comparing and testing asymmetries in conditional correlations. Conditional on both the downside and the upside, correlations in the data differ from the conditional correlations implied by a normal distribution by around 8.5%. We find that conditional asymmetric correlations are fundamentally different from other measures of asymmetry like skewness and co-skewness, and are inversely related to beta. We find that small stocks, value stocks and past loser stocks have more asymmetric movements. Controlling for size, we find that stocks with lower betas exhibit greater correlation asymmetries and we find no relationship between leverage and correlation asymmetries. Correlation asymmetries in the data overwhelmingly reject the null of multivariate normal distributions at daily, weekly and monthly frequencies. However, several empirical models with greater flexibility, particularly regime-switching models, perform much better at capturing correlation asymmetries.

1 Introduction

Correlations conditional on "downside" movements (when both an equity portfolio and the market fall) are on average over 11.6% higher than implied by a normal distribution, while correlations conditional on "upside" movements (when both an equity portfolio and the market rise) cannot be statistically distinguished from those implied by a normal distribution. Asymmetric correlations are important for several applications. For example, in optimal portfolio allocation problems, if all stocks tend to fall together as the market falls, the value of diversification will be overstated by naïve investors holding portfolios constructed without taking the increase in downside correlations into account. Asymmetric correlations have similar implications on risk management. In this paper we examine this correlation asymmetry in several ways.

We begin by formally defining downside (upside) correlations to be correlations where both the equity portfolio and the market return are below (above) a pre-specified level. Downside correlations are much larger than those on the upside, as shown by Longin and Solnik (2001) plots of downside and upside correlations. These graphs dramatically demonstrate that on the downside, portfolios are much more likely to move together with the market than on the upside.

Second, we measure this asymmetry by developing a summary statistic H. The H statistic quantifies the degree of asymmetry in correlations across downside and upside markets relative to a particular model or distribution. This measurement of asymmetry is different from that established in the literature. Covariance asymmetry has usually been interpreted within a particular GARCH model, where covariance asymmetry is defined to be an increase in covariance resulting from past negative shocks in returns.[1] In contrast, our statistic measures correlation asymmetry by looking at behavior in the tails of the distribution. Our statistic is not model-specific, hence we can apply the statistic to evaluate several different models. We show that these conditional correlations are different from other measures of third moments (skewness and co-skewness), and risk measured by beta.

The H statistic corrects for a bias induced by the conditioning information. Boyer, Gibson and Loretan (1999), Forbes and Rigobon (1999) and Stambaugh (1995) note that calculating correlations conditional on high or low returns, or high or low volatility, induces a conditioning bias in the correlation estimates. For example, for a bivariate normal distribution with a given unconditional correlation, the conditional correlations calculated on joint upside or downside moves will be different from the unconditional correlation. Ignoring these conditioning biases may lead to spurious findings of correlation asymmetry.

Third, we establish several empirical facts about asymmetric correlations in the US equity market. We find the extent of the asymmetry at the daily, weekly and monthly frequencies

[1] Authors such as Cho and Engle (2000), Bekaert and Wu (2000), Kroner and Ng (1998) and Conrad, Gultekin and Kaul (1991) document the covariance asymmetry of domestic stock portfolios using multivariate asymmetric GARCH models.

and overwhelmingly reject the null of a normal distribution. To investigate the nature of these asymmetric movements we examine the magnitudes of correlation asymmetries using portfolios sorted on various characteristics. Returns on portfolios of small firms, value firms and those of past loser stocks exhibit more correlation asymmetry. We find significant correlation asymmetry in traditional 'defensive' sectors such as petroleum and utilities and that riskier stocks measured by higher beta have lower correlation asymmetry than lower beta stocks. The magnitude of correlation asymmetry is unrelated to the leverage of a firm after controlling for size. Previous work has analyzed asymmetric movements only of leverage-sorted portfolios of Japanese stocks (Bekaert and Wu (2000)), and US size-sorted portfolios (Kroner and Ng (1998) and Conrad, Gultekin and Kaul (1991)).

Finally, we try to explain asymmetric correlations by asking if several reduced-form empirical models of stock returns can reproduce the asymmetric correlations found in the data. These candidate models have been used by various authors to capture the increase in covariances on the downside. We discuss four models which allow asymmetric movements between upside and downside movements in returns: an asymmetric GARCH-M model, a Poisson Jump model where jumps are layered on a bivariate normal distribution, a regime-switching Normal model, and a regime-switching GARCH model. We find the most successful models in replicating the empirical correlation asymmetry are regime-switching models. However, all models leave unexplained some amount of asymmetries in correlations.

Our study of asymmetric correlations is related to several areas of finance. There is a long literature documenting the negative correlation between a stock's return and its volatility of returns.[2] Other studies analyze patterns of asymmetries in the covariances of stock returns in domestic equity portfolios.[3] This literature has concentrated on documenting covariance asymmetry within a GARCH framework. Our approach uses a different methodology to document asymmetric correlations, interpreting asymmetries more broadly than just within the class of GARCH models. We examine a much wider range of portfolio groups to characterize the nature of asymmetric correlations than previously used in the literature. We also seek to explain empirical correlation asymmetry with other classes of empirical models.

Our approach of creating portfolios sorted by different firm characteristics obtains a very different view of the determinants of conditional correlations than previously done in the literature. The H statistic uses the full sample (time-series) of observations to calculate the cor-

[2] For example see, among others, French, Schwert and Stambaugh (1987), Schwert (1989), Cheung and Ng (1992), Campbell and Hentschel (1992), Glosten, Jagannathan and Runkle (1993), Engle and Ng (1993), Hentschel (1995), and Duffee (1995). Bekaert and Wu (2000) provide a summary of recent GARCH model applications with asymmetric volatility.

[3] Some papers documenting asymmetric betas are Ball and Kothari (1989), Braun, Nelson and Sunier (1995) and Cho and Engle (2000). Conrad, Gultekin and Kaul (1991), Kroner and Ng (1998) and Bekaert and Wu (2000) document asymmetric covariances in multivariate GARCH models.

relation at the extreme tails of the joint-distribution. This means we allow the use of as many observations as possible to calculate correlations for events where there are naturally relatively few observations. We focus on the cross-sectional determinants of correlation asymmetry in stock returns whereas Erb, Harvey and Viskanta (1994) and Dumas, Harvey and Ruiz (2000) use conditioning not on the distribution, but on instrumental variables such as business cycle indicators, to determine the characteristics of time-varying correlations.

Work in international markets has found that the correlations of international stock markets tend to increase conditional on large negative (or bear market) returns.[4] Longin and Solnik (2001) use extreme value theory to show that the correlation of large negative returns is much larger than the correlation of positive returns. However, Longin and Solnik do not give distribution-specific characterizations of downside and upside correlations. Our paper highlights that strong correlation asymmetries exist in domestic markets and is not just an international phenomenon in aggregate markets. In our domestic focus we examine which indivdiual firm characteristics are most related to the magnitude of correlation asymmetry.

Other related studies by Campbell et als (2001), Bekaert and Harvey (2000) and Duffee (1995) examine cross-sectional dispersion of individual stocks, which has increased in recent periods. Duffee (2000) and Stivers (2000) document an asymmetric component in the cross-sectional dispersion. Chen, Hong and Stein (2000) and Harvey and Siddique (2000) analyze cross-sectional differences in conditional skewness of stock returns. However, these authors have not examined the relationships between firm characteristics and asymmetric movements, measured by correlations. We find stocks which are smaller, have higher book-to-market ratios and have low past returns exhibit greater asymmetric movements. Stocks with higher beta risk show fewer correlation asymmetries.

The remainder of this paper is organized as follows. Section 2 demonstrates the economic significance of asymmetries in correlations within a portfolio allocation framework. Section 3 visually shows that correlation asymmetries exist in domestic US equity data. We characterize the conditional correlations of a bivariate normal in closed-form and discuss how to correct for conditioning bias. Section 4 measures the correlation asymmetries, and analyzes their cross-sectional determinants. Here, we develop the H statistic measure of correlation asymmetry and present overwhelming evidence of asymmetric correlations in equity portfolios using the normal distribution as the benchmark. In Section 5 we ask if several models incorporating asymmetry into the conditional covariance structure can replicate the asymmetry found empirically in the data. Section 6 concludes. Proofs and the construction of the data are found in the Appendix.

[4] See Erb, Harvey and Viskanta (1994), Lin, Engle and Ito (1994), Longin and Solnik (1995, 2001), Karolyi and Stulz (1996), De Santis and Gerard (1997), Das and Uppal (1999), Forbes and Rigobon (1999), Boyer, Gibson and Loretan (1999), Stărică (2000), Ang and Bekaert (2000) and Bae, Karolyi and Sultz (2000).

2 Economic Significance of Asymmetric Correlations

In this section we demonstrate the economic significance of asymmetric correlations using a simple asset allocation problem. Appendix B details the solution and the calibration method used in this example. Suppose an investor can hold amounts α_1 and α_2 of two assets with expected excess returns of x and y. The remainder of her wealth is held in a riskless asset. Let $\tilde{x}$ and $\tilde{y}$ denote the standardized transformations of x and y, respectively.[5] The agent maximizes her expected end-of-period CRRA utility:

$$\max_{\alpha_1,\alpha_2} \mathrm{E}\left[\frac{W^{1-\gamma}}{1-\gamma}\right]$$

where end-of-period wealth is given by $W = 1 + r_f + \alpha_1 x + \alpha_2 y$, $r_f = 0.05$ is a constant risk-free rate and γ is the agent's coefficient of risk aversion. We set $\gamma = 4$.

To abstract from the effects of means and variances on portfolio weights, suppose both assets have the same mean and volatility. Denote the expected excess return of both x and y as $\mu = 0.07$ and the volatility of the excess return as $\sigma = 0.15$. We set the unconditional correlation of x and y to be $\rho = 0.50$.

Suppose that the agent believes x and y are normally distributed. Since each asset has the same mean and volatility, the investor holds equal amounts of either asset. Let $\alpha^\dagger$ denote this portfolio position. With normal distributions, lower unconditional correlations imply greater benefits from diversification.

We examine the joint behavior of the two assets conditional on bear-market or downside moves. We define this downside move to be a draw that is below each asset's mean by more than one standard deviation. If x and y are normally distributed with unconditional correlation $\rho = 0.5$, the correlation conditional on $x < \mu - \sigma$ and $y < \mu - \sigma$ is:

$$\bar{\rho} = \mathrm{corr}(x, y | x < \mu - \sigma,\ y < \mu - \sigma) = \mathrm{corr}(\tilde{x}, \tilde{y} | \tilde{x} < -1,\ \tilde{y} < -1) = 0.1789.$$

Note that the downside correlation for a normal distribution is less than the unconditional correlation. Appendix A shows how to calculate this conditional correlation in closed-form.

Suppose the actual distribution of x and y is a Regime-Switching (RS) Model while the agent believes erroneously that x and y are normally distributed. Under the RS Model, returns $X = (x,\, y)$ are given by:

$$X \sim N(\mu_{s_t}, \Sigma_{s_t}), \quad s_t \in \{1, 2\}. \tag{1}$$

[5] To standardize a variable x, we perform the transformation $\tilde{x} = (x - \mu)/\sigma$ where μ is the unconditional mean of x, and σ is the unconditional standard deviation of x. We use tildes to denote standardized returns throughout the paper; variables without tildes are not standardized.

We denote μ_i as the mean returns in regime $s_t = i$ and Σ_i as the covariance matrix in regime $s_t = i$. The transitions between the regimes $s_t = 1$ and $s_t = 2$ are given by a Markov chain with transition probabilities:

$$\begin{pmatrix} P & 1-P \\ 1-Q & Q \end{pmatrix}$$

where $P = Pr(s_t = 1|s_{t-1} = 1)$ and $Q = Pr(s_t = 2|s_{t-1} = 2)$. We calibrate the RS Model to have the same unconditional mean μ, the same unconditional volatility σ and the same unconditional correlation ρ as the normal distribution.

Instead of the downside correlation $\bar{\rho}$ being 0.1789, suppose that the true downside correlation $\bar{\rho}$ is H percent higher. That is:

$$\text{corr}(\tilde{x}, \tilde{y}|\tilde{x} < -1,\ \tilde{y} < -1) = 0.1789 + H.$$

This magnitude H reflects the statistic we later develop. This is an effect which cannot be captured by using the normal distribution which is determined only by its first two moments. However, the increase in correlation on downside moves relative to the normal distribution can be captured by the RS Model.

Let the RS Model have parameters $\mu_1 = \mu_2 = (0.14,\ 0.14)'$, $P = 2/3$, and $Q = 2/3$. This implies that the stable probabilities of the Markov Chain $\pi = Pr(s_t = 1) = \frac{1}{2}$. We express the covariance matrices Σ_i as:

$$\Sigma_i = \sigma_i^2 \begin{pmatrix} 1 & \rho_i \\ \rho_i & 1 \end{pmatrix}, \quad i = 1, 2$$

where ρ_i is the correlation of returns in regime i. We set $\sigma_1 = \sigma_2 = 0.15$ to concentrate only on the effect of regime-dependent correlations. The correlations ρ_1 and ρ_2 are chosen so that the RS Model has the same unconditional correlation ρ as the normal distribution. We choose $\rho_1 > \rho_2$ such that $\frac{1}{2}(\rho_1 + \rho_2) = \rho$. Hence the RS Model has the same first two unconditional moments as the normal distribution, but its correlation conditional on downside moves is higher than what the normal distribution implies.

The asset allocations from the RS Model are regime-dependent. Since x and y have the same moments, the optimal holdings in each asset will be the same but the proportion held in x and y will differ across the regimes. Denote the optimal portfolio holdings in each asset as $\alpha^*_{s_t}$ for regime s_t. In regime 1 with the higher $\rho_1 > \rho$, the investor, who holds weights $\alpha^\dagger$ thinking x and y are normally distributed, will be holding too high a proportion of equity compared to the optimal holding α^*_1. In regime 2 with $\rho_2 < \rho$ the investor is holding too little equity compared to the optimal α^*_2. The higher ρ_1 in the first regime will cause downside correlations to increase relative to the normal distribution. Since the normal distribution cannot incorporate the asymmetries in conditional correlations, the investor overestimates the benefits of diversification on the downside in regime 1 and overinvests in risky assets. Similarly, she underestimates the benefits of diversification in regime 2 and underinvests in risky assets.

We can calculate the utility loss, or monetary compensation required for an investor to use the non-optimal normal weights $\alpha^\dagger$ instead of the optimal RS Model weights $\alpha^*_{s_t}$. This is the ex-ante compensation in cents per dollar of wealth the investor should have received in order to hold $\alpha^\dagger$ instead of $\alpha^*_{s_t}$. This is given by $w = 100 \times (\bar{w} - 1)$ where:

$$\bar{w} = \left(\frac{Q^*_{s_t}}{Q^\dagger_{s_t}} \right)^{\frac{1}{1-\gamma}}.$$

$Q^*_{s_t}$ is the indirect CRRA utility under the RS Model with optimal weights $\alpha^*_{s_t}$ conditional on being in regime s_t and $Q^\dagger_{s_t}$ is the indirect CRRA utility under the RS Model distribution with sub-optimal weights $\alpha^\dagger$ conditional on being in regime s_t. That is:

$$Q^*_{s_t} = \mathrm{E}[(W^*_{s_t})^{1-\gamma}|s_t], \quad Q^\dagger_{s_t} = \mathrm{E}[(W^\dagger)^{1-\gamma}|s_t]$$

where $W^*_{s_t} = 1 + r_f + \alpha^*_{s_t} x + \alpha^*_{s_t} y$, $W^\dagger = 1 + r_f + \alpha^\dagger x + \alpha^\dagger y$, and both expectations are taken under the RS Model.

Figure (1) graphs the ex-ante monetary compensation the investor should have received for using the sub-optimal normal distribution weights instead of the optimal RS Model weights. The compensation required per dollar of wealth is not small: for $H = 0.10$ over 100 basis points in compensation is required in regime 1, while regime 2 requires a compensation of 70 basis points.

This simple example shows that potential ex-post losses and ex-ante utility losses is economically significant if correlations increase on the downside relative to a standard normal distribution. In Figure (1), the H measures the difference between the true downside correlation and what a normal distribution implies. We now formally develop the H statistic, show and correct for a bias in measuring it, and use it to characterize the nature of asymmetric correlations in stock portfolios.

3 Calculating Upside and Downside Correlations and Betas

3.1 Upside and Downside Correlations

Conditioning on upside or downside moves and calculating correlations induces a “conditioning bias”. For a bivariate normal with unconditional correlation ρ, the correlation calculated conditioning on a subset of observations (for example taking observations above or below a certain level) will differ from the unconditional correlation. Appendix A calculates this bias in closed form for a bivariate normal distribution.[6] In this section we show that the conditioning bias for

[6] Related work by Forbes and Rigobon (1999) looks at the correlation of returns conditioning on different volatilities. Boyer, Gibson and Loretan (1999) derive correlations for a bivariate normal conditioning on events for one variable. Stambaugh (1995) in an NBER discussion of Karolyi and Stulz (1996) demonstrates by simulation this conditioning bias.

a bivariate normal distribution is non-negligible, and hence not taking the conditioning bias into account can result in incorrect inference when testing for correlation asymmetry.

Take a pair of standardized returns $(\tilde{x},\, \tilde{y}) \sim N(0, \Sigma)$, where Σ has unit variances and unconditional correlation ρ. We define:

$$\hat{\rho}(h_1, h_2, k_1, k_2) = \text{corr}(\tilde{x}, \tilde{y} | h_1 < \tilde{x} < h_2, k_1 < \tilde{y} < k_2; \rho) \tag{2}$$

as the correlation between $\tilde{x}$ and $\tilde{y}$ conditional on observations for which $h_1 < \tilde{x} < h_2$ and $k_1 < \tilde{y} < k_2$, where $\tilde{x}$ and $\tilde{y}$ have unconditional correlation ρ. (This is the correlation of a doubly truncated bivariate normal.)

There are several special cases of this setup. When $h_2 = \infty$ and $k_2 = \infty$ we obtain Rosenbaum (1961)'s one-sided truncation case. Another special case is the Longin and Solnik (2001) exceedance correlation. An exceedance correlation at an exceedance level ϑ is defined as the correlation between two variables when both variables have made moves of more than (or more negative than) ϑ standard deviations away from their means:

$$\bar{\rho}(\vartheta) = \begin{cases} \hat{\rho}(\vartheta, \infty, \vartheta, \infty) & = & \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} > \vartheta, \tilde{y} > \vartheta; \rho) \text{ if } \vartheta \geq 0 \\ \hat{\rho}(-\infty, \vartheta, -\infty, \vartheta) & = & \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} < \vartheta, \tilde{y} < \vartheta; \rho) \text{ if } \vartheta \leq 0 \end{cases} . \tag{3}$$

For a bivariate normal these are the same, by symmetry. Longin and Solnik discuss the limiting behavior of exceedance correlations using extreme value theory, but do not give distribution-specific characterizations of exceedance correlations.

For an exceedance level ϑ, we calculate the empirical exceedance correlation $\bar{\rho}(\vartheta)$ as follows. For pairs of standardized observations $\{(\tilde{x},\, \tilde{y})\}$ we select a subset of observations $\{(\tilde{x},\, \tilde{y}) | \tilde{x} > \vartheta \text{ and } \tilde{y} > \vartheta\}$ for $\vartheta \geq 0$ and $\{(\tilde{x},\, \tilde{y}) | \tilde{x} < \vartheta \text{ and } \tilde{y} < \vartheta\}$ for $\vartheta \leq 0$. The correlation of the observations in this subset is the empirical exceedance correlation at ϑ. For $\vartheta = 0$, we calculate both $\text{corr}(\tilde{x}, \tilde{y} | \tilde{x} > 0, \tilde{y} > 0)$ and $\text{corr}(\tilde{x}, \tilde{y} | \tilde{x} < 0, \tilde{y} < 0)$. In calculating the exceedance correlation, ϑ determines how many standard deviations from the empirical means observations must lie in order to be included in the conditioning sample. For $\vartheta = -1$, the exceedance correlation of an equity portfolio and the market is calculated on a subset where both the equity portfolio and the market return are more than 1 standard deviation below their empirical means.

The top plot of Figure (2) shows graphs of conditional correlations of a bivariate normal, where the conditioning is done on returns above or below a certain level. The top plot shows exceedance correlations $\bar{\rho}(\vartheta)$ for various unconditional ρ.[7] For a given ρ, the exceedance correlations are tent-shaped. Intuitively, the exceedance correlations tend to zero as $\vartheta \to \pm\infty$ because the tails of the bivariate normal are very flat. The exceedance correlations are calculated conditioning on a quadrant of $\tilde{x}$ and $\tilde{y}$ with origin at the point $(\vartheta,\, \vartheta)$. As ϑ increases the quadrant is pushed further into the tails of the bivariate normal where the distribution becomes flatter. Hence the exceedance correlations tend to zero as $\vartheta \to \pm\,\infty$.

[7] These are calculated using Equation A-4 in Appendix A.

The exceedance correlation $\bar{\rho}(\vartheta)$ is always less than the unconditional correlation ρ. Table (1) reports the difference $\bar{\rho}(\vartheta) - \rho$. Even when conditioning on moves of only 1 standard deviation, the difference between the exceedance correlation and the unconditional correlation can be over 20%. One way to determine if correlations are different for upside ($\vartheta > 0$) or downside ($\vartheta < 0$) moves is to compare positive or negative exceedance correlations in data with those implied from a particular distribution, such as the normal distribution. Figure (2) and Table (1) show that comparing correlations conditioning on high or low absolute returns cannot be done without taking into account the conditioning bias.

We can also construct correlations conditioning only on levels of one variable, $\tilde{x}$. The bottom plot in Figure (2) shows conditional correlations $\hat{\rho}(h_1, h_2, -\infty, +\infty) = \text{corr}(\tilde{x}, \tilde{y} | h_1 < \tilde{x} < h_2; \rho)$ over different intervals $(h_1,\, h_2)$. The truncation points h_1 and h_2 are chosen to correspond to abscissae from an inverse cumulative normal, which we denote by $\Phi^{-1}(\cdot)$. In Figure (2) h_1 and h_2 correspond to the abscissae intervals of probabilities [0 0.2 0.4 0.6 0.8 1]. That is, the intervals $(h_1,\, h_2)$ correspond to:

$$\begin{aligned}
(\Phi^{-1}(0.0),\, \Phi^{-1}(0.2)) &= (-\infty,\, -0.8146) \\
(\Phi^{-1}(0.2),\, \Phi^{-1}(0.4)) &= (-0.8146,\, -0.2533) \\
&\vdots \\
(\Phi^{-1}(0.8),\, \Phi^{-1}(1.0)) &= (0.8146,\, +\infty). \qquad (4)
\end{aligned}$$

The conditional correlations $\hat{\rho}(h_1, h_2, -\infty, +\infty)$ are plotted at the inverse cumulative normal abscissae corresponding to the midpoints [0.1 0.3 0.5 0.7 0.9]. The conditional correlations produced this way lie in a U-shape.[8] Hence, comparing conditional correlations constructed from samples where one variable has large (absolute) returns to conditional correlations constructed from samples where the same variable has small (absolute) returns must also be done taking into account the conditioning bias. In particular, calculating conditional correlations when the conditioning information set consists of exogenous instrumental variables such as macroeconomic variables may also induce a bias, if these conditioning variables are correlated with returns.

In our empirical work, we take $\tilde{x}$ to be standardized returns of a stock portfolio and $\tilde{y}$ to be standardized market returns. We can look at movements in $\tilde{x}$ and $\tilde{y}$ conditional on large movements in both the market and the stock portfolio (Longin and Solnik (2001)'s analysis), or look at movements in $\tilde{x}$ and $\tilde{y}$ conditional only on large market moves (see Butler and Joaquin (2000)). In both cases we cannot naïvely compare conditional correlations of high return, or

[8] A similar exercise in showing conditional correlation bias over different intervals is done by Boyer, Gibson and Loretan (1999). If we were to show a plot of conditional correlations $\text{corr}(x, y | h_1 < x < h_2; \rho)$ where h_1 and h_2 values are chosen with equal intervals, we will produce a plot very similar to the first Longin-Solnik plot, which has a tent-shape. This also applies if we show correlations conditioning only on x, such as $\text{corr}(x, y | x > \vartheta; \rho)$. In this case a tent similar to the top plot of Figure (2) will be produced.

low return periods. We concentrate on the analysis based on the exceedance correlations of Longin and Solnik (2001). This characterization has the advantage of succinctly describing the conditional correlations with one parameter, the exceedance level ϑ, rather than a series of truncation intervals, as is done in the bottom plot of Figure (2). The exceedance conditioning of both $\tilde{x}$ and $\tilde{y}$ also focuses attention on joint "downside" and "upside" moves. This is particularly relevant given past episodes of market crashes when stocks have made simultaneous extreme moves on the downside.

3.2 Asymmetric Correlations in the Returns Data

3.2.1 Data

We focus on portfolio returns of stocks sorted by industry classifications, market capitalizations (size), book-to-market ratios (value) and past returns (momentum). The industry groupings are miscellaneous, petroleum, finance, durables, basic industries, food and tobacco, construction, capital goods, transportation, utilities, textile and trade, service and leisure. Stocks are sorted on market capitalization, book-to-market ratios and lagged past 6-month returns and grouped into quintiles to form size, book-to-market and momentum portfolios (smallest to largest, growth to value and losers to winners, respectively). We focus on these portfolio groups because industries have varyings exposures to economic factors (see Ferson and Harvey (1991)), the popularity of the Fama and French (1993) model using size and value-based factors, and the recent focus on the momentum effect, which cannot be explained by the Fama and French model (see Fama and French (1996)). We also study portfolios formed by other cross-sectional characteristics, such as beta and co-skewness, and portfolios formed by other firm characteristics such as leverage. These portfolios are also divided into quintiles. To control for possible interaction between market capitalization (size) and other characteristics, we also construct two sets of doubly sorted portfolios : one on size and beta, and another on size and leverage.

We use daily US equity returns from CRSP NYSE/AMEX and Nasdaq files, covering the period July 1963 to December 1998. We construct log-returns of value-weighted portfolios at the weekly frequency (from the close of Wednesday to the close of next Wednesday). In the case of momentum portfolios, we construct equal-weighted portfolios following Jegadeesh and Titman (1993) to be consistent with the momentum literature. The market portfolio is the value-weighted portfolio of all stocks in our sample. Portfolios are reformed monthly. At a weekly frequency this yields 1852 observations. We also study returns at daily and monthly frequencies. All returns are calculated in excess of the one-month Treasury Bill rate. Appendix C details the construction of portfolios

Table (2) presents the summary statistics of the market, industry, size, book-to-market and momentum portfolios at the weekly frequency.[9] The mean and standard deviation of the excess portfolio returns are annualized by multiplying the mean (standard deviation) by 52 ($\sqrt{52}$). The size effect, value effect, and momentum effect are clearly depicted by the mean returns of these portfolios across quintiles.

Non-synchronous trading can cause a bias in the estimation of covariance, and hence correlation. Our portfolio constructions rebalance portfolios at the end of every month, minimizing micro-structure bid-ask bounce effects. We focus on the weekly frequency since this frequency represents the best trade-off to avoid the market microstructure biases at daily frequencies yet provide a large number of observations. We also focus on value-weighted portfolios to avoid putting too much weight on small illiquid stocks. As a check, the last two columns of Table (2) list the sample unconditional correlation with the market portfolio at both the weekly and the monthly frequencies. The unconditional correlations calculated using weekly data and monthly data are very similar. This evidence suggests our results are not plagued by errors in the estimation of correlations induced by non-synchronous trading.

Table (3) lists the ten largest positive and negative excess weekly returns (not annualized) of the market portfolio. The table shows that besides the large negative 19% return due to the October 1987 crash, the top ten largest weekly moves in absolute magnitude of the market are approximately the same for both positive and negative moves. This suggests that the results on asymmetric correlations we present are not due to under-sampling of either the downside or upside movements relative to each other at the weekly frequency. Our results of asymmetric correlations are also robust to excluding the October 1987 crash.[10]

3.2.2 Plots of Exceedance Correlations

If equity and market returns are normally distributed, their exceedance correlations would exhibit the tent-shapes of Figure (2). To construct plots of empirical exceedance correlations we take $\tilde{x}$ to be the standardized excess return of an equity portfolio, and $\tilde{y}$ to be the standardized excess return of the market. Figure (3) shows the exceedance correlations for the equity portfolios at the weekly frequency.[11] The figure provides clear pictorial representations of the asymmetric movements between the equity portfolios and the market. There are two main stylized features of the plots. First, we observe that far from being symmetric, the exceedance correlations for negative exceedance levels are always greater than the exceedance correlations

[9] We do not report the statistics of other portfolios for brevity. Additional summary statistics of the other portfolios and other frequencies are available from the authors.

[10] In Section 3.2.2 the plots of exceedance correlations are almost unchanged when excluding the October 1987 crash, as are the H statistics measuring the correlation asymmetry we present in Section 4

[11] Plots for daily and monthly frequencies and for equal-weighted market returns are available on request. Both the daily and the monthly frequencies exhibit the same highly asymmetric patterns as documented here for the weekly frequency. The H statistic in the legend is the measure of this asymmetry we develop in Section 4.

for positive exceedances. There is a sharp break at $\vartheta = 0$ where the conditioning changes from calculating corr$(\tilde{x}, \tilde{y} | \tilde{x} > 0, \tilde{y} > 0)$ using the positive quadrant to corr$(\tilde{x}, \tilde{y} | \tilde{x} < 0, \tilde{y} < 0)$ using the negative quadrant. Second, instead of tapering off to zero as in the case of a bivariate normal, the negative exceedances are either flat, or tend to increase as ϑ becomes more negative. The positive exceedance correlations are more variable than the negative ones, but there is some evidence that these taper off to zero for some portfolios as ϑ increases.

Figure (4) shows the exceedance correlations of two representative equity portfolios. It plots the exceedance correlations of the first and fifth quintiles of the size portfolios with the market. On the same plot are the implied exceedance correlations from a bivariate normal with the same unconditional correlation as the equity portfolio and market pairs. The plot shows that the negative exceedance correlations for both portfolios do not tend towards zero and are substantially greater than the exceedance correlations of the bivariate normals. This pattern indicates that correlations between the market and the portfolios are significantly higher in falling markets than a normal distribution would imply. The positive exceedances for the fifth size quintile are approximated fairly well by the implied bivariate normal, while for the first size quintile the empirical exceedances lie above those implied by the bivariate normal. Figure (4) suggests that while a bivariate normal distribution cannot match the negative exceedances from the data, it may approximate positive exceedances for some portfolios.

The exceedance plots in Figures (3) and (4) give strong graphical representation of the asymmetric movements in equity portfolios. They clearly show that correlation asymmetries exist in the data. Section 4 develops a summary statistic to measure correlation asymmetry, which the exceedance plots can show only in a qualitative way.

3.3 Upside and Downside Betas

Analogous to the upside and downside exceedance correlations, we can define upside and downside betas.[12] For simplicity, we measure upside and downside betas relative to the means μ_x and μ_y of the portfolio excess return x and market excess return y, or relative to zero for the standardized excess returns $\tilde{x}$ and $\tilde{y}$. We define an upside β^+ as:

$$\beta^+ = \frac{\text{cov}(x, y | x > \mu_x, y > \mu_y)}{\text{var}(y | x > \mu_x, y > \mu_y)} = \frac{\sigma_x^+}{\sigma_y^+} \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} > 0, \tilde{y} > 0) \tag{5}$$

where $\sigma_x^+ = \sqrt{\text{var}(x | x > \mu_x, y > \mu_y)}$ and $\sigma_y^+ = \sqrt{\text{var}(y | x > \mu_x, y > \mu_y)}$. Similarly, we can define a downside β^- as:

$$\beta^- = \frac{\text{cov}(x, y | x < \mu_x, y < \mu_y)}{\text{var}(y | x < \mu_x, y < \mu_y)} = \frac{\sigma_x^-}{\sigma_y^-} \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} < 0, \tilde{y} < 0) \tag{6}$$

[12] We thank an anonymous referee for suggesting this analysis.

where $\sigma_x^- = \sqrt{\text{var}(x|x < \mu_x, y < \mu_y)}$ and $\sigma_y^- = \sqrt{\text{var}(y|x < \mu_x, y < \mu_y)}$.

Denoting $k^+ = \sigma_x^+/\sigma_y^+$ and $k^- = \sigma_x^-/\sigma_y^-$ we can write β^+ and β^- as:

$$\begin{aligned} \beta^+ &= k^+ \bar{\rho}(0)^+ \\ \beta^- &= k^- \bar{\rho}(0)^- \end{aligned} \tag{7}$$

where $\bar{\rho}(0)^+ = \text{corr}(\tilde{x}, \tilde{y}|\tilde{x} > 0, \tilde{y} > 0)$ is the positive exceedance correlation at $\vartheta = 0$ and $\bar{\rho}(0)^- = \text{corr}(\tilde{x}, \tilde{y}|\tilde{x} < 0, \tilde{y} < 0)$ is the negative exceedance correlation at $\vartheta = 0$. The term k^+ (k^-) is the ratio of upside (downside) portfolio volatility to market volatility.

For a bivariate normal $\beta^+ = \beta^-$ by symmetry. The Proposition in Appendix A can be used to calculate β^+ and β^- in closed form. Note that for a bivariate normal $k^- = k^+$. Betas can increase on the downside if the downside exceedance correlation increases, or if portfolios become more volatile on the downside relative to the market. In order for the latter condition to hold, the conditional $\text{var}(\tilde{x}|\tilde{x} < 0, \tilde{y} < 0)$ must increase relative to $\text{var}(\tilde{y}|\tilde{x} < 0, \tilde{y} < 0)$ compared to their upside counterparts.

The upside and downside betas examined here are related to, but different from, the asymmetric betas defined by Ball and Kothari (1989), Braun, Nelson and Sunier (1995) and Cho and Engle (2000). Our upside and downside betas condition on the tails of the distribution, or the entire time series, analogous to the upside and downside exceedance correlations.

3.4 Asymmetric Betas in the Returns Data

Under the normal distribution, upside and downside betas will be equal. Table (4) reports β^- and β^+ for industry, size and book-to-market portfolios. The first column of Table (4) lists the unconditional beta of each portfolio. The second column gives the (equal) theoretical value of $\beta^- = \beta^+$ assuming the null of a normal distribution. In all portfolios, $\beta^- > \beta > \beta^+$ where β is the unconditional beta and in only one case we cannot reject that β^- is equal to its theoretical value implied by a normal distribution. However, on the upside we usually fail to reject that β^+ is equal to its theoretically implied value.

Volatility is well-known to be asymmetric and increasing on the downside, and for the market $\sigma_y^- = 0.0148$ and $\sigma_y^+ = 0.0129$.[13] However, the ratio of the downside portfolio volatility to the market $k^- = \sigma_x^-/\sigma_y^-$ is roughly the same as the ratio of upside portfolio volatility to the market $k^+ = \sigma_x^+/\sigma_y^+$. The last three columns of Table (4) show k^-, k^+ and a p-value of the test $k^- = k^+$. The table shows that in most cases we cannot reject that $k^- = k^+$. Hence, the very statistically significant increase in downside betas is largely driven by the increase in downside correlations relative to upside correlations from equation (7). The next section formally measures and examines the characteristics of these asymmetric correlations.

[13] The theoretical value implied by a normal distribution is $\sigma_y^- = \sigma_y^+ = 0.0122$. We reject that the observed σ_y^- equals this value at a 1% confidence interval, but fail to reject that σ_y^+ equals this value at a 5% confidence interval.

4 A Formal Characterization of Asymmetric Correlations

4.1 A Statistic to Measure Correlation Asymmetries

In this section, we develop a summary H statistic of correlation asymmetries which quantitatively measures asymmetric correlations. We construct the statistic by quantifying the differences between the exceedance plots implied by the data and those implied under a distributional assumption (the null distribution). This has several advantages over graphical approaches. First, the statistic formally summarizes the magnitudes of correlation asymmetries by providing a succinct numerical measure. This means that the degree of asymmetry can be measured and compared across different portfolios and different frequencies. The H statistic can be used to rank portfolios and examine if various characteristics of equity portfolios contribute to the degree of correlation asymmetry. Second, we can numerically compare empirical exceedance correlations with those implied by a null distribution. Hence we can directly incorporate the conditioning bias in the exceedance correlations. Finally, we can formally test if exceedance correlations in the data can be produced by candidate null distributions.

4.1.1 Description of the H Statistic

As in Equation (3), we denote the exceedance correlation for a given exceedance level ϑ_i as $\bar{\rho}(\vartheta_i)$ for standardized data $(\tilde{x},\, \tilde{y})$. We choose N exceedance levels $\theta = (\vartheta_1,\, \vartheta_2,\, \ldots \vartheta_N)$. These exceedance levels are set a priori. Suppose we wish to test if a distribution $\xi(\phi)$ characterized by parameters ϕ can produce the empirical exceedances $\bar{\rho}(\vartheta_i)$ in the data. Denote the exceedance correlations implied by distribution $\xi(\phi)$ as $\breve{\rho}(\vartheta_i, \phi)$.

If $\xi(\phi)$ were to perfectly explain the degree of correlation asymmetry in the data then we would have $\bar{\rho}(\vartheta_i) - \breve{\rho}(\vartheta_i, \phi) = 0$ on average. We create a quadratic statistic based on this difference. Define the statistic $H = H(\phi)$ as:

$$H = \left[\sum_{i=1}^{N} w(\vartheta_i) \cdot (\breve{\rho}(\vartheta_i, \phi) - \bar{\rho}(\vartheta_i))^2 \right]^{\frac{1}{2}} \tag{8}$$

where the weights $w(\vartheta_i) \geq 0$ satisfy:

$$\sum_{i=1}^{N} w(\vartheta_i) = 1.$$

This statistic measures a weighted average difference of the exceedance correlations implied by a model and those given by data. For example, an $H = 0.10$ means that, on average, the difference between the exceedance correlations in the data and those implied by the model is 10%. Note that H is a (non-linear) function of ϕ for a fixed set of ϑ. The weights are exogenously set and will be related to how much sampling error is associated with a particular exceedance correlation. The more accurately estimated the exceedance correlation for exceedance level ϑ_i, the higher we will set $w(\vartheta_i)$. We discuss various choices for the weights $w(\vartheta_i)$ below.

The H statistic can be written in matrix notation. We denote $\bar{\rho}(\theta)$ as the N vector of exceedances from data, and $\breve{\rho}(\theta, \phi)$ as the N vector of exceedances implied by distribution $\xi(\phi)$:

$$\bar{\rho}(\theta) = \begin{pmatrix} \bar{\rho}(\vartheta_1) \\ \bar{\rho}(\vartheta_2) \\ \cdots \\ \bar{\rho}(\vartheta_N) \end{pmatrix} \qquad \breve{\rho}(\theta, \phi) = \begin{pmatrix} \breve{\rho}(\vartheta_1, \phi) \\ \breve{\rho}(\vartheta_2, \phi) \\ \cdots \\ \breve{\rho}(\vartheta_N, \phi) \end{pmatrix}$$

Then H can be expressed as:

$$H = \sqrt{(\bar{\rho} - \breve{\rho}(\phi))' \Omega^{-1} (\bar{\rho} - \breve{\rho}(\phi))}, \tag{9}$$

where we suppress the dependence on θ and $\Omega = \Omega(\theta)$ is a fixed diagonal weighting matrix dependent only on θ which takes the form:

$$\Omega = \begin{pmatrix} w(\vartheta_1)^{-1} & 0 & \cdots & 0 \\ 0 & w(\vartheta_2)^{-1} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & w(\vartheta_N)^{-1} \end{pmatrix} \tag{10}$$

If the sample exceedances $\bar{\rho}(\theta)$ are considered fixed, then the δ-method can be used to obtain standard errors for H, since the exceedance correlations $\breve{\rho}(\theta, \phi)$ implied by the model are a function of the parameters ϕ of distribution ξ. Suppose ϕ can be estimated with covariance matrix Γ. We can find, using the δ-method:

$$\text{var}(H) = D_1' \Gamma D_1 \tag{11}$$

where D_1 is:

$$D_1 = \frac{\partial}{\partial \phi} H.$$

The square root of equation (11) is the standard error of H. Under the null hypothesis that the distribution is ξ, by the δ-method H will be asymptotically normally distributed with mean zero and variance var(H).

We first take ξ to be normally distributed. In this case, the exceedance correlations of $\xi(\phi)$ are characterized by a single parameter $\phi = \rho$, the unconditional correlation. Moreover, the implied exceedance correlations $\breve{\rho}(\theta, \phi)$ are closed-form and can be calculated using the Proposition in Appendix A. The standard errors of H can also be derived analytically. Later in Section 5, we will consider more complex distributions for which $\breve{\rho}(\theta, \phi)$ will not be closed-form and must be calculated by simulation. We discuss the calculation of H for these cases in Appendix D.

4.1.2 Choices of Weights

The H statistic can be interpreted as the square root of a quadratic statistic. For the quadratic statistic J, where J is (suppressing θ):

$$J = (\bar{\rho} - \breve{\rho}(\phi))'\Omega^{-1}(\bar{\rho} - \breve{\rho}(\phi)),$$

the efficient choice for Ω, Ω_E, is:

$$\Omega_E = \text{var}(\bar{\rho}) - 2\text{cov}(\bar{\rho}, \breve{\rho}(\phi)) + \text{var}(\breve{\rho}(\phi)),$$

for the case that N is less than the number of parameters in ϕ. We choose not to use the efficient weighting matrix for two reasons.

First, if the data are fixed, or we estimate $\bar{\rho}(\theta)$ without error, then $\Omega_E = \text{var}(\breve{\rho}(\theta, \phi))$ and J would have a conventional χ^2_N distribution. For a normal distribution, there is only one degree of freedom $\phi = \rho$ in the parameters of the bivariate normal which determines the exceedance correlation, so this approach would mean only one exceedance correlation can be incorporated in J. In the case of a normal with $N > 1$, $\Omega_E = D_2'\Gamma D_2$, where $D_2 = \frac{\partial}{\partial\phi}\breve{\rho}(\theta, \phi)$, is singular because there are more restrictions (exceedance correlations) than degrees of freedom allowed by the parameters. However, we can capture the notion of using weights inversely proportional to the sample variance of $\breve{\rho}(\theta, \phi)$, $\sigma^2(\breve{\rho}(\vartheta_i, \phi))$ by using a standardized measure of the inverse of $\sigma^2(\breve{\rho}(\vartheta_i, \phi))$:

$$w(\vartheta_i) = \frac{\sigma^{-2}(\breve{\rho}(\vartheta_i, \phi))}{\left(\sum_{j=1}^{N} \sigma^{-2}(\breve{\rho}(\vartheta_j, \phi))\right)}. \tag{12}$$

The larger the sampling variance of $\breve{\rho}(\vartheta_i, \phi)$, the smaller the weight placed on that exceedance. We calculate $\sigma^2(\breve{\rho}(\vartheta_i, \phi))$ using the δ-method:

$$\sigma^2(\breve{\rho}(\vartheta_i, \phi)) = D_{2i}'\Gamma D_{2i}$$

where D_{2i} is:

$$D_{2i} = \frac{\partial}{\partial\phi}\breve{\rho}(\vartheta_i, \phi).$$

The difference between this choice of Ω and the efficient GMM choice is that Ω is diagonal to avoid singularities and is normalized to unity.

The second reason we choose not to use the efficient weighting matrix is that each different model or distribution ξ implies a different weighting matrix. The first choice of weights above is not immune to this critique. Since each model implies a different set of weights, the H statistics are not directly comparable across models. Like the constant weighting matrix of Hansen and Jagannathan (1997) used to compare different pricing kernels with the same data, we would like to use a constant weighting matrix to compare different models with the same data. The next two choices of weights do not depend on the distribution, so can be used to compare how different models replicate the correlation asymmetry in data.

The next set of weights is held constant across models and takes into account some notion of sampling error. We note that increasing the number of observations will increase the accuracy of the estimate. For the normal distribution, covariance sampling error is of the order $1/\sqrt{T}$ where T is the sample size. An *ad hoc* way to account for sampling error is to set the weights proportional to the number of observations used to calculate the exceedance correlations. Hence a second choice for $w(\vartheta_i)$ uses weights:

$$w(\vartheta_i) = \frac{T_i}{\left(\sum_{j=1}^{N} T_j\right)} \tag{13}$$

where T_i is the number of observations used in calculating $\bar{\rho}(\vartheta_i)$, the sample exceedance correlation at the exceedance level ϑ_i. This choice of weights places more emphasis on exceedance correlations for which more data are available.

Finally, equal weights may be used:

$$w(\vartheta_i) = \frac{1}{N}. \tag{14}$$

for N exceedances. This choice places greater weight on observations in the extreme tails of the distribution than the previous choice of weights.

Our preferred form of the H statistic uses the weights in equation (13). It assumes that the point estimates of sample exceedances $\bar{\rho}(\theta)$ are fixed. It takes into account the sampling error of the estimates of ϕ to produce a standard error for H using the implied model exceedance correlations $\breve{\rho}(\phi, \theta)$ via the δ-method. We address the sampling error of the sample estimate of $\bar{\rho}(\theta)$ using larger weights for exceedance correlations calculated on more data. However, we find all of our results to be robust to different choices of weights.

4.2 Magnitudes and Tests of Asymmetic Correlations

For various pairs of standardized excess returns of the market and stock portfolios $(\tilde{x},\ \tilde{y})$ we estimate the unconditional correlation ρ, and calculate H under the null hypothesis of a bivariate standard normal distribution with unconditional correlation ρ. We set the exceedance levels $\theta = [-1.5,\ -1.0,\ -0.5,\ 0.0,\ 0.0,\ 0.5,\ 1.0,\ 1.5]$. For the repeated zero, we calculate exceedance correlations corr$(\tilde{x}, \tilde{y}|\tilde{x} < 0, \tilde{y} < 0)$ and corr$(\tilde{x}, \tilde{y}|\tilde{x} > 0, \tilde{y} > 0)$. To enable a comparison of downside versus upside correlations, we also use the sets of exceedance levels $\theta^- = [-1.5,\ -1.0,\ -0.5,\ 0.0]$ and $\theta^+ = [0.0,\ 0.5,\ 1.0,\ 1.5]$. The H statistics calculated from

θ^- and θ^+ we denote as H^- and H^+ respectively.[14] We estimate the standard errors of H, H^-, $H+$ by GMM using 6 Newey-West (1987) lags.

The H statistics capture the same features as the exceedance plots. H statistics are reported in the legends of Figure (3) corresponding to the various portfolios. The bigger the difference in positive and negative exceedance correlations, the larger the H statistic. With the H statistic, a numerical measure of the correlation asymmetry can now be assigned to each portfolio.

4.2.1 Impact of Weights and Frequencies

Table (5) presents H statistics using the three choices of weights for the five size-sorted portfolios at daily, weekly and monthly frequencies. We present the size portfolios as they are representative; similar results hold for the other portfolios. Columns 1-2 present the H statistics weighted by the variances in the normal distribution (equation (12)). Columns 3-4 are weighted by the number of observations used to construct the sample exceedances (equation (13)). The last two columns present equal-weighted H statistics (equation (14)).

There are two major results of Table (5):

Empirical Fact 1 *Asymmetric correlations in the data overwhelmingly reject the null of a normal distribution.*

Empirical Fact 2 *The magnitude of the correlation asymmetries is unrelated to the horizon.*

In Table (5) the p-values of the H statistics are all less than 0.1% across all choices of weights and across all frequencies (and so are not reported). There is also no discernable pattern across the sampling frequencies. For the smallest and the largest size portfolios, correlation asymmetries with the market portfolio are the greatest at the monthly frequency with all three weight choices.

The equally-weighted H statistic is always larger than the other two choices of weights. This is because the largest sampling error in the normal distribution and the smallest number of observations occur at the largest absolute value exceedance levels ($\vartheta = \pm 1.0,\ 1.5$). At these exceedances, in particular for the negative exceedances, the largest discrepancies between the normal distribution and the sample exceedance correlations arise (see Figures (3) to (4)). These discrepancies are given more weight in the equal-weighted H statistic.

These results extend to other portfolios. Since the rejection of the normal distribution and the patterns of asymmetries are robust to the weighting choice and the frequency of observations, we concentrate on using weights proportional to the number of observations in each sample exceedance (equation (13)) and analyze in depth the weekly frequency for the rest of the paper.

[14] Note that $H^2 = (H^+)^2 + (H^-)^2$, so H is a non-linear average of H^+ and H^-.

4.2.2 Characterizing Asymmetric Correlations

In order to further characterize the asymmetric correlations in equity portfolios we examine the relationship between different portfolio sortings and their H statistics. Estimating extreme correlations requires using observations lying in the tails where there are relatively few data points. The H statistic uses the full sample (time-series) of returns to measure correlation asymmetries. To maintain the use of the full sample, we sort portfolios of stocks by various cross-sectional characteristics and examine their correlation asymmetries.

Table (6) presents the H statistics across a wide selection of portfolios assuming the null of a bivariate normal distribution. Panel A examines the properties of portfolios formed by industry classifications, size, book-to-market and momentum. Panel B investigates the asymmetry properties of portfolios formed by past beta, co-skewness and leverage. The first three columns of Table (6) show the H, H^+ and H^- statistics. The H statistics for all portfolios have p-values smaller than the 1% level of significance, just as Table (5) showed for the size portfolios. The average H^- statistic is 0.1161, while the average H^+ statistic is 0.0300:

Empirical Fact 3 *Correlation asymmetries are greater for extreme downward moves.*

Six industry portfolios cannot reject at the 5% significance level that the upside correlations can be reproduced by a bivariate normal. In contrast, all H^- statistics reject at the 1% level of significance. In calculating the average H statistics across portfolios in Table (6), we observe:

Empirical Fact 4 *Conditional on downside and upside moves, on average the observed correlations between a portfolio and the market differ from the correlations implied by a normal distribution by 8.48%.*

The next two columns of Table (6) report standardized measures of skewness and co-skewness, and their standard errors, where skewness and co-skewness are defined as:

$$\text{skewness} = \frac{\mathrm{E}[\dot{x}^3]}{(\mathrm{E}[\dot{x}^2])^{3/2}}, \qquad \text{co-skewness} = \frac{\mathrm{E}[\dot{x}\dot{y}^2]}{\sqrt{\mathrm{E}[\dot{x}^2]}\mathrm{E}[\dot{y}^2]},$$

where $\dot{x}$ is the demeaned excess return of the portfolio x, $\dot{x} = x - \mathrm{E(x)}$ and $\dot{y}$ is the demeaned excess return of the market y, $\dot{y} = y - \mathrm{E(y)}$. All standard errors are calculated by GMM using 6 Newey-West lags.

Table (6) also shows that at the weekly frequency, each of the portfolios are negatively skewed and are negatively co-skewed with the market. This may indicate that there is some common component among all three asymmetry statistics. To ensure that we are not capturing the same information in H as skewness and co-skewness, we present the correlation among these statistics across the 43 portfolios in Table (7). The correlation of H with skewness is

0.2433, and with co-skewness is 0.1498. This indicates that H is capturing something fundamentally different from co-skewness and skewness. Skewness and co-skewness are much more highly correlated at 0.9510.

The final column of Table (6) reports the betas of the portfolios with the market. The correlation between beta, H, skewness and co-skewness is reported in Table (7). All measures of return asymmetries appear to have little positive relation with systematic risk. In particular, the H statistics are negatively correlated (-0.2744) with beta.

Table (6) reveals that certain portfolios exhibit greater asymmetric correlations than others:

Empirical Fact 5 *Petroleum and utility industries have the most asymmetric correlations relative to a normal distribution, while financials and basic industries exhibit the lowest asymmetric correlations.*

Among industries, petroleum ($H = 0.1801$) and utilities ($H = 0.1454$) are the most asymmetric, while financials and basic industries exhibit the least asymmetric correlations. Petroleum and utilities have low betas (0.8394 and 0.6302 respectively), suggesting that investing in these traditional "defensive" sectors may not be as beneficial as one would initially think. Note these industries have the least negative skewness and co-skewness, and would not appear by these measures to be the most non-normal.

Among size-sorted stock portfolios:

Empirical Fact 6 *Decreasing size increases the correlation asymmetry.*

Similar results hold for asymmetry measure by skewness and by co-skewness. This pattern has been previously documented in a GARCH-specification by Kroner and Ng (1998) and Conrad, Gultekin and Kaul (1991). The book-to-market portfolios also display an increasing pattern of H statistics going from growth to value stocks:

Empirical Fact 7 *Value stocks are more asymmetric than growth stocks.*

While Fama and French (1993) observe a size and value premium, portfolios formed on these characteristics may be more risky by their greater correlation asymmetry than by measuring risk only by second moments. In both the size and book-to-market portfolio sortings, the H statistics are monotonic, unlike the point statistics of the skewness and co-skewness measures. Moreover the latter two measures do not display any discernable pattern.[15]

Turning to the momentum quintiles we observe:

Empirical Fact 8 *The past loser portfolio has greater correlation asymmetry than the past winner quintile.*

[15] We also sorted on Scholes-Williams (1977) betas to alleviate potential concerns over non-synchronous trading. We found slightly lower H statistics but the qualitative results were unchanged.

In the Jegadeesh and Titman (1993) momentum strategy, investors short past losers and go long past winners. In periods of extreme downside moves, the loser portfolio will likely lose much more money than estimated with constant correlations, thus affording the momentum players even greater rewards in down markets. This effect would tend to exacerbate the puzzle posed by the momentum effect. Like Chen, Hong and Stein (2000) and Harvey and Siddique (2000), we also find that the past winner portfolio are more negatively skewed than the past loser portfolio. However, the relationship between H and skewness or co-skewness goes in the opposite direction: the past losers are the least skewed or co-skewed but are the most asymmetric.

In Panel B of Table (6), we search for additional determinants of asymmetries.[16] We first characterize the correlation asymmetries of portfolios sorted by systematic risk. The portfolio of lowest beta stocks is the portfolio that exhibits the greatest correlation asymmetry. It is lower risk, not higher risk firms, which have more correlation asymmetries. Note that co-skewness monotonically increases with beta, while skewness has no discernable pattern.

The relationship is clearer with size controls. We examine the interaction between size and risk with correlation asymmetries in Table (8), where we perform a double sort. Each month, we first sort stocks in our universe into quintiles by size. Then within each size quintile, we perform a second sort of stocks into quintiles by past estimates of beta. Based on this 5×5 portfolio grouping, we find that controlling for size, riskier firms have fewer correlation asymmetries than less risky firms. In Table (8) we observe that H statistics decrease going down the rows, where we sort by size. Going across the columns, where we control for size and sort by beta, the lowest beta stocks have the highest H statistics. Thus we conclude that:

Empirical Fact 9 *Increasing beta decreases correlation asymmetry.*

When the sorting criteria is individual stock's past co-skewness in Panel B of Table (6), we do not find any pattern between past co-skewness and correlation asymmetry. This portfolio sort suggests that Harvey and Siddique (2000)'s co-skewness measure is not related to the degree of correlation asymmetry in the data. There is also no pattern in the skewness or the co-skewness of portfolios formed by past conditional co-skewness. The risk (beta) of stocks sorted by past co-skewness is near market risk across all quintiles.

Finally, we observe that the most leveraged stocks have the greatest correlation asymmetry. This effect is weakly monotonic, and not reflected in the skewness or the co-skewness measure. Bekaert and Wu (2000) find that the leverage-effect accounts for only a small proportion of asymmetric covariance. In Table (9) we examine the effect of leverage on correlation asymmetry when controlling for size. We observe, as expected from Empirical Fact 6, that H statistics

[16] We also calculated H statistics for portfolios sorted by volatility (no relationship), skewness (results similar to co-skewness), turnover (lower H for low turnover stocks) and earnings yield (results similar to book-to-market). These statistics are available upon request.

decrease as stocks become larger (going down the rows). However, when size is held constant, there is no discernible pattern across debt levels and correlation asymmetries:

Empirical Fact 10 *There is no relationship between leverage and correlation asymmetries once size has been controlled for.*

The lack of patterns within size groups may account for Bekaert and Wu's weak support for the leverage-effect as an explanation behind covariance asymmetry.

4.2.3 Summary of Empirical Facts

We find that correlation asymmetries in equity portfolios are not fully explained by traditional skewness and co-skewness measures. These correlation asymmetries persist across daily, weekly and monthly frequencies and are greatest for downside moves. Correlation asymmetries are larger for small, high book-to-market and past loser portfolios. This suggests that size, and value strategies are exposed to more contemporaneous downside moves with the market, which is not reflected in a measure of only second moments such as volatility. Momentum strategies are more profitable than they first appear because in times of market distress, loser stocks, which investors are short, are more likely to fall with the market than past winners, which investors are long. High beta portfolios are less asymmetric than low beta portfolios. Once we have controlled for size, there is no discernable pattern between correlation asymmetries and the leverage of firms.

5 Empirical Models of Asymmetric Correlations

The previous section examines the characteristics of asymmetric correlations relative to a normal distribution. We now seek to explain the correlation asymmetries in the data by using richer models of stock returns which can potentially capture the asymmetric movements. Using only criteria based on how closely each model match the correlation asymmetries in the data, we investigate the performance of four empirical reduced-form models of stock returns. Section 5.1 describes the models, Section 5.2 presents the empirical results of the H statistics using these models as the null distribution, and Section 5.3 provides some intuition behind the rejection patterns.

5.1 Description of Models

Our choice of models is motivated by examining several popular models which previous authors have used to capture asymmetries between upside and downside movements in stock returns.

We use weekly data, and following Braun, Nelson and Sunier (1995), Cho and Engle (2000) and others, we work with independent pairs of (stock portfolio, aggregate market).

The first model is a GARCH-M Model with asymmetry. The GARCH-M Model uses a time-varying expected returns model (where volatility risk is priced in the expected return) with the conditional covariances following a GARCH process. The GARCH process incorporates asymmetry which allows covariances to increase on the downside. The second model is the Jump Model. This layers negative jumps, which are perfectly correlated in time for both returns, on top of a bivariate normal distribution to produce larger downside correlations. The last two models are regime-switching (RS) models. The RS Normal Model mixes two different bivariate normal distributions. This allows returns to switch to a regime with lower conditional means, higher volatility and higher correlations. Transitioning into this regime increases downside correlations. The RS-GARCH Model combines elements of the switching behavior of pure RS Normal Models with the volatility persistence of GARCH processes. We outline each model in turn below.[17]

We note that these four models are not the only empirical models capable of producing asymmetric correlations. One large class of models which we do not pursue here are continuous-time stochastic volatility models where shocks to conditional mean and conditional volatility factors may be correlated, with possible jumps in prices or volatility. This class of models is very hard to estimate (see Pan (2000), for example) particularly on multivariate series, and it is not clear that they will produce markedly different results from the discrete-time models on weekly sampled data. Our Jump Model captures jumps in returns but without stochastic volatility. The regime-switching models we estimate can both capture stochastic volatility and jump effects through regime switches.

Other models we do not investigate involve residuals drawn from distributions which have higher moments. One such model is Harvey and Siddique (1999) which draws from a non-central t-distribution to capture skewness and kurtosis. In a multivariate application, this model is extremely computationally intensive because maximum likelihood methods cannot be used. However, the mixture of normal distributions we employ can also match any degree of conditional skewness and kurtosis, as noted by Bekaert, Erb, Harvey and Viskanta (1998).

[17] We do describe the details of the the estimation methods behind each model, but we give references in the text where estimation algorithms, or maximum likelihood functions, can be obtained. For each model we obtain maximum likelihood estimates with White (1980) standard errors for use in constructing H standard errors. Parameter estimates are available upon request.

5.1.1 Asymmetric GARCH-M Model

As before, we denote the excess returns of the equity portfolio by x, and the excess market returns by y. We model the pair $(x_t,\, y_t)$ as:

$$\begin{aligned} x_t &= \delta \, \text{cov}_{t-1}(x_t, y_t) + \epsilon_t^1 \\ y_t &= \delta \, \text{var}_{t-1}(y_t) + \epsilon_t^2. \end{aligned} \tag{15}$$

We take $\epsilon_t = (\epsilon_t^1,\, \epsilon_t^2)'$, with $\epsilon_t \sim N(0, H_t)$. The coefficient δ is the price of risk and is positive in the CAPM. We can model the conditional covariances H_t of $(x_t,\, y_t)$ as a GARCH model but introduce asymmetry using a multivariate version of Glosten, Jagannathan and Runkle (1993):

$$H_t = C'C + A'H_{t-1}A + B'\epsilon_{t-1}\epsilon_{t-1}'B + D'\eta_{t-1}\eta_{t-1}'D, \tag{16}$$

where

$$\eta_{t-1} = \epsilon_t \odot \mathbf{1}_{\{\epsilon_{t-1}<0\}}.$$

The symbol $\odot$ is a Hadamard product representing element by element multiplication, and $\mathbf{1}_{\{\epsilon_{t-1}<0\}}$ is a vector of individual indicator functions for the sign of the errors for x and y. The matrices A, B, C and D are symmetric to ensure H_t is positive definite. Shocks on the downside increase the variance and covariance, through the asymmetric term in H_t, but also increase the conditional mean, by allowing H_t to enter the conditional mean in equation (15). Equation (16) is the asymmetric BEKK model of Engle and Kroner (1995), and its multivariate form of asymmetry is a special case of the nonmenclature system of Kroner and Ng (1998). Similar GARCH-M Models with asymmetry are estimated by Bekaert and Wu (2000), De Santis, Gerard and Hillion (1999), and Bekaert and Harvey (1997).

5.1.2 Bivariate Normal with Poisson Jumps

Das and Uppal (1999) recommend a model where returns are drawn from a bivariate normal but with negative jumps. The jumps occur simultaneously in time for both variables, but the size of the jumps can differ.[18] This jump induces higher correlation with downward moves. The model is given by:

$$X_t = \mu + \Sigma^{\frac{1}{2}}\epsilon_t + \sum_{i=1}^{n_t} Y_t \tag{17}$$

with $X_t = (x_t,\, y_t)'$. There is a Poisson jump process with intensity λ, with jump distribution $Y_t \sim N(\mu_j, \Sigma_j)$ which is independent of X_t. There are n_t actual jumps during each period. Das and Uppal discuss how this model can produce unconditional skewness and kurtosis which match equity data.

[18] This model was proposed in an old version of Das and Uppal (1999). The current version of this paper uses a jump model of this form, except the Poisson parameter can switch between two separate states. This model has more of the flavor of a RS model, which is outlined next.

5.1.3 Regime Switching Bivariate Normal

The Regime Switching Bivariate Normal (RS Normal) Model draws the portfolio returns $X_t = (x_t,\, y_t)'$ from one of two bivariate normal distributions of returns, depending on the prevailing regime $s_t = 1, 2$ at time t:

$$X_t = \mu(s_t) + \Sigma^{\frac{1}{2}}(s_t)\epsilon_t, \tag{18}$$

where $\epsilon_t \overset{\text{IID}}{\sim} N(0, I)$. Following Hamilton (1989), s_t follows a Markov Chain with transition probability matrix Π given by:

$$\Pi = \begin{pmatrix} P & 1-P \\ 1-Q & Q \end{pmatrix}. \tag{19}$$

where $P = Pr(s_t = 1|s_{t-1} = 1)$ and $Q = Pr(s_t = 2|s_{t-1} = 2)$.

This model has been used by Ang and Bekaert (2000) to look at international asset allocation under higher correlations with downside moves in country returns. Ang and Bekaert show that this model captures a large part of the asymmetric correlations in international equity markets of developed countries. In this model asset returns are allowed to switch into a regime with higher correlations and volatility and potentially lower means.

5.1.4 Regime-Switching GARCH

In the Regime-Switching GARCH (RS-GARCH) Model, portfolio returns $X_t = (x_t,\, y_t)'$ follow the process:

$$X_t = \mu(s_t) + \epsilon_t \tag{20}$$

with two regimes $s_t = 1, 2$ and the error term $\epsilon_t \sim N(0, H_t(s_t))$. The regime variable s_t follows the same Markov Chain with transition probability matrix Π given by equation (19). The conditional covariance $H_t(s_t)$ is given by:

$$H_t(s_t) = C(s_t)'C(s_t) + A(s_t)'H_{t-1}A(s_t) + B(s_t)'\epsilon^*_{t-1}\epsilon^{*'}_{t-1}B(s_t), \tag{21}$$

where the forecast error ϵ^*_{t-1} is given by:

$$\begin{aligned} \epsilon^*_{t-1} &= X_{t-1} - \mathrm{E}_{t-2}(X_{t-1}) \\ &= X_{t-1} - (p_{t-1}\mu_1 + (1 - p_{t-1})\mu_2), \end{aligned} \tag{22}$$

where $\mu_i = \mu(s_t = i)$, and p_{t-1} is the ex-ante probability $p_{t-1} = p(s_{t-1} = 1|\mathcal{I}_{t-2})$. Following Gray (1996) H_{t-1} is given by:

$$\begin{aligned} H_{t-1} &= \mathrm{E}_{t-2}(X_{t-1}X'_{t-1}) - \mathrm{E}_{t-2}(X_{t-1})\mathrm{E}_{t-2}(X_{t-1})' \\ &= p_{t-1}(\mu_1\mu_1' + H_{t-2,1}) + (1 - p_{t-1})(\mu_2\mu_2' + H_{t-2,2}) \\ &\quad - [p_{t-1}\mu_1 + (1 - p_{t-1})\mu_2][p_{t-1}\mu_1 + (1 - p_{t-1})\mu_2]' \end{aligned} \tag{23}$$

where $H_{t-2,i} = H_{t-2}(s_t = i)$. The matrix $C(s_t)$ is symmetric but for reasons of parsimony we restrict $A(s_t)$ and $B(s_t)$ to be diagonal.

This RS-GARCH Model uses a RS version of the Engle and Kroner (1995) BEKK multivariate GARCH Model (equation (21)). It uses a multivariate generalization of Gray (1996)'s algorithm in equation (23) to re-combine the lagged RS conditional covariance term. The model combines the switching character of the RS Normal Model, with the volatility persistence of GARCH. One of the features of this model is that the volatility can also switch to a higher volatility, less persistent regime together with a switch in the conditional mean. Glosten, Jagannathan and Runkle (1993) discuss that pure asymmetric GARCH specifications cannot easily capture this feature.

5.2 Model Performance

In this section we use the H statistic as a criterion to judge the adequacy of a model to match the asymmetric correlation found in data. We consider a model to do an "adequate" job of capturing the correlation asymmetry in the data if that model's H statistic cannot be statistically rejected. As a second measure, since the H statistic measures the difference between the empirical conditional correlations and the conditional correlations implied by the models, we consider the average magnitude of H statistics across portfolio pairs. We calculate the H statistics from the models using fixed weights from equation (13), which place more weight on sample exceedance correlations which have been calculated with more observations. These weights ensure that the same weighting matrix is used across all four models. In this section, we focus our analysis on the portfolios formed by industry classifications, by size, by book-to-market and by past returns.

Table (10) summarizes the rejections across the portfolios. Of these 28 portfolios, the GARCH-M Model is rejected by 17 portfolios at the 5% level, the Jump Model is rejected by 25 portfolios and the RS Normal Model is rejected by 12 portfolios. At the 5% level, the RS-GARCH Model is rejected by 8 out of 28 portfolios, giving the RS-GARCH model the best performance by this criterion. However, the model still leaves some amount of the correlation asymmetry unexplained. The full details of the H statistics on the four empirical models are listed in Table (11).

Table (11) shades in gray the model which produces the lowest H statistic for each portfolio. In all cases the normal distribution's H statistic is higher than the best-performing empirical model of Section 5.1. To summarize each model's performance in producing the lowest H statistic, we tabulate how many times a particular model produces the smallest H statistic out of all five models:

Normal	0
GARCH-M	1
Jump Model	0
RS-Normal	22
RS-GARCH	5
Total	28

By the criterion of producing the smallest H statistic, the RS-Normal Model provides the best performance.

Looking at the magnitudes of the H statistics in Table (11), we find that while the RS-GARCH Model rejects the null hypothesis the least, it can be a very poor fit of the data for some portfolios. The H statistic for the RS-GARCH Model is greater than 0.13 for 6 out of 28 portfolios. The average H statistic across all 28 portfolios for the RS Normal Model is 0.0564, while the averages for the GARCH-M Model, the Jump Model, and the RS-GARCH Model are larger: 0.0850, 0.0899, and 0.0960, respectively. In comparison, the average H statistic for the normal distribution is 0.0954. Hence, while the RS-GARCH Model rejects the least number of times, the RS Normal Model provides the best fit across all the portfolios.

The same portfolios which proved difficult to match their empirical correlation asymmetry using the normal distribution tend to be difficult to fit across all four models. In general, the petroleum and utility industries have the highest H statistics across models. Portfolios formed of small stocks, value stocks and past loser stocks also tend to have the highest H statistics across models.

5.3 Explaining the Model Performance

In explaining the performance in matching the correlation asymmetries by each model it is instructive to examine a portfolio where all models are rejected by the data. Figure (5) shows the exceedance correlations from the third momentum quintile, which rejects all four empirical models. The sample exceedance correlations are given by the solid line. Taking each model in turn, Figure (5) shows that the GARCH-M Model produces exceedance correlations which are asymmetric but go the wrong way. That is, the sample exceedance correlations increase on the downside (for negative ϑ), but the GARCH-M Model exceedance correlations are higher on the upside (for positive ϑ). The Poisson Jump Model produces exceedance correlations which have a tent-shape, much like the normal distribution. The RS Normal Model produces exceedance correlations with the correct asymmetry, but decay too quickly on the downside. Finally, the RS-GARCH Model produces a tiny bit of correlation asymmetry in the right direction, but is too persistent on both the downside and the upside.

The wrong direction of exceedance correlation asymmetry for the GARCH-M Model in Figure (5) is shared by all other portfolios. Although this model allows the conditional covariance to increase in response to an unanticipated shock in returns, the expected return of both the market and the portfolio also increase in this model. Equation (15) shows that for a positive price of risk δ, both the conditional mean of the market and the stock portfolio may increase when the conditional covariance increases. So although the conditional covariance increases through a negative shock in expected returns, the expected return also increases making it more likely to draw returns on the upside. However, the GARCH effect does induce persistence in the exceedance correlations across increasing (or decreasing) ϑ which the normal distribution cannot capture.

To illustrate what happens when a negative price of risk is used in the GARCH-M Model we turn to Figure (6). This figure shows exceedance correlations for the smallest size portfolio for all four models in each panel, against the sample exceedance correlations. The top left hand panel shows the exceedance correlations for the GARCH-M Model. The estimated exceedance correlations implied by the model are given by circles. If we change the price of risk to be negative the GARCH-M Model can closely match the sample exceedance correlations. This is the main failing of the GARCH-M Model: asymmetric exceedance correlations can be produced, but the asymmetry goes the wrong way unless a negative price of risk is employed. Economic models do not necessarily rule out negative prices of risk, but the economic plausibility of negative prices of risk and empirical estimates of the Sharpe ratio of the US market weigh heavily against this assumption.

As in Figure (5), the Poisson Jump Model in Figure (6) produces a tent-shape. This is a general result, and is the reason behind the poor performance of this model. The Jump Model performs poorly because it fails to capture the persistence in volatility. The other three models do capture this feature of the data. The Jump Model can be interpreted to be a special case of the RS Normal Model where one regime can be interpreted as a jump regime, and the probability of entering this regime is positive but the probability of remaining in this regime is zero. Ang and Bekaert (2000) find that in international data, this crash-like regime is persistent, which cannot be captured in a jump model as it assumes an immediate exit from this regime.

The intuition behind why the Jump Model produces mostly tent-shapes in the exceedance plot is that ordinarily the returns are drawn from a normal distribution, which has a tent-shape, and only occasionally drawn from another normal distribution when a jump occurs. These jumps are not persistent, and the effect is to mirror the tent-shapes of an ordinary normal distribution. The model will produce a correlation asymmetry, but it is very small and not persistent across exceedance levels. Changing the parameters of the Jump Model has little effect on the tent-shape of its exceedance correlations. The top right hand panel of Figure (6) shows what happens when the correlation of the market and stock portfolio increases in the jump distribution. In this case the tent-shape has moved upwards but has hardly changed shape. A similar effect occurs when increasing the jump intensity.

Figure (5) shows the RS Normal Model may produce exceedance correlations that die too fast when the exceedance levels $\vartheta \rightarrow \pm\infty$. Why this model occasionally fails is that the exceedance correlations can be too persistent across ϑ for the RS Normal Model to mimic. Empirical estimates of this model produce a "normal regime" with high expected returns, lower volatilities and correlations and a "downside" regime with lower (or negative) expected returns, higher volatilities and correlations.

The bottom left panel of Figure (6) shows that merely increasing the probability of staying in the down-regime (Q in the Markov Chain in equation (19) if the down-regime corresponds to $s_t = 2$) does not necessarily increase the degree of asymmetry. What is more important is the extent of persistence in the downside and normal regimes relative to each other. Too much difference, so that the down-regime is not at all persistent ($Q = 0$), will make the RS Normal Model act like a Jump Model and produce tent-shapes. However, the case of $P = 1 - Q$ is a simple-switching model, where the regimes have no persistence. This case produces produces less correlation asymmetry than when both regimes are persistent. The bottom left panel of Figure (6) shows what happens when Q is increased but comes very close to $1 - P$. The persistence through time of the two regimes drives the persistence across exceedance levels ϑ of the exceedance correlations. Unfortunately sometimes the persistence across the exceedances ϑ when $\vartheta \rightarrow \pm\infty$ cannot be matched by the RS Normal Model.

The final model, the RS-GARCH can capture the correct direction of asymmetry as the RS Normal Model, but the extra covariance persistence in the RS-GARCH process (equation (23)) means that this model can more successfully match the persistence in the exceedance correlations across the exceedance levels. Figure (5) shows that when the RS-GARCH Model fails the implied exceedance correlations can be "too persistent" across ϑ, and have the wrong level. The bottom right panel of Figure (6) shows the RS-GARCH Model exceedance correlations against the sample exceedance correlations. The panel also shows what happens to the exceedance correlation when the probability of staying in the normal regime given that we are in the normal regime (P in equation (19) if the normal regime is $s_t = 1$) increases. In this case the exceedance correlations switch sign so that correlations increase on the upside. In general, the superior performance of this model arises from being able to produce asymmetries of the right direction, as the RS Normal Model, and adding the ability to match exceedance correlation persistence across ϑ.

In summary, of the four empirical discrete-time models we considered – an asymmetric GARCH-M model, a Poisson Jump model, a RS Normal model, and a RS-GARCH model – no single model captures all of the asymmetry in correlations observed in the data. The GARCH-M Model produces correlation asymmetry which is persistent across the exceedance levels, but this correlation asymmetry goes the wrong way unless a negative price of risk is estimated. The Jump Model is rejected almost uniformly across all the portfolios, showing the importance of allowing for persistent volatility and covariance effects. Volatility persistence cannot be captured in a pure Jump Model. The RS Normal Model can produce the correct sign of correlation asymmetry and provides the best fit with the data. It generally produces the lowest H statistics across all the models considered here. However, this model does match the persistence of the asymmetries across exceedance levels. The RS-GARCH Model is rejected by the data least frequently, and is able to match the persistence of the asymmetries across exceedance levels. Our results point to the need for the development of more sophisticated empirical models to capture the empirical asymmetric correlations. These models must capture persistent volatility effects as well as capture more asymmetric correlation patterns than the models presented here.

6 Conclusion

Correlations between domestic equity portfolios and the aggregate market are greater in downside markets than in upside markets. Graphs of correlations conditional on upside or downside movements dramatically illustrate this point. To quantify the effects shown in these plots, we develop an H statistic to measure the asymmetries in correlations. Unlike previous literature which examines covariance asymmetry in the context of the class of asymmetric GARCH models, we can assess the extent of correlation asymmetry in the data relative to any particular model. Moreover, the statistic we develop has the advantage of allowing us to succinctly measure correlation asymmetries, easily compare the degree of asymmetries across portfolios, frequencies and null distributions, and formally conduct statistical tests of asymmetries.

Asymmetries between upside and downside correlations exist between stocks in a single market, as well as across markets internationally. We find that correlation asymmetries are fundamentally different from other measures of asymmetries such as skewness and co-skewness, and tend to be inversely related to systematic beta risk. We examine the sources of correlation asymmetries and find greater asymmetries among smaller stocks, value stocks and recent losers. Correlation asymmetry is the largest among traditional ‘defensive’ sectors such as petroleum and utilities. We find that riskier stocks measured by a higher beta have lower correlation asymmetry, and controlling for size, the degree of correlation asymmetry is unrelated to leverage. Overall, a typical portfolio exhibits correlations conditional on the downside and upside that differ from those of a normal distribution by around 8.5%.

We examine several empirical models to see if they can account for the correlation asymmetries in the data. Normal distributions are overwhelmingly rejected by the data. We estimate asymmetric GARCH-M models, Poisson Jump models, regime-switching Normal models and regime-switching GARCH models. Of these, the regime-switching Normal model is the most able to match the magnitude of empirical correlation asymmetries, while the regime-switching

GARCH model is statistically rejected least often. The popular CAPM-based GARCH-M models can produce asymmetric correlations but these go the wrong way unless a negative price of risk is used. Poisson Jump models fail to capture the persistence of covariance dynamics in the data and capture almost no asymmetric correlation effects. While regime-switching models perform best in empirically explaining the amount of correlation asymmetry in the data, these models still leave a significant amount of correlation asymmetry in the data unexplained.

One implication of our results is for empirical and theoretical asset-pricing. Harvey and Siddique (2000) demonstrate that non-linearities in third moments are priced. Since asymmetric correlations are different from skewness or co-skewness, asymmetric correlations may also play a role in an asset-pricing model. One example is an economy with a representative agent with First Order Risk Aversion (see Ang, Bekaert and Liu (2000)) or Loss Aversion (see Barberis, Huang and Santos (2001)) preferences. Such an investor asymmetrically treats gains and losses and is very averse to downside risks. Our H statistic quantifies asymmetric correlation risk on the downside which also may be priced. Our results also have implications on portfolio allocations and risk management.

Our work raises the question: why do asymmetric movements in asset returns arise in the first place? They may reflect some particular structure of the macro-economy or some intricate interactions of economic agents in equilibrium. While Dumas, Harvey and Ruiz (2000) show that aggregate characteristics affect returns across countries, we show that cross-sectional firm characteristics are related to the magnitudes of asymmetric correlations within a domestic market. Modern equilibrium models with noise traders and frictions in the economy, like Kyle and Xiong (1999), are a long way from explaining the relation between firm characteristics and asymmetric movements. Other explanations for the asymmetric movements may be due to the interaction of disparately informed agents with market frictions as modeled by Hong and Stein (1999). These authors do not model cross-sectional differences between individual asset characteristics. Our work shows that these differences in firm characteristics are related to the asymmetries in asset returns.

Appendix

A Proposition

Let $X = (x,\ y) \sim N(0, \Sigma)$, where Σ has unit variances and unconditional correlation ρ. We define:

$$\hat{\rho}(h_1, h_2, k_1, k_2) = \text{corr}(x, y | h_1 < x < h_2, k_1 < y < k_2; \rho) \tag{A-1}$$

as the correlation of x and y conditional on observations for which $h_1 < x < h_2$ and $k_1 < y < k_2$, where x and y have unconditional correlation ρ.

Let $L(\cdot)$ denote the cumulative density of a doubly truncated bivariate normal distribution :

$$L(h_1, h_2, k_1, k_2) = \int_{h_1}^{h_2} \int_{k_1}^{k_2} g(x, y; \rho) dx dy, \tag{A-2}$$

where

$$g(x, y; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2) \right)$$

is the density function of X. $L(\cdot)$ can be evaluated by numerical methods.

The following Proposition allows us to obtain a closed-form solution for $\hat{\rho}$:

Proposition A.1 *Let* $m_{ij} = \text{E}(x^i y^j | h_1 < x < h_2, k_1 < y < k_2)$. *Then*

$$\begin{aligned}
m_{10} &= \left(\frac{1}{L(\cdot)} \right) [\psi(h_1, h_2, k_1, k_2; \rho) + \rho\psi(k_1, k_2, h_1, h_2; \rho)] \\
m_{20} &= \left(\frac{1}{L(\cdot)} \right) \left[L(\cdot) + \chi(k_1, k_2, h_1; \rho) - \chi(k_1, k_2, h_2; \rho) + \rho^2\chi(h_1, h_2, k_1; \rho) - \rho^2\chi(h_1, h_2, k_2; \rho) \right] \\
m_{11} &= \left(\frac{1}{L(\cdot)} \right) [\rho L(\cdot) + \rho\Upsilon(h_1, h_2, k_1; \rho) - \rho\Upsilon(h_1, h_2, k_2; \rho) + \rho\Upsilon(k_1, k_2, h_1; \rho) \\
&\qquad - \rho\Upsilon(k_1, k_2, h_2; \rho) + \Lambda(h_1, h_2, k_1; \rho) - \Lambda(h_1, h_2, k_2; \rho)]
\end{aligned} \tag{A-3}$$

where $\psi(\cdot)$, $\chi(\cdot)$, $\Upsilon(\cdot)$ *and* $\Lambda(\cdot)$ *are given in the proof. The moments* m_{01} *and* m_{02} *are obtained by interchanging* $(h_1,\ h_2)$ *and* $(k_1,\ k_2)$ *in the formulae for* m_{10} *and* m_{20}.

From the Proposition:

$$\begin{aligned}
\text{var}(x | h_1 < x < h_2, k_1 < y < k_2) &= m_{20} - m_{10}^2 \\
\text{var}(y | h_1 < x < h_2, k_1 < y < k_2) &= m_{02} - m_{01}^2 \\
\text{cov}(x, y | h_1 < x < h_2, k_1 < y < k_2) &= m_{11} - m_{10}m_{01}
\end{aligned} \tag{A-4}$$

which allows us to calculate $\hat{\rho}(h_1, h_2, k_1, k_2)$ as

$$\hat{\rho}(h_1, h_2, k_1, k_2) = \frac{\text{cov}(x, y | h_1 < x < h_2, k_1 < y < k_2)}{\sqrt{\text{var}(x | h_1 < x < h_2, k_1 < y < k_2)}\sqrt{\text{var}(y | h_1 < x < h_2, k_1 < y < k_2)}}$$

A.1 Proof of Proposition

Let

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$$

denote the $N(0, 1)$ density and

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-\frac{w^2}{2}) dw$$

denote the cumulative distribution function of $N(0, 1)$.

A.2 First moment

The first moment m_{10} is obtained from the definition:

$$m_{10} = \frac{1}{2\pi\sqrt{1-\rho^2}L(\cdot)} \int_{k_1}^{k_2} \int_{h_1}^{h_2} x \exp\left(-\frac{1}{2}\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dxdy. \tag{A-5}$$

The equation for m_{01} is similar, by symmetry. Make the change of variable $z = (x - \rho y)/(\sqrt{1-\rho^2})$, and let $v_1 = (h_1 - \rho y)/(\sqrt{1-\rho^2})$ and $v_2 = (h_2 - \rho y)/(\sqrt{1-\rho^2})$. We re-write m_{10} as:

$$\begin{aligned} m_{10}L(\cdot) = \frac{\sqrt{1-\rho^2}}{2\pi} \int_{k_1}^{k_2} \left[-\exp\left(-\frac{1}{2}(z^2+y^2)\right)\right]_{z=v_1}^{z=v_2} dy \\ + \frac{\rho}{2\pi}\int_{k_1}^{k_2} y\exp\left(-\frac{y^2}{2}\right)\left[\int_{v_1}^{v_2}\exp\left(-\frac{z^2}{2}\right)dz\right]dy \end{aligned} \tag{A-6}$$

The second term of equation (A-6) is $\rho\, m_{01}L(\cdot)$, and the first term can be written, after a further change of variable and integration by parts as $(1-\rho^2)\psi(h_1, h_2, k_1, k_2)$, where

$$\begin{aligned} \psi(h_1, h_2, k_1, k_2; \rho) = \phi(h_1)\left[\Phi\left(\frac{k_2 - \rho h_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{k_1 - \rho h_1}{\sqrt{1-\rho^2}}\right)\right] \\ - \phi(h_2)\left[\Phi\left(\frac{k_2 - \rho h_2}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{k_1 - \rho h_2}{\sqrt{1-\rho^2}}\right)\right]. \end{aligned} \tag{A-7}$$

By symmetry we have:

$$\begin{aligned} m_{10}L(\cdot) &= (1-\rho^2)\psi(h_1, h_2, k_1, k_2; \rho) + \rho m_{01}L(\cdot) \\ m_{01}L(\cdot) &= (1-\rho^2)\psi(k_1, k_2, h_1, h_2; \rho) + \rho m_{10}L(\cdot), \end{aligned} \tag{A-8}$$

hence

$$m_{10}L(\cdot) = \psi(h_1, h_2, k_1, k_2; \rho) + \rho\psi(k_1, k_2, h_1, h_2; \rho) \tag{A-9}$$

and m_{01} is given by interchanging the order of h_1, h_2, and k_1 and k_2.

A.3 Second Moment

By definition:

$$m_{20} = \frac{1}{2\pi\sqrt{1-\rho^2}L(\cdot)} \int_{k_1}^{k_2} \int_{h_1}^{h_2} x^2 \exp\left(-\frac{1}{2}\frac{x^2 - 2\rho xy + y^2}{(1-\rho^2)}\right) dxdy. \tag{A-10}$$

Using the same change of variables as above, we have:

$$\begin{aligned} m_{20}L(\cdot) = \frac{1}{2\pi}\int_{k_1}^{k_2}\left[(z(1-\rho^2) + 2\rho\sqrt{1-\rho^2}y)\exp\left(-\frac{z^2}{2}\right)\right]_{z=v_1}^{z=v_2}\exp\left(-\frac{y^2}{2}\right)dy \\ + \frac{1}{2\pi}\int_{k_1}^{k_2}[(1-\rho^2)+]\rho^2y^2]\int_{v_1}^{v_2}\exp\left(-\frac{z^2}{2}\right)dz\exp\left(-\frac{y^2}{2}\right)dy \end{aligned} \tag{A-11}$$

The first term equals $(1-\rho^2)L(\cdot) + \rho^2 m_{02}L(\cdot)$ and the second term, after a further change of variables and integration by parts can be written as:

$$(1-\rho^4)\left(\chi(k_1, k_2, h_1; \rho) - \chi(k_1, k_2, h_2; \rho)\right),$$

where:

$$\chi(k_1,k_2,h_1;\rho) = h_1\phi(h_1)\left[\Phi\left(\frac{k_2-\rho h_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{k_1-\rho h_1}{\sqrt{1-\rho^2}}\right)\right] + \frac{\rho\sqrt{1-\rho^2}}{\sqrt{2\pi}(1+\rho^2)}\left[\phi\left(\frac{\sqrt{k_1^2-2\rho k_1 h_1+h_1^2}}{\sqrt{1-\rho^2}}\right) - \phi\left(\frac{\sqrt{k_2^2-2\rho k_2 h_1+h_1^2}}{\sqrt{1-\rho^2}}\right)\right]. \quad \text{(A-12)}$$

By symmetry we have:

$$\begin{aligned} m_{20}L(\cdot) &= L(\cdot)((1-\rho^2)+\rho^2 m_{02}) + (1-\rho^4)\left(\chi(k_1,k_2,h_1;\rho)-\chi(k_1,k_2,h_2;\rho)\right) \\ m_{02}L(\cdot) &= L(\cdot)((1-\rho^2)+\rho^2 m_{20}) + (1-\rho^4)\left(\chi(h_1,h_2,k_1;\rho)-\chi(h_1,h_2,k_2;\rho)\right), \end{aligned} \quad \text{(A-13)}$$

and solving these gives:

$$m_{20}L(\cdot) = L(\cdot) + \chi(k_1,k_2,h_1;\rho) - \chi(k_1,k_2,h_2;\rho) + \rho^2\chi(h_1,h_2,k_1;\rho) - \rho^2\chi(h_1,h_2,k_2;\rho). \quad \text{(A-14)}$$

A.4 Cross Moment

By definition:

$$m_{11} = \frac{1}{2\pi\sqrt{1-\rho^2}L(\cdot)}\int_{k_1}^{k_2}\int_{h_1}^{h_2} xy\exp\left(-\frac{1}{2}\frac{x^2-2\rho xy+y^2}{(1-\rho^2)}\right)dxdy. \quad \text{(A-15)}$$

Using the same change of variables as above, we have:

$$m_{11}L(\cdot) = \frac{1}{2\pi}\int_{k_1}^{k_2}\int_{v_1}^{v_2}\rho y^2\exp\left(-\frac{1}{2}(y^2+z^2)\right)dzdy + \frac{1}{2\pi}\int_{k_1}^{k_2}(\sqrt{1-\rho^2}y)\left[-\exp\left(-\frac{z^2}{2}\right)\right]_{z=v_1}^{z=v_2}\exp\left(-\frac{y^2}{2}\right)dy. \quad \text{(A-16)}$$

The first term in equation (A-16) is $\rho\, m_{02}L(\cdot)$, and the second term can be written as (after a change of variables and integration by parts):

$$\rho(1-\rho^2)(\Upsilon(k_1,k_2,h_1;\rho)-\Upsilon(k_1,k_2,h_2;\rho)) + \frac{(1-\rho^4)}{(1+\rho^2)}(\Lambda(k_1,k_2,h_1;\rho)-\Lambda(k_1,k_2,h_2;\rho)),$$

where:[19]

$$\begin{aligned} \Upsilon(k_1,k_2,h_1;\rho) &= h_1\phi(h_1)\left[\Phi\left(\frac{k_2-\rho h_1}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{k_1-\rho h_1}{\sqrt{1-\rho^2}}\right)\right] \\ \Lambda(k_1,k_2,h_1;\rho) &= \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}}\left[\phi\left(\frac{\sqrt{k_1-2\rho k_1 h_1+h_1^2}}{\sqrt{1-\rho^2}}\right) - \phi\left(\frac{\sqrt{k_2-2\rho k_2 h_1+h_1^2}}{\sqrt{1-\rho^2}}\right)\right]. \end{aligned} \quad \text{(A-17)}$$

After simplification we can write m_{11} as:

$$m_{11}L(\cdot) = \rho L(\cdot) + \rho\Upsilon(h_1,h_2,k_1;\rho) - \rho\Upsilon(h_1,h_2,k_2;\rho) + \rho\Upsilon(k_1,k_2,h_1;\rho) - \rho\Upsilon(k_1,k_2,h_2;\rho) + \Lambda(h_1,h_2,k_1;\rho) - \Lambda(h_1,h_2,k_2;\rho). \quad \text{(A-18)}$$

[19] Note that $\chi(a,b,c;\rho) = \Upsilon(a,b,c;\rho) + \frac{\rho}{1+\rho^2}\Lambda(a,b,c;\rho)$. Also note that $(\Lambda(k_1,k_2,h_1;\rho)-\Lambda(k_1,k_2,h_2;\rho)) = (\Lambda(h_1,h_2,k_1;\rho)-\Lambda(h_1,h_2,k_2;\rho))$.

B Solution of the Asset Allocation Problem in Section 2

The first order conditions (FOC) of the investor's investment problem are:

$$\mathrm{E}_t(W^{-\gamma}x) = 0, \tag{B-1}$$

where $W = 1 + r_f + \alpha x + \alpha y$. Since x and y have the same distribution (but are correlated) the portfolio holding in each asset is identical. This expectation can be computed by numerical quadrature (as described in Tauchen and Hussey (1991)):

$$\sum_{s=1}^{M}(W_s^{-\gamma}x_s p_s) = 0 \tag{B-2}$$

where the M values of the risky asset returns ($\{x_s\}_{s=1}^M$ and $\{y_s\}_{s=1}^M$) and associated probabilities are chosen by an optimal quadrature rule. W_s represents the investor's terminal wealth when the risky asset returns are x_s and y_s. Tauchen and Hussey (1991) demonstrate that quadrature is very accurate with very few optimally chosen points. The FOC in equation (B-2) can be solved over α by a non-linear root solver.

When x and y are bivariate normally distributed, Gaussian quadrature is used with 5 points to approximate the distribution of x and y. Hence we use $M = 5 \times 5 = 25$ quadrature points. Correlation is achieved by using a Cholesky decomposition transformation.

When $X = (x, y)'$ is drawn from the RS Model, we approximate the joint distribution as follows. For regime $s_t = 1$ we approximate the normal distribution $N(\mu_1, \Sigma_1)$ using 25 quadrature points, constructed as per the case of the bivariate normal distribution. For regime $s_t = 2$ another 25 quadrature points are used. Then conditional on regime $s_t = 1$ we use weights P and $1 - P$, where $P = Pr(s_t = 1|s_{t-1} = 1)$, to mix the associated probabilities of the quadrature points of regimes 1 and 2 to give give an $M = 50$ quadrature point approximation to the RS Model conditional on regime 1. Conditional on regime $s_t = 2$ we use weights $1 - Q$ and Q, where $Q = Pr(s_t = 2|s_{t-1} = 2)$, to mix the associated probabilities of the quadrature points of regimes 1 and 2.

To match the first and second moments of the RS Model to the unconditional means, volatilities and correlation of the normal distribution, we note that the unconditional mean of the RS Model is given by:

$$\pi\mu_1 + (1 - \pi)\mu_2 \tag{B-3}$$

where $\pi = Pr(s_t = 1)$ is the stable probability of the RS Model which is given by

$$\pi = \frac{1 - Q}{2 - P - Q}$$

and the unconditional covariance is given by:

$$\pi(\Sigma_1 + \mu_1\mu_1') + (1 - \pi)(\Sigma_2 + \mu_2\mu_2') - (\pi\mu_1 + (1 - \pi)\mu_2)(\pi\mu_1 + (1 - \pi)\mu_2)' \tag{B-4}$$

By exogenous choices of $P = Q = 2/3$, $\mu_1 = \mu_2 = (0.07, 0.07)'$, $\sigma_1 = \sigma_2 = 0.15$, and the stable probability $\pi = 1/2$, the unconditional means (volatilities) of x and y using the RS Model are both 0.07 (0.15) and we can choose ρ_1 and ρ_2 to produce the unconditional correlation ρ by setting $\frac{1}{2}(\rho_1 + \rho_2) = \rho$.

We produce a particular H as follows. We choose ρ_2 (say $\rho_2 = 0.35$), which determines $\rho_1 = 0.65$. This gives the RS Model the same unconditional means, volatilities and correlation as the bivariate normal distribution. Then we calculate corr$(\tilde{x}, \tilde{y}|\tilde{x} < -1,\ \tilde{y} < -1)$ for x and y drawn from the RS Model by using simulation with 100,000 draws. This will be greater than the correlation with the same conditioning calculated from the bivariate normal (given in Appendix A in closed form). The difference between corr$(\tilde{x}, \tilde{y}|\tilde{x} < -1,\ \tilde{y} < -1)$ calculated from the RS Model and from the bivariate normal gives H. To produce Figure (1) we choose $\rho_2 \in \{0.19, 0.20, \ldots 0.48\}$.

C Data Construction

For our empirical analysis, we use data from the Center for Research in Security Prices (CRSP) and Standard & Poor's COMPUSTAT to construct portfolios based on various firm and distributional characteristics. We use both daily and monthly returns from CRSP for the period covering July 1st, 1963 to December 31st, 1998. We use COMPUSTAT's annual files to obtain information about book values and financial leverage. We follow standard conventions and restrict our universe to common stocks listed on NYSE, AMEX or NASDAQ of companies incorporated in the United States. For the risk-free rate, we use the one-month Treasury Bill rate provided by Ibbotson Associates.

We first construct a set of value-weighted industry portfolios grouped by their two-digit Standard Industrial Classification (SIC) codes. The classification of these industries follow exactly that of the SIC grouping used in Ferson and Harvey (1991). In addition, we group all stocks that do not fall into this classification scheme into a 'miscellaneous' industry. The industries analyzed are miscellaneous, petroleum, finance, durables, basic industries, food and tobacco, construction, capital goods, transportation, utilities, textile and trade, service and leisure.

Within each month, for each portfolio, we calculate daily returns of a buy-and-hold strategy using the CRSP daily file. At the beginning of every month, each portfolio is rebalanced and reformed according to the strategy. The returns are aggregated into weekly frequency by calculating the total buy-and-hold return of each strategy from the end of every Wednesday to the end of the following Wednesday. The monthly returns are calculated directly from the CRSP monthly file, and are also rebalanced and reformed at the beginning of every month. Finally, all returns are converted into continuously compounded yields and expressed as returns in excess of the one-month T-bill rate.

The second set of portfolios we construct are value-weighted size-sorted portfolios. At the beginning of every month, we determine the breakpoints on market capitalization for our stocks based on the quintile breakpoints of stocks listed on the NYSE. Hence, our first size-sorted portfolio contains all the stocks listed on the combined NYSE/AMEX/NASDAQ that are smaller than the 20^{th} percentile NYSE stock.

The third set of portfolios we construct are value-weighted book-to-market portfolios. At the beginning of every month, our universe of stocks is once again sorted based on quintile breakpoints of stocks listed on the NYSE. The sorting variable is the book-to-market calculated using the most recently available fiscal year-end balance sheet data on COMPUSTAT. Following Fama and French (1993), we define 'book value' as the value of common stock holders' equity, plus deferred taxes and investment tax credit, minus the book value of preferred stock. The book value is then divided by the market value on the day of the firm's fiscal year-end.

The next set of portfolios consists of the '6-6' momentum strategy portfolios of Jagadeesh and Titman (1993). This time, we sort our stocks based upon the past six-months returns of all stocks in our universe, rather than just on NYSE stocks. Each month, an equal-weighted portfolio is formed based on six-months returns ending one month prior. To avoid market microstructural effects, we require a one month lag between when the returns are realized and when the portfolios are formed. Similarly, equal-weighted portfolios are formed based on past returns that ended two months prior, three months prior, and so on up to six months prior. We then take the simple average of six such portfolios. Hence, our first momentum portfolio consists of $1/6$ of the returns of the worst performers one month ago, plus $1/6$ of the returns of the worst performers two months ago, etc.

The next two sets of portfolios are based on distributional characteristics of past returns. Beta with respect to the market is estimated for each month as the regression coefficient of last 60 months market monthly excess returns on the prior 60 months portfolio monthly excess returns. Standardized coskewness is calculated every month for every stock using past one year daily stock returns. The final set of portfolios are formed according to firm leverage. Leverage is calculated annually as total assets divided by book value, where book value is defined as above. Leverage for a given month is defined as the mostly recently reported value. As with size and book-to-market portfolios, we compute quintile breakpoints based on stocks listed on NYSE and value-weighted portfolios are formed.

In addition, we create two sets of doubly sorted portfolios: one sorted on size and beta and the another sorted on size and leverage. For both sets, we first sort every stock in our universe by size into quintiles using NYSE breakpoints. Then within each size quintile group, we further sort stocks into quintiles based on beta. The breakpoints for beta within each size quintile are also calculated using only NYSE stocks. We then form value-weighted portfolios according to the 5×5 groupings. Size and leverage portfolios are formed the same way, except that we use leverage rather than beta.

Finally, we take CRSP's value-weighted return of all stocks to be used as the 'market' portfolio.

D Calculating H Statistics for Non-Normal Distributions

To calculate the H statistics using the null distribution of the empirical models presented in Section 5 we need to calculate the implied exceedance correlations $\breve{\rho}(\theta, \phi)$ by simulation. Denote the distribution under the null as $\xi(\phi)$, where ξ represents one of the models from Section 5 with parameters ϕ. For each equity portfolio, we estimate the parameters ϕ of the model. Then at the estimated parameters, we create a simulated time-series with 100,000 observations. We take the exceedance correlations of the simulated time-series as the exceedance correlation implied by the distribution, $\breve{\rho}(\theta, \phi)$.

As with the case with a normal distribution, we would like to calculate the standard error of H. Suppose the estimated parameters $\hat{\phi}$ have covariance matrix Γ. Using equation (11) we can calculate the sample variance of H if the derivative $D_1 = \frac{\partial}{\partial \phi} H$ can be numerically computed. To do this, we note that $\breve{\rho}(\vartheta, \phi)$, the exceedance correlation implied by $\xi(\phi)$, is just a function of ϕ which can be computed by simulation. Hence we construct

H (equation (9)) by simulating $\breve{\rho}(\vartheta, \phi)$. Holding fixed the simulated errors involved in the simulation, we then numerically compute the derivative D_1. Using the same simulated errors, we change the i-th parameter in $\hat{\phi}$ by $\epsilon = 0.0001$, and re-compute the simulated time series at the new parameters. This new time series is used to calculate a new implied exceedance correlation which we denote $H_i(\theta, \phi)$. The i-th element of D_1 can be estimated with the directional derivative (Gateaux derivative) for an increment of ϵ in the i-th parameter of $\hat{\phi}$, given by $(H_i(\theta, \phi) - H(\theta, \phi))/\epsilon$. Hence, this statistic is completely analogous to the one defined for the normal distribution. The only difference is that the statistic and its standard errors are calculated by simulation.

References

[1] Ang, A., and G. Bekaert, 2000, "International Asset Allocation with Time-Varying Correlations," working paper Columbia University.

[2] Ang, A., G. Bekaert, and J. Liu, 2000, "Why Stocks May Disappoint," working paper Columbia University.

[3] Bae, K. H., G. A. Karolyi, and R. M. Stulz, 2000, "A New Approach to Measuring Financial Cotagion," NBER working paper 7913.

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

[5] Banz, R., 1981, "The Relation between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, 3-18.

[6] Barberis, N., M. Huang, and J. Santos, 2001, "Prospect Theory and Asset Prices," forthcoming *Quarterly Journal of Economics*.

[7] Bekaert, G., C. R. Erb, C. R. Harvey, and T. E. Viskanta, 1998, "Distributional Characteristics of Emerging Market Returns and Asset Allocation," *Journal of Portfolio Management*, 24, 2, 102-116.

[8] Bekaert, G., and C. R. Harvey, 2000, "Foreign Speculators and Emerging Equity Markets," *Journal of Finance*, 55, 2, 565-613.

[9] Bekaert, G., and C. R. Harvey, 1997, "Emerging Equity Market Volatility," *Journal of Financial Economics*, 43, 29-77.

[10] Bekaert, G., and G. Wu, 2000, "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, 1, 1-42.

[11] Boyer, M. S., B. H. Gibson, and M. Loretan, 1999, "Pitfalls in Tests for Changes in Correlations," International Finance Discussion Paper 597, Board of Governors of the Federal Reserve System.

[12] Braun, P. A., D. N. Nelson and A. M. Sunier, 1995, "Good News, Bad News, Volatility and Betas," *Journal of Finance*, 50, 5, 1575-1603.

[13] Butler, K. C., and D. C. Joaquin, 2000, "Are the Gains from International Portfolio Diversification Exaggerated? The Influence of Downside Risk in Bear Markets," working paper.

[14] Campbell, C. J., and L. Hentschel, 1992, "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns," *Journal of Financial Economics*, 31, 281-318.

[15] Campbell, C. J., M. Lettau, B. G. Malkiel, and Y. Xu, 2001, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," forthcoming *Journal of Finance*, 56, 1.

[16] Chen, J., H. Hong, and J. Stein, 2000, "Forecasting Crashes: Trading Volume, Past Returns and Conditional Skewness in Stock Prices," forthcoming *Journal of Financial Economics*.

[17] Cheung, Y. W., and L. K. Ng, 1992, "Stock Price Dynamics and Firm Size: An Empirical Investigation," *Journal of Finance*, 47, 5, 1985-1997.

[18] Cho, Y. H., and R. F. Engle, 2000, "Time-Varying Betas and Asymmetric Effects of News: Empirical Analysis of Blue Chip Stocks," NBER working paper 7330.

[19] Conrad, J., M. Gultekin, and G. Kaul, 1991, "Asymmetric Predictability of Conditional Variances," *Review of Financial Studies*, 4, 597-622.

[20] Das, S., and R. Uppal, 1999, "The Effect of Systemic Risk on International Portfolio Choice," working paper, University of British Columbia.

[21] De Santis, G., B. Gerard and P. Hillion, 1999, "The Relevance of Currency Risk in the EMU," working paper, UCLA.

[22] Duffee, G. R., 1995, "Stock Returns and Volatility: A Firm-Level Analysis," *Journal of Financial Economics*, 37, 399-420.

[23] Duffee, G. R., 2000, "Asymmetric Cross-Sectional Dispersion in Stock Returns: Evidence and Implications," working paper, U.C. Berkeley.

[24] Dumas, B., C. R. Harvey, and P. Ruiz, 2000, “Are Correlations of Stock Returns Justified by Subsequent Changes in National Outputs,” working paper, Duke University.

[25] Engle, R. F., and K. F. Kroner, 1995, “Multivariate Simultaneous Generalized ARCH,” *Econometric Theory*, 11, 122-150.

[26] Engle, R. F., and V. Ng, 1993, “Measuring and Testing the Impact of News on Volatility,” *Journal of Finance*, 48, 1749-1778.

[27] Erb, C., C. Harvey, and T. Viskanta, 1994, “Forecasting International Equity Correlations,” *Financial Analysts Journal*, Nov/Dec, 32-45.

[28] Fama, E. F., and K. R. French, 1993, “Common risk factors in the returns on stocks and bonds,” *Journal of Financial Economics*, 33, 3-56.

[29] Fama, E. F., and K. R. French, 1996, “Multifactor Explanations of Asset Pricing Anomalies,” *Journal of Finance*, 51, 1, 55-84.

[30] Ferson, W. E., and C. R. Harvey, 1991, “The Variation of Economic Risk Premiums,” *Journal of Political Economics*, 99, 385-415.

[31] Forbes, K. and R. Rigobon, 1999, “No Contagion, Only Interdependence: Measuring Stock Market Co-Movements,” NBER working paper 7267.

[32] French, K. R., G. W. Schwert and R. Stambaugh, 1987, “Expected Stock Returns and Volatility,” *Journal of Financial Economics*, 19, 3-29.

[33] Glosten, L. R., R. Jagannathan and D. E. Runkle, 1993, “On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks,” *Journal of Finance*, 48, 5, 1770-1801.

[34] Gray, S. F., 1996, “Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process,” *Journal of Financial Economics*, 42, 27-62.

[35] Hamilton, J. D., 1989, “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,” *Econometrica*, 57, 357-384.

[36] Hansen, L., and R. Jagannathan, 1997, “Assessing Specification Errors in Stochastic Discount Factor Models,” *Journal of Finance*, 52, 2, 557-590.

[37] Harvey, C. R., and A. Siddique, 2000, “Conditional Skewness in Asset Pricing Tests,” *Journal of Finance*, 55, 3, 1263-1295.

[38] Harvey, C. R., and A. Siddique, 1999, “Autoregressive Conditional Skewness,” *Journal of Financial and Quantiative Analysis*, 34, 4, 465-488.

[39] Hentschel, L., 1995, “All in the Family: Nesting Symmetric and Asymmetric GARCH Models,” *Journal of Financial Economics*, 39, 71-104.

[40] Hong, H., and J. Stein, 1999, “Differences of Opinion, Rational Arbitrage and Market Crashes,” working paper, NBER.

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

[42] Karolyi, A., and R. Stulz, 1996, “Why do Markets Move Together? An Investigation of US-Japan Stock Return Comovements,” *Journal of Finance*, 51, 3, 951-986.

[43] Kroner, K. F., and V. K. Ng, 1998, “Modeling Asymmetric Comovements of Asset Returns,” *Review of Financial Studies*, 11, 4, 817-844.

[44] Kyle, A. S., and W. Xiong, 1999, “Contagion as a Wealth Effect of Financial Intermediaries,” working paper, Duke University

[45] Lin, W. L., R. F. Engle, and T. Ito, 1994, “Do Bulls and Bears Move Across Borders? International Transmission of Stock Returns and Volatility,” *Review of Financial Studies*, 7, 507-538.

[46] Longin, F., and B. Solnik, 2001, “Extreme Correlation of International Equity Markets” forthcoming *Journal of Finance*.

[47] Longin, F., and B. Solnik, 1995, “Is the Correlation in International Equity Returns Constant: 1960-1990?,” *Journal of International Money and Finance*, 14, 3-26.

[48] Newey, W. K., and K. D. West, 1987, “A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55, 703-8.

[49] Pan, J., 2000, “Integrated Time-Series Analysis of Spot and Option Prices,” working paper, M.I.T.

[50] Rosenbaum, S., 1961, “Moments of a Truncated Bivariate Normal Distribution,” *Journal of the Royal Statistical Society* (Ser. B) 23, 2, 405-8.

[51] Scholes, M, and J. T. Williams, 1977, “Estimating Betas from Nonsyncronous Data,” *Journal of Financial Economics*, 5, 3, 309-27.

[52] Schwert, G. W., 1989, “Why Does Stock Market Volatility Change Over Time?,” *Journal of Finance*, 44, 1115-1153.

[53] Stambaugh, R., 1995, Unpublished discussion of Karolyi and Stulz (1996) at the May 1995 NBER Universities Research Conference on Risk Management.

[54] Stărică, 2000, “Multivariate Extremes for Models with Constant Conditional Correlations,” *Journal of Empirical Finance*, 6, 515-553.

[55] Stivers, C., 2000, “Firm-Level Return Dispersion and the Future Volatility of Aggregate Stock Market Returns,” working paper.

[56] Tauchen, G., and R. Hussey, 1991, “Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models,” *Econometrica*, 59, 2, 371-396.

[57] White, H., 1980, “A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,” *Econometrica*, 48, 817-838.

Table 1: Bias in Correlations Conditioning on ϑ

$\lvert\vartheta\rvert$	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$
0.0	-0.2373	-0.2035	-0.1312
0.5	-0.3009	-0.2675	-0.1818
1.0	-0.3597	-0.3306	-0.2367
1.5	-0.4114	-0.3897	-0.2932
2.0	-0.4555	-0.4428	-0.3488

The table reports the difference $\bar{\rho}(\vartheta) - \rho$, where $\bar{\rho}(\vartheta) = \text{corr}(\tilde{x}, \tilde{y} | \tilde{x} > \vartheta, \tilde{y} > \vartheta; \rho)$, and $\tilde{x}$ and $\tilde{y}$ are drawn from a standardized bivariate normal distribution with correlation ρ.

Table 2: Summary Statistics

Portfolio	Mean	Std Dev	Auto 1	Auto 2	Unconditional Corr with the Market weekly freq	Unconditional Corr with the Market monthly freq
Market Portfolio	0.0657	0.1461	0.0682	0.0035		
Industry Portfolios (value-weighted)						
Misc	0.0314	0.1876	0.1372	0.0457	0.8596	0.8834
Petroleum	0.0535	0.1736	0.0332	-0.0014	0.7063	0.7078
Finance	0.0615	0.1595	0.1198	0.0242	0.9204	0.9170
Durables	0.0520	0.1771	0.0842	0.0254	0.9346	0.9341
Basic Ind	0.0555	0.1566	0.0492	0.0044	0.9459	0.9443
Food/Tobacco	0.0812	0.1429	0.0335	0.0503	0.8656	0.8674
Construction	0.0503	0.1854	0.1081	0.0110	0.8949	0.9015
Capital Goods	0.0483	0.1786	0.0638	0.0002	0.9185	0.9051
Transportation	0.0322	0.2003	0.1015	0.0118	0.8543	0.8507
Utilities	0.0476	0.1154	0.0923	0.0230	0.7975	0.7922
Textile/Trade	0.0602	0.1810	0.1008	0.0348	0.8718	0.8594
Service	0.0735	0.2089	0.1437	0.0334	0.8950	0.8961
Leisure	0.0654	0.1866	0.1358	0.0965	0.8910	0.8881
Size Portfolios (value-weighted)						
1 Smallest	0.0537	0.1633	0.3325	0.1538	0.8227	0.8305
2	0.0660	0.1690	0.2422	0.0799	0.8942	0.9028
3	0.0645	0.1620	0.1934	0.0540	0.9314	0.9336
4	0.0633	0.1558	0.1400	0.0261	0.9664	0.9666
5 Largest	0.0526	0.1452	0.0129	-0.0087	0.9878	0.9869
Book-to-Market Portfolios (value-weighted)						
1 Growth	0.0483	0.1673	0.0366	0.0036	0.9606	0.9545
2	0.0492	0.1534	0.0855	0.0092	0.9660	0.9706
3	0.0495	0.1408	0.0989	0.0037	0.9378	0.9425
4	0.0727	0.1324	0.0824	0.0205	0.9166	0.9133
5 Value	0.0911	0.1427	0.1161	0.0580	0.8751	0.8588
Momentum Portfolios (equal-weighted)						
1 Past Losers	0.0226	0.1706	0.3709	0.1686	0.7828	0.7429
2	0.0603	0.1395	0.3189	0.1447	0.8644	0.8308
3	0.0782	0.1308	0.2912	0.1260	0.9005	0.8734
4	0.0918	0.1360	0.2521	0.1029	0.9109	0.8867
5 Past Winners	0.1114	0.1633	0.2237	0.0862	0.8821	0.8573

Summary statistics for the market and equity portfolios. Frequency is weekly (except for the last column). The number of observations is 1852 (426 for the last column). Data is sampled from July 1963 to December 1998. The mean and the standard deviation have been annualized by multiplying the mean and standard deviation in the data by 52 and $\sqrt{52}$ respectively. The columns Auto 1 and Auto 2 give the first and the second autocorrelations. The last two columns show the unconditional correlation of the portfolios with the market at weekly and monthly frequencies. All returns are log-returns in excess of the annualized 1-month T-bill risk-free rate.

Table 3: Ten Largest Weekly Negative and Positive Moves of the Market Portfolio

Largest Negative Moves		Largest Positive Moves	
21-Oct-87	-0.1937	3-Jun-70	0.0966
28-Oct-87	-0.1078	13-Oct-82	0.0892
2-Sep-98	-0.0888	25-Aug-82	0.0863
20-Nov-74	-0.0666	29-Jan-75	0.0855
22-Aug-90	-0.0656	21-Oct-98	0.0747
29-Oct-97	-0.0617	4-Nov-87	0.0704
14-Aug-74	-0.0615	1-Dec-71	0.0670
31-Jul-74	-0.0603	26-Aug-70	0.0641
10-Dec-80	-0.0597	7-Jan-87	0.0638
7-Oct-98	-0.0583	9-Oct-74	0.0618

We present the ten largest positive and negative moves for the value-weighted market portfolio in excess of the risk-free rate. Data is sampled weekly, from July 1963 to December 1998. Dates reported are end-of-period. Returns are not annualized.

Table 4: Asymmetries in Beta and Volatility

Industry Portfolios (value-weighted)

Portfolio	Unconditional β	Theoretical $\beta^- = \beta^+$	Observed β^-	Observed β^+	Observed k^-	Observed k^+	$k^- = k^+$ p-value
Misc	1.1044	0.8908	1.1552**	0.8131	1.4179	1.2056	0.0286*
Petroleum	0.8394	0.5588	0.7437**	0.6115	1.1553	1.2161	0.5503
Finance	1.0055	0.8852	0.9693**	0.8690	1.1050	1.0756	0.6163
Durables	1.1336	1.0196	1.0916**	0.9525**	1.2064	1.1596	0.3070
Basic Ind	1.0141	0.9281	1.0252**	0.8797*	1.1044	1.0095	0.0026**
Food/Tobacco	0.8471	0.6890	0.8484**	0.6561	1.0158	0.9316	0.2309
Construction	1.1360	0.9632	1.1852**	0.9148	1.3734	1.2236	0.0099**
Capital Goods	1.1233	0.9859	1.0755**	0.9661	1.2312	1.1796	0.2629
Transportation	1.1716	0.9382	1.1207**	0.8646	1.3715	1.3110	0.3237
Utilities	0.6302	0.4683	0.5687**	0.4641	0.7536	0.8196	0.3037
Textile/Trade	1.0803	0.8864	1.0621**	0.8789	1.2843	1.2133	0.3002
Service	1.2799	1.0854	1.2451**	1.0349	1.4653	1.3515	0.1471
Leisure	1.1382	0.9598	1.1314**	0.9018	1.3460	1.2157	0.1454

Size Portfolios (value-weighted)

Portfolio	Unconditional β	Theoretical $\beta^- = \beta^+$	Observed β^-	Observed β^+	Observed k^-	Observed k^+	$k^- = k^+$ p-value
1 Smallest	0.9200	0.7064	1.0093**	0.6471	1.2723	1.0806	0.0232*
2	1.0349	0.8767	1.1094**	0.8055*	1.2991	1.0682	0.0032**
3	1.0333	0.9250	1.1088**	0.8436*	1.2356	1.0151	0.0003**
4	1.0312	0.9751	1.0705**	0.9083**	1.1307	0.9992	0.0002**
5 Largest	0.9820	0.9616	0.9818	0.9213*	1.0085	0.9473	0.0766

Book to Market Portfolios (value-weighted)

Portfolio	Unconditional β	Theoretical $\beta^- = \beta^+$	Observed β^-	Observed β^+	Observed k^-	Observed k^+	$k^- = k^+$ p-value
1 Growth	1.1005	1.0307	1.0761*	1.0091	1.1614	1.1034	0.3015
2	1.0145	0.9584	1.0354**	0.8890	1.0909	0.9845	0.0003**
3	0.9042	0.8172	0.9500**	0.7514	1.0333	0.8987	0.0000**
4	0.8312	0.7275	0.8172**	0.6670	0.9263	0.8615	0.1304
5 Value	0.8548	0.7044	0.8134**	0.6447	0.9848	0.9427	0.5116

The first column of this table shows the unconditional beta observed in the data. The second column shows the beta conditional on an upside or downside move under the normal distribution. The third and fourth columns show the beta conditional on upside and downside moves observed in the data. The fifth and sixth columns show k^- and k^+ where $k^- = \sigma_x^-/\sigma_y^-$, $k^+ = \sigma_x^+/\sigma_y^+$ and $\sigma_x^- = \sqrt{\text{var}(x|x < \mu_x, y < \mu_y)}$, $\sigma_y^- = \sqrt{\text{var}(y|x < \mu_x, y < \mu_y)}$, $\sigma_x^+ = \sqrt{\text{var}(x|x > \mu_x, y > \mu_y)}$ and $\sigma_y^+ = \sqrt{\text{var}(y|x > \mu_x, y > \mu_y)}$. The last column shows the p-value of testing $k^- = k^+$. A '$*$' indicates rejection of a test that the observed value equal the theoretical value at the 5% confidence level, while a '$**$' indicates rejection at the 1% confidence level. Tests for the observed β^- and β^+ are tests if β^- or β^+ equal the theoretical value implied by a normal distribution. P-values for the test of $k^- = k^+$ are done by bootstrap with 1000 simulated samples.

Table 5: H Statistics for the Size Portfolios with the Market

Portfolio	Weighted by Normal Distn $\sigma^2(\breve{\rho})$		Weighted by number of observations		Equally Weighted	
Daily Frequency						
	H Stat	SE	H Stat	SE	H Stat	SE
1 Smallest	0.1520	0.0085	0.1493	0.0083	0.1848	0.0088
2	0.1408	0.0092	0.1323	0.0090	0.1780	0.0097
3	0.1215	0.0080	0.1089	0.0077	0.1508	0.0086
4	0.0814	0.0052	0.0677	0.0047	0.0964	0.0057
5 Largest	0.0232	0.0026	0.0168	0.0022	0.0259	0.0028
Weekly Frequency						
	H Stat	SE	H Stat	SE	H Stat	SE
1 Smallest	0.1499	0.0170	0.1471	0.0161	0.1989	0.0195
2	0.0984	0.0137	0.0913	0.0128	0.1261	0.0154
3	0.0742	0.0108	0.0654	0.0098	0.0922	0.0122
4	0.0492	0.0071	0.0401	0.0060	0.0586	0.0080
5 Largest	0.0125	0.0020	0.0097	0.0017	0.0141	0.0021
Monthly Frequency						
	H Stat	SE	H Stat	SE	H Stat	SE
1 Smallest	0.2138	0.0154	0.2145	0.0142	0.2516	0.0209
2	0.1457	0.0096	0.1353	0.0090	0.1751	0.0119
3	0.1198	0.0061	0.0955	0.0075	0.1615	0.0040
4	0.0566	0.0096	0.0459	0.0085	0.0667	0.0106
5 Largest	0.0430	0.0087	0.0371	0.0080	0.0468	0.0091

We present the H statistics under the null of a bivariate normal distribution for the value-weighted size-sorted portfolios. A different bivariate normal is fitted for each pair of (x, y) where x is the normalized excess market return and y is a normalized excess stock portfolio return. The first two columns use weights constructed using the variances of the exceedance correlations $\breve{\rho}(\vartheta, \phi)$ implied by a bivariate normal distribution, as in equation (12). In columns 3-4 the weights are proportional to how many observations are used to construct each $\bar{\rho}(\vartheta)$, the sample exceedance, as in equation (13). The last two columns use equal weights (equation (14)). The null of a bivariate normal is rejected at the 0.1% confidence level for every portfolio at all frequencies by the H statistics (p-values are not reported). All standard errors are calculated using GMM and 6 Newey-West (1987) lags.

Table 6: H Statistics from a Bivariate Normal Distribution

Panel A : Industry and Momentum Portfolios

Industry Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
Misc	0.1249	0.1742	0.0327	-0.8809	-0.6239	1.1044
Petroleum	0.1801	0.2372	0.0941	-0.1484	-0.3527	0.8394
Finance	0.0737	0.0994	0.0326	-0.3777	-0.4506	1.0055
Durables	0.0778	0.1089	$0.0164^{\ddagger}$	-0.5275	-0.5396	1.1336
Basic Ind	0.0718	0.1014	$0.0154^{\ddagger}$	-0.6702	-0.5798	1.0141
Food/Tobacco	0.1174	0.1631	$0.0325^{\dagger}$	-0.5002^{*}	-0.4770	0.8471
Construction	0.1191	0.1650	0.0370	-0.8748	-0.6489	1.1360
Capital Goods	0.0874	0.1164	0.0433	-0.5092	-0.5146	1.1233
Transportation	0.1321	0.1852	$0.0201^{\ddagger}$	-0.5715	-0.5734	1.1716
Utilities	0.1454	0.2023	$0.0265^{\ddagger}$	-0.1147	-0.4008	0.6302
Textile/Trade	0.1251	0.1651	0.0637	-0.5682	-0.5270	1.0803
Service	0.0944	0.1315	$0.0269^{\dagger}$	-0.5832	-0.5218	1.2799
Leisure	0.0784	0.1098	0.0188	-0.5392^{*}	-0.4993	1.1382

Size Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Smallest	0.1471	0.1977	0.0656	-0.8928	-0.6535	0.9200
2	0.0913	0.1242	0.0386	-0.9526	-0.6292	1.0349
3	0.0654	0.0878	0.0311	-0.9353	-0.6226	1.0333
4	0.0401	0.0542	0.0181	-0.7166	-0.5763	1.0312
5 Largest	0.0097	0.0120	0.0070	-0.5301	-0.5018	0.9820

Book to Market Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Growth	0.0368	0.0449	0.0268	-0.4544	-0.4626	1.1005
2	0.0450	0.0631	0.0117	-0.6623	-0.5690	1.0145
3	0.0795	0.1078	0.0347	-0.9033	-0.6518	0.9042
4	0.0896	0.1215	0.0375	-0.5309	-0.5422	0.8312
5 Value	0.0995	0.1357	$0.0376^{\dagger}$	-0.3980	-0.4946	0.8548

Momentum Portfolios (equal-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Past Losers	0.1653	0.2239	0.0575	-0.1122	-0.4861	0.9144
2	0.1186	0.1626	0.0414	-0.5895	-0.5678	0.8254
3	0.0930	0.1291	0.0282	-1.0381	-0.6758	0.8065
4	0.0772	0.1099	0.0121	-1.3448	-0.7433	0.8482
5 Past Winners	0.0917	0.1303	$0.0161^{\dagger}$	-1.3477^{*}	-0.7451	0.9864

Panel B : Beta, Co-skewness and Leverage Portfolios

Beta Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Low Beta	0.1232	0.1714	0.0378	-0.7616	-0.6334	0.6123
2	0.0527	0.0738	0.0152 †	-0.7331	-0.5834	0.8638
3	0.0566	0.0784	0.0189	-0.7186	-0.5862	0.9767
4	0.0564	0.0762	0.0251	-0.7621	-0.5878	1.1092
5 High Beta	0.0682	0.0907	0.0327	-0.4992	-0.5009	1.3300

Co-skewness Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Low/Neg. Coskew	0.0661	0.0916	0.0224	-0.6864 *	-0.5762	0.9941
2	0.0544	0.0747	0.0204	-0.7592	-0.5962	1.0051
3	0.0574	0.0771	0.0263	-0.6137	-0.5605	0.9899
4	0.0527	0.0721	0.0202	-0.8721	-0.6229	0.9879
5 High/Pos. Coskew	0.0672	0.0903	0.0314	-0.4397	-0.4823	0.9242

Leverage Portfolios (value-weighted)

Portfolio	H Statistic	H^-	H^+	Skewness	Co-skewness	β
1 Low Debt	0.0629	0.0858	0.0250	-0.8207	-0.6184	1.0025
2	0.0432	0.0587	0.0181	-0.4717	-0.4889	0.9568
3	0.0504	0.0684	0.0211	-0.6006	-0.5436	0.9665
4	0.0586	0.0806	0.0217	-0.7351	-0.5989	0.9639
5 High Debt	0.1009	0.1400	0.0299	-0.6131	-0.5768	1.0235

This table presents H, H^+ and H^- statistics for equity portfolios assuming the null of a bivariate normal distribution. Frequency of the data is weekly. Weights proportional to the number of observations in each sample exceedance are used (equation (13)) to construct the H statistics. The null of a bivariate normal is rejected at the 0.1% confidence level for every portfolio at all frequencies by the H statistics (p-values are not reported). For H^+ and H^- statistics '‡' and '†' indicate that the model *cannot* be rejected at the 5% and 1% confidence levels, respectively. For skewness and co-skewness '$*$' indicates rejection of the statistic from zero at the 5% confidence level. All standard errors are calculated using GMM and 6 Newey-West (1987) lags.

Table 7: Correlations Among Asymmetry Statistics

	H Statistic	Skewness	Co-skewness	β
H Statistic	1.0000	0.2433	0.1498	-0.2744
Skewness		1.0000	0.9510	-0.0539
Coskewness			1.0000	-0.0030
β				1.0000

We present the correlations among the estimates of asymmetry statistics calculated in Table (6). The correlations are calculated using the 43 estimates of H statistic, skewness, co-skewness and beta.

Table 8: H Statistics Across Size/Beta Portfolios

Size × Beta Portfolios (value-weighted)

	1 Low Beta	2	3	4	5 High Beta
1 Smallest	0.1848	0.1621	0.1674	0.1588	0.1446
2	0.1680	0.1183	0.1191	0.1182	0.0998
3	0.1491	0.0998	0.0944	0.0878	0.0831
4	0.1567	0.0878	0.0773	0.0577	0.0666
5 Largest	0.1271	0.0818	0.0566	0.0562	0.0562

This table presents the H statistics for equity portfolios assuming the null of a bivariate normal distribution. Frequency of the data is weekly. Weights proportional to the number of observations in each sample exceedance are used (equation (13)) to construct the H statistic. The null of a bivariate normal is rejected at the 0.1% confidence level for every portfolio at all frequencies by the H statistic (p-values are not reported).

Table 9: H Statistics Across Size/Leverage Portfolios

Size × Leverage Portfolios (value-weighted)

	1 Low Debt	2	3	4	5 High Debt
1 Smallest	0.1579	0.1511	0.1558	0.1559	0.1624
2	0.1152	0.1160	0.1099	0.1108	0.1171
3	0.1036	0.0838	0.0848	0.1064	0.0861
4	0.0828	0.0642	0.0720	0.0919	0.0949
5 Largest	0.0791	0.0573	0.0648	0.0660	0.0737

This table presents the H statistics for equity portfolios assuming the null of a bivariate normal distribution. Frequency of the data is weekly. Weights proportional to the number of observations in each sample exceedance are used (equation (13)) to construct the H statistic. The null of a bivariate normal is rejected at the 0.1% confidence level for every portfolio at all frequencies by the H statistic (p-values are not reported).

Table 10: Summary of Rejections from Table (11)

Rejections at 5% confidence level

	GARCH-M	Jump Model	RS Normal	RS-GARCH
Industry	6/13	13/13	3/13	6/13
Size	4/5	4/5	2/5	1/5
Book to Market	4/5	3/5	3/5	0/5
Momentum	3/5	5/5	4/5	1/5
Overall	17/28	25/28	12/28	8/28

Rejections at 1% confidence level

	GARCH-M	Jump Model	RS Normal	RS-GARCH
Industry	4/13	13/13	3/13	5/13
Size	4/5	3/5	2/5	1/5
Book to Market	1/5	3/5	2/5	0/5
Momentum	3/5	5/5	3/5	1/5
Overall	12/28	24/28	10/28	7/28

We present a summary of rejections from Table (11). We list the number of rejections, M, out of a possible N number of portfolios as M/N in the Table.

Table 11: H Statistics from Other Distributions

Industry Portfolios (value-weighted)

Portfolio	GARCH-M		Jump Model		Regime Switching		RS-GARCH	
	H Stat	SE	H Stat	SE	H Stat	SE	H Stat	SE
Misc	0.1141	0.2485	0.1168 **	0.0170	0.0760	0.0850	0.0439 *	0.0219
Petroleum	0.1067	0.2585	0.2120 **	0.0457	0.1272 **	0.0116	0.1682 **	0.0648
Finance	0.0604 *	0.0295	0.0674 **	0.0200	0.0339	0.0926	0.0695	0.1270
Durables	0.0729 *	0.0293	0.0738 **	0.0155	0.0564	0.0721	0.0764	0.1495
Basic Ind	0.0555	0.0399	0.0767 **	0.0147	0.0483 **	0.0067	0.0456	0.1495
Food/Tobacco	0.1301	0.3135	0.1269 **	0.0287	0.0731	0.0458	0.1062 **	0.0370
Construction	0.0867 **	0.0180	0.1107 **	0.0177	0.0667	0.1183	0.0762	0.0858
Capital Goods	0.0511 **	0.0143	0.0575 **	0.0173	0.0369 **	0.0091	0.0302	0.0799
Transportation	0.1138	0.0873	0.1023 **	0.0069	0.0819	0.0771	0.1871 **	0.0677
Utilities	0.1560	0.0931	0.1380 **	0.0052	0.1091	0.1028	0.1973 **	0.0738
Textile/Trade	0.0723	0.0529	0.1080 **	0.0284	0.0605	0.1224	0.1017	0.1437
Service	0.0794 **	0.0189	0.0645 **	0.0177	0.0350	0.0292	0.1284	0.0937
Leisure	0.0679 **	0.0207	0.0612 **	0.0166	0.0387	0.0537	0.0541 **	0.0110

Size Portfolios (value-weighted)

Portfolio	GARCH-M		Jump Model		Regime Switching		RS-GARCH	
	H Stat	SE	H Stat	SE	H Stat	SE	H Stat	SE
1 Smallest	0.1239 **	0.0261	0.1265 **	0.0248	0.0841 **	0.0285	0.0692	0.0535
2	0.0734 **	0.0236	0.0747 **	0.0175	0.0392	0.0305	0.0917	0.0938
3	0.0456 **	0.0176	0.0440 **	0.0165	0.0259 **	0.0077	0.0642	0.0799
4	0.0297 **	0.0095	0.0300	0.0601	0.0158	0.0508	0.0546	0.0735
5 Largest	0.0077	0.0043	0.0162 *	0.0070	0.0028	0.0015	0.0298 **	0.0022

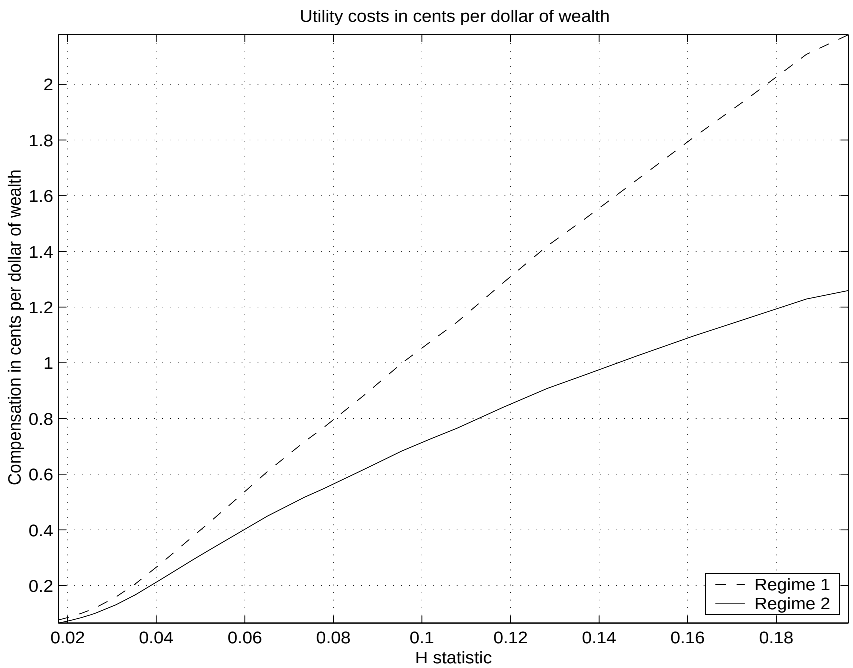
Book-to-Market Portfolios (value-weighted)

Portfolio	GARCH-M		Jump Model		Regime Switching		RS-GARCH	
	H Stat	SE	H Stat	SE	H Stat	SE	H Stat	SE
1 Growth	0.0228 *	0.0101	0.0425	0.0220	0.0093	0.0062	0.0597	0.0487
2	0.0391 **	0.0099	0.0363 **	0.0078	0.0241 **	0.0078	0.0355	0.0264
3	0.0592	0.1045	0.0829 **	0.0101	0.0479 **	0.0111	0.1097	0.0696
4	0.0803 *	0.0362	0.0932	0.0808	0.0611	0.1759	0.1219	0.1639
5 Value	0.1138 *	0.0517	0.1153 **	0.0258	0.0775 *	0.0302	0.1754	0.0921

Momentum Portfolios (equal-weighted)

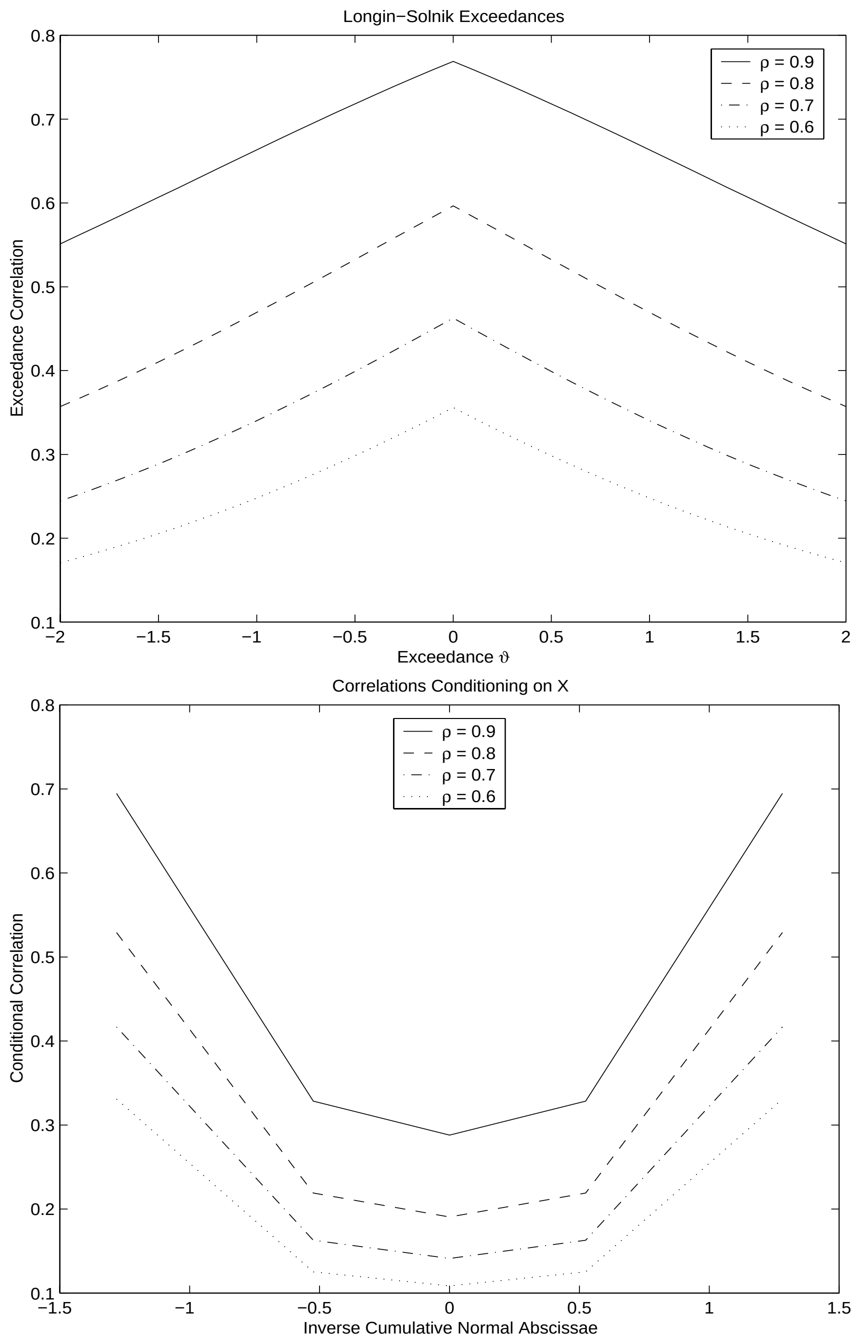
Portfolio	GARCH-M		Jump Model		Regime Switching		RS-GARCH	
	H Stat	SE	H Stat	SE	H Stat	SE	H Stat	SE
1 Past Losers	0.2387	0.1261	0.1574 **	0.0177	0.1009 **	0.0146	0.0895	0.1970
2	0.1054 **	0.0196	0.1167 **	0.0180	0.0804 *	0.0408	0.1003	0.0887
3	0.0937 **	0.0218	0.0959 **	0.0214	0.0606 **	0.0183	0.1377 **	0.0481
4	0.0806 **	0.0225	0.0738 **	0.0188	0.0459 **	0.0117	0.1153	0.0856
5 Past Winners	0.0994	0.0582	0.0968 **	0.0227	0.0593	0.0365	0.1492	0.1035

This table contains the H statistics for equity portfolios under the null of other distributions: a GARCH-M model, a Poisson Jump model, a regime-switching Normal model and a regime-switching GARCH model. The weights used are proportional to the number of observations used to calculate the sample exceedance correlations (equation (13)). Frequency of the data is weekly. A '$*$' indicates rejection of the model at the 5% confidence level, and '$**$' indicates rejection at the 1% confidence level. For each portfolio the statistic shaded in gray indicates the distribution, among normal, GARCH-M, Poisson Jump, regime-switching Normal and regime-switching GARCH, which produces the smallest H statistic.



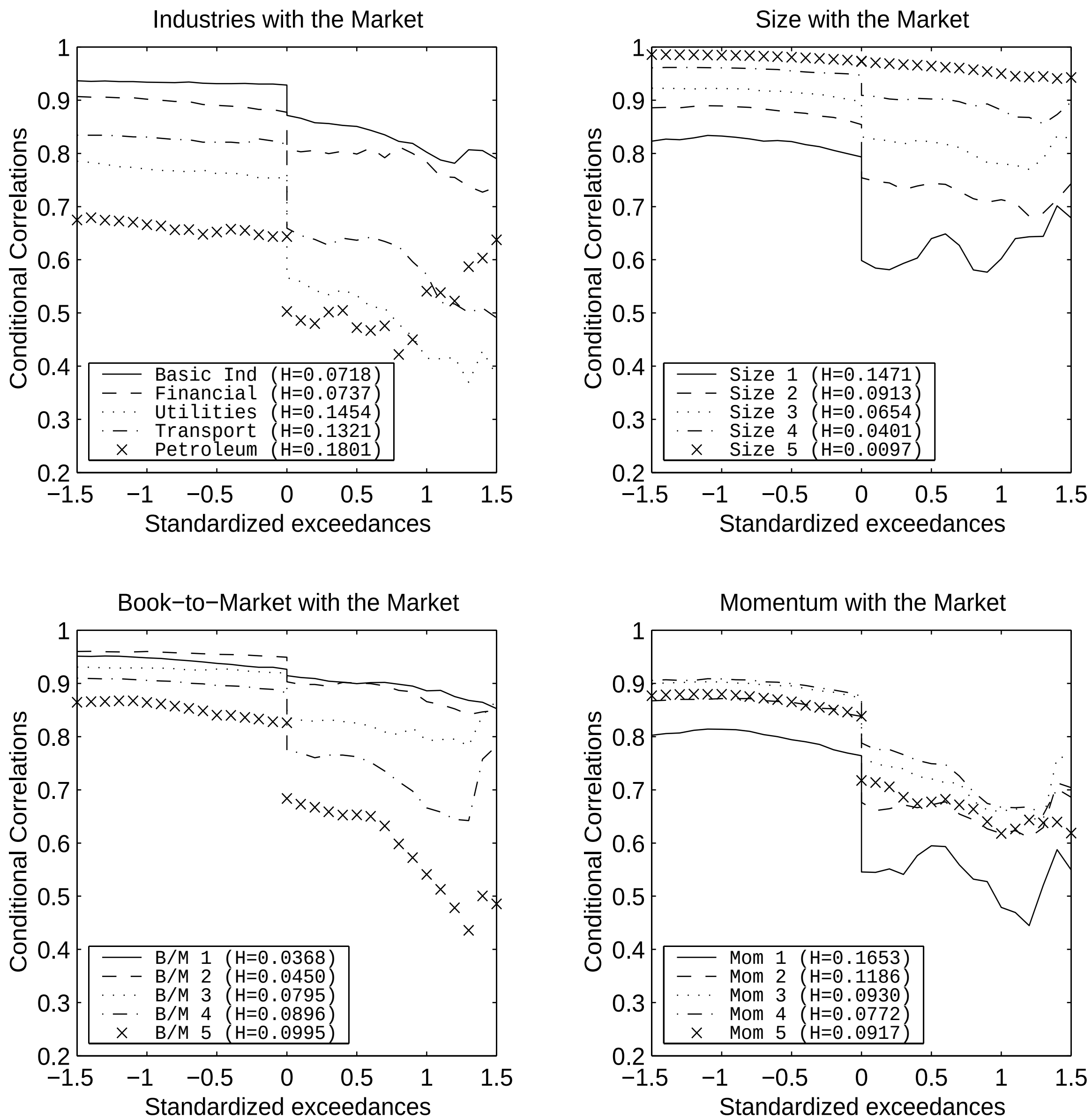
The plot shows the effects of ignoring increasing correlation on the downside in a hypothetical portfolio allocation problem. A CRRA investor with risk aversion $\gamma = 4$ allocates her portfolio among two risky assets and a riskless asset. She believes the assets are normally distributed and chooses asset holdings $\alpha^{\dagger}$. Under the normal distribution, the correlation conditional on downside move of both assets by more than 1 standard deviation from the mean is given by $\bar{\rho}$. The true distribution is given instead by a RS Model with identical unconditional means, variances and correlation. This distribution instead produces a true correlation of $\bar{\rho} + H$ conditional on a downside move of more than 1 standard deviation from the mean, where $H > 0$. The optimal portfolio weights implied by the RS Model (which the investor does not hold) are given by $\alpha^*_{s_t}$ for regime $s_t = 1, 2$. The regime-dependent correlations of the RS Models are chosen to produce various H statistics. The plot shows ex-ante utility losses in cents per dollar of wealth for the investor to hold the sub-optimal weights $\alpha^{\dagger}$ instead of $\alpha^*_{s_t}$ for regime s_t.

Figure 1: Economic Costs of Downside Asymmetric Correlations



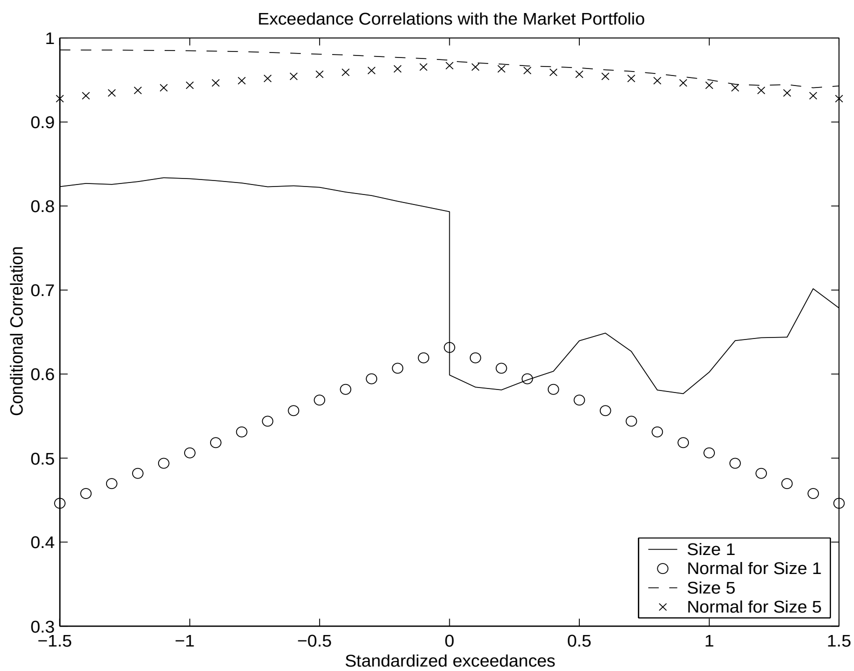
The top plot shows the exceedance correlations, corr$(\tilde{x}, \tilde{y} | \tilde{x} > \vartheta, \tilde{y} > \vartheta; \rho)$, for exceedance $\vartheta > 0$ of $\tilde{x}$ and $\tilde{y}$ drawn from a bivariate normal with zero mean, unit variances and unconditional correlation ρ. For $\vartheta < 0$ the exceedance correlation is corr$(\tilde{x}, \tilde{y} | \tilde{x} < \vartheta, \tilde{y} < \vartheta; \rho)$. The bottom plot gives conditional correlations corr$(\tilde{x}, \tilde{y} | h_1 < \tilde{x} < h_2; \rho)$ where h_1 and h_2 are chosen to correspond to absicssae from an inverse cumulative normal. We choose h_1 and h_2 to correspond to the absicssae intervals of probabilities [0 0.2 0.4 0.6 0.8 1]. That is, the first $(h_1,\, h_2) = (\Phi^{-1}(0),\, \Phi^{-1}(0.2))$ where $\Phi^{-1}(\cdot)$ is an inverse cumulative normal. We plot these at the inverse cumulative normal absicssae corresponding to the midpoints [0.1 0.3 0.5 0.7 0.9], that is the x-axis points are $\Phi^{-1}(0.1)$, $\Phi^{-1}(0.3)$, etc.

Figure 2: Conditional Correlations of a Bivariate Normal



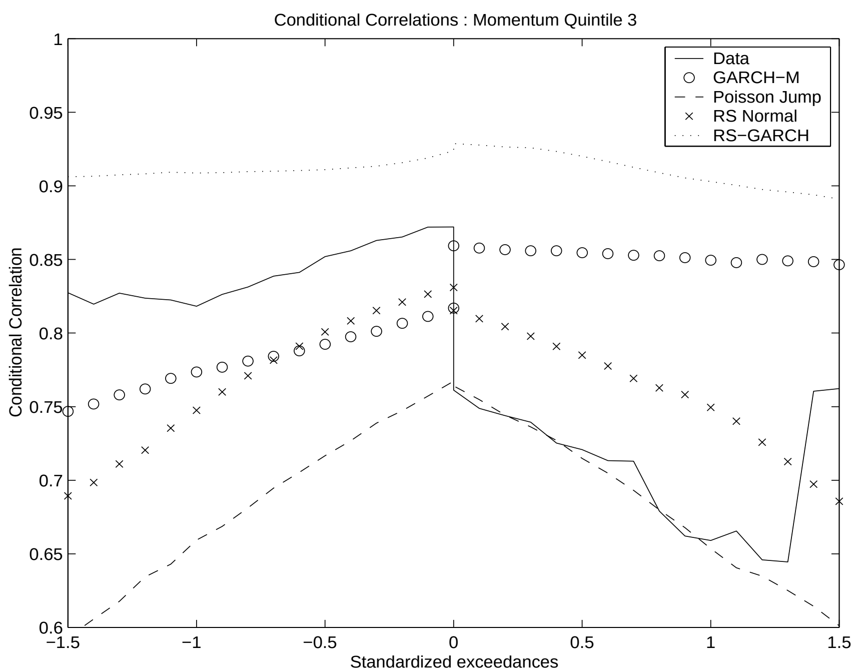
We plot exceedance correlations with the market portfolio for selected industry, size, book-to-market and momentum portfolios. These are the conditional correlations corr$(\tilde{x}, \tilde{y} | \tilde{x} > \vartheta, \tilde{y} > \vartheta; \rho)$ for exceedance $\vartheta > 0$ for normalized portfolio $\tilde{x}$ and the normalized market portfolio $\tilde{y}$. For $\vartheta < 0$ the exceedance correlation is defined as corr$(\tilde{x}, \tilde{y} | \tilde{x} > \vartheta, \tilde{y} > \vartheta; \rho)$. Exceedance correlations are calculated at the weekly frequency. The H statistic in the legend is the measure of correlation asymmetry developed in Section 4.

Figure 3: Exceedance Correlations of Industry, Size, B/M and Momentum Portfolios



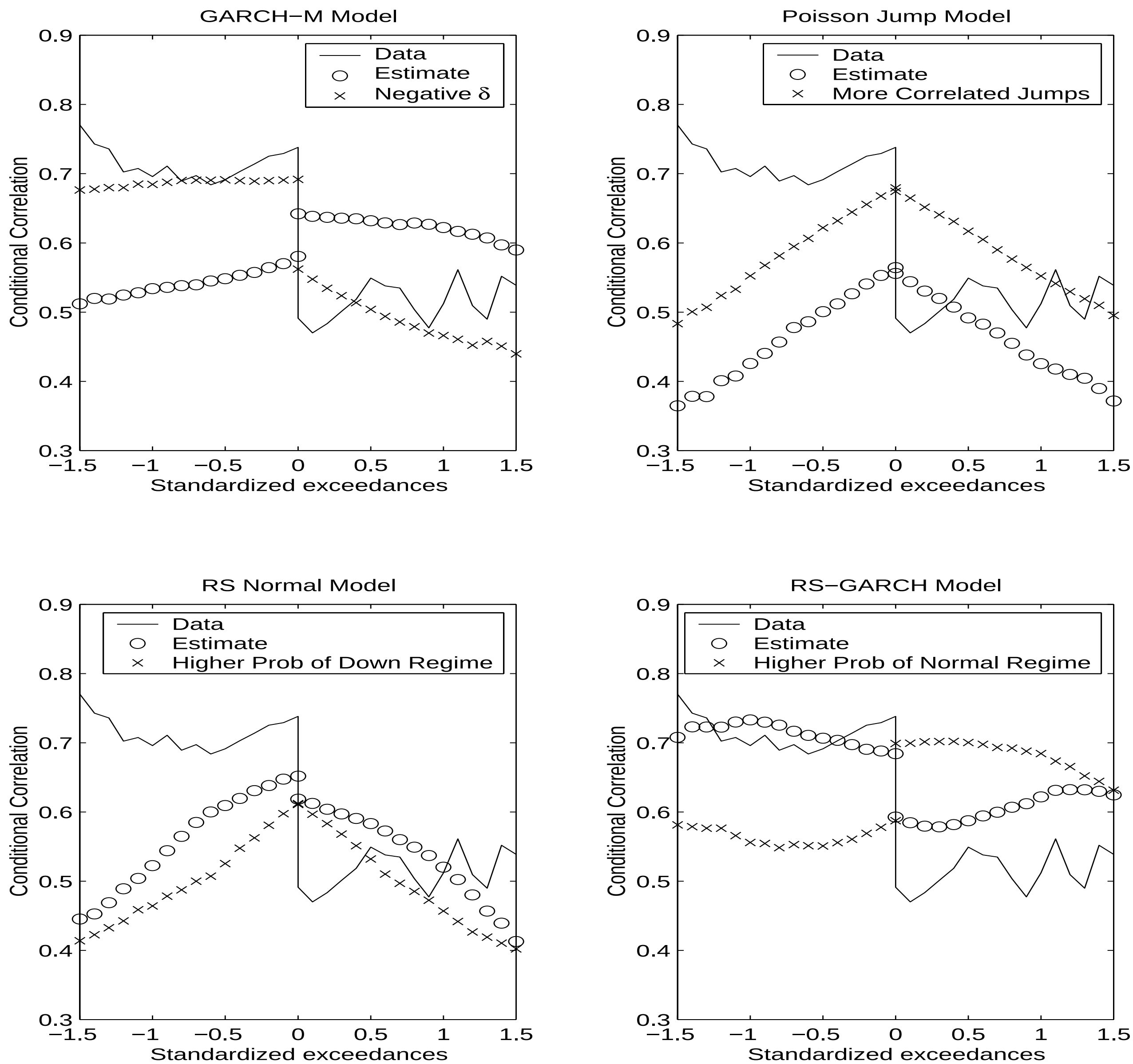
This figure shows the exceedance correlations with the market portfolio for the quintile 1 and quintile 5 size portfolios. Frequency is weekly. The theoretical exceedance correlations from a bivariate normal with the same unconditional correlation is also shown on the plot for each portfolio.

Figure 4: Exceedance Correlations: Empirical versus Bivariate Normal



This figure shows the empirical exceedance correlations for the third momentum quintile portfolio with the value-weighted market portfolio. Frequency is weekly. The theoretical exceedance correlations from an asymmetric GARCH-M, a Poisson Jump model, a RS Normal and a RS-GARCH model are presented on the same plot together with the empirical exceedance correlations found in the data.

Figure 5: Exceedance Correlations for the 3rd Momentum Portfolio: Empirical vs Models



We plot the exceedance correlations for the smallest size portfolio with the value-weighted market at the weekly frequency. We show the exceedance correlations from the data (solid lines) and those implied by various models. From top left clockwise, we have a GARCH-M model, a Jump model, a RS-GARCH model and a RS Normal model. Within each panel, we also plot an exceedance correlation of a comparative static, that is, altering one parameter of the models and re-calculating the exceedance correlations.

Figure 6: Exceedance Correlations for the Smallest Size Portfolio