Second Paper

Portfolio Selection: An Extreme Value Approach[☆]

Francis J. DiTraglia[a], Jeffrey R. Gerlach[b]

[a]*Faculty of Economics, University of Cambridge*
[b]*Senior Financial Economist, Federal Reserve Bank of Richmond*

Abstract

We show that lower tail dependence (χ), a measure of the probability that a portfolio will suffer large losses given that the market does, contains important information for risk-averse investors. We then estimate χ for a sample of DJIA stocks and show that it differs systematically from other risk measures including variance, semi-variance, skewness, kurtosis, beta, and coskewness. In out-of-sample tests, portfolios constructed to have low values of χ outperform the market index, the mean return of the stocks in our sample, and portfolios with high values of χ. Our results indicate that χ is conceptually important for risk-averse investors, differs substantially from other risk measures, and provides useful information for portfolio selection.

Keywords: Portfolio selection, Extreme value theory, Tail dependence
JEL: C58, G11

1. Introduction

In this paper, we apply portfolio selection techniques to a sample of large-cap stocks using "lower tail dependence" as the measure of risk. The concept of tail dependence comes from Extreme Value Theory (EVT), which provides a method to describe the tails of a distribution without specifying the underlying distribution. Specifically, χ describes the dependence between the extreme left tail of the distribution of a portfolio of stocks and the left tail of the distribution of the market return. Let X represent losses (returns multiplied by negative one) for some portfolio and Y losses for the market index. Then, the lower tail dependence between this portfolio and the market is given by:

$$\chi = \lim_{s \uparrow 1} \mathbb{P}\left\{X > F_X^{-1}(s) | Y > F_Y^{-1}(s)\right\} \tag{1}$$

where F_X^{-1} and F_Y^{-1} are the quantile functions of X and Y. Intuitively, χ is a kind of limiting conditional Value at Risk, capturing the limit of the probability that a portfolio suffers losses beyond its sth quantile, $F_X^{-1}(s)$, given that the market has suffered

[☆]The views expressed in this article are solely those of the authors. They do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

Email address: fjd26@cam.ac.uk (Francis J. DiTraglia)
URL: http://www.ditraglia.com (Francis J. DiTraglia)

Current Version: January 2, 2012

equivalently large losses.[1] When $\chi = 0$, X and Y are asymptotically independent; when $\chi = 1$, they are perfectly asymptotically dependent.[2] Unlike correlation, lower tail dependence is an asymmetric measure of dependence, examining only one region of the distribution. And unlike other asymmetric risk measures such as the downside β of Ang et al. (2006) or conditional coskewness of Harvey and Siddique (2000), lower tail dependence is not a moment. Consequently it is guaranteed to exist for all distributions, even those with extremely heavy tails.

Intuitively, lower tail dependence is a reasonable measure of risk if, *ceteris paribus*, investors prefer to hold portfolios that perform relatively well when market returns are extremely poor. Because χ measures dependence in the left tail of the distribution only, we would expect that during extreme bear markets low-χ portfolios will tend to have higher returns than high-χ portfolios. By construction χ provides no information about the right tail of the distribution so we might expect that during other market conditions, the low-χ portfolios would have returns similar to the market return. Alternatively, if χ is priced, low-χ portfolios would show lower returns in up markets to compensate for the downside protection they afford. Our out-of-sample results show that low-χ portfolios have relatively high cumulative returns over the ten-year test period, and relatively high returns in eight of those ten years, including both bull and bear markets.

The main contributions of this paper are as follows: First, we illustrate that χ contains valuable information for risk-averse investors that is not adequately captured by standard portfolio optimization. Second, as far as we know, this is the first paper to estimate the tail dependence of individual stocks and stock portfolios with the market return. Previous studies using EVT tail dependence, including Longin and Solnik (2001), Poon et al. (2004), and Chollete et al. (2009) use international stock market indexes. Third, we show χ differs systematically from other risk measures including variance, semi-variance, skewness, kurtosis, CAPM β, coskewness, and γ, an EVT measure of univariate tail thickness. Finally, in out-of-sample tests, portfolios that exhibit low tail dependence with the S&P 500 index have higher returns than the market index, the mean return of stocks in the sample, and portfolios with high tail dependence.

The paper is organized as follows. Section 2 reviews related work, showing why EVT tail dependence may provide a better measure of risk than commonly used risk measures, and explaining how our paper fits with the growing literature applying concepts from EVT to finance. Section 3 uses simulations to illustrate why χ may be useful for portfolio selection. Section 4 describes the EVT model used to estimate lower tail dependence. Section 5 describes our data and empirical results, and Section 6 concludes.

[1] Since χ depends only on the copula between X and Y, not their respective marginal distributions, it is equivalent to define it by $\lim_{s\to\infty} \mathbb{P}(Z_X > s | Z_Y > s)$ where Z_X and Z_Y are versions of X and Y that have been transformed to have the same marginal distribution. We use this alternative definition in Section 4.

[2] Here, the term "asymptotically" refers not to an increasing sample size, but an increasing threshold s above which the conditional probability defining χ is calculated.

2. Literature Review

In his path-breaking work on portfolio selection, Markowitz (1952, 1959) considers how investors can maximize expected return for a given risk level, or equivalently, minimize risk for a given expected return. In Markowitz's original formulation and many later refinements, including the CAPM of Sharpe (1964), Lintner (1965), Mossin (1966), and Treynor (1961), risk is measured by the variance of returns on an investor's overall portfolio.[3] In this setting, the mean return vector and covariance matrix of individual asset returns tells us everything we need to know to carry out portfolio selection.

There are effectively two ways to justify the mean-variance approach to portfolio selection. The first is to assume that investors maximize an expected utility function that is well approximated by a second-order Taylor expansion. When utility is exactly quadratic, investors only care about the mean and variance of final portfolios, regardless of the underlying distribution of returns. Yet even if investors' utility takes a more plausible functional form, it may still be the case that the portfolios obtained by explicit utility maximization are similar to those constructed using the much simpler mean-variance analysis (see, e.g. Kroll et al. (1984)). The quality of the approximation, however, depends not only on investors' true utility function but on the distribution of returns. If this distribution is sufficiently "tame," nearly always generating realizations in the region where the local approximation is good, the resulting portfolios will be similar. As we discuss in Section 3, however, mean-variance approximations can give misleading results in the presence of tail dependence.

The second way to justify the mean-variance approach is to assume that returns are normally distributed. Under normality, the mean vector and covariance matrix of returns fully describe their joint behavior, and all portfolios have marginal normal distributions. In this setting, because variance alone controls the thickness of the tails, *any* sensible measure of risk over final portfolios will be strictly increasing in the variance. Thus, investors will always find it sufficient to minimize the variance of their holdings when they seek to minimize risk. Normality also justifies the use of the CAPM β as a measure of an asset's systematic risk: under normality, correlation is the only dependence information relevant for diversification. There is considerable evidence, however, that asset returns are not normally distributed.[4] Real-world asset returns exhibit heavier tails than a normal distribution (Mandelbrot, 1963; Fama, 1965) as well as gain/loss asymmetry: large losses are more common than equivalently large gains (Cont, 2001; Premaratne and Bera, 2005).[5]

The non-normality of asset returns and problematic nature of quadratic approximations to investor preferences create scope for the use of additional information

[3] Although his name is practically synonymous with mean-variance analysis, Markowitz considers several risk measures, settling on variance merely as a pragmatic compromise (Markowitz, 1959, pp. 193-194). As Markowitz knew, the question of which risk measure to use depends both on the preferences of investors and the statistical properties of returns.

[4] For a detailed discussion of the non-normality of asset returns, see Rachev et al. (2005).

[5] Asymmetry is somewhat more controversial than heavy tails. Peiró (1999) and Kearney and Lynch (2007), for example, find limited evidence of asymmetry.

beyond means, variances and correlations in portfolio selection. Markowitz (1959) himself noted that while investors fear only losing money, not gaining it, the variance treats unusually large gains the same as unusually large losses: both contribute to an asset's measured risk. This incongruity led him to devote an entire chapter of his book to an alternative measure of risk, the semi-variance. Whereas the variance is defined as the expected squared deviation *from* the mean, the semi-variance is the expected squared deviation *below* the mean. Markowitz concludes that portfolios constructed using semi-variance are preferable because they concentrate on reducing losses, while those based on variance sacrifice too much expected return by minimizing extreme positive returns as well as extreme negative returns.[6] In a similar vein, Harvey and Siddique (2000) suggest measuring an asset's risk by its coskewness, the component of its (negative) skewness that is related to the skewness of the market portfolio, and find that this measure is associated with a significant risk premium in asset-pricing tests. Harvey et al. (2010) also emphasize the relevance of multivariate skewness for portfolio selection, proposing a Bayesian procedure that incorporates parameter uncertainty into the choice of portfolio. In the presence of asymmetric returns, the CAPM β is a potentially misleading measure of dependence. Responding to this concern, Ang et al. (2006) estimate "downside β," a measure of covariance with the market during periods of decline. They find that assets with high downside β command a risk premium that is not accounted for by ordinary β, coskewness, size, or momentum effects.

A potential problem with moment-based portfolio selection in the presence of heavy-tailed returns, especially methods based on higher moments such as skewness, is that the moments in question may be undefined or infinite. An early approach that does not fall subject to this critique is the "safety first" idea of Roy (1952) and Arzac and Bawa (1977), who suggest that investors minimize the probability of losing more than a pre-specified amount of money. Intuitively, safety first is an extremely appealing idea, but in practice it is nearly impossible to estimate the needed probabilities using classical techniques. It is precisely this problem of estimating tail probabilities that has led to the growing popularity of EVT in Finance. Danielsson and De Vries (1997), for example, use EVT to improve estimates of large losses in foreign exchange markets. More recently, Jansen et al. (2000), Jansen (2001), Susmel (2001), and Hyung and De Vries (2007) take the same approach to improve the safety-first model of Arzac and Bawa (1977), while McNeil and Frey (2000) incorporate stochastic volatility in an EVT framework to improve estimates of Value-at-Risk.

Besides its use in estimating tail probabilities, EVT has been applied to study the underlying distribution of asset returns. In an early study, Jansen and De Vries (1991) estimate γ, an EVT measure of univariate tail thickness described in Section 4, for a sample of 10 stocks and two stock market indexes from 1962–1986. Their estimates indicate that return distributions are consistent with the Student-t and ARCH classes, but not stable distributions and mixtures of normal distributions.

[6] As a measure of risk, semi-variance is only superior to variance if the distribution of returns is asymmetric. If the distribution of returns is symmetric, the semi-variance is simply half the variance.

In related work, Longin (2005) uses EVT to help determine the distribution of the underlying daily returns for S&P 500 index, rejecting the normal and stable Paretian distributions, but not the Student-t distribution and ARCH processes. Given its relationship to the tail thickness of returns, and thus the probability of large losses, it seems plausible that γ itself could be relevant for asset pricing. The only study we know of along these lines is Wu et al. (2008), who show there is a positive risk premium associated with the left-tail γ of stock market returns even after controlling for size, value, reversals, momentum, and liquidity.

Studies like these have motivated further interest in γ itself. Longin (1996) models extreme returns on an index of NYSE stocks and finds that γ is stable over time and under daily, weekly, and monthly temporal aggregation. Kearns and Pagan (1997) compare a number of methods of estimating γ, concluding that so-called "Hill estimator" works best. In a study of stock market indices for 5 developed and 15 developing countries, Jondeau and Rockinger (2003) find no significant differences between the left and right tail γ of the distribution of asset returns.

Most studies that apply EVT to financial data study lower tail thickness, via γ, for individual series, but recently several have considered the implications of EVT for dependence and diversification. Hyung and De Vries (2005) show that idiosyncratic risk decreases more quickly when assets with fat-tailed return distributions are added to a portfolio, while Ibragimov and Walden (2007) show that diversification may increase value at risk for heavy-tailed risky assets when potential losses are large. Their results illustrate how the presence of fat-tailed return distribution can dramatically affect portfolio selection. In a more applied setting, Longin and Solnik (2001) use a bivariate EVT model to study stock market indexes in the U.S., U.K., France, Germany and Japan and reject bivariate normality for the left-tail of the distribution and conclude the correlation across markets increases during bear markets. Poon et al. (2004) use multivariate extreme value theory to model extreme dependence in the same five markets. They find international stock markets tend to be asymptotically independent and conclude EVT models that assume asymptotic dependence will overstate systemic risk. Chollete et al. (2009) use returns from 14 national stock market indexes from 1990 – 2006 to compare Pearson correlation, Spearman correlation, and tail dependence for bivariate combination of the 14 countries. They report three main results: First, the three dependence measures provide different signals about risk; Second, dependence in general has increased over time; and Third, all regions studied show asymmetric dependence.

This paper uses a bivariate EVT model to study the extreme dependence of individual stocks and portfolios with the market index. It is not only the probability of large losses, described by the tail index, that concerns investors, but the timing of these losses. An asset that performs well when the market as a whole is doing poorly provides consumption insurance. For this reason, knowledge of the marginal distribution of asset returns is not enough: dependence on the market matters. This basic idea underlies the traditional CAPM model, which measures dependence by a scaled version of the correlation between a given asset's returns and those of the market: the well-known CAPM beta. In a world of normally distributed asset returns, or investors who measure risk by the variance of their wealth or consumption streams, correlation would be the appropriate measure of dependence. In the real world of heavy-tailed returns, and investors who treat losses and gains asymmetrically, however, it can be seriously misleading. In contrast, lower tail dependence directly concerns the probability that a given asset will suffer large losses, given that the market has. For investors concerned about large losses during bear markets, lower tail dependence describes the risks they face.

This approach is similar in spirit to safety first because it assumes that investors are not concerned with a general notion of variability, but primarily large negative returns. Unlike Arzac and Bawa (1977), though, we do not assume investors only care about preserving wealth once the probability of a large loss exceeds a critical level, or that investors maximize expected wealth when the probability of a large loss is below the critical level. Instead, we assume that, given a set of risky assets among which they can choose, investors will select portfolios that minimize lower tail dependence with the market, all other things equal. Our main hypothesis is that portfolios with low estimated tail dependence will outperform those with high tail dependence out-of-sample in years when the market portfolio has suffered large losses.

3. Motivating Tail Dependence: A Simple Example

Lower tail dependence is an intuitively appealing risk measure as it captures both investors' aversion to large losses and their sensitivity to the timing of those losses. An asset that performs poorly when the market as a whole has suffered large losses is particularly unattractive. In this section we further motivate lower tail dependence in a standard theoretical example: the Consumption Asset Pricing Model (C-CAPM) with power (CRRA) utility. Using log-normal margins calibrated to real-world asset returns, we induce varying degrees of lower tail dependence via the Clayton Copula, a relatively simple model of asymmetric extreme dependence. Even in this textbook setting, tail dependence affects risk premia in a way that standard measures fail to capture, showing its relevance for portfolio selection.

The risk premium for asset i under the C-CAPM is given by

$$\mathbb{E}_t\left[R^i_{t+1}\right] - R^f = \frac{-Cov_t\left[u'(c_{t+1}), R^i_{t+1}\right]}{\mathbb{E}_t\left[u'(c_{t+1})\right]}. \tag{2}$$

where R^i_{t+1} is the gross return on asset i in period $t+1$, R^f is the risk-free rate, c_{t+1} is consumption in period $t+1$, u is an increasing, concave function, and t subscripts indicate conditioning on time t information. The intuition for this expression is straightforward: because marginal utility is decreasing in consumption, an extra unit of consumption is most valuable when consumption is low. Accordingly, if an investor is to hold an asset whose returns are low when marginal utility is high, she will require a positive risk premium.

Assume that investors have standard power (CRRA) utility, so that

$$u(c) = \frac{c^{1-\rho} - 1}{1-\rho} \tag{3}$$

where ρ is the coefficient of relative risk aversion. Then, $u'(c) = c^{-\rho}$ and the C-CAPM equation becomes

$$\mathbb{E}_t\left[R^i_{t+1}\right] - R^f = \frac{-Cov_t\left[c^{-\rho}_{t+1}, R^i_{t+1}\right]}{\mathbb{E}_t\left[c^{-\rho}_{t+1}\right]}. \tag{4}$$

We can now see why lower tail dependence could influence risk-premia: $c^{-\rho}_{t+1}$ approaches infinity as c_{t+1} approaches zero. Thus, states of the world in which asset i performs poorly *and* aggregate consumption is low receive extra weight in determining the risk premium.

In spite of this observation, it is frequently suggested that, because CRRA utility is sufficiently smooth, a mean-variance approximation to Equation 4 is sufficient for all practical purposes. Kroll et al. (1984), for example, argue that even an expected utility maximizer whose utility function is not quadratic stands to gain very little by using anything more complicated than mean-variance analysis. Similarly, Herings and Kubler (2007) present a wide range of simulations showing that the CAPM makes relatively small pricing errors even if investors have CRRA utility. In a recent study of downside risk, broadly similar to ours in its aims, Ang et al. (2006) assert that any downside risk premium that does arise under CRRA utility will be "economically negligible." It is indeed true that, for a great many potential distributions of asset returns, mean-variance ideas, of which the CAPM is an example, perform remarkably well. However in the presence of lower tail dependence, this may not be the case.

The key to ensuring that Equation 4 is approximately mean-variance is to rule out situations in which asset i suffers very large losses at the same time as aggregate consumption, or the return on the market portfolio, is extremely low. This is because, as mentioned above, it is precisely these situations which receive the most weight under CRRA utility. If recent events have taught us anything, however, it is that such states of the world are far from impossible. Indeed, there are sound theoretical reasons to suppose that asset returns should show far more dependence on the extreme downside than the upside. Factors that cause extremely poor returns, for example credit crises, wars, or natural disasters, tend to affect all firms at once, while those that cause extremely high returns are more frequently firm-specific. It is this phenomenon that our example is designed to capture.

We compute cumulative annual returns over all consecutive one-year (250 trading day) periods from 30 October, 1986 through 31 December, 2008 for the S&P 500 and Microsoft Corp.[7] We then calculate the mean and standard deviation of each collection of rolling-block returns.[8] For the S&P 500, this yields an average annual return of approximately 8% with a standard deviation of 16%; for Microsoft an average annual return of 31% with a standard deviation of approximately 44%.

We specify the marginal distributions of the market return and our simulated

[7]This is the full sample period used in the empirical tests in Section 5.

[8]Our rolling-block calculations for average returns are approximately equal to those found by taking the arithmetic mean of daily returns and multiplying by 250. However, calculating the standard deviation of average returns in the same fashion would lead to a massive overestimate of the volatility faced by an investor with a one-year horizon. Our method yields more realistic values.

stock returns to be lognormal and choose parameter values to match the empirical annual gross returns and standard deviations. For the market return, the log-normal parameters are $\mu = 0.07$ and $\sigma = 0.16$. For the stock return they are $\mu = 0.22$ and $\sigma = 0.33$.[9] We then specify the dependence structure between the two margins according to a Clayton Copula with parameter θ:

$$C_\theta(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}, \text{ for } \theta > 0 \tag{5}$$

This gives a lower tail dependence parameter of $\chi = 2^{-1/\theta}$ and an upper tail dependence parameter of zero (Embrechts et al., 2002), capturing our intuition regarding the asymmetry of real-world dependence.

Using the distribution specified above, we calculate the risk premia implied by the C-CAPM with CRRA utility (Equation 4) and compare them to those given by a second-order Taylor approximation around the mean market return.[10] We repeat this exercise over a grid of values for ρ, the risk-aversion parameter, and χ, the lower tail dependence parameter. Results appear in Table 1 and Figure 1. All calculations are carried out via Monte Carlo integration with $N = 1,000,000$. Our procedure for drawing from a Clayton Copula appears in Appendix B.

We first note from Figure 1 that the stock's risk premium under the C-CAPM is strictly increasing in lower tail dependence for all values of ρ. Since this premium captures investors' risk preferences, it follows that lower tail dependence is relevant for portfolio selection. However, it could still be the case that standard measures, like the CAPM β, capture most of the same information as χ. We see from Table 1, however, that this is not the case. Indeed, for $\rho = 1$, the CAPM/mean-variance approximation to CRRA utility is nearly perfect.[11] However, for realistic values of ρ and χ, the CAPM/mean-variance approximation substantially underestimates the true annual risk premium. Indeed, for $\rho = 10$ and $\chi = 0.45$, our estimated lower tail dependence value for the *real* Microsoft returns from the empirical section of this paper, the true risk premium is ten points higher than the approximation.

The point of this exercise is not to suggest that expected utility maximization with constant relative risk aversion is an accurate model of investor behavior, nor that the market return is necessarily a good proxy for aggregate consumption. Instead, the point is simply that exotic assumptions are not necessary to motivate tail dependence. A textbook asset pricing model combined with standard marginal distributions and a widely-used utility specification is enough to make χ relevant for portfolio selection in the presence of downside risk.

[9]These parameters give the mean and standard deviation of gross returns on a logarithmic scale.

[10]That is, we use a mean-variance approximation to CRRA utility. This is equivalent to a CAPM model in which the risk-free rate is calculated from a mean-variance approximation to the underlying utility function.

[11]Since log utility is sufficient to turn the C-CAPM into the CAPM, the small discrepancy is a computational artifact.

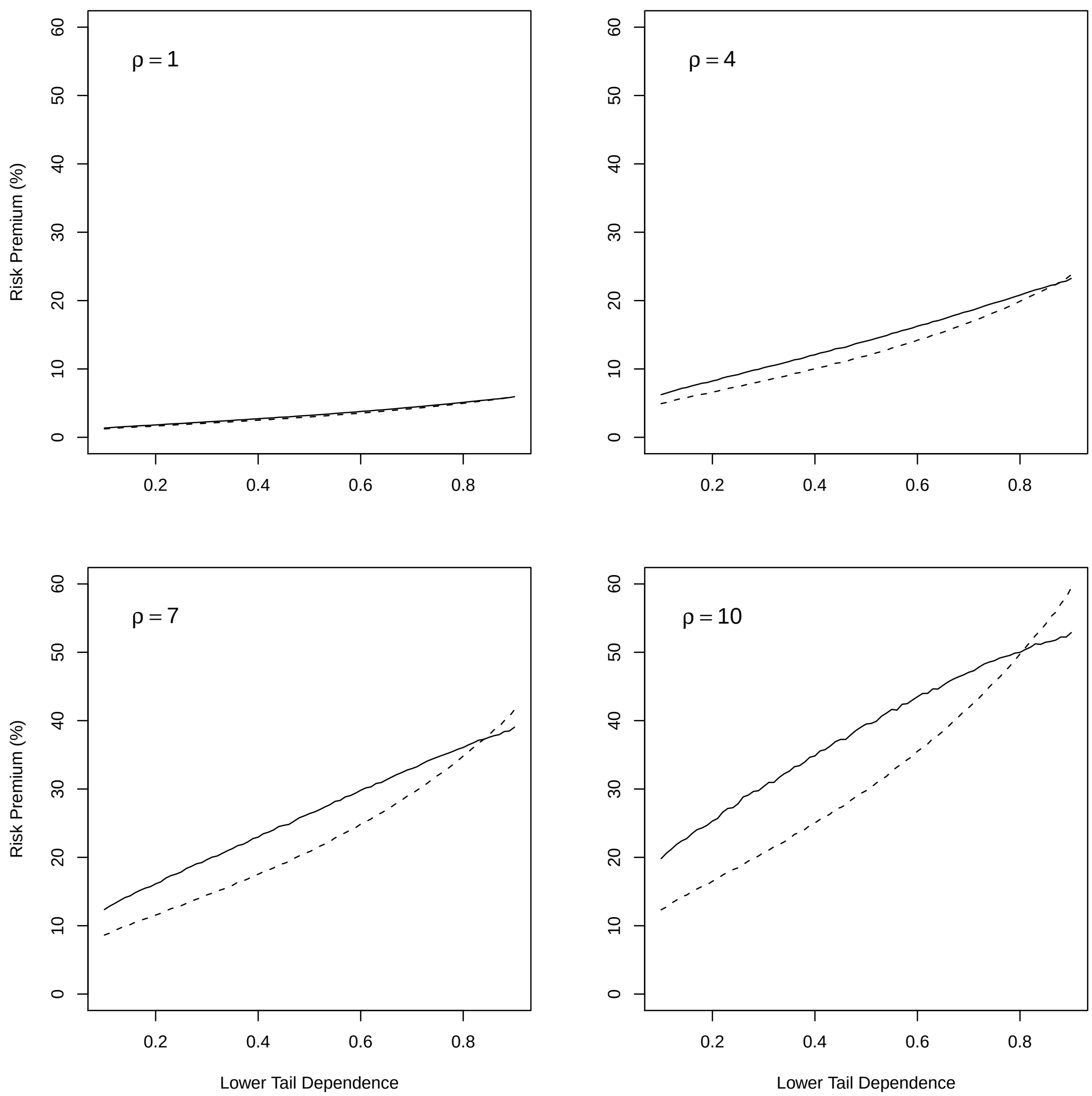


Figure 1: A mean-variance approximation to CRRA makes substantial pricing errors under left tail dependence. Solid lines give the C-CAPM risk premium, dashed lines the quadratic approximation.

Table 1: Exact (CRRA) minus approximate (mean-variance) risk premia in percentage points.

χ / ρ	1	2	3	4	5	6	7	8	9	10
0.15	0.1	0.4	0.9	1.5	2.2	3.2	4.2	5.4	6.8	8.3
0.20	0.2	0.5	1.0	1.6	2.5	3.4	4.6	5.9	7.3	8.8
0.25	0.2	0.5	1.1	1.8	2.7	3.7	4.9	6.3	7.8	9.4
0.30	0.2	0.6	1.1	1.9	2.8	3.9	5.2	6.5	8.0	9.6
0.35	0.2	0.6	1.2	2.0	3.0	4.1	5.4	6.8	8.3	9.9
0.40	0.2	0.6	1.3	2.1	3.0	4.2	5.4	6.8	8.3	9.8
0.45	0.2	0.7	1.3	2.1	3.1	4.3	5.6	7.0	8.4	10.0
0.50	0.2	0.7	1.3	2.2	3.2	4.3	5.6	6.9	8.3	9.7
0.55	0.2	0.7	1.3	2.1	3.1	4.2	5.3	6.6	7.8	9.0
0.60	0.2	0.7	1.3	2.1	2.9	3.9	4.9	6.0	7.0	8.0
0.65	0.2	0.6	1.2	1.9	2.7	3.6	4.4	5.3	6.0	6.7
0.70	0.2	0.6	1.1	1.7	2.3	3.0	3.7	4.2	4.8	5.2
0.75	0.2	0.5	0.9	1.4	1.9	2.3	2.7	3.0	3.1	3.1
0.80	0.1	0.4	0.6	0.9	1.1	1.3	1.3	1.1	0.8	0.2
0.85	0.1	0.2	0.3	0.3	0.3	0.1	-0.3	-0.9	-1.7	-2.6
0.90	0.0	-0.1	-0.2	-0.5	-1.0	-1.7	-2.6	-3.7	-5.0	-6.6

4. Estimating Tail Dependence

There are many approaches for estimating tail dependence.[12] At one extreme there are fully parametric methods. Tail dependence, upper or lower, is a property of the copula joining a pair of random variables. With a parametric model of the marginal distributions and dependence structure, it is possible to estimate χ by maximum likelihood. If the distributional specification is correct, this method is consistent, and asymptotically efficient. Unfortunately, there is no consensus on the underlying distribution of returns. If we specify the wrong distribution, our answers could be severely biased. For a particularly stark example, consider the widespread use of the Gaussian copula to model risk in the run-up to the Financial Crisis. As was known well before 2008 (Embrechts et al., 2002), this model rules out the possibility of tail dependence, lower or upper.

At the other extreme, we might prefer a fully non-parametric model, using the empirical CDF to transform X and Y to common marginals S and T, and taking a sample analogue of $\lim_{s\to\infty} \mathbb{P}(S > s|T > s)$ based on a sufficiently high threshold s^*. The problem with this approach is that tail dependence, by definition, concerns a region of the joint distribution from which we observe very few datapoints. Although the empirical CDF is a uniformly consistent estimator of the true marginal distribution, for any fixed sample size it must necessarily underestimate the thickness of the tails. To put it another way, any finite sample must have a lowest and highest value,

[12] For a detailed discussion of the many possibilities, their advantages and disadvantages, see Frahm et al. (2005).

but this does not imply that the underlying distribution is bounded. Even ignoring the problem of the marginals, any threshold s^* that seems high enough to give a plausible estimate of χ must exclude the vast majority of observations.

This is a classic bias-variance tradeoff. Fully parametric methods make much more efficient use of the information contained in the sample but are sensitive to the assumed distribution. Nonparametric methods, on the other hand, make few assumptions but yield wildly variable estimates as they try to study a region of the empirical distribution that contains practically no observations. We take a different approach. By working with monthly minimum returns, we exploit an important limit result from Extreme Value Theory (EVT) that provides a convenient and flexible estimator for the lower tail dependence between a given portfolio and the market. The approach used here was first described by Coles et al. (1999) in the context of meteorological data, and is related in greater length by Coles (2001). Beirlant et al. (2004) and Reiss and Thomas (1997) provide comprehensive treatments of EVT that discuss statistical issues.

The idea behind EVT is to describe the probabilistic behavior of unusually large or small observations. This can be achieved in two different ways: by studying block maxima or threshold exceedances.[13] The block approach analyzes the distribution of the maximum of a sequence of random variables $X_1, \ldots, X_n$ as n approaches infinity, while the threshold approach concerns the distribution of $X|X > s$ as the threshold s approaches infinity. These alternatives are in fact two ways of looking at the same question: how does the tail of the distribution of X behave? In practice, however, there may be reasons to favor one approach over the other. For our purposes, block methods are more appropriate as they should be more robust to volatility clustering, a well-documented feature of asset returns (Cont, 2001).[14]

The main advantage of EVT is that it is not necessary to know the underlying distribution of asset returns to describe the distribution of univariate extremes. In much the same way as the Central Limit Theorem establishes the asymptotic normality of sample averages, its EVT analogue, the Extremal Types Theorem, fully characterizes the limiting distribution of univariate sample maxima. Thus, EVT offers a way to study the tails of a distribution based on the observed extremes. Combining this with an assumption about the behavior of joint extremes yields an estimator of lower tail dependence. First, the key result:

Theorem 4.1 (Extremal Types Theorem)**.** *Let $X_1, X_2, \ldots, X_n$ be weakly dependent, identically distributed, scalar random variables and define $M_n \equiv \max\{X_1, \ldots, X_n\}$. If there exist sequences of constants $\{a_n > 0\}$, $\{b_n\}$ such that $(M_n - b_n)/a_n$ has a proper limit distribution, that distribution is of the form*

$$G(z) = \exp\left\{-\left[1 + \gamma\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\gamma}\right\} \tag{6}$$

[13] It is traditional to phrase results in terms of maxima. To study minima we simply need to multiply by negative one: converting negative returns to positive losses.

[14] Coles (2001) provides an intuitive discussion of the differing effects of temporal dependence on block versus threshold methods.

for $1 + \gamma(z - \mu)/\sigma > 0$*, with parameters satisfying* $-\infty < \mu < \infty$*,* $\sigma > 0$ *and* $-\infty < \gamma < \infty$*.*

Equation 6 gives the Generalized Extreme Value Distribution, a three-parameter family with centrality parameter μ, scale parameter σ, and shape parameter γ, which describes the tail thickness of the underlying distribution. Positive values of γ yield the Fréchet Distribution, indicating a very thick upper tail (polynomial decay). As γ approaches zero, the GEV reduces to the Gumbel Distribution and the upper tail exhibits exponential decay. Negative values of γ yield the reversed Weibull distribution and indicate a bounded upper tail. When γ is non-negative and the underlying distribution is symmetric, the shape parameter γ shares an important relationship with the tail index, κ, of the underlying distribution, namely $\kappa = 1/\gamma$. For example, if $\gamma = 0.5$, the underlying distribution has two finite moments.

Recall that the definition of lower tail dependence assumes common marginal distributions. Although strictly speaking the Extremal Types Theorem is a limit result, it immediately suggests a practical estimation procedure for the margins. Let $x_1, x_2, \ldots, x_N$ denote daily returns for some portfolio and $y_1, y_2, \ldots, y_N$ denote daily returns for the market index. Our strategy is to split the sample into non-overlapping blocks of sufficient length that we can model the maximum *loss* in each block as GEV and use this to transform the *extremes* to common marginals without specifying the underlying distribution of returns.

Consider blocks of length k where M denotes the total number of blocks of length k, that is $M = \lfloor N/k \rfloor$. If extra observations remain, simply discard those at the beginning of the series. Now let $\underline{x}_i$ denote the minimum return on portfolio x in block i and $\underline{y}_i$ denote the minimum return on the market index in the same block, where $i = 1, \ldots, M$. Note that although $\underline{x}_i$ and $\underline{y}_i$ come from the same block, they may not correspond to the same day. To convert negative block minimum returns to positive block maximum losses, we multiply through by negative one, yielding $\{(-\underline{x}_i, -\underline{y}_i)\}_{i=1}^{M}$. Now, if k is sufficiently large, the marginal distribution of block maximum losses should be *approximately* GEV by the Extremal Types Theorem. Thus, we can model the univariate series $\{-\underline{x}_i\}_{i=1}^{M}$ and $\{-\underline{y}_i\}_{i=1}^{M}$ according to Equation 6 and estimate parameters $(\hat{\mu}_x, \hat{\sigma}_x, \hat{\gamma}_x)$ and $(\hat{\mu}_y, \hat{\sigma}_y, \hat{\gamma}_y)$ via maximum likelihood. Details appear in Appendix C.1.[15]

Using the maximum likelihood estimates from the univariate models, we can transform the block maximum losses for asset x and the market index y to common marginals. Although the particular distribution to which we transform does does not affect estimated tail dependence, a convenient choice is the unit Fréchet, given by

$$F(z) = \exp(-1/z) \tag{7}$$

[15] It may seem as though we have ignored the normalizing constants from Extremal Types Theorem when asserting that the un-normalized block maximum losses may be modeled according to Equation 6. This does not present a problem: when we assume that the limit result holds in a finite sample, the normalizing constants can simply be absorbed into the parameters. That is, if $(Z - b_n)/a_n$ follows a generalized extreme value distribution, so does Z, only with different parameter values.

for $z > 0$. Because this is a GEV distribution with $\mu = \sigma = \gamma = 1$ (see Equation 6), the transformation is straightforward (see Appendix C.2). Define the block maximum losses after transformation to unit Fréchet by $\{-\tilde{\underline{x}}_i\}_{i=1}^{M}$ and $\{-\tilde{\underline{y}}_i\}_{i=1}^{M}$.

We can now use the transformed block maximum losses to estimate the tail dependence between asset x and the market y. Just as the shape parameter γ of the limiting distribution of univariate maxima fully characterizes the tail of the underlying distribution of returns, the tail dependence parameter of the limiting distribution of transformed *joint* maxima $\{(-\tilde{\underline{x}}_i, -\tilde{\underline{y}}_i)\}_{i=1}^{M}$ corresponds to that between the portfolio returns x and the market index y. Unfortunately, while there is a single limiting distribution for normalized block maxima, the same is not true of joint maxima. Instead, there is a *family* of limiting distributions characterized by a fairly complicated integral equation. To proceed any further, we need a parametric assumption. For convenience, we work with the simplest possible limit, the bivariate logistic distribution. Two random variables X and Y with unit Frećhet margins are said to follow a bivariate logistic distribution if their joint distribution is given by

$$G(x, y) = \exp\left\{-\left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha}\right\} \tag{8}$$

for $x, y > 0$ and $\alpha \in [0, 1]$. Because the margins are already specified, this is a *one parameter distribution* and the single parameter α controls the strength of dependence. When $\alpha = 1$ the two margins are asymptotically independent; when $\alpha = 0$ they are perfectly asymptotically dependent. The lower tail dependence parameter χ is related to α according to

$$\chi = 2 - 2^{\alpha} \tag{9}$$

When $\alpha = 0$ we have $\chi = 1$ and when $\alpha = 1$, $\chi = 0$. Thus, the final step of our estimation procedure is to fit the bivariate logistic model given by Equation 8 to the transformed block maximum losses $\{(-\tilde{\underline{x}}_i, -\tilde{\underline{y}}_i)\}_{i=1}^{M}$ by maximum likelihood. The resulting estimate $\hat{\alpha}$ can be converted to $\hat{\chi} = 2 - 2^{\hat{\alpha}}$ using the invariance property of maximum likelihood estimators. Details appear in Appendix C.3.

In the analysis that follows, we estimate the lower tail dependence between individual portfolios and the market index using the three-step maximum likelihood procedure described above with blocks of 22 trading days.[16] Although one-step estimation is theoretically more efficient, it is much more difficult to find good starting values for seven parameters at once than for blocks of three, three and one parameter. As we ultimately compute χ for tens of thousands of portfolios, starting values were a major concern. Under relatively weak dependence assumptions our procedure should produce good estimates of lower tail dependence. However, we do not report standard errors, as correcting them for temporal dependence would require much stronger assumptions. Appendix D reports robustness tests for our estimator.

To summarize, our procedure for estimating lower tail dependence is as follows:

[16]Robustness tests described in Appendix D show that our results are not sensitive to the choice of block length.

1. Partition the daily returns on the market index and the individual stock or portfolio into 22-day blocks and select the minimum return in each block for each series.
2. Estimate the GEV parameters μ, σ, and γ as described in Appendix C.1 separately for the block minima of the market index and those of the individual stock or portfolio.
3. Use the GEV parameter estimates to transform the block minima for both series to unit Frećhet, as described in Appendix C.2.
4. Estimate α as described in Appendix C.3 and convert it to χ using Equation 9.

5. Empirical Results

We estimate the lower tail dependence between portfolios formed from constituents of the Dow Jones Industrial Average (DJIA) and the market portfolio, proxied by the S&P 500 index, using the method described in Section 4. We employ blocks of 22 trading days throughout and our data consist of daily returns from the CRSP database for constituents of the DJIA and the S&P 500 index from October 30, 1986 – December 31, 2008.[17] When calculating the CAPM β for purposes of comparison, we use daily returns on a three-month U.S. Government T-bill for the risk-free rate.

We begin by estimating χ, the extreme value shape parameter γ, and a variety of other risk measures for the individual assets listed in Table 2 over the full October 30, 1986– December 31st, 2008 sample period. This gives 5,592 trading days, and 254 blocks of 22 trading days each after discarding the first four observations. Table 3 reports descriptive statistics and risk measures for all of the firms in the sample and Figure 2 shows how the risk measures are correlated. The numbers in each square in the upper left of the figure show the correlation coefficient for the two risk measures that intersect in that square. For example, the value of .25 for β and χ is based on estimating β and χ for each stock using the entire sample period, and then calculating the correlation between these estimates. Panels with a shaded background indicate a relationship that is significant at the 10% level. The size of each number is directly proportional to the absolute magnitude of the correlation coefficient. Each square in the lower right of the figure graphs each combination of the two risk measures for each stock in the sample after centering and scaling. Thus, the slope of the line gives the correlation between one risk measure and another.

We see that χ is not significantly correlated with γ, the Sharpe ratio, β, semi-variance, or variance. It is positively correlated with kurtosis and negatively correlated with coskewness with correlation coefficients of +.35 and -.36, respectively. Figure 2 shows that χ is not simply a proxy for one of the other risk measures, but

[17]October 30, 1986 is the first trading day for which we have data for all of the firms in our sample. CitiGroup, formed as a merger of Citicorp and Travelers Group 1998, has permno 70519 in the CRSP database. Prior to the merger, that permno is associated with Travelers Group, Travelers, Primerica, Commercial Credit Group, and Commercial Credit. The first available observation in the CRSP database for Commercial Credit is October 30, 1986.

Table 2: Key to DJIA constituents.

Table 2 lists the firms included in this study, their permnos from the CRSP database, and ticker symbols.

	Permno	Ticker
3M	22592	MMM
AIG	66800	AIG
Alcoa	24643	AA
Altria	13901	MO
American Express	59176	AXP
AT&T Corp	10401	T
AT&T Inc	66093	T
Boeing	19561	BA
Caterpillar	18542	CAT
Chevron	14541	CVX
Citigroup	70519	C
Coca-Cola	11308	KO
Disney	26403	DIS
Dupont	11703	DD
Exxon Mobil	11850	XOM
GE	12060	GE
GM	12079	GM
Goodyear	16432	GT
Home Depot	66181	HD
Honeywell Intl.	10145	HON
HP	27828	HPQ
IBM	12490	IBM
Intel	59328	INTC
International Paper	21573	IP
Johnson & Johnson	22111	JNJ
JP Morgan Chase	47896	JPM
Kodak	11754	EK
McDonalds	43449	MCD
Merck	22752	MRK
Microsoft	10107	MSFT
Pfizer	21936	PFE
Procter & Gamble	18163	PG
Sears	14322	S
Union Carbide	15659	UK
United Tech.	17830	UTX
Verizon	65875	VZ
Wal Mart	55976	WMT

Table 3: Summary Statistics and Risk Measures

Table 3 presents summary statistics and risk measures for those assets with full sample of returns (30 October, 1986 – 31 December 2008). The first four columns give the mean, standard deviation, skewness, and kurtosis of daily returns. The following four present Markowitz's semi-variance, the Sharpe Ratio, CAPM β, and coskewness. The rightmost columns display the extreme value tail index γ and parameter χ. Both extreme value statistics are calculated by maximum likelihood by the method of block minima, using 254 blocks of 22 days each. All other statistics are based on a sample size of 5592 daily returns, except for Altria which has one observation fewer (see Data Notes in Appendix A). The dependence statistics, *Coskew*, β, and χ use daily returns for the S&P 500 as the market portfolio. Both *Coskew* and β use the one month rate on a U.S. Government T-bill as the risk-free rate. Values appear rounded to two significant digits.

	μ	σ	μ_3	μ_4	S	*Sharpe*	β	*Coskew*	γ	χ
EK	7.6e-05	0.021	-0.450	25.0	2.3e-04	0.057	0.92	-2.2e-06	0.41	0.34
INTC	1.0e-03	0.028	-0.096	8.9	3.9e-04	0.590	1.40	-1.5e-06	0.31	0.34
GM	1.0e-04	0.026	0.450	27.0	3.1e-04	0.064	1.20	-1.3e-06	0.34	0.37
HPQ	7.6e-04	0.025	0.120	9.3	3.1e-04	0.470	1.20	-1.6e-06	0.23	0.37
MO	7.7e-04	0.019	-0.340	15.0	1.8e-04	0.660	0.67	-1.3e-06	0.29	0.38
PFE	6.2e-04	0.019	-0.160	7.4	1.7e-04	0.530	0.88	-1.4e-06	0.21	0.39
GT	1.8e-04	0.026	-0.150	12.0	3.4e-04	0.110	1.20	-2.7e-06	0.33	0.40
MRK	5.7e-04	0.018	-0.690	16.0	1.7e-04	0.500	0.84	-1.1e-06	0.27	0.40
CAT	6.9e-04	0.020	-0.160	9.8	2.0e-04	0.540	0.98	-1.7e-06	0.27	0.41
MCD	6.4e-04	0.017	-0.085	8.3	1.5e-04	0.590	0.75	-1.4e-06	0.24	0.41
IBM	4.4e-04	0.019	-0.057	13.0	1.7e-04	0.370	0.96	-1.7e-06	0.29	0.42
JNJ	6.7e-04	0.016	-0.240	12.0	1.2e-04	0.680	0.75	-1.4e-06	0.18	0.42
WMT	7.4e-04	0.019	0.170	6.4	1.7e-04	0.620	0.96	-6.2e-07	0.17	0.42
HD	1.0e-03	0.022	-0.380	13.0	2.5e-04	0.710	1.20	-1.6e-06	0.24	0.43
MSFT	1.2e-03	0.024	-0.160	13.0	2.7e-04	0.810	1.20	-2.0e-06	0.18	0.45
MMM	4.9e-04	0.016	-0.720	21.0	1.2e-04	0.500	0.81	-2.0e-06	0.25	0.45
AA	5.2e-04	0.023	0.150	14.0	2.5e-04	0.360	1.10	-1.9e-06	0.25	0.45
CVX	6.4e-04	0.016	0.120	15.0	1.3e-04	0.610	0.79	-1.4e-06	0.22	0.46
BA	4.9e-04	0.019	-0.058	9.9	1.9e-04	0.400	0.86	-1.1e-06	0.27	0.46
PG	6.9e-04	0.016	-1.600	50.0	1.4e-04	0.660	0.74	-2.0e-06	0.25	0.47
VZ	4.6e-04	0.017	0.240	11.0	1.4e-04	0.420	0.82	-1.1e-06	0.23	0.48
T	5.4e-04	0.018	0.140	14.0	1.6e-04	0.470	0.88	-1.4e-06	0.23	0.48
KO	6.2e-04	0.017	-0.073	20.0	1.3e-04	0.600	0.79	-1.5e-06	0.24	0.50
UTX	6.6e-04	0.018	-0.650	19.0	1.6e-04	0.590	0.91	-1.5e-06	0.21	0.51
DIS	5.7e-04	0.020	-0.230	18.0	2.1e-04	0.440	1.10	-2.0e-06	0.28	0.51
XOM	6.7e-04	0.016	-0.030	22.0	1.3e-04	0.660	0.84	-1.6e-06	0.21	0.52
HON	5.4e-04	0.021	0.260	29.0	2.2e-04	0.400	1.10	-2.4e-06	0.33	0.53
DD	4.0e-04	0.018	-0.140	8.8	1.6e-04	0.350	0.94	-1.6e-06	0.24	0.53
JPM	6.1e-04	0.024	0.120	14.0	2.8e-04	0.400	1.40	-2.4e-06	0.25	0.54
C	6.7e-04	0.026	1.800	58.0	3.1e-04	0.400	1.50	-1.9e-06	0.25	0.55
IP	2.4e-04	0.020	-0.320	16.0	2.0e-04	0.190	0.98	-2.7e-06	0.26	0.56
AIG	1.9e-04	0.027	-2.400	110.0	3.9e-04	0.110	1.30	-1.5e-06	0.33	0.58
AXP	5.0e-04	0.023	-0.210	12.0	2.6e-04	0.350	1.40	-2.2e-06	0.22	0.59
GE	5.5e-04	0.018	-0.055	11.0	1.5e-04	0.490	1.20	-1.7e-06	0.30	0.62

Figure 2: Correlation between risk measures.

Figure 2 presents pairwise plots and correlation coefficients between risk measures for all firms for which a full sample of data is available (30 October 1986 – 31 December 2008). The panels above the diagonal give the numerical value of the correlation between risk measures, with the size of each value proportional to the strength of correlation between the corresponding measures. Shaded panels indicate a correlation that is significantly different from zero at the 10% level. The panels below the diagonal are centered and scaled so that the slope of the least squares fit, given as a solid line, is equal to the correlation between the two risk measures and all plots are on a common scale. Numerical values for the risk measures are given in Table 3.

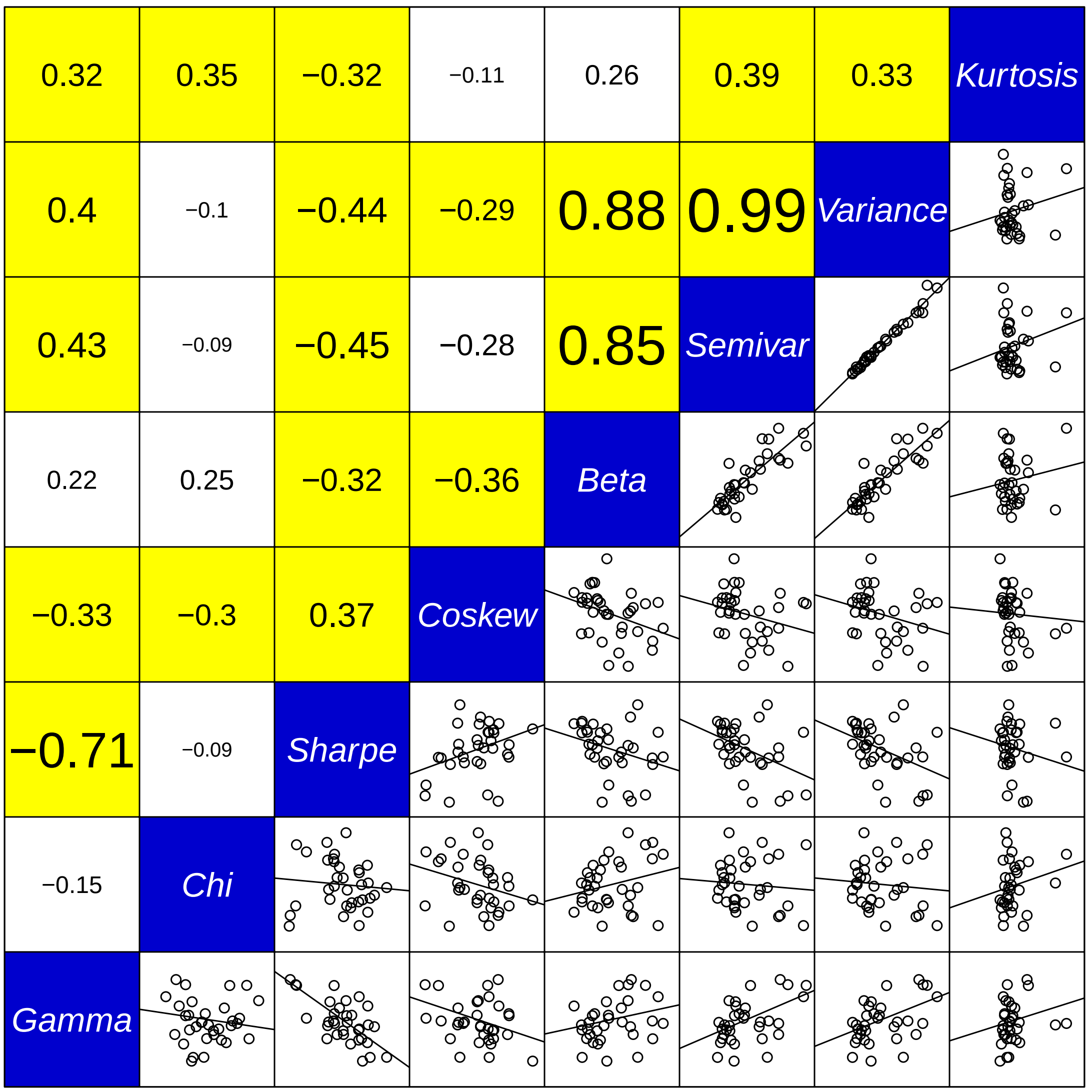
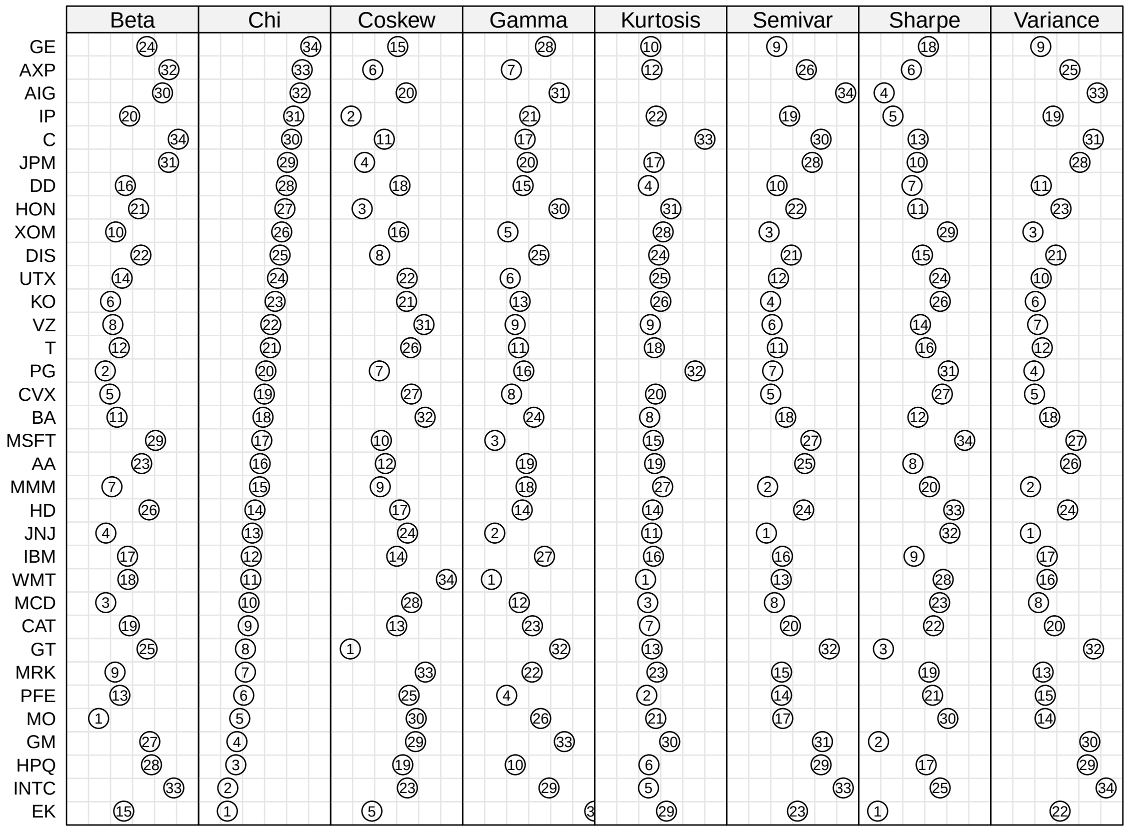


Figure 3: Centered and scaled risk measures and rankings

Figure 3 presents normalized risk measures and corresponding rankings for all firms for which a full sample of data is available (30 October 1986 – 31 December 2008). Each column corresponds to a risk measure and each row to a DJIA constituent. The position of a circle shows how far a firm's risk measure lies from the average of that risk measure over all firms, measured in standard deviations. Risk rankings, from low to high, appear within the circles. Numerical values for the risk measures appear in Table 3.



incorporates substantially different information from the others. In the two cases in which χ is significantly correlated with other risk measures, kurtosis and coskewness, the correlation coefficients are not high enough to suggest that χ provides the same information as those two measures. Figure 3 gives a graphical version of Table 3 that also includes rankings for each risk measure. Firms are sorted in descending order by their values of χ, ranging from General Electric ($\chi = 0.62$) at the upper right of the first column to Eastman Kodak ($\chi = 0.34$) at the bottom left of the first column. If any two of the risk measures provided similar information about the riskiness of stock market returns, they would have similar patterns in Figure 3, but that is not the case: χ differs substantially from all of the other risk measures.

We now examine the performance of χ in portfolio selection. We specify a one-year test period, and form 10,000 randomly selected, equally weighted portfolios from the DJIA constituents the day before the start of the test period. For each portfolio, we then estimate χ using data from a formation period beginning on 30 October 1986 and ending the day before the start of the test period. Using the estimated values, we sort portfolios into three groups on the basis of χ: low-χ (lowest 10%), middle-χ (middle 80%), and high-χ (highest 10%). Finally, we calculate the annualized return for each portfolio in the test period, and average the results across each χ-sorted group. This allows us to address the question of how portfolios with different estimated values of lower tail dependence perform out-of-sample.

We carry out this procedure for ten one-year test periods, 1999–2008, for 10,000 five and ten-stock portfolios. In forming portfolios we exclude firms that were listed on the DJIA fewer than 5 years before the start of the test period to avoid potential selection problems.[18] Table 4 shows the annualized mean return for the formation period and the one-year test period for the five-, and ten-stock χ-sorted portfolios from 1999–2008. In general, the low-χ portfolios provide higher returns than the high-χ portfolios. For example, the five-stock, low-χ portfolios have higher annual returns than the five-stock, high-χ portfolios for eight of the ten years in the sample. The cumulative return for the five-stock, low-χ portfolios is 44%, compared to -8% for the five-stock, high-χ portfolios, and -25%, +19%, and +19% for the S&P 500 Index, DJIA, and an equal-weighted index of the stocks in our sample, respectively. From 1999–2008, there are four years in which the annual return on the DJIA and the S&P 500 was negative, namely 2000–2002, and 2008. For three of the four years, the five-stock, low-χ portfolios have higher mean returns than the indexes, and middle 80% and upper 10% χ portfolios. The last row of Table 4 shows the mean cumulative return from an investment strategy of going long the low-χ portfolios and funding the long position with an offsetting short position in the high-χ portfolios. The mean cumulative return is 20% for the five-stock portfolios and 13% for the ten-stock portfolios.

Note that during 2008, when the market indexes fell 38%, the five-stock, and ten-stock low-χ portfolios were down 32% and 33% while the high-χ portfolios fell 46% and 42%. Thus, the mean return on the low-χ portfolios was better than the market index during 2008 while the mean return on the high-χ portfolios was worse than the market index during 2008. The low-χ portfolios generally provide higher returns than the high-χ portfolios when the market returns are low, as hypothesized. However, the low-χ portfolios outperformed the high-χ portfolios during most of the sample period. For example, the five-stock, low-χ portfolios had higher mean returns than the high-χ portfolios for five of the six years in which the S&P 500 and DJIA had positive annual returns.

[18]The selection criterion that firms were listed on the DJIA for five years prior to inclusion in the sample eliminates Hewlett-Packard, Citigroup, Johnson & Johnson, and Wal-Mart from 1999 - 2002, and Microsoft, Intel, SBC, and Home Depot from 2000 – 2004. As an example of the potential selection problem, the S&P 500 index had a total return of about 190% in the five years before Microsoft and Intel were added to the DJIA in 1999. Those two firms, though, had total returns over 1,000% during the same period. The purpose is to estimate the tail dependence of individual firms with the market, and the high growth prior to index inclusion could distort the estimate of tail dependence.

Table 4: Choosing Portfolios with χ

Table 4 compares the out-of-sample performance of low- and high-χ portfolios from 1999–2008 for a sample of 10,000 five- and ten-stock portfolios composed of DJIA constituents. Test periods are the years given in the first column of the table, while formation periods run from 31 October 1986 through the day before the first day of the corresponding test period. The leftmost panel gives annual returns for the DJIA, the S&P 500, and the subset of the DJIA constituents that were included in the index at least five years prior to the beginning of the test period (Restrict). The three middle panels give summary statistics for portfolios grouped by formation period χ, the extreme value dependence parameter. The final column presents p-values for the two-sided t-test of null hypothesis that there is no difference between the test-period returns of the bottom and top 10% of portfolios ranked by χ. All returns are annualized except for Cumulative and High/Low which are calculated for the entire period from 1999–2008.

					Bottom 10%				Middle 80%				Top 10%				P
					Formation		Test		Formation		Test		Formation		Test		Hi vs.
Year	S&P	Dow	Restrict	Size	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Lo
1999	20	20	21	5	18	0.63	18	0.5	19	0.70	17	0.2	20	0.75	21	0.4	0.00
				10	19	0.73	19	0.3	19	0.77	17	0.1	20	0.81	18	0.3	0.00
2000	-10	2	1	5	19	0.64	-4	0.4	20	0.70	1	0.2	21	0.75	3	0.4	0.00
				10	19	0.74	-1	0.3	20	0.78	1	0.1	20	0.81	0	0.2	0.00
2001	-13	-4	-3	5	17	0.63	-1	0.2	18	0.69	-7	0.1	19	0.74	-9	0.2	0.00
				10	18	0.72	-5	0.1	19	0.75	-7	0.1	19	0.79	-7	0.2	0.00
2002	-23	-12	-10	5	16	0.63	-9	0.2	16	0.68	-11	0.1	17	0.73	-12	0.2	0.00
				10	16	0.72	-11	0.1	17	0.75	-11	0.0	17	0.79	-12	0.1	0.00
2003	26	27	29	5	15	0.65	33	0.4	16	0.70	31	0.1	17	0.75	26	0.3	0.00
				10	16	0.74	32	0.2	16	0.78	31	0.1	16	0.80	28	0.2	0.00
2004	9	8	7	5	17	0.65	9	0.2	16	0.70	8	0.1	17	0.75	8	0.2	0.00
				10	17	0.74	8	0.1	17	0.78	8	0.0	17	0.81	8	0.1	0.00
2005	3	0	0	5	21	0.64	4	0.2	19	0.70	2	0.1	18	0.74	0	0.2	0.00
				10	20	0.74	3	0.1	19	0.77	2	0.1	19	0.80	1	0.1	0.00
2006	14	21	21	5	20	0.63	21	0.3	18	0.69	22	0.1	18	0.74	20	0.2	0.00
				10	19	0.73	22	0.1	18	0.77	22	0.1	18	0.79	21	0.1	0.00
2007	4	9	9	5	19	0.65	14	0.2	18	0.69	9	0.1	18	0.74	5	0.3	0.00
				10	19	0.74	12	0.2	18	0.77	9	0.1	18	0.80	7	0.2	0.00
2008	-38	-38	-38	5	19	0.62	-32	0.4	18	0.68	-37	0.1	17	0.73	-46	0.4	0.00
				10	18	0.72	-33	0.3	18	0.76	-37	0.1	18	0.79	-42	0.3	0.00
Cumulative				5			44				18				-8		
				10			34				18				1		
High/Low				5							20						
				10							13						

Table 5 shows the mean value of the CAPM β, return standard deviation, and the Sharpe Ratio for the low-χ, middle-χ, and high-χ portfolios for the formation period and out-of-sample test period for each year from 1999–2008. Since the low-χ portfolios tend to have lower βs than the high-χ portfolios in both the formation period and the test period, it is possible that their superior returns in down markets may simply reflect this fact rather than their lower tail dependence. To rule this out, we recalculated Table 5, using β-adjusted returns instead of raw returns. The results appear in Table 6. Even after adjusting for β, our main conclusions from above continue to hold: low-χ portfolios consistently outperform high-χ portfolios. The cumulative mean return for the ten-year test period for the low-χ, five-stock portfolios is 79% compared to 23% for the high-χ portfolios. The long/short strategy yields a cumulative mean return of 20% for the five-stock portfolios. For the years in which the indexes had negative returns, the low-χ portfolios had returns at least as high as the high-χ portfolios for three of the four years.

Because of its unusually low returns, 2008 is of particular interest for investors concerned about minimizing downside risk. To examine 2008 in more detail, we construct χ-sorted portfolios on a monthly basis for this year. For each month, we estimate χ using all available data prior to the first of the month, sort the portfolios according to χ as above, and calculate the mean return for each group. Table 7 shows that once again low-χ portfolios consistently outperform the high-χ portfolios and the indexes. For the eight months when the DJIA had negative returns, the low-χ portfolios had higher mean returns in seven of the eight. The cumulative mean return for the five-stock, low-χ portfolios is -30% compared to -44% for the high-χ portfolios, and -38% for the indexes. The long/short strategy for the five-stock portfolios yields a 10% return during 2008.

The results suggest χ is not fully priced as one would expect the high-χ portfolios to have higher returns on average to compensate for the higher risk of extreme losses. Instead, we find just the opposite with the low-χ portfolios outperforming the high-χ portfolios. Our results for χ-sorted portfolios are consistent with previous research showing low-beta portfolios generally outperform high-beta portfolios. Baker et al. (2011) find portfolios of low-beta stocks outperform high-beta stocks by a wide margin in U.S. stocks from 1968 - 2008. They attribute the results to a combination of behavioral factors and limits on arbitrage. Frazzini and Pedersen (2010) find portfolios that are long low-beta assets and short high-beta assets earn significant excess returns across many asset classes and international markets. They attribute the results to limits on leverage that force investors seeking high returns to bid up the price of high-beta assets. Those institutional and behavioral factors may explain our results that low-χ portfolios outperform high-χ portfolios.

An alternative explanation for our results is that most portfolio optimization methods do not account for extreme left-tail dependence. The benefits of reducing large losses on a portfolio by using EVT would not be fully priced because not enough investors are using the necessary techniques.

Table 5: The CAPM β, σ and the Sharpe Ratio for χ-sorted Portfolios

Table 5 presents average values of the CAPM β, the standard deviation of returns, and the Sharpe ratio for portfolios grouped by χ (see Table 4). The Sharpe ratio is based on an arithmetic average of daily returns, annualized by multiplying by 250. The standard deviation is annualized accordingly and then multiplied by 100 to give a result in percentage points. Test periods are the years from 1999–2008 as specified in the leftmost column of the table, while Formation Periods are the period from 31 October 1986 through the day before the start of the Test Period.

		Bottom 10%						Middle 80%						Top 10%					
		Formation Period			Test Period			Formation Period			Test Period			Formation Period			Test Period		
Year	Size	β	σ	Sharpe	β	σ	Sharpe	β	σ	Sharpe	β	σ	Sharpe	β	σ	Sharpe	β	σ	Sharpe
1999	5	1.00	20	0.93	0.59	22	0.79	1.02	19	1.00	0.66	20	0.82	1.04	19	1.05	0.73	20	1.02
	10	1.01	18	1.02	0.62	18	1.01	1.02	18	1.07	0.66	17	0.97	1.03	18	1.10	0.70	17	1.01
2000	5	0.96	19	0.96	0.63	27	-0.05	1.00	19	1.01	0.63	25	0.14	1.04	20	1.07	0.63	25	0.22
	10	0.98	18	1.07	0.63	23	0.07	1.00	18	1.09	0.63	22	0.15	1.02	18	1.10	0.64	21	0.11
2001	5	0.93	20	0.88	0.92	25	0.08	0.95	20	0.93	0.94	25	-0.18	0.99	20	0.96	0.99	26	-0.27
	10	0.93	18	0.99	0.94	23	-0.10	0.95	18	1.01	0.94	23	-0.19	0.97	18	1.02	0.96	24	-0.18
2002	5	0.91	20	0.81	0.93	28	-0.22	0.95	20	0.84	0.98	29	-0.27	1.00	21	0.85	1.06	30	-0.27
	10	0.93	19	0.90	0.95	27	-0.28	0.95	19	0.91	0.98	27	-0.29	0.98	19	0.91	1.04	29	-0.31
2003	5	0.94	21	0.77	0.97	19	1.53	0.99	21	0.77	1.00	20	1.46	1.04	22	0.80	1.01	19	1.27
	10	0.95	19	0.84	0.97	18	1.60	0.99	20	0.84	1.00	18	1.54	1.02	20	0.84	1.00	18	1.44
2004	5	0.92	21	0.83	0.93	13	0.74	0.99	21	0.81	0.97	13	0.67	1.04	22	0.80	0.96	13	0.63
	10	0.94	19	0.90	0.95	12	0.74	0.99	20	0.88	0.97	12	0.72	1.03	20	0.86	0.97	12	0.66
2005	5	1.03	22	0.94	0.95	13	0.36	1.02	21	0.89	0.96	12	0.19	1.04	21	0.87	0.94	12	0.04
	10	1.01	20	1.00	0.96	11	0.28	1.02	20	0.97	0.95	11	0.21	1.04	20	0.95	0.94	11	0.12
2006	5	1.03	22	0.90	0.97	13	1.51	1.02	21	0.86	0.96	12	1.69	1.04	21	0.85	0.92	11	1.62
	10	1.02	19	0.96	0.96	11	1.80	1.02	19	0.94	0.96	11	1.86	1.04	19	0.93	0.95	11	1.81
2007	5	1.00	21	0.91	0.89	16	0.92	1.02	21	0.88	0.93	17	0.59	1.05	21	0.87	0.96	17	0.36
	10	0.99	19	0.99	0.89	15	0.80	1.02	19	0.96	0.93	16	0.60	1.04	19	0.94	0.95	16	0.52
2008	5	1.01	21	0.91	0.91	41	-0.75	1.01	20	0.87	1.02	46	-0.79	1.04	21	0.85	1.15	52	-0.93
	10	1.00	19	0.98	0.94	40	-0.80	1.01	19	0.94	1.02	44	-0.84	1.04	19	0.93	1.10	47	-0.91

Table 6: Choosing Portfolios with χ

Table 6 presents the same comparisons as Table 4 with CAPM-adjusted returns replacing raw test period returns. Formation period values of β are used for the adjustment.

					Bottom 10%				Middle 80%				Top 10%				P
					Formation		Test		Formation		Test		Formation		Test		Hi vs.
Year	S&P	Dow	Restrict	Size	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Lo
1999	20	20	21	5	18	0.63	-3	0.4	19	0.70	-4	0.1	20	0.75	-1	0.3	0.00
				10	19	0.73	-2	0.2	19	0.77	-3	0.1	20	0.81	-3	0.2	0.00
2000	-10	2	1	5	19	0.64	4	0.4	20	0.70	10	0.2	21	0.75	13	0.5	0.00
				10	19	0.74	8	0.3	20	0.78	11	0.1	20	0.81	10	0.2	0.00
2001	-13	-4	-3	5	17	0.63	12	0.3	18	0.69	6	0.1	19	0.74	4	0.3	0.00
				10	18	0.72	8	0.2	19	0.75	6	0.1	19	0.79	7	0.2	0.00
2002	-23	-12	-10	5	16	0.63	15	0.3	16	0.68	14	0.1	17	0.73	15	0.2	0.67
				10	16	0.72	14	0.1	17	0.75	15	0.1	17	0.79	14	0.1	0.27
2003	26	27	29	5	15	0.65	7	0.3	16	0.70	4	0.1	17	0.75	-1	0.2	0.00
				10	16	0.74	5	0.2	16	0.78	4	0.1	16	0.80	1	0.1	0.00
2004	9	8	7	5	17	0.65	1	0.2	16	0.70	-1	0.1	17	0.75	-2	0.2	0.00
				10	17	0.74	0	0.1	17	0.78	-1	0.0	17	0.81	-2	0.1	0.00
2005	3	0	0	5	21	0.64	1	0.2	19	0.70	-1	0.1	18	0.74	-3	0.2	0.00
				10	20	0.74	0	0.1	19	0.77	-1	0.0	19	0.80	-2	0.1	0.00
2006	14	21	21	5	20	0.63	6	0.2	18	0.69	7	0.1	18	0.74	5	0.2	0.00
				10	19	0.73	7	0.1	18	0.77	7	0.0	18	0.79	6	0.1	0.00
2007	4	9	9	5	19	0.65	10	0.2	18	0.69	5	0.1	18	0.74	1	0.3	0.00
				10	19	0.74	7	0.2	18	0.77	5	0.1	18	0.80	3	0.2	0.00
2008	-38	-38	-38	5	19	0.62	8	0.5	18	0.68	3	0.2	17	0.73	-8	0.6	0.00
				10	18	0.72	7	0.4	18	0.76	3	0.1	18	0.79	-3	0.4	0.00
Cumulative				5			79				51				23		
				10			68				55				34		
High/Low				5							20						
				10							12						

Table 7: Choosing Portfolios with χ – Monthly Windows 2008

Table 7 compares the out-of-sample performance of low- and high-χ portfolios during each month of 2008 for a sample of 10,000 five- and ten-stock portfolios composed of DJIA constituents. The leftmost panel gives monthly returns for the DJIA and S&P 500, while the three middle panels give summary statistics for portfolios grouped by formation period χ, the extreme value dependence parameter. The test periods correspond to particular months of 2008, while the formation periods correspond to the period from 31 October 1986 through the day before the beginning of the test period. The final column presents p-values for the two-sided t-test of null hypothesis that there is no difference between the test-period returns of the bottom and top 10% of portfolios ranked by χ. Test period mean returns are cumulative over the test month, while formation period returns are annualized. Both Cumulative and High/Low returns are cumulative over 2008.

				Bottom 10%				Middle 80%				Top 10%				P
				Formation		Test		Formation		Test		Formation		Test		Hi vs.
Month	S&P	Dow	Size	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Mean	χ	Mean	S.E.	Lo
Jan	-6	-4	5	1	0.62	-5	0.1	1	0.68	-4	0.0	1	0.73	-4	0.1	0.00
			10	1	0.72	-5	0.1	1	0.76	-4	0.0	1	0.79	-3	0.1	0.00
Feb	-3	-4	5	1	0.63	-2	0.1	1	0.69	-4	0.0	1	0.74	-5	0.1	0.00
			10	1	0.73	-3	0.1	1	0.77	-4	0.0	1	0.80	-5	0.1	0.00
Mar	-1	0	5	1	0.63	0	0.1	1	0.69	0	0.0	1	0.73	0	0.1	0.12
			10	1	0.73	0	0.1	1	0.76	0	0.0	1	0.79	0	0.1	0.00
Apr	5	4	5	1	0.63	3	0.1	1	0.69	4	0.0	1	0.73	4	0.1	0.00
			10	1	0.73	3	0.1	1	0.76	4	0.0	1	0.79	4	0.1	0.00
May	1	-2	5	1	0.63	0	0.1	1	0.68	-2	0.0	1	0.73	-3	0.1	0.00
			10	1	0.73	-1	0.1	1	0.76	-2	0.0	1	0.79	-3	0.1	0.00
June	-9	-12	5	1	0.63	-9	0.1	1	0.69	-12	0.0	1	0.73	-13	0.1	0.00
			10	1	0.73	-10	0.1	1	0.77	-12	0.0	1	0.79	-12	0.1	0.00
July	-1	1	5	1	0.64	1	0.1	1	0.69	1	0.0	1	0.73	2	0.1	0.00
			10	1	0.74	1	0.1	1	0.77	1	0.0	1	0.79	2	0.1	0.00
Aug	1	2	5	1	0.63	3	0.1	1	0.69	2	0.0	1	0.73	1	0.1	0.00
			10	1	0.74	3	0.0	1	0.77	2	0.0	1	0.79	1	0.0	0.00
Sept	-9	-9	5	1	0.64	-6	0.1	1	0.69	-9	0.1	1	0.73	-9	0.3	0.00
			10	1	0.74	-8	0.1	1	0.77	-9	0.1	1	0.80	-8	0.2	0.03
Oct	-17	-16	5	1	0.65	-15	0.2	1	0.70	-16	0.1	1	0.74	-17	0.1	0.00
			10	1	0.75	-15	0.1	1	0.78	-16	0.0	1	0.80	-16	0.1	0.00
Nov	-7	-6	5	1	0.65	-5	0.1	1	0.70	-6	0.0	1	0.74	-7	0.1	0.00
			10	1	0.75	-5	0.1	1	0.78	-6	0.0	1	0.80	-6	0.1	0.00
Dec	1	-1	5	1	0.65	1	0.1	1	0.70	-1	0.1	1	0.75	-3	0.1	0.00
			10	1	0.75	0	0.1	1	0.78	-1	0.0	1	0.80	-1	0.1	0.00
Cumulative	-38	-38	5			-30				-39				-44		
			10			-35				-39				-39		
High/Low			5							10						
			10							4						

6. Conclusion

Our main conclusions are as follows: First, we show intuitively that χ can provide important information for investors who seek to limit their downside risk. Second, we show that χ differs systematically from other risk measures including variance, semi-variance, skewness, kurtosis, CAPM β, coskewness, and the univariate EVT scale parameter γ. Although χ is positively, significantly correlated with kurtosis and negatively, significantly correlated with coskewness, the correlation coefficients of +.35 and -.30 show χ is not simply a proxy for those two measures. Third, in out-of-sample tests using data from 1999–2008, low-χ portfolios outperform the market index, the mean return of the stocks in our sample, and high-χ portfolios. In summary, we show χ is conceptually important for portfolio optimization, different from other risk measures, and provides practical information for selecting portfolios.

References

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. The Review of Financial Studies 19 (4), 1191–1239.

Arzac, E., Bawa, V., 1977. Portfolio choice and equilibrium in capital markets with safety-first investors. Journal of Financial Economics 4 (3), 277–288.

Baker, M., Bradley, B., Wurgler, J., 2011. Benchmarks as limits to arbitrage: understanding the low volatility anomaly. Financial Analysts Journal 67 (1), 40–54.

Beirlant, J., Goegebeur, Y., Teugels, J., Segers, J., Waal, D. D., Ferro, C., 2004. Statistics of Extremes: Theory and Applications. Wiley, Hoboken.

Chollete, L., de la Pena, V., Lu, C., 2009. International diversification: An extreme value approach. UiS Working Papers in Economics and Finance.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. Springer, London.

Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extreme value analyses. Extremes 2 (4), 339–265.

Cont, R., 2001. Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance 1, 223–236.

Danielsson, J., De Vries, C., 1997. Tail index and quantile estimation with very high frequency data. Journal of Empirical Finance 4 (2-3), 241–257.

Embrechts, P., Lindskog, F., McNeil, A., 2003. Modelling dependence with copulas and applications to risk management. In: Rachev, S. (Ed.), Handbook of Heavy Tailed Distributions in Finance. Elsevier, Ch. 8, pp. 329–384.

Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. In: Dempster, M. A. H. (Ed.), Risk Management: Value at Risk and Beyond. Cambridge University Press, pp. 176–223.

Fama, E., 1965. The behavior of stock-market prices. Journal of Business 38 (1), 34–105.

Frahm, G., Junker, M., Schmidt, R., 2005. Estimating the tail-dependence coefficient: Properties and pitfalls. Insurance: Mathematics and Economics 37, 80–100.

Frazzini, A., Pedersen, L., 2010. Betting against beta.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55 (3), 1263–1295.

Harvey, C. R., Liechty, J. C., Liechty, M. W., Müller, P., May 2010. Portfolio selection with higher moments. Quantitative Finance 10 (5), 469–485.

Herings, P. J.-J., Kubler, F., 2007. Approximate capm when preferences are crra. Computational Economics 29, 13–31.

Hyung, N., De Vries, C., 2005. Portfolio diversification effects of downside risk. Journal of Financial Econometrics 3 (1), 107.

Hyung, N., De Vries, C., 2007. Portfolio selection with heavy tails. Journal of Empirical Finance 14 (3), 383–400.

Ibragimov, R., Walden, J., 2007. The limits of diversification when losses may be large. Journal of Banking & Finance 31 (8), 2551–2569.

Jansen, D., 2001. Limited downside risk in portfolio selection among us and pacific basin equities. International Economic Journal 15 (4), 1–22.

Jansen, D., De Vries, C., 1991. On the frequency of large stock returns: Putting booms and busts into perspective. The Review of Economics and Statistics 73 (1), 18–24.

Jansen, D., Koedijk, K., de Vries, C., 2000. Portfolio selection with limited downside risk. Journal of Empirical Finance 7 (3-4), 247–269.

Jondeau, E., Rockinger, M., 2003. Testing for differences in the tails of stock-market returns. Journal of Empirical Finance 10 (5), 559–581.

Kearney, C., Lynch, M., 2007. Are international equity markets really skewed? Applied Financial Economics 17, 399–411.

Kearns, P., Pagan, A., 1997. Estimating the density tail index for financial time series. Review of Economics and Statistics 79 (2), 171–175.

Kroll, Y., Levy, H., Markowitz, H. M., March 1984. Mean-variance versus direct utility maximization. The Journal of Finance 39 (1), 47–61.

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

Longin, F., 1996. The asymptotic distribution of extreme stock market returns. Journal of Business 69 (3), 383.

Longin, F., 2005. The choice of the distribution of asset returns: How extreme value theory can help? Journal of Banking and Finance 29 (4), 1017–1035.

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

Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36 (4), 394.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7 (1), 77–91.

Markowitz, H., 1959. Portfolio selection: efficient diversification of investments. Cowles Foundation for Research in Economics at Yale University.

McNeil, A., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance 7 (3-4), 271–300.

Mossin, J., 1966. Equilibrium in a capital asset market. Econometrica 34 (4), 768–783.

Peiró, A., 1999. Skewness in financial returns. Journal of Banking and Finance 23 (6), 847–862.

Poon, S., Rockinger, M., Tawn, J., 2004. Extreme value dependence in financial markets: Diagnostics, models, and financial implications. Review of Financial Studies 17 (2), 581.

Premaratne, G., Bera, A., 2005. A test for symmetry with leptokurtic financial data. Journal of Financial econometrics 3 (2), 169.

Rachev, S. T., Menn, C., Faboozzi, F. J., 2005. Fat-tailed and Skewed Asset Return Distributions: implications for Risk Management, Portfolio Selection and Option Pricing. Wiley, Hoboken.

Reiss, R., Thomas, M., 1997. Statistical analysis of extreme values: from insurance, finance, hydrology, and other fields. Boston: Basel.

Roy, A., 1952. Safety first and the holding of assets. Econometrica 20 (3), 431–449.

Sharpe, W., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The Journal of Finance 19 (3), 425–442.

Susmel, R., 2001. Extreme observations and diversification in latin american emerging equity markets. Journal of International Money and Finance 20 (7), 971–986.

Treynor, J., 1961. Toward a theory of market value of risky assets. Unpublished manuscript.

Wu, F., Huang, W., Liu, Q., Rhee, S., 2008. Extreme downside risk and expected stock returns, working Paper.

Appendix A. Data Notes

Missing Observation for Altria:. The NYSE suspended trading in Altria's (Philip Morris at the time) stock on May 25, 1994 while the company's board of directors met to decide whether to spin-off the company's tobacco business. As trading resumed on May 26, 1994 the CRSP return of -5.81% on this date represents the percentage change in the stock price from May 24 to May 26. This explains the missing observation.

Appendix B. Sampling from the Clayton Copula

The general algorithm for simulating a random vector (X, Y) with (strictly increasing) marginal distribution functions F_X, F_Y and dependence given by a Clayton Copula is as follows[19]

1. Generate independent Uniform$(0, 1)$ random variables U_1 and V
2. Let $U_2 = \left[U_1^{-\theta}\left(V^{-\theta/(1+\theta)} - 1\right) + 1\right]^{-1/\theta}$
3. Set $X = F_1^{-1}(U_1)$, $Y = F_2^{-1}(U_2)$

Appendix C. Estimating Tail Dependence – Technical Details

Appendix C.1. Likelihood for the Block Maximum Losses

Consider a series of block maxima $\{z_i\}_{i=1}^{M}$. According to Equation 6, the *individual* block maxima are approximately GEV if the blocks are sufficiently long. Provided the series is stationary and satisfies a weak dependence condition, estimation based on a likelihood that assumes independence will still produce consistent estimates, although the usual standard errors are invalid. Thus we write the log-likelihood as

$$\begin{aligned}\ell(\mu, \sigma, \gamma) = -M \log \sigma - (1 + 1/\gamma) \sum_{n=i}^{M} \log \left[1 + \gamma \left(\frac{z_i - \mu}{\sigma}\right)\right] \\ - \sum_{n=i}^{M} \left[1 + \gamma \left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\gamma}\end{aligned} \tag{C.1}$$

defined wherever $1 + \gamma(z_n - \mu)/\sigma > 0$ for all $i = 1, \ldots, M$. This expression follows from Equation 6 by differentiating with respect to z to yield the probability density function, taking the product of this result over all i and using the natural logarithm to convert products into sums.[20]

[19] For details, see Embrechts et al. 2003.

[20] For the case of $\gamma = 0$ the likelihood requires a slightly different treatment. However real market data, in our experience, tend to have $\gamma > 0$, corresponding to very fat tails. In particular, all of our estimates of γ from this study are well above zero.

Appendix C.2. Transforming to Standard Fréchet Marginals

Suppose that a random variable Z follows a generalized extreme value distribution. Then,

$$\tilde{Z} \equiv f(Z) = \left[1 + \gamma\left(\frac{Z-\mu}{\sigma}\right)\right]^{1/\gamma} \tag{C.2}$$

is unit Fréchet. To see why this is the case, we use the C.D.F. technique for transforming a random variable. Thus, using Equation 6,

$$\begin{aligned}
\mathbb{P}(\tilde{Z} \le z) &= \mathbb{P}\left(Z \le f^{-1}(z)\right) = G\left(f^{-1}(z)\right) \\
&= \exp\left\{-\left[1 + \frac{\gamma}{\sigma}\left(f^{-1}(z) - \mu\right)\right]^{-1/\gamma}\right\} \\
&= \exp\left\{-\left[1 + \frac{\gamma}{\sigma}\left(\left[\left(\frac{\sigma}{\gamma}\right)(z^{\gamma} - 1) + \mu\right] - \mu\right)\right]^{-1/\gamma}\right\} \\
&= \exp(-1/z)
\end{aligned}$$

Appendix C.3. Bivariate Logistic Likelihood

The pdf corresponding to Equation 8 is given by cross partial derivative of G, that is

$$g(x,y) = \frac{\partial}{\partial x \partial y} G(x,y)$$

Thus,

$$g(x,y) = e^{-V}(V_x V_y - V_{xy}) \tag{C.3}$$

where

$$V = \left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha} \tag{C.4}$$

$$V_x = -\left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha-1} x^{-(\alpha+1)/\alpha} \tag{C.5}$$

$$V_y = -(x^{-1/\alpha} + y^{-1/\alpha})^{\alpha-1} y^{-(\alpha+1)/\alpha} \tag{C.6}$$

$$V_{xy} = \frac{\alpha - 1}{\alpha}\left(x^{-1/\alpha} + y^{-1/\alpha}\right)^{\alpha-2}(xy)^{-(\alpha+1)/\alpha} \tag{C.7}$$

Under weak conditions we may ignore temporal dependence[21] and write the likelihood for the entire sample as the product of the likelihoods of each observation:

$$\ell(\alpha) = \sum_{i=1}^{M} \log g(x_i, y_i; \alpha) \tag{C.8}$$

where g is defined as in Equation C.3, and x, y have been transformed to unit Fréchet as described in Appendix C.2.

[21] From the perspective of parameter estimation our limiting results allow us to ignore temporal dependence, but it complicates the calculation of appropriate standard errors. The usual MLE standard errors are too small, but correcting them would require further assumptions.

Appendix D. Robustness Tests

In this section we test the robustness of our estimates of χ by examining how the values for individual assets, given in Table 3, change when we alter our estimation procedure.

We begin by shifting block starting dates through time: first estimating χ for blocks shifted by one day, then two days, and so on. Because we use blocks of 22 trading days, we repeat this process through 21 days to cover all possible block shifts. We then compare all 22 estimates of χ: one for the original blocks and 21 corresponding to the shifted blocks. The results are clear: estimates of χ are virtually unchanged for any of the block shifts. In particular, the correlations and rank correlations between our original estimates of χ and those based on shifted blocks are all greater than 0.99. We then considered the effect of changing block lengths on our estimation procedure, using lengths ranging from 8 to 36 trading days, and then calculating the correlation and rank correlation between the resulting estimates. When the block lengths change by only a few days, the correlation and rank correlation of the estimates of χ across the block lengths are extremely high, generally exceeding 0.95. Yet even for block lengths that differ considerably, the correlations remain relatively high: the lowest correlation is 0.85 (block lengths of 9 and 33 trading days) and the lowest rank correlation is 0.84 (block lengths of 9 and 34 trading days). Thus, neither the block shifts nor changes in the block lengths substantially change our estimates of χ.

As a final robustness test, we compared our maximum likelihood block method of estimating χ to a non-parametric threshold method used by Poon et al. (2004). Specifically, we estimated χ for each stock in the sample using ten thresholds, ranging from the lowest 1% of returns to the lowest 10% of returns. Unlike the block shifts and changes to the block length, the nonparametric threshold method often provides estimates of χ that vary substantially from the method described in Section 4. Table D.8 shows how the resulting estimates of χ depend on the threshold used for the estimation. As the threshold for estimating χ becomes more extreme, moving from the worst 10% of returns to the worst 1% of returns, the variation in the estimates increases. The χ estimates from the 1% threshold have a standard deviation and range more than twice as large as the χ estimates from the 10% threshold. Table D.9 shows the correlation between the χ estimates from the threshold method and our block maximum likelihood method. The block method produces estimates of χ that are most closely correlated with those from the 1% threshold. The pattern is similar for the rank correlation with the block minima method most closely correlated with the 1% threshold.

Table D.8: Estimates of χ Using Block Minima and Threshold Methods

	10%	9%	8%	7%	6%	5%	4%	3%	2%	1%	mle
MICROSOFT	0.47	0.46	0.45	0.45	0.44	0.43	0.40	0.38	0.35	0.34	0.45
HONEYWELL INTERNATIONAL	0.47	0.47	0.46	0.47	0.46	0.45	0.44	0.42	0.42	0.39	0.53
AT&T	0.44	0.43	0.43	0.43	0.41	0.39	0.37	0.36	0.35	0.31	0.50
COCA COLA	0.47	0.46	0.46	0.45	0.45	0.44	0.44	0.42	0.40	0.40	0.53
DU PONT	0.44	0.44	0.43	0.43	0.42	0.42	0.40	0.37	0.36	0.35	0.34
KODAK	0.45	0.45	0.44	0.44	0.42	0.41	0.39	0.38	0.37	0.37	0.52
EXXON MOBIL	0.58	0.58	0.59	0.60	0.59	0.58	0.56	0.54	0.55	0.61	0.62
GENERAL ELECTRIC	0.44	0.43	0.43	0.41	0.40	0.39	0.37	0.35	0.37	0.36	0.37
GENERAL MOTORS	0.47	0.46	0.46	0.45	0.46	0.45	0.45	0.42	0.40	0.36	0.42
IBM	0.42	0.41	0.40	0.40	0.38	0.36	0.34	0.32	0.31	0.26	0.38
ALTRIA	0.41	0.41	0.40	0.39	0.39	0.38	0.36	0.37	0.35	0.43	0.46
SEARS	0.44	0.43	0.43	0.41	0.40	0.38	0.39	0.38	0.35	0.30	0.40
CHEVRON	0.47	0.45	0.44	0.43	0.44	0.43	0.42	0.41	0.40	0.43	0.51
UNION CARBIDE	0.43	0.42	0.41	0.40	0.39	0.38	0.35	0.33	0.31	0.30	0.47
GOODYEAR	0.44	0.43	0.42	0.40	0.40	0.39	0.38	0.38	0.35	0.33	0.41
UNITED TECHNOLOGIES	0.43	0.43	0.42	0.41	0.40	0.41	0.39	0.38	0.37	0.35	0.46
PROCTER & GAMBLE	0.44	0.44	0.44	0.43	0.42	0.40	0.39	0.37	0.36	0.42	0.56
HONEYWELL	0.45	0.44	0.44	0.43	0.42	0.42	0.39	0.38	0.38	0.35	0.39
CATERPILLAR	0.44	0.43	0.43	0.42	0.42	0.41	0.40	0.37	0.36	0.34	0.42
BOEING	0.47	0.46	0.46	0.45	0.44	0.44	0.43	0.42	0.39	0.38	0.45
INTERNATIONAL PAPER	0.45	0.45	0.45	0.43	0.42	0.41	0.39	0.38	0.36	0.35	0.40
PFIZER	0.44	0.44	0.43	0.42	0.42	0.41	0.40	0.38	0.38	0.39	0.45
JOHNSON & JOHNSON	0.48	0.47	0.47	0.45	0.44	0.44	0.43	0.42	0.39	0.37	0.51
3M	0.46	0.46	0.45	0.44	0.44	0.41	0.41	0.39	0.36	0.31	0.37
MERCK	0.43	0.42	0.41	0.39	0.38	0.37	0.36	0.33	0.31	0.30	0.41
ALCOA INC	0.50	0.50	0.49	0.49	0.48	0.47	0.46	0.44	0.43	0.42	0.54
DISNEY WALT CO	0.47	0.46	0.46	0.45	0.43	0.43	0.41	0.38	0.36	0.33	0.42
HEWLETT PACKARD	0.52	0.51	0.52	0.50	0.50	0.50	0.47	0.49	0.49	0.49	0.59
MCDONALDS	0.47	0.47	0.46	0.46	0.45	0.44	0.43	0.41	0.37	0.32	0.34
JPMORGAN CHASE	0.44	0.44	0.42	0.42	0.41	0.40	0.40	0.39	0.37	0.36	0.48
WAL MART	0.44	0.44	0.42	0.43	0.42	0.40	0.40	0.38	0.37	0.34	0.48
AMERICAN EXPRESS	0.46	0.46	0.46	0.45	0.44	0.46	0.45	0.42	0.39	0.35	0.43
INTEL	0.48	0.48	0.48	0.47	0.46	0.45	0.44	0.45	0.45	0.44	0.58
VERIZON	0.50	0.50	0.50	0.49	0.48	0.47	0.47	0.49	0.46	0.47	0.55

Table D.9: Correlation of χ Estimates from Block Minima and Threshold Methods

	10%	9%	8%	7%	6%	5%	4%	3%	2%	1%	Block
10%	1	0.99	0.99	0.98	0.98	0.96	0.94	0.93	0.91	0.77	0.58
9%	0.99	1	0.99	0.99	0.98	0.97	0.95	0.93	0.91	0.78	0.59
8%	0.99	0.99	1	0.99	0.99	0.97	0.95	0.93	0.92	0.8	0.6
7%	0.98	0.99	0.99	1	0.99	0.98	0.95	0.92	0.92	0.8	0.59
6%	0.98	0.98	0.99	0.99	1	0.98	0.97	0.94	0.93	0.82	0.6
5%	0.96	0.97	0.97	0.98	0.98	1	0.97	0.95	0.94	0.82	0.59
4%	0.94	0.95	0.95	0.95	0.97	0.97	1	0.96	0.93	0.78	0.56
3%	0.93	0.93	0.93	0.92	0.94	0.95	0.96	1	0.97	0.85	0.64
2%	0.91	0.91	0.92	0.92	0.93	0.94	0.93	0.97	1	0.9	0.68
1%	0.77	0.78	0.8	0.8	0.82	0.82	0.78	0.85	0.9	1	0.76
Block	0.58	0.59	0.6	0.59	0.6	0.59	0.56	0.64	0.68	0.76	1