Market Crashes, Market Booms and the Cross-Section of Expected Returns*

Jonathan F. Spitzer†

January 25, 2006

Abstract

Large market-wide price movements cannot be completely diversified away by investors. Consequently, expected returns should be higher for stocks that crash heavily during market crashes or stocks that have their best payoffs during market booms. Over the period 1967 to 2004, a zero-investment strategy isolating crash risk returns 6% per annum and a strategy isolating boom risk delivers 7.5%. These returns are not reward for bearing overall market, size, book-to-market or momentum risk. Intriguingly, these effects are concentrated in the period following the crash of Oct 19th, 1987 and the returns to the crash strategy can be predicted by the skew in implied volatilities for S&P 500 index put options.

*I am grateful for comments from Wally Boudry, Stephen Brown, Prachi Deuskar, Ned Elton, Rob Engle, Steve Figlewski, Marty Gruber, Amrut Nashikkar, Robert Whitelaw and seminar participants at NYU. I also thank Steve Figlewski for providing data. Financial support from the Nasdaq Derivatives Research Project at the Salomon Center is gratefully acknowledged.

†Leonard N. Stern School of Business, New York University, Department of Finance, Suite 9-180, 44 West 4th Street, New York, New York 10012. email: jspitzer@stern.nyu.edu Tel: 212-998-0327 Fax: 212-998-4233.

A distinguishing feature of financial markets is the relative frequency of large, market-wide price movements. Markets have a tendency to crash and, to a lesser extent, to boom. Even investors holding well-diversified portfolios cannot completely escape such movements, for the very reason that the price movements are market-wide. This paper explores whether these undiversifiable risks play a role in the determination of expected returns.

The dependence structure between an individual stock and the broader market should play a role in determining the expected return of that individual stock, since it determines when an asset pays off. State prices for payoffs during market crashes should be high since this is when marginal utility is highest. Conversely, payoffs are easy to obtain during market booms so the state prices for such payoffs are low.

These ideas are at the heart of the CAPM. With reasonable utility specifications, the distributional assumptions required for the CAPM to hold are such that the stock has linear conditional dependence upon the market. Beta then determines the amount of undiversifiable risk that the stock would add to a well-diversified portfolio and so is the relevant measure of risk. Yet despite strong theoretical foundations and undeniable intuitive appeal, the model has failed empirically.

Using a Taylor series expansion of the utility function, Rubinstein (1973) broadened the definition of market dependence by showing that expected returns should be a function of *all* of the co-moments of the stock and the market. Kraus and Litzenberger (1976) examine the third co-moment, co-skewness, and find that asset returns reflect this risk, as do Harvey and Siddique (2000) using a conditional definition of co-skewness. Dittmar (2002) examines the fourth co-moment and finds that this too is priced. Chung, Johnson, and Schill (2006) extend the analysis to higher co-moments and demonstrates their ability to subsume the pricing implications of size and book-to-market (Fama and French (1993)).

Another approach has been to examine some subset of the joint return distribution. The idea goes back as far as Markowitz (1959) where the idea of semi-variance was introduced. Bawa and Lindenberg (1977) develop a formal equilibrium model where lower partial moments determine expected returns. More recently, Ang, Chen, and Xing (2005) consider betas conditioned on down market returns and find a contemporaneous relation between returns and downside beta. Our approach also examines a subset of the joint return distribution, but we focus on payoffs during the extreme states of the market. This allows us to focus solely on times when state prices should be most different, ie. high for payoffs during market crashes and low during for payoffs market booms.

It is well known that low frequency, high impact events may play an important role in the way investors asses risk. Reitz (1988) and, more recently, Barro (2005) suggest that the equity premium puzzle may be the result of "peso problem" where a low-probability calamitous event is part of the perceived market return distribution but the event has not been observed in the historical record. Further, Brown, Goetzmann, and Ross (1995) demonstrate the pervasive nature of the statistical difficulties that such a scenario entails, since only the surviving markets are examined and these are precisely the markets that are less likely to have experienced the calamitous decline. These studies focus on the impact of large crashes upon the market premium, while in this study we will examine effects of exposure to such events in the cross-section of stocks.

Our main finding is that the market rewards investors for bearing both crash risk and boom risk. While crash risk is easily understood, the intuition behind boom risk is not as immediate. The key issue is that investors would prefer to have the best payoffs from a stock at a time when payoffs are not so plentiful. A stock that offers a large payoff will be more valuable if that payoff is made during normal times (or even during a crash) than if that payoff occurs during a market boom.

The zero investment factor portfolio designed to capture the risk of holding stocks that crash when the market crashes has excess returns over the Fama-French-Carhart model of 6% per year over the period 1967 to 2004. Likewise, the zero investment portfolio constructed to capture the risk of holding stocks that have their best payoffs when the market booms delivers 7.5% per year. These returns are not compensation for bearing overall market, size, book-to-market or momentum risk.

The time series of these returns reveal that the premia increase substantially following the 1987 crash. In fact, the premia can not be statistically distinguished from zero preceding the crash. Tests for a structural break in the premia with either a pre-specified breakpoint or using a data-driven choice of breakpoint each point to the 1987 crash as the pivotal date.

Such a shift in asset pricing following the crash is also evident in post-1987 skew in implied volatilities for out-of-the-money S&P 500 index put options (Rubinstein (1994)). While understanding the skew is made difficult by our lack of the "true" option pricing model, the skew nonetheless suggests that state prices for payoffs during crashes are quite high. Since option prices embed forecasts, the skew can be interpreted as a barometer of investors perception of future crash risk. We show that when crashes are perceived to be more likely (as measured by a larger skew), subsequent returns to the strategy isolating crash risk are also higher. Additionally, when the skew becomes unexpectedly larger, contemporaneous returns are lower as prices drop on stocks with greater exposure to crash risk. The relationship is particularly noteworthy since the appearance of the skew represents a notable shift in investors' perception of the return distribution and/or their tastes. It would suggest an unrealistic degree of market segmentation for these same effects to not be apparent in the stock market itself.

This paper is structured as follows. In Section 1 we present a simple example of an economy where tail dependence generates cross-sectional return premia. In Section 2 we discuss the technology we will employ to measure extreme co-movements: Extreme Value Theory. Pricing implications are examined in Section 3 and evidence of a structural break in tail dependence risk premia following the October 1987 crash is presented in Section 4. Section 5 documents the relationship between the returns to bearing crash risk and the skew in implied volatilities of S&P 500 index put options. Section 6 concludes.

1 A Simple Example of Tail Premia

To illustrate the importance of booms and crashes, we will consider a very simple economy. The main idea flows naturally from the Reitz (1988) analysis of the equity premium puzzle, which has also been revisited more recently by Barro (2005). The central insight is that the riskfree asset can be made much more valuable if there is a low probability of a large crash, because the riskfree asset will maintain its value during a crash. Investors are so desperate for payoffs during a crash that state prices are very high, which makes the riskfree asset far more desirable and so the riskfree rate is much lower. The equity premium puzzle is then solved as there is a larger equity premium without the need for unrealistically high risk aversion.

Despite being a well-known solution to the equity premium puzzle, the important cross-sectional implication of the Reitz idea has not yet been explored. Just as the possibility of a crash can create a large difference between the equity market return and the riskfree return, precisely the same logic implies that stocks with differences in just their payoffs during crashes can also have large spreads between their expected returns. We also extend the idea to market booms, where payoffs are so plentiful that the state prices become very low.

The model we consider is a very simple economy. There is a representative investor who maximizes power utility with coefficient of relative risk aversion of $\gamma = 5$ and subjective discount factor $\beta = 0.98$. This investor chooses current consumption and consumption today and at one date in the future, after which the world promptly ends. The investor has current endowment normalized to one and faces four possible endowments in the future, which define the four possible states of nature. In the 'good' state of nature, the endowment is 1.07 and in the 'bad' state of nature the endowment is 0.985. There is also a 'crash' state where the endowment falls to 0.6 and a 'boom' state where the endowment becomes 1.5. Both the 'crash' and 'boom' state occur with probability $p = 0.01$ or about once every century, while the 'good' and 'bad' states each occur with equal probability $(1 - 2p)/2 = 0.49$.

The level of risk aversion, subjective discount factor and probabilities were chosen to deliver a riskfree rate around 1% and an equity premium around 6%. The growth rates in the 'good' and 'bad' state were chosen to approximately match the growth rate and standard deviation of log consumption in the US.

By assuming complete markets, we get a unique state price density that we can use to price any set of payoffs, allowing us to isolate the differences in payoffs in the crash and boom states without imposing specific distributional assumptions upon the assets. The unique stochastic discount factor is given by:

$$\beta \frac{u'(c_{t+1})}{u'(c_{t+1})} = 0.98 \left(\frac{c_{t+1}}{c_t} \right)^{-5}. \tag{1}$$

Since we have a representative investor, all of the endowment must be consumed so the state prices are then:

$$
\begin{aligned}
SP\,(Boom) &= 0.01 \times 0.98 \times 1.5^{-4} \\
SP\,(Good) &= 0.49 \times 0.98 \times 1.07^{-4} \\
SP\,(Good) &= 0.49 \times 0.98 \times 0.985^{-4} \\
SP\,(Crash) &= 0.01 \times 0.98 \times 0.6^{-4}.
\end{aligned}
\tag{2}
$$

The riskfree rate is then given by $1/(\sum SP) = 1.06\%$. The endowment can also be priced, giving a expected equity return of 7.75%, which yields an equity premium of 6.49%.

Crash risk premia can be understood by considering two hypothetical stocks, one risky and one safe. Each stock pays out 10% of whatever the endowment turns out to be in the 'boom', 'good' and 'bad' states. In the 'crash' state, the safe stock is worth 15% of the 'crash' endowment while the risky stock is worth only 5%. Since we have the state prices and we know each stocks payoff, we can calculate expected returns. The risky stock commands 11.86%, while the safe stock commands only 3.94%, a premium of 7.92% generated from differential exposure to a possible market crash.

Declining marginal utility does make generating a boom premium a little more difficult. We again examine two hypothetical stocks, one risky and one safe. Each stock pays 10% of endowments in the 'good', 'bad' and 'crash' states. In the 'boom' state, the safe stock pays 1% of the endowment, while the risky stock is worth 20%. The expected return on the risky stock is 9.77% while the expected return on the safe stock is 6.52% for a more moderate spread of 3.24%.

The model is a very simple one and certainly not a complete description of reality. Yet it does demonstrate that differing exposures during crashes can generate a large cross-sectional risk premia within a pretty standard model. The model is not as successful in generating a large enough boom premia, suggesting that some additional extension to the standard model will be required explain the premium we identify in Section 3.

2 Extreme Value Theory

Since the behavior of a stock during extreme market movements can be an important driver of the expected return of that stock, we need an econometric apparatus to capture that behavior. Extreme value theory is custom built to understand the extremes of distributions. Rather than examine the whole distribution, in extreme value theory just the largest or smallest observations are examined so that 'the tails may speak for themselves'.[1]

The univariate techniques of extreme value theory are designed to understand the behavior of the tail of a distribution. Since this determines the relative frequency of rare events, these techniques are a natural technology to employ to gain insight into the return distribution. Jansen and de Vries (1991) and Longin (1996) examine the returns of US stock indices and find that the extreme returns are completely consistent with the extreme value distributions. Gabaix, Gopikrishnan, Plerou, and Stanley (2003) provide a model in which this behavior is generated by the trades of large participants.

Our focus is on the dependence of individual stocks upon the market and so we will use bivariate extreme value theory. In contrast to the univariate theory, the bivariate theory is specifically concerned with measuring the likelihood that the extremes of two variables coincide. These techniques are newer to the financial literature, although

[1]One possible alternative is to use correlations conditioned upon extreme market movements. However Boyer, Gibson, and Loretan (1999) show that the very act of conditioning can cause the conditional correlation to differ from the correlation measured over the full distribution, making inference difficult.

Poon, Rockinger, and Tawn (2004) have explored their application to the dependence among international equity indices using the same techniques as we use. Longin and Solnik (2001) also use extreme value techniques to examine international indices, although their approach is more parametric.

The central result, due to Ledford and Tawn (1996), is that just as rescaling sums of random variables delivers the Central Limit Theorem, rescaling the joint extremes of bivariate distributions also delivers a convenient limiting representation.

To obtain the limiting representation, the marginal distributions X and Y are both transformed to a Fréchet distribution with tail index one. These marginal distributions both satisfy

$$Pr\left(Z > z\right) \sim z^{-1} \quad \text{as } z \rightarrow \infty. \tag{3}$$

and have a fat right tail. In order to obtain a limiting representation of the dependence, we construct a new series by taking the lesser observation from each observation pair, ie. $T \equiv \min\left(X, Y\right)$. Ledford and Tawn (1996) show that this new variable follows:

$$Pr\left(T > z\right) = Pr\left(X > z, Y > z\right) \sim \mathcal{L}\left(z\right) z^{-\frac{1}{\eta}} \quad \text{as } z \rightarrow \infty. \tag{4}$$

where $\frac{1}{\eta} \in (0, 1]$ and $\mathcal{L}\left(z\right)$ is a slowly varying function that essentially behaves like a constant as z becomes large[2]. The parameter η is the tail index of the transformed series[3] which defines the joint tail behavior of the original variables. As the transformed series is univariate, η can be estimated using standard techniques from

[2]The precise definition is that $\mathcal{L}\left(zt\right)/\mathcal{L}\left(z\right) \rightarrow 1$ as $z \rightarrow \infty$ for any fixed $t > 0$

[3]The tail index of a distribution defines the tail behavior of a univariate distribution. Fat-tailed distributions have positive tail indices, light-tailed distributions such as the normal have tail indices of zero while distributions with a finite endpoint have negative tail indices. See Coles (2001).

univariate extreme value theory. Combining (4) and (3) we have:

$$Pr\left(X > z | Y > z\right) \sim \mathcal{L}\left(z\right) z^{\frac{\eta-1}{\eta}} \quad \text{as } z \to \infty. \tag{5}$$

Equation 5 is a fundamental characterization of tail dependence. It provides the correct approach for extrapolation, allowing insight into the joint behavior at previously unobserved levels. Such a feature is clearly advantageous in a financial setting.

Equation 5 also demonstrates that there are two distinct classes of tail dependence, delineated by the parameter η. If $\eta = 1$ and $\mathcal{L}\left(z\right) \to 0$, then the limiting probability (5) is non-zero so the variables are said to be asymptotically dependent, with the strength of that dependence determined by $\mathcal{L}\left(z\right)$. Alternatively, the limiting probability (5) may equal to zero if $\eta < 1$ or $\mathcal{L}\left(z\right) \to 0$. This situation is asymptotic independence, since conditioning on the extremes of one variable provides no information about the extremes of the other variable[4].

However, asymptotic independence does not imply that the probability (5) is zero for large z, only that the probability is zero for $z = \infty$. For large but finite z and $\mathcal{L}\left(z\right) \to 0$, there can still be sub-asymptotic dependence and understanding joint tail behavior can still be informative. Asymptotic independence implies that the *most* extreme observations will not occur together, large values can still occur together. In such cases, if $\eta > \frac{1}{2}$ there is positive sub-asymptotic dependence in the sense that large observations are positively associated and if $\eta < \frac{1}{2}$ there is negative sub-asymptotic dependence. Sub-asymptotic independence occurs if $\eta = \frac{1}{2}$.

Bivariate normal distributions with correlation $|\rho| < 1$ are the classic example of

[4]The unconditional probability of outcomes above a threshold for a single variable also limits to zero as the threshold increases, so the conditional and unconditional probabilities are the same and hence asymptotic independence.

asymptotically independent variables. Sibuya (1960) proves that the limiting probability in equation (5) is zero, since the probability (5) decays to zero as the threshold z is increased. Clearly, for positive ρ large values will coincide, but the very largest observations will not coincide. At any finite level, however, the probability is greater than zero but it still decaying towards zero. For these cases, the rate of decay is the important driver of tail dependence. For bivariate normal distributions, this sub-asymptotic dependence is given by $2\eta - 1 = \rho$, which provides a convenient benchmark for interpreting η.

In fact, we follow Coles, Heffernan, and Tawn (1999) and define $\bar{\chi} \equiv 2\eta - 1$, which can then be interpreted in the same fashion as the standard correlation coefficient of a bivariate normal distribution. Also define $\chi \equiv \lim_{z\to\infty} \mathcal{L}(z)$. Together, the pair $(\chi, \bar{\chi})$ fully describe the extremal dependence structure. If $\bar{\chi} < 1$ the variables are asymptotically independent with any sub-asymptotic dependence quantified by $\bar{\chi}$. On the other hand, if $\bar{\chi} = 1$ the variables are asymptotically dependent with the strength of that dependence measured by χ.

An example of a distribution with asymptotically dependence is the bivariate Student t distribution with $|\rho| < 1$ and with each marginal distribution having common degrees of freedom, $\nu < \infty$. In this case, $\eta = \bar{\chi} = 1$, and Demarta and McNeil (2005) show that

$$\lim_{s\to\infty} \mathcal{L}(s) = 2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right), \tag{6}$$

which is always greater than zero, implying asymptotic dependence for zero and even negative correlations. The example demonstrates the generality of the tail dependence approach, since the tail dependence is still well defined when $\nu = 1$ although none of the moments are well defined.

Correctly determining the form of extremal dependence is crucial since it guides model selection when using parametric models or the choice of particular estimator in the non-parametric approach. Bruun and Tawn (1998) present evidence that correctly identifying the form of extremal dependence is more important than choosing the correct parametric model, albeit in the setting of predicting coastal flooding. For the purposes of assessing financial risk, the distinction between classes can also have a great impact. If one assumes asymptotic independence, as is implicit in any model which assumes multivariate normality of returns, and the true structure is asymptotic dependence then the expected drop in the portfolio value in a crash will be too conservative.

In Appendix A we demonstrate that the dependence structures of stocks upon the market is one of asymptotic dependence and so we narrow our focus to estimating χ.

2.1 Estimation

Estimation of asymptotic tail dependence is remarkably simple. Let R(X) rank the observations of X from 0 to 1 and I(.) be the indicator function. To estimate upper tail asymptotic dependence, denoted χ_{BOOM}, we choose a high threshold above which we expect equation 5 to hold. Let k denote the number of joint observations for which the market return exceeds the threshold out of n total observations. The estimator is then:

$$\chi_{\widehat{BOOM}} = \frac{\sum_{i=1}^{n} I\left(R\left(Mkt\right) > 1 - \frac{k}{n}, R\left(Stk\right) > 1 - \frac{k}{n}\right)}{k/n}, \tag{7}$$

which is simply the fraction of observations for which both variables exceed the threshold, divided by the fraction of market observation is above the threshold[5].

[5] In fact, since the observations have all been transformed to ranks, we could divide through by the fraction of stock observations exceeding the threshold as both are just $1 - \frac{k}{n}$. Dividing through by the fraction of the market observation above the threshold is easier to interpret.

Lower tail asymptotic dependence, denoted χ_{CRASH} is estimated similarly, although now the threshold is very low (but still results in examining k joint observations). The estimator is then:

$$\widehat{\chi_{CRASH}} = \frac{\sum_{i=1}^{n} I\left(R\left(Mkt\right) < \frac{k}{n}, R\left(Stk\right) < \frac{k}{n}\right)}{k/n}. \tag{8}$$

The choice of threshold is a difficult issue. Setting a low threshold will increase the number of observations exceeding that threshold and the estimators will have greater precision. However, bias will also increase, since equation (5) will be a less accurate representation of the distribution at less extreme levels. On the other hand, setting a very high threshold will reduce the bias, but the resulting estimates will be based on very few observations and will lack precision.

For identification purposes, we will consider a range of thresholds. The choice of threshold is also discussed further in Appendix A.

2.2 Full sample tail dependence estimates

Applications of extreme value theory naturally require large samples, so it is natural to employ daily returns. All of the stock return data used is drawn from the CRSP daily tape from July 1962 through December 2004. All stocks considered are common stocks (CRSP code 10 or 11). The value-weighted CRSP market index is used as market proxy.

There are well-known difficulties with employing daily data. In particular, the observed daily returns are tainted by bid-ask bounce and stale prices.

Bid-ask bounce is caused by non-zero returns induced by trades occurring at the bid and then at the ask, when the fundamental price is unchanged. This problem is not of great concern for extreme value analysis deals with only the largest returns

Table I

Stale prices and the October 19, 1987 Crash

NYSE, Amex and Nasdaq stocks are sorted into quintile portfolios according to market capitalization at the end of the preceding June using breakpoints constructed from the NYSE stocks. Each portfolio is value weighted. The market portfolio is the value-weighted CRSP index.

	Market	Small	2	3	4	Large
Oct. 19, 1987	-17.14	-10.60	-11.79	-12.77	-15.07	-19.26
Oct. 20, 1987	0.41	-9.15	-8.77	-6.73	-4.68	4.99

which will almost surely not be driven by the bid-ask spread.

A more important issue is the problem of stale prices. Returns are calculated using the last transaction price of the day, which for less liquid stocks may have occurred much earlier during the day. Consider a hypothetical stock that trades once a day at noon. If there is a substantial market movement in the afternoon, that movement will not show up in the stock's return until the next day when the stock trades again. A careless econometrician examining the daily returns may conclude that the dependence is much weaker than it is.

The problem of stale prices is especially apparent on a day of particular interest here - October 19th, 1987 aka “Black Monday”. On that day, the value-weighted CRSP market index fell 17.14%. Table I documents the returns for five size-based portfolios on the day of the crash and the next day. Size-based portfolios are informative here since small stocks trade less frequently and the problem of stale prices is more prevalent. The smaller stocks fell by less on the day of the crash and were still falling the next day, consistent with the prices or bid-ask averages used to calculate the Monday return for the smaller stocks not fully reflecting the events of the day.

Dealing with stale prices is difficult, and perhaps even more so in the current context where we are only examining the largest returns. Our approach will be to be cognisant of the problem and exploit the fact that the problem is more severe among small stocks (Lo and MacKinlay (1990)).

Table II presents the cross-sectional averages of crash and boom dependence and the stale price problem is particularly apparent. The averages are presented for various thresholds and with stocks grouped by the length of their daily return history[6]. The averages in Table II vary by sample size and threshold, with longer histories and higher thresholds having higher averages. The variation with the threshold is the effect of bias caused by higher order terms in the slowly varying function. The averages using fewer observations will have the least bias and since we are averaging across many stocks, these will still be precise numbers. The lower averages for the shorter histories are attributed to these being generally smaller firms where stale prices are more likely to attenuate the dependence.

While there is some variation associated with the differing thresholds, the probability that the worst crash of the market and stock coincide averages around one in five. The chance that the biggest boom of the market and of the stock coincide averages around one in six.

Such asymmetry between the crash dependence and the boom dependence is important since it precludes the CAPM from holding under any acceptable utility specification. Ingersoll (1987) proves that the mean-variance analysis is optimal[7] (and hence the CAPM holds) only if the multivariate return distribution is a member of the elliptical class, which are distributions with ellipsoidal probability contours. Since ellipsoidal contours are inconsistent with differing crash and boom dependences, market beta is an incomplete measure of risk and the CAPM will not hold.

[6]Grouping by other characteristics such as size is difficult since these will vary over the life of the firm and any breakpoints used will also vary through time.

[7]The utility assumptions required are that higher mean return is preferred and return dispersion is disliked.

Table II

Average Boom and Crash Dependence for Individual Stocks

All NYSE, Amex and Nasdaq stocks with at least one thousand valid daily returns included. Estimation of χ is done at a number of different thresholds, where k is the number of joint observations used. Reported numbers are cross-sectional averages of the estimated χ.

Crash Dependence

Obs.	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k/n=1\%$	$k/n=2.5\%$	$k/n=5\%$	$k/n=10\%$
1,000-2,499	0.08	0.10	0.12	0.13	0.13	0.18	0.25	0.30	0.36
2,500-4,999	0.09	0.11	0.12	0.13	0.13	0.21	0.27	0.32	0.38
5,000-7,499	0.18	0.18	0.19	0.20	0.20	0.26	0.32	0.36	0.40
7,500-9,999	0.20	0.21	0.21	0.22	0.21	0.27	0.32	0.36	0.41
Over 10,000	0.22	0.20	0.21	0.21	0.20	0.27	0.32	0.37	0.41

Boom Dependence

Obs.	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k/n=1\%$	$k/n=2.5\%$	$k/n=5\%$	$k/n=10\%$
1,000-2,499	0.07	0.09	0.11	0.12	0.12	0.07	0.17	0.23	0.28
2,500-4,999	0.07	0.10	0.11	0.12	0.13	0.11	0.20	0.25	0.30
5,000-7,499	0.10	0.12	0.13	0.15	0.15	0.16	0.24	0.29	0.34
7,500-9,999	0.11	0.13	0.14	0.14	0.15	0.17	0.25	0.30	0.35
Over 10,000	0.10	0.12	0.14	0.15	0.15	0.17	0.26	0.31	0.35

3 Tail Dependence and the Cross-Section of Expected Returns

Stocks with higher tail dependence are riskier stocks. State prices are high when the market crashes, so investors will charge higher expected returns to hold stocks whose values collapse during crashes. Likewise, state prices are low during booms, so stocks that have their best payoffs during such periods also command higher expected returns. In this section we show that these relationships are borne out in the data.

Since any informative test for pricing implications will use past information to predict the future, crash dependence χ_{CRASH} and boom dependence χ_{BOOM} are each estimated from the preceding five years of daily data. The CRSP daily tape begins in July of 1962, so the first month that we will have an estimate of the cross section of tail dependences is July of 1967, while the last month is December 2004 for a total of 450 cross-sections. All NYSE, Amex and Nasdaq common stocks are considered each month. Any stocks with invalid returns comprising more than half of the estimation period (around two and a half years) or missing a price on the last day of the estimation window are excluded from that month. Also, a stock must have sufficient Compustat data to calculate the book-to-market ratio and stocks with negative ratios are dropped[8].

The criteria determining the choice of threshold in the pricing context are somewhat different from those used when characterizing tail dependence. While there will be bias in any estimate of χ, this common bias in the cross-section is far more benign since we are concerned now with determining which stocks have greater tail dependence. As such, any common bias will simply wash out in any comparison. Precision

[8]Note that since we require the stock to have been on the CRSP tape for five years, discarding the first two years of book-to-market data to avoid the bias associated with backfilling is unnecessary.

becomes important since sorting noisy estimates will capture more measurement error than underlying truth, suggesting that lower thresholds may be more acceptable. Additionally, the threshold needs to be low enough so that there are sufficiently many observations that the resulting set of estimates has a decent amount of cross-sectional variation. In light of these issues, the threshold for both stock and market are set at 5%, so that the χ measures pick up the number of observations that exceed *both* the 95th percentile of the stock distribution and the 95th percentile of the market distribution.

3.1 Do historical estimates predict future behavior?

Before considering whether tail dependences are related to expected returns, we must first check whether the historical estimates contain any information about subsequent returns during market crashes or booms. In particular, we examine the cross-section of realized returns for the five largest daily declines in the value-weighted CRSP market index as well as the five largest daily increases. The sample for each date consists of all stocks that traded on the day of the particular crash[9] and also fulfilled the estimation requirements. For each date, the cross-section of individual returns are regressed upon either χ_{BOOM} or χ_{CRASH} estimated from the five years ending the month preceding the crash.

The results for the crashes are in Panel A of Table III. The univariate regressions demonstrate that the historical tail crash dependences are highly informative about future crash behavior, with coefficients in all five univariate regressions being significantly negative. The $\hat{R}^2$ for the October 19th, 1987 crash is noticeably large at 0.140, although χ_{CRASH} has far less explanatory power in the other cases.

[9] Stocks that did not trade are identified by CRSP recording the stock price as negative and the recorded volume being zero or missing.

Table III

Predicting Stock Returns During Extreme Market Movements

Each date is one of the five largest negative or largest positive daily returns for the CRSP value-weighted market index. Joint crash probabilities, χ_{CRASH}, and joint boom probabilities, χ_{BOOM} are estimated from five years of daily data ending the month preceding the crash. Market betas are estimated using monthly data over the same period. Size is the logarithm of equity capitalization at the end of the preceding month. The dependent variable is the return for stocks that traded on the particular day. Each row reports results for a regression of returns upon just dependence and results from another regression with additional controls. The t-statistics in parenthesis are heteroskedasticity consistent.

Panel A. Days With Largest Decrease in CRSP Market Index

Date	Mkt/Obs	Int	χ_{CRASH}	$\hat{R}^2$	Int	χ_{CRASH}	Beta	Size	$\hat{R}^2$
Oct 19,	-17.14	-5.935	-40.138	0.140	6.599	-14.743	-3.674	-1.058	0.221
1987	2,858	(-18.83)	(-21.78)		(6.38)	(-5.62)	(-12.71)	(-10.49)	
Oct 26,	-8.26	-7.381	-6.440	0.004	-4.203	-1.407	-2.788	-0.081	0.050
1987	2,839	(-23.30)	(-4.01)		(-3.71)	(-0.56)	(-10.43)	(-0.76)	
Apr 14,	-6.63	-4.944	-5.240	0.003	-3.961	1.564	-3.542	0.091	0.101
2000	3,621	(-19.03)	(-4.18)		(-5.11)	(0.82)	(-13.74)	(1.23)	
Aug 31,	-6.59	-4.784	-5.405	0.004	-4.828	-5.019	-1.571	0.115	0.032
1998	3,700	(-22.36)	(-4.75)		(-5.36)	(-2.64)	(-8.29)	(1.40)	
Oct 27,	-6.53	-4.388	-8.502	0.014	-0.543	-0.678	-1.804	-0.248	0.080
1997	3,670	(-25.51)	(-8.29)		(-0.67)	(-0.44)	(-9.51)	(-3.57)	

Panel B. Days With Largest Increase in CRSP Market Index

Date	Mkt/Obs	Int	χ_{BOOM}	$\hat{R}^2$	Int	χ_{BOOM}	Beta	Size	$\hat{R}^2$
Oct 21,	8.66	5.141	17.992	0.016	-1.518	7.373	3.693	0.361	0.049
1987	2,611	(10.63)	(7.35)		(-0.80)	(1.80)	(8.32)	(1.95)	
Jul 29,	5.32	1.350	17.886	0.032	-4.436	4.564	0.661	0.576	0.051
2002	3,462	(4.33)	(10.66)		(-4.29)	(2.20)	(3.66)	(6.67)	
May 27,	5.30	3.356	19.041	0.032	0.224	7.454	4.999	-0.122	0.167
1970	1,440	(5.72)	(7.14)		(0.15)	(2.44)	(11.41)	(-1.04)	
Jan 3,	5.29	1.249	21.438	0.030	4.928	22.657	3.787	-0.598	0.148
2001	3,430	(3.56)	(10.53)		(4.21)	(7.69)	(12.43)	(-5.57)	
Jul 24,	5.23	0.267	16.489	0.026	-9.046	-2.542	0.351	0.951	0.077
2002	3,481	(0.94)	(10.76)		(-9.34)	(-1.14)	(1.87)	(11.08)	

In order to test whether the crash dependences contain information over and above existing measures, we also control for beta and size[10] in a second set of regressions. Beta is included as an obvious measure of overall dependence while size is included in order to control for effect of small stocks lagging larger stocks (Lo and MacKinlay (1990)). For four of the five crashes, the coefficient remains negative and significantly so for two of the cases.

Panel B of Table III repeats the exercise for the five largest increases. In the univariate regressions, the boom dependences are significantly positive in every case. When we control for beta and size, three of the coefficients on boom dependence are significantly positive and there is one case with a negative point estimate.

These results indicate that the historical estimates are informative about the future returns during large market movements of stocks, but perhaps not in every case. However, the regressions in Table III are not exact tests of the predictive power of the estimated tail dependences, although they are the more economically interesting tests. Recall that the tail dependence measures the probability that the *worst* returns of the stock and the market will coincide. The worst return for a particular stock may not be large relative to the returns of other stocks, so a more direct test of the predictive abilities of the tail dependence estimates is to take the return of each stock standardized by the daily standard deviation over returns over the preceding five years as the dependent variable. Table IV repeats Table III with the new dependent variable and correlation taking the place of beta.

For both booms and crashes, the results are stronger with the all of the coefficients in the univariate regressions having the correct sign and highly significant. In the regressions controlled for correlation and size, each of the χ_{CRASH} coefficients are

[10]Size is not being used here as a variable related to expected returns. Tables III and IV are presented as evidence that the estimated tail dependences are capturing the intended behavior and that such behavior can be predicted.

Table IV

Predicting Standardized Stock Returns During Extreme Market Movements

Each date is one of the five largest negative or largest positive daily returns for the CRSP value-weighted market index. Joint crash probabilities, χ_{CRASH}, and joint boom probabilities, χ_{BOOM} are estimated from five years of daily data ending the month preceding the crash. Market correlations are estimated using monthly data over the same period. Size is the logarithm of equity capitalization at the end of the preceding month. The dependent variable is the return for stocks that traded on the particular day divided by standard deviation of daily returns over the preceding five years. The t-statistics in parenthesis are heteroskedasticity consistent.

Panel A. Days With Largest Decrease in CRSP Market Index

Date	Mkt/Obs	Int	χ_{CRASH}	$\hat{R}^2$	Int	χ_{CRASH}	Corr	Size	$\hat{R}^2$
Oct 19,	-17.14	-0.705	-30.348	0.35	7.851	-10.028	-1.416	-0.941	0.46
1987	2,858	(-5.78)	(-35.23)		(21.45)	(-8.68)	(-3.37)	(-22.50)	
Oct 26,	-8.26	-1.885	-10.126	0.09	1.300	-2.050	-1.180	-0.335	0.13
1987	2,839	(-18.49)	(-17.69)		(4.13)	(-2.19)	(-3.44)	(-9.74)	
Apr 14,	-6.63	-0.452	-6.704	0.13	0.782	-4.334	0.172	-0.133	0.14
2000	3,621	(-9.15)	(-24.66)		(4.78)	(-9.77)	(0.87)	(-7.84)	
Aug 31,	-6.59	-0.696	-8.689	0.12	1.229	-4.477	-0.303	-0.196	0.15
1998	3,700	(-13.19)	(-23.94)		(5.96)	(-7.88)	(-1.23)	(-9.27)	
Oct 27,	-6.53	-0.556	-11.487	0.26	2.667	-4.082	-0.857	-0.313	0.34
1997	3,670	(-13.42)	(-34.60)		(15.71)	(-8.57)	(-4.57)	(-18.23)	

Panel B. Days With Largest Increase in CRSP Market Index

Date	Mkt/Obs	Int	χ_{BOOM}	$\hat{R}^2$	Int	χ_{BOOM}	Corr	Size	$\hat{R}^2$
Oct 21,	8.66	0.973	15.468	0.09	-3.868	5.475	0.576	0.514	0.12
1987	2,611	(6.26)	(16.90)		(-6.97)	(3.73)	(1.04)	(8.62)	
Jul 29,	5.32	0.220	6.618	0.08	-2.372	1.335	0.257	0.266	0.17
2002	3,462	(3.44)	(17.80)		(-13.60)	(2.51)	(1.41)	(15.75)	
May 27,	5.30	1.098	8.661	0.06	-0.228	7.319	1.103	0.092	0.07
1970	1,440	(6.12)	(10.21)		(-0.53)	(7.83)	(1.72)	(2.31)	
Jan 3,	5.29	-0.028	7.164	0.07	0.891	6.754	3.077	-0.137	0.13
2001	3,430	(-0.40)	(15.92)		(4.31)	(10.95)	(11.05)	(-6.61)	
Jul 24,	5.23	0.075	5.634	0.05	-3.561	-0.090	-1.022	0.383	0.20
2002	3,481	(1.07)	(14.12)		(-19.55)	(-0.15)	(-4.76)	(20.65)	

significantly negative, while the χ_{BOOM} coefficient is significantly positive in four cases, but negative and insignificant in the last case.

These results are encouraging, especially considering the fact that the tail dependences are generated regressors and these regressions each considered the cross-section for just a solitary day.

3.2 Do historical estimates predict future expected returns?

Having determined that the historical tail dependences are able to forecast stock returns during large market moves, we now address the real question - are tail dependences related to future expected returns? The expected relationship for the crash dependence is clear from standard state dependent utility arguments, stocks that crash with the market should have higher expected returns since these stocks fall in value just when marginal utility is highest. The expected relationship for boom dependence also follows from utility intuition, although the intuition is perhaps a little less immediate. Stocks that deliver their best payoffs when the market booms do so when marginal utility is lowest and so should also command higher expected returns.

Table V documents Fama-Macbeth tests for return premia carried out on a monthly frequency. Each month, the cross-section of realized returns is regressed upon the variables we hypothesize to determine expected returns as well as variables suggested by other research. Alongside the historical tail dependences, the variables used in the test are historical market beta estimated using five years of monthly data, size which is the logarithm of equity market capitalization at previous month end, lagged book-to-market ratio as defined by Fama and French (1992) and lagged return which is the stocks' return over the previous year excluding the last month. Each of the variables is public information at the end of the month prior to the return realization.

The basis of the Fama-Macbeth test is then to examine whether the mean of the time series of coefficients for any variable is significantly positive.

In the presence of log(size) and beta, the coefficients on χ_{CRASH} and χ_{BOOM} take their expected sign and are significant, but without size and beta the coefficients take the wrong sign. There are solid economic reasons why this would be the case.

Small stocks trade less frequently and lag the returns of larger stocks, which will attenuate the observed dependence structure. The dependence measures are strongly positively related to size - large stocks have higher average tail dependences for both crashes and booms. The time series average $\hat{R}^2$ from regressing χ_{BOOM} upon log(size) is 0.46 with a minimum of 0.22 and the coefficient is always positive. Similarly the average $\hat{R}^2$ for χ_{CRASH} is 0.41, although the minimum is only 0.07 and the coefficient is again always positive. Beta allows us to control for the overall dependence structure.

Adding lagged book-to-market ratio and historical return do not affect the tail dependence results.

One difficulty with this approach is that many of the variables in the regression are measured with differing scales, leaving the Fama-Macbeth coefficients difficult to interpret. Additionally, the amount of cross-sectional variation in the monthly returns itself varies considerably from month to month, so the magnitude of coefficients will have different return implications in different months. Economically, we are interested in the spread in return which can be attributed to the spread in the variable of interest, in this case the tail dependences. As such, Table V also includes the mean of the implied return series for χ_{CRASH} and χ_{BOOM}. These are calculated by taking the fitted return value for the 95th empirical percentile and subtracting the fitted return value for the 5th percentile each month, holding all other variables constant.

The implied returns suggest that the returns to bearing tail dependence are not just statistically significant, but also economically large. For boom dependence risk

Table V

Cross-Sectional Tests for Risk Premia

Tests are Fama-Macbeth tests conducted on monthly data from July 1967 until December 2004. Joint crash and joint boom probabilities, χ_{CRASH} and χ_{BOOM} are estimated from five preceding years of daily data. Market betas are estimated using monthly data over the same period, winsorised at the 99th percentile. Size is the logarithm of equity capitalization at the end of the previous month. B/M is the book-to-market ratio calculated according to Fama and French (1992) and values greater than the 99th percentile are set to the 99th percentile. PastRet is the return over the previous year except for the last month. The estimated premia are the annualized average of each month's return premia obtained by taking the fitted value for the 95th empirical percentile and subtracting the fitted value for the 5th percentile. Newey-West t-statistics in parenthesis use twelve lags.

Fama-Macbeth Premia						Return Premia	
χ_{BOOM}	χ_{CRASH}	Beta	Size	B/M	PastRet	χ_{BOOM}	χ_{CRASH}
-2.62						-8.60	
(-3.19)							
	-2.71						-9.30
	(-3.96)						
-1.16	-1.93					-4.17	-6.62
(-1.69)	(-3.35)						
2.23	1.29	-0.16	-0.29			6.31	3.99
(3.73)	(2.51)	(-1.13)	(-4.65)				
2.14	1.34	-0.14	-0.26	0.20	0.06	6.08	4.24
(3.63)	(2.66)	(-1.05)	(-4.25)	(3.34)	(0.42)		
July 1967 - March 1986							
0.19	0.10	-0.25	-0.16	0.27	-0.01	0.40	-0.14
(0.35)	(0.15)	(-1.45)	(-1.86)	(2.77)	(-0.04)		
April 1986 - December 2004							
4.10	2.57	-0.02	-0.36	0.12	0.12	12.09	8.80
(5.05)	(4.09)	(-0.12)	(-4.56)	(2.04)	(0.89)		

the returns are 6% per annum, while the crash strategy returns over 4%. It is surprising that the boom dependence has a larger premium.

A further concern is that the crash and boom dependences are not that different for each stock. The time series average correlation between the two measures is high at 0.69, however the correlation between the Fama-Macbeth coefficients is not significantly different from zero in the full sample or in either of the subsamples. Similarly, the implied returns are also uncorrelated, suggesting that the two measures are capturing different effects.

These results are not robust to splitting the sample period. Since there are 450 months in total, the sample is split into two equal subsamples of 225 months. Tail premia, as measured by either the Fama-Macbeth coefficients or the implied returns, are insignificant in the first subsample but highly significant in the second subsample. This is a result that we will continue to see throughout the pricing analysis and is examined in further detail in Section 4.

Both χ_{BOOM} and χ_{CRASH} are correlated with size and, to a lesser extent, beta. One possible explanation for the results in Table V is that χ_{BOOM} and χ_{CRASH} are significant because of this correlation. To demonstrate that this is not so, we remove the effect of size and beta from χ_{BOOM} and χ_{CRASH}. We define χ^*_{BOOM} as the residuals from a regression of χ_{BOOM} upon historical beta and log(size) and define χ^*_{CRASH} similarly[11]. Stocks can then be sorted into decile portfolios according to χ^*_{BOOM} and also according to χ^*_{CRASH}.

Since this is the first cross-sectional study of tail dependence in the equity market, we first examine the industry composition of each of the high and low dependence decile portfolios. To determine if an industry is over- or under-represented in a portfo-

[11] The variables could also be orthogonalized to book-to-market ratios and lagged return, but the economic rationale for doing so is not clear. In any case, the results are not changed from doing so although the alphas for the factors are slightly lower, but still highly significant.

lio in a particular month, the proportion of stocks in an industry within that portfolio is divided by the proportion of all candidate stocks in that industry. Thus, a number greater than one indicates an industry that is over-represented in that portfolio. The numbers reported in Table VI are time-series averages of these relative weightings.

The portfolio industry composition fit well with intuition. Consumer non-durables, telecommunications and utilities are more represented in the low dependence portfolios for both crash and boom dependence. On the other hand, consumer durables and technology stocks are more likely to be assigned to the high dependence portfolios. The manufacturing industry has more dispersion in tail dependence over the full sample since these stocks are over-represented in both the high and low dependence portfolios. However, manufacturing firms are more likely to be assigned to the high dependence portfolios in the second half of the sample. Energy stocks are over-represented in both the low dependence portfolios, while they are heavily over-represented in the high dependence stocks in the first half of the sample and heavily under-represented in the second half of the sample. Stocks in the health and shops industries are generally under-represented in all portfolios, although health stocks are over-represented in the high crash dependence and low boom dependence stocks in the second half of the sample.

Returns for the tail dependence portfolios are documented in Table VII. The portfolios constructed on the basis of boom dependence show near monotonic increase in average return from 1.14% per month to 1.64% per month as we move from low to high over the full sample. The difference of 0.50% per month is statistically significant. Once again, the results are concentrated in the second half of the sample where the spread is 0.93% per month. Results are very similar for the sorts based on crash dependence and mirror the magnitude of the return premia inferred from the Fama-MacBeth test.

Table VI

Industry Composition of Tail Dependence Portfolios

Each month, the proportion of stocks from an industry grouping within a decile portfolio is divided by the proportion of stocks in that grouping amongst all the candidate stocks. Reported numbers are time-series averages with numbers greater then one indicating an industry which is over-represented in that portfolio. Industry groupings are based on CRSP SIC codes and industry definitions are taken from Kenneth French's website.

Panel A. Crash Portfolios

	Low Dependence			High Dependence		
Industry	Full Sample	7/1967 -3/1986	4/1986 -12/2004	Full Sample	7/1967 -3/1986	4/1986 -12/2004
Cons. NonDur.	1.19	1.43	0.95	0.95	0.82	1.09
Cons. Dur.	0.89	1.08	0.70	1.07	1.11	1.03
Manufacturing	1.04	1.25	0.82	1.42	1.61	1.22
Energy	1.42	1.05	1.80	1.52	2.52	0.51
Technology	0.63	0.50	0.76	1.04	1.06	1.02
Telecom.	1.18	1.07	1.28	0.80	0.50	1.09
Shops	0.89	0.95	0.82	0.83	0.77	0.90
Health	0.67	0.37	0.97	0.93	0.81	1.06
Utilities	2.56	3.32	1.80	0.72	0.59	0.84
Miscellaneous	0.97	0.84	1.11	0.83	0.70	0.96

Panel B. Boom Portfolios

	Low Dependence			High Dependence		
Industry	Full Sample	7/1967 -3/1986	4/1986 -12/2004	Full Sample	7/1967 -3/1986	4/1986 -12/2004
Cons. NonDur.	1.26	1.52	1.01	0.92	0.90	0.94
Cons. Dur.	0.86	1.09	0.63	1.43	1.66	1.20
Manufacturing	1.06	1.25	0.87	1.41	1.63	1.19
Energy	1.88	1.74	2.03	0.86	1.29	0.43
Technology	0.55	0.41	0.69	1.15	1.19	1.11
Telecom.	1.33	1.05	1.61	0.74	0.70	0.78
Shops	0.88	1.00	0.76	0.82	0.68	0.96
Health	0.81	0.46	1.16	0.78	0.88	0.67
Utilities	1.77	1.81	1.73	0.97	0.93	1.01
Miscellaneous	1.00	0.93	1.06	0.83	0.65	1.02

Table VII

Tail Dependence Decile Portfolios Returns

Stocks are sorted based upon the residuals of the tail dependence measure regressed upon the historical logarithm of equity capitalization and beta, denoted χ^*_{CRASH} for crash dependence and χ^*_{BOOM} for boom dependence. Rows labeled Jan and Non-Jan report the average High-Low spread for January and Non-January months, respectively. Reported numbers are arithmetic average monthly returns.

	χ^*_{BOOM} sorts			χ^*_{CRASH} sorts		
Decile	Full Sample	7/1967 -3/1986	4/1986 -12/2004	Full Sample	7/1967 -3/1986	4/1986 -12/2004
Low	1.140	1.250	1.048	1.130	1.189	1.071
2.	1.226	1.374	1.096	1.251	1.351	1.151
3.	1.283	1.366	1.219	1.371	1.512	1.230
4.	1.395	1.546	1.259	1.350	1.512	1.189
5.	1.398	1.489	1.322	1.412	1.462	1.361
6.	1.319	1.365	1.291	1.436	1.484	1.387
7.	1.508	1.423	1.605	1.518	1.457	1.579
8.	1.558	1.508	1.626	1.586	1.521	1.650
9.	1.678	1.452	1.919	1.514	1.352	1.676
High	1.643	1.318	1.981	1.576	1.247	1.906
High-Low	0.503	0.069	0.932	0.447	0.058	0.836
	(4.057)	(0.496)	(4.633)	(4.040)	(0.412)	(4.982)
Jan	3.826	1.973	5.781	2.050	0.224	3.977
	(5.380)	(2.337)	(5.880)	(3.298)	(0.261)	(6.036)
Non-Jan	0.205	-0.107	0.516	0.303	0.042	0.562
	(1.899)	(-0.869)	(2.952)	(2.905)	(0.321)	(3.515)

Breaking the high-low spreads out into January and non-January months also demonstrates that the effect is not due to the well-known January seasonal, although the spreads are much larger for the January months. The full-sample high-low spread for χ^*_{BOOM} is insignificant for the non-January months, but this can be attributed to the first half of the sample. In the second half of the sample, the spread is significant in both January and non-January months. For χ^*_{CRASH} spreads, the full-sample results are significant across months and this effect is again concentrated in the second half of the sample.

The high-low spreads are statistically significantly positive and economically large, but it remains to to examine risk-adjusted returns to ensure the differential returns are not attributable loadings upon other factors. Denote the zero-investment portfolio that goes long the high χ^*_{CRASH} portfolio and short the low χ^*_{CRASH} portfolio as *CRASH*. Define *BOOM* similarly. We examine the risk adjusted return by regressing these portfolios on the factors from the CAPM, the Fama-French model and the Fama-French-Carhart model in Table VIII.

For the full sample, the excess returns are significantly positive and economically large no matter which model is used to control for risk. Depending on the model used, the monthly excess returns for BOOM range from 0.34% per month to 0.62%, while excess returns for CRASH range from 0.30% to 0.50% per month. The inclusion of the momentum factor, MOM, in particular induces much larger alphas for both factors.

The factor loadings in Table VIII are obtained from the four-factor Fama-French-Carhart specification. Each of the factors has positive market beta of around 0.20, despite the sort variables, χ^*_{CRASH} and χ^*_{BOOM}, being orthogonalized with respect to historical market beta[12]. Such a sensitivity to the market can arise if the errors in

[12]While the full set of tail dependence being uncorrelated with the full set of market betas does not

Table VIII

Excess Returns and Factor Loadings

Tail dependence factors are obtained by sorting stocks based upon the residuals obtained from regressing tail dependence upon historical beta and logarithm of equity capitalization. Factors are then formed by going long the decile portfolio with the highest dependence and short the decile portfolio with the lowest decile dependence. Risk-adjusted excess returns are obtained from the CAPM, three factor Fama-French (FF) and four factor Fama-French-Carhart (FFC) models. Factor loadings are obtained from the Fama-French-Carhart model. t-statistics in parenthesis are calculated using Newey-West standard errors with twelve lags.

	Excess Returns			Factor Loadings			
	CAPM	FF	FFC	β_{MKT}	β_{SMB}	β_{HML}	β_{MOM}
BOOM 67-04	0.39	0.34	0.62	0.19	-0.08	0.03	-0.28
	(3.43)	(2.90)	(4.57)	(5.87)	(-1.56)	(0.49)	(-7.01)
CRASH 67-04	0.30	0.31	0.50	0.20	-0.08	-0.05	-0.19
	(2.40)	(2.51)	(3.81)	(8.26)	(-2.10)	(-1.19)	(-6.82)
BOOM 67-86	0.02	-0.06	0.16	0.17	0.02	0.05	-0.20
	(0.20)	(-0.52)	(1.20)	(3.15)	(0.30)	(0.58)	(-3.48)
CRASH 67-86	-0.02	-0.01	0.12	0.15	0.06	-0.07	-0.12
	(-0.09)	(-0.08)	(0.64)	(3.76)	(1.12)	(-1.25)	(-2.88)
BOOM 86-04	0.76	0.72	1.02	0.17	-0.13	-0.00	-0.31
	(4.95)	(4.98)	(5.67)	(4.66)	(-1.45)	(-0.00)	(-7.11)
CRASH 86-04	0.61	0.61	0.82	0.22	-0.14	-0.05	-0.22
	(4.22)	(4.48)	(5.60)	(7.71)	(-3.01)	(-0.84)	(-6.65)

the estimated betas are correlated with our own grouping procedure. More precisely, high beta stocks with estimation error that makes them appear to have lower beta still seem to have relatively high tail dependence and vice-versa. However, historical beta still takes a negative Fama-Macbeth coefficient alongside the tail dependences in Table V and the tail factors have positive CAPM alpha in Table VIII, suggesting that any measurement error picked up by the tail dependences is not responsible for the empirical failings of the CAPM.

The orthogonalizing procedure is more effective with size since both tail factors have generally insignificant loadings on the size factor, SMB. If anything, the point estimates suggest the factors load like large stocks. The loadings on the book-to-market factor, HML, are insignificant in all cases. Both factors load negatively on momentum and the relationship is highly statistically significant.

Splitting the sample once again shows that all of the action is occurring in the second half of the sample. In the first half of the sample, the alphas for each factor are insignificant against each of the pricing models.

A lingering question is that of how much difference there is between the CRASH and the BOOM return series. When each measure is orthogonalized to size and beta, there remains commonality between each tail dependence and the BOOM and CRASH returns have a 0.66 correlation. The orthogonalization process can be easily extended so that the sort variable for the CRASH factor is also orthogonal to the χ_{BOOM} and vice-versa for the BOOM factor. Returns to these new factors (unreported for brevity) are uncorrelated with each other, but retain their positive return characteristics, albeit with lower point estimates. Since each of these factors earns a positive risk premium while paying off at different times, this demonstrates that

imply that the beta of the tenth portfolio will necessarily be equal to the beta of the first portfolio, the highly significant loading suggests this is not due to sampling variability.

each set of tail dependences is capturing separate pricing information. Since such a procedure discards the pricing information that is common to both tail dependences, we will continue to orthogonalize against just size and beta.

Another dimension to consider is determining how much of her portfolio a mean-variance investor would devote to the factors. This approach is considered in Pastor and Stambaugh (2003). While multi-factor pricing is not necessarily consistent with mean-variance optimization it is still interesting to examine how much a Markowitz investor would gain by having knowledge about the additional factors.

As a first cut, we follow Pastor and Stambaugh (2003) and examine the portfolio weights that deliver the *ex post* maximum monthly Sharpe ratio. Since the candidate assets are all long-short portfolios, we do not impose short sales constraints. Results are contained in Table IX.

The tail dependence factors are very attractive to mean-variance investor. When CRASH is made available in the Fama-French universe, over one-third of the portfolio is allocated to the new security. When BOOM is made available the allocation is 30%. When both are added, over 38% of the portfolio is allocated to the tail dependence factors. When the momentum factor is added, the tail factors are even more attractive with the allocation to each of the factors on their own rises to over 35%, the largest allocation of all of the factors. With both tail factors alongside momentum, the total allocation to tail factors exceeds 40%, while the market premium constitutes just 3.1% of the optimal portfolio.

The difficulty with this approach is that our investor knew the future return history for each factor when she chose her weights back in 1967. The marked increases in Sharpe ratio from adding extra factors need to be interpreted with caution since allowing her access to *any* new asset will almost certainly allow her to increase her Sharpe ratio. If taking a position in the new asset causes the Sharpe ratio to fall,

Table IX

Ex Post Maximum Sharpe Ratio Portfolio Weights

Portfolios weights maximize the ex post Sharpe ratio. Data is monthly from July 1967 until December 2004.

Panel A. Full Sample

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe
100.0						0.099
27.9	15.8	56.3				0.247
11.8	13.9	39.7		34.6		0.288
14.5	14.5	40.2			30.7	0.284
10.5	13.9	37.3		22.7	15.6	0.295
21.3	10.7	42.4	25.7			0.348
5.5	9.3	27.0	22.3	35.9		0.437
5.9	9.3	24.1	25.1		35.6	0.455
3.2	9.1	22.7	23.6	18.3	23.1	0.473

Panel B. July 1967 - March 1986

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe
100.0						0.063
18.7	17.8	63.5				0.260
19.5	18.4	64.7		-2.6		0.260
22.3	20.0	69.7			-11.9	0.262
21.6	19.4	68.8		3.5	-13.3	0.262
10.8	16.8	41.3	31.1			0.403
8.0	14.3	37.5	28.9	11.4		0.408
6.5	14.0	34.2	29.5		15.8	0.413
5.9	13.3	33.6	28.8	5.3	13.1	0.414

Panel C. April 1986 - December 2004

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe
100.0						0.135
34.8	15.0	50.3				0.255
7.3	16.3	28.8		47.6		0.384
13.8	16.2	30.8			39.2	0.367
7.2	16.4	27.5		32.7	16.2	0.395
29.4	8.5	40.9	21.2			0.332
2.4	10.2	20.2	19.8	47.4		0.552
6.9	9.3	19.7	23.0		41.1	0.551
1.9	9.9	17.7	21.3	27.6	21.6	0.596

she can always choose a weight of zero. Thus the benchmark needs to reflect the average gain from randomly adding any new asset, and appropriately determining this average gain is a non-trivial task.

Fortunately, a better test can be devised to minimize such an ex post problem. Specifically, each month we use the preceding ten years[13] of data to obtain the weights that would have delivered the maximum Sharpe ratio over the preceding ten years. Then use these weights going forward to obtain a portfolio return for the subsequent month and then measure the Sharpe ratio of these returns. This approach circumvents any problem with the additional factor automatically increasing the Sharpe ratio, since this investor is denied knowledge of future returns when choosing her weights. Table X documents the results from this modified approach.

In the first half of the sample, the investor is devotes a large fraction of her wealth to the tail factors, but pays a large cost in terms of Sharpe ratio, particularly when the CRASH factor is available. For the second half of the sample, the investor again invests heavily in the tail dependence factors but now realizes much higher Sharpe ratios (although not quite as high as the ex post optimal ratios, but not too far off). The weights for the first ten years of the second subsample are formulated using returns which include returns from the first subsample where the factor returns are much weaker. If we examine just the years where the weights are calculated from second subsample returns, ie. 1997 to 2004, then the portfolio allocations to the tail factors are larger still.

The increases in Sharpe ratio from adding the tail factors are striking, but a formal test of the differences is useful. This is particularly true as Lo (2002) points out that most of the uncertainty in an estimated Sharpe ratio comes from uncertainty about

[13] The test were initially carried out using five years of data, but these would occasionally deliver nonsensical allocations.

Table X

Ex Ante Maximum Sharpe Ratio Portfolio Weights

Each month portfolio weights that would have delivered the optimal monthly Sharpe ratio over the preceding ten years are determined. These weights are then used over the subsequent month. Reported portfolio weights are time-series averages and Sharpe ratios characterize the performance of the strategy over the sample period excluding the first ten years. P-values are from one-sided tests for increases in Sharpe ratios, where models including tail factors are tested against the same model excluding tail factors and models excluding tail factors are tested against the next most parsimonious model.

Panel A. Full Sample

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe	p-value
100.0						0.137	
32.2	7.1	60.7				0.193	0.00000
13.7	7.3	42.7		36.4		0.177	1.00000
16.5	7.8	40.8			34.8	0.242	0.00000
11.8	6.9	38.3		27.7	15.3	0.200	0.00789
18.4	10.3	41.2	30.1			0.336	0.00000
5.6	8.7	29.7	24.5	31.5		0.381	0.00000
5.7	9.4	24.0	27.5		33.3	0.450	0.00000
3.1	8.6	24.6	24.5	19.7	19.5	0.419	0.00000

Panel B. July 1977 - March 1986

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe	p-value
100.0						0.141	
10.9	23.4	65.7				0.194	0.00130
-2.5	9.2	41.3		52.1		0.085	1.00000
4.7	18.6	53.5			23.2	0.106	1.00000
-1.8	9.2	43.0		57.5	-7.8	0.049	1.00000
6.9	20.6	39.9	32.5			0.421	0.00000
-2.5	9.7	27.8	24.5	40.5		0.352	1.00000
-0.8	14.9	25.6	29.7		30.5	0.402	0.99999
-3.9	9.2	24.7	24.4	34.9	10.6	0.338	1.00000

Panel C. April 1986 - December 2004

MKT	SMB	HML	MOM	CRASH	BOOM	Sharpe	p-value
100.0						0.135	
42.2	-0.5	58.3				0.194	0.00000
21.2	6.4	43.3		29.1		0.211	0.00002
22.0	2.8	34.9			40.3	0.292	0.00000
18.2	5.8	36.1		13.8	26.0	0.256	0.00000
23.8	5.5	41.7	29.0			0.299	0.00000
9.4	8.2	30.6	24.5	27.3		0.394	0.00000
8.7	6.9	23.3	26.5		34.6	0.471	0.00000
6.4	8.3	24.5	24.5	12.6	23.6	0.455	0.00000

the mean excess return, rather than uncertainty about the standard deviation. Higher Sharpe ratios, such as those we obtain, are estimated less precisely.

Lo (2002) derives standard errors for estimated Sharpe ratios and in an extension communicated privately by Andrew Lo, a test for the difference in Sharpe ratio is described. Given two return series, firstly use GMM to estimate the means and variances of the series and the parameter variance-covariance matrix. The difference in Sharpe ratios is a function of the estimated parameters, so the delta method can be used to determine standard errors for testing whether an increase in Sharpe ratio is statistically significant. Additionally, heteroskedasticity and autocorrelation can be taken into account within the GMM procedure. The null hypothesis is that the Sharpe ratio is maximized by setting the weight on the CRASH and/or BOOM factor to zero. As can be seen in the leftmost column of Table X, in the second half of the sample the realized Sharpe ratio is always improved by the availability of the tail factors.

4 Was there a structural break following the October 1987 crash?

There are two possible explanations for the lack of significant risk premia in the first half of the sample. The first is that prices became more informative over the sample period (perhaps due to increasing trading volume or technological innovation mitigating the stale prices problem, or even just due to the increasing size of the cross-section), but the pricing phenomenon was present all along. Alternatively, there is the possibility that there was a structural break in the risk premia as investors re-evaluate the future following a large market movement.

There have been developments in the market that suggest that the quality of the data has improved over the sample period. Increased trading will mitigate the stale prices problem and may have decreased the lag in small stock returns. The sheer number of stocks on the CRSP tapes has also increased markedly, allowing a much wider cross-section. Such developments are, however, gradual and one would expect them to manifest themselves in the form of a trend in the observed risk premia series.

On the other hand, there is an "elephant in the room" in the form of the crash of October 19th, 1987. The midpoint of the sample, April 1986, is sufficiently close to that date that the reader may have already considered the possibility that the differences in risk premia over the two subsamples may be related. The story is that investors were surprised by the magnitude of the crash and subsequently demanded a risk premium to hold portfolios that were exposed to this risk[14]. Since the distance between assumption and conclusion will be perilously small for any rigorous model making such a prediction, such a proposition is purely an empirical question. In such a situation, we should observe increase in the risk premia following the event, which can easily be modeled as a dummy variable. The proposition is a sharp one, given the low signal-to-noise ratio of the mean of asset return series.

To test whether the risk premia is gradually revealed or suddenly appears can be achieved by testing for a time trend or structural break in the factor return series. The philosophical question becomes whether we pre-specify the date of the change or allow the data to determine the breakpoints. We explore both approaches.

[14] The representative investor may have updated her belief about the return distribution or simply learnt how much a crash hurt. Distinguishing between these two ideas is difficult and not explored in this paper.

4.1 Tests with a pre-specified breakpoint

We first examine tests where the break-date is pre-specified as October 1987 in Table XI. While the loadings are not documented in the table, each regression includes MKT, SMB, HML and MOM and the loadings are allowed to change following October 1987. Specifications where the loadings were fixed over the whole sample period and also without the Fama-French-Carhart factors were also examined and the conclusions drawn were unchanged.

Regressing the CRASH factor upon a dummy equal to one after October 1987 produces a statistically significant coefficient, but the regression with a linear time trend also produces a significant coefficient. However, when the regression includes both the post-87 dummy and the time trend only the dummy remains significant suggesting that there was a structural break rather than a gradual shift in the CRASH factor premia.

Another possibility remains that the premia is only evident while the October 1987 crash is within the historical five year estimation window[15]. To address this, the regression is run with a dummy equal to one while the crash is within the estimation window and another dummy equal to one once the crash is no longer within the window. Each dummy is statistically significant and the $\hat{R}^2$ is lower than for the regression with just the post-1987 dummy, demonstrating that the change in the premia did persist. Additionally, this test shows that the the performance of the factors is not an artefact of the bull market of the late 90's, since the two dummies are both both statistically significant and not significantly different from each other.

When the same tests are conducted on the BOOM factor, the post-87 dummy and

[15]Such an issue is more problematic in the moment-based approach since the outlier is raised to a higher power and dominates the summation used for the estimator. The estimator used here is robust to outlying observations since it is a count variables and all of the magnitude information is discarded.

Table XI
Structural Breaks in Factor Premia Following the October 1987 Crash

Each row documents a regression designed to explain the non-constancy in the returns of tail factors. The dependent variables are monthly factor returns from July 1967 until December 2004. Trend is a linear time trend. D_{87} is a dummy variable that is one every month after October 1987. D_{87+5} is a dummy variable equal to one for the five years following October 1987 while D_{92} is a dummy variable equal to one once the five year period is over. Each regression includes MKT, SMB, HML and MOM and each coefficient is allowed to change following October 1987. The t-statistics in parenthesis use Newey-West standard errors with twelve lags.

	Intercept	D_{87}	Trend	D_{87+5}	D_{92}
CRASH	0.137	0.758			
	(0.792)	(3.317)			
CRASH	0.011		0.002		
	(0.044)		(2.415)		
CRASH	0.281	1.022	-0.001		
	(0.962)	(2.245)	(-0.643)		
CRASH	0.137			1.079	0.618
	(0.792)			(3.589)	(2.484)
BOOM	0.236	0.814			
	(1.810)	(3.350)			
BOOM	-0.053		0.003		
	(-0.298)		(3.546)		
BOOM	0.091	0.548	0.001		
	(0.370)	(1.021)	(0.615)		
BOOM	0.236			0.883	0.784
	(1.810)			(1.963)	(2.946)

the time trend are again statistically significant when alone. However, when both are included in the regression, neither produces a statistically significant coefficient.

4.2 Tests with unknown breakpoints

While the 1987 crash is a natural date to consider when testing for shifts in the premia, there is additional value in a more agnostic, data-driven approach. Bai and Perron (1998) and Bai and Perron (2003), henceforth BP, develop an econometric framework for testing for structural breaks in linear time series models when the break dates are unknown and provide a method for obtaining confidence intervals for the break dates.

The tests are modified F-tests based on the sum of squared residuals which are calculated for all possible combinations of break points. The breakpoints themselves are restricted to have at least %15 of the sample, or 67 months, between any two breakpoints. The F-tests are adjusted for autocorrelation and heteroskedasticity and the residual variance is also allowed to change at each breakpoint.

With the sums of squared residuals for all of the possible sets of break points calculated, the primary issue is to test whether the null hypothesis of no breaks can be rejected. BP develop two so-called double maximum tests that test this hypothesis against the alternative hypothesis of up to M structural breaks. Both tests take a weighted average of F-statistics from the optimal partitions for each possible number of breakpoints from 1 to M. The test statistic is then the largest F-statistic from the set of F-statistics that are themselves the largest F-statistics for each possible number of breakpoints, hence the name. The unweighted double maximum test, or UDmax, test uses equal weights to determine the average optimal F-statistic. The weighted double maximum test, or WDmax test, applies weights such that the marginal p-values are equal across the different values of M. Critical values for the tests provided by BP reflect the fact that each set of the breakpoints has been selected after testing all possible combinations of breakpoints as well as the fact that up to M breakpoints are being allowed for.

In the case where the null hypothesis of no structural breaks is rejected, the task of selecting the optimal number of breaks remains. To this end, BP use $\text{supF}(i+1|i)$ tests which test the hypothesis of $i+1$ breaks under the hypothesis of i breaks. Once again the critical values are adjusted to reflect the estimation of the breakpoints and are provided by BP. The optimal number of breaks can then be determined sequentially by adding additional breaks until the hypothesis of an additional break is rejected.

Table XII documents the BP tests upon the CRASH and BOOM factors where the regression includes the Fama-French-Carhart factors. Turning first to Panel A where only the mean return parameter is allowed to change, the hypothesis of no structural breaks is rejected for both the CRASH and BOOM series since the test statistics for both the UDmax and WDmax tests exceed the 5% critical values. This implies that there is at least one and up to five structural breaks in the means of the factors. Applying the $supF$ tests sequentially, the $supF(2|1)$ for the existence of two structural breaks against the existence of just one structural break is rejected for both series, leading us to conclude that there is only one structural break in the mean of each series.

The date of the break for the CRASH return series is highly suggestive - November 1987 being the month immediately after the crash. This would be *exactly* when we expect the increase in premia to occur as only the October end prices would reflect any new premia. However, the 95% confidence interval covers fifteen years from August 1981 through August 1996. The width of the confidence interval is not surprising when we recall that we are testing for a shift in the mean of a return series which are notoriously noisy and the implied prior belief of when any break

Table XII

Detecting Structural Breaks at Unknown Dates

Each column presents the results from test for structural breaks following Bai and Perron (1998) based upon reductions in the sum of squared residuals. Each regression has a tail factor as the dependent variable and MKT, SMB, HML and MOM factor returns as explanatory variables. In Panel A, only the intercept of the regression is allowed to change, while in Panel B the factor loadings are also allowed to change. The minimum portion of the sample that can fall between break dates is 15% or 67 months, allowing a maximum of five possible breaks. The UDmax and WDmax tests test the null hypothesis of no structural breaks against the alternative of at least one and no more than five structural breaks. The supF($i+1|i$) test the hypothesis of $i+1$ breaks against the null of i breaks and are robust to heteroskedasticty and autocorrelation. The supF tests are applied sequentially to determine the number of breaks present in the data. Data is monthly from July 1967 to December 2004.

	Panel A. Break only in mean			Panel B. All parameters break		
	CRASH	BOOM	5% C.V.	CRASH	BOOM	5% C.V.
UDmax	13.82	16.08	8.88	30.91	26.28	18.42
WDmax	20.00	16.08	9.91	39.94	31.39	19.96
supF(1|0)	13.82	16.08	8.58	30.91	23.41	18.23
supF(2|1)	5.11	2.62	10.13	15.35	15.54	19.91
# Breaks	1	1		1	1	
Date	11/1987	12/1988		9/1987	4/1978	
95% C.I.	[8/1981 - 8/1996]	[11/1980 - 10/1992]		[2/1984 - 5/1991]	[4/1970 11/1980]	

might occur is completely uninformative. The estimated break date for the BOOM series is December 1988 while the confidence interval spans November 1980 through to October 1992, which does include October 1987.

Panel B. repeats the test for the situation where the factor loadings as well as the mean of the series are allowed to change. Once again, the UDmax and WDmax tests reject the hypothesis of no structural breaks and the sequential supF tests imply one structural break in each series. The structural break for the CRASH series is estimated at September 1987 with a narrower 95% confidence interval from February 1984 through to May 1991.

The structural break for the BOOM series is estimated to occur in April of 1978, with confidence interval spanning April 1970 until November 1980. Clearly this structural break is not attributable to the 1987 crash. We attribute the break to something far more mundane. The Nasdaq stocks are included on the CRSP tapes from December 1972 and enter the factor portfolios five years later - January of 1978. The break in the series is due to shifts in the size and book-to-market loadings, which both decrease significantly. The break is sufficiently close to the break in the mean series that this "soaks up" any effect.

5 Crash Risk Returns and the Skew in Implied Volatilities

A shift in the risk premia for the tail factors following the 1987 crash is not the only instance of a shift in asset pricing after the crash. Rubinstein (1994) identified a break in the behavior of the index option markets following the October 1987 crash. Before the crash, implied volatilities from the Black and Scholes (1973) model were basically flat across different strikes. Following the crash, implied volatilities for out-of-the-money puts became much higher than for their at-the-money counterparts. While the skew may be in part due to failures of the assumptions of the model, it remains that deep-out-of-the money options are expensive relative to their at-the-money counterparts and that this shift coincided exactly with the 1987 crash. Such a skew reflects more probability in the left tail of the risk-neutral density, or as Rubinstein colorfully describes it: "crash-o-phobia".

If the skew does represent crash-o-phobia, then we should expect a relationship with the CRASH factor returns. Writing out-of-the money puts on the index would be an exceptionally effective way of maximizing the exposure of any portfolio to a market crash. Higher implied volatilities for out-of-the-money puts on the index suggest that state prices for payoffs during crashes are quite high.

Since option prices embed forecasts, we can infer that the market forecasts that a crash is more likely when the implied volatilities are more skewed. If the market forecasts a higher probability of a crash, then the required rate of return on stocks that crash heavily should also be higher. A natural test is to see whether returns to the crash-prone stocks are higher following larger skews in the implied volatilities.

In order to test the hypothesis, we can use the returns on the CRASH factor for the realized returns but we require a proxy for the skew. Data is from the SPX options written on the S&P500 index, which represents a narrower definition of the market portfolio than used previously but is the best optionable proxy. The SPX contract has European style exercise, so the implied volatilities can be obtained from inverting the Black-Scholes formula with the option price taken to be the midpoint of the bid-ask spread. The actual dividends paid over the life of the option are used to adjust for the dividends paid by constituent stocks.

Data is daily from 1991 until 2004 with data from 1991 through 1995 from the

Figlewski (2002) dataset (originally from the Berkeley Options Database) and the data for 1996 through 2004 is from OptionMetrics. It is worth noting that the analysis was originally carried out just on the OptionMetrics data so that subsequent analysis of the Figlewski data provided an out-of-sample confirmation of the earlier findings.

The factor returns are measured over calendar months, so we require estimates of the skew over the preceding month. The expiry date for SPX options is always the third Friday of the month. We want to measure the forecast of crash likelihood over the next month, so we choose the first option series that matures after the end of the next month, ie. the series that matures approximately a month and a half after the portfolio formation date.

The approach to quantifying the skew is as follows. Each day, we require at least three out-of-the-money options so that the skew can be estimated. For such days, we obtain the at-the-money forward implied volatility by linear interpolation of the two options with strikes nearest-to-the-money. This implied volatility, combined with the days remaining, yields a forecast of the remaining volatility that accounts for both the heteroskedasticity of the index returns and the variation in the number of days remaining. We can determine a strike level that is one standard deviation[16] out of the money. Next we measure the skew by taking all out-of-the-money options and regressing their implied volatilities upon $\log(Strike/Spot)$. If the regression fit is so poor that the $\bar{R}^2$ is less than 0.9 we drop the observation[17]. We then use the estimated function to determine the implied volatility of the option at the strike one standard deviation below the current index level. The measure of the skew is determined as this implied volatility minus the implied volatility of the at-the-money option. The

[16] An one standard deviation movement is not a crash, but the measure we obtain is designed to be a measure of extra implied volatility per standard deviation.

[17] In general, the $\bar{R}^2$ is above 0.99 so the linearity assumption is not the issue. Examining days where the fit is poor, the function is highly non-monotonic and the problem appears to be driven by stale quotes or data errors.

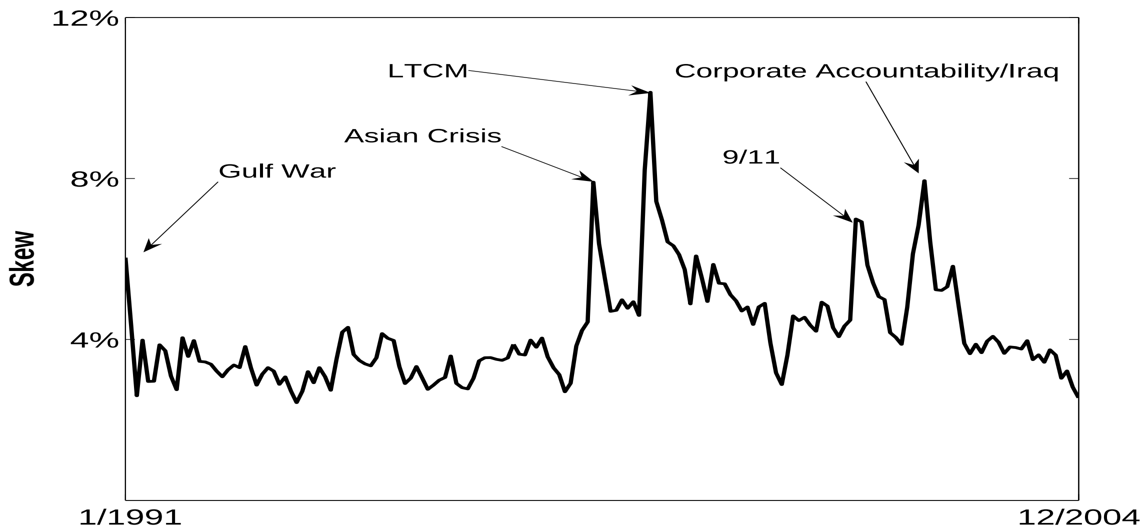


Figure 1. Time Series of S&P 500 Index Put Option Skew The skew is measured as the implied volatility of an out out-of-the-money put option minus the implied volatility of the at-the-money put option, averaged over each day in the month using the option series expiring after the end of the subsequent month. The out-of-the-money option has a strike one standard deviation less than spot, inferred from the at-the-money implied volatility and the time-to-maturity of the option.

monthly measure is then the average of the daily measures.

Figure 5 plots the time series of the implied volatility skew measure. The measure shows pronounced spikes during periods of financial turmoil which are excellent candidates for periods when crashes may be more probable. The first Iraq War, the Asian Crisis, Long Term Capital Management, the terrorist attacks of 9/11 and the corporate accountability scandals/second Iraq war all coincide with large increases in the skew.

Predictive regression can run afoul is the persistence of the predictive variable. Stambaugh (1999) demonstrates that when the predictive variable follows an AR(1) process and the errors in the AR(1) process are correlated with the errors from the predictive relationship, then the OLS coefficient on the predictive variable will be biased. Such a problem is of relevance here as the skew measure follows an AR(1) process with autoregressive parameter 0.83 (t=19.12), which is quite persistent although much less persistent than some other variables used in the return prediction literature. We first present results using a baseline OLS specification and then demonstrate that these results are not driven by a Stambaugh problem by using a more careful econometric method.

Panel A of Table XIII documents regressions of subsequent returns of the CRASH factor upon the skew from the preceding month. The coefficient on the skew is significantly positive for a one-tailed test, but not so for a more standard two-tailed test. However the economic significance is large since the skew variable has a standard error of 1.2% so that a one-standard deviation increase in the skew is associated with approximately 35 basis points of extra return in the subsequent month. However, lagging the skew by an additional month (ie. January skew predicting March return) produces a highly significant result with the skew strongly predicting future returns to the CRASH factor. This suggests that the relationship may be occurring at a lower frequency than monthly.

The next row of Table XIII uses the skew to predict the return on the CRASH factor over the subsequent two months (ie. January skew predicts CRASH return over February and March), and the result is that when the skew is one standard deviation higher than average, the CRASH returns are an average of 50 basis points higher for each of the two subsequent months. Of course, the dependent variable now consists of overlapping observations so we need to be wary of results being induced by the interaction between persistence in the dependent variable and the overlapping observations (see Boudoukh, Richardson, and Whitelaw (2005)).

To avoid such problems with overlapping observations, we split the sample into observations where the skew is measured in an odd month (January, March, etc.) and observations where the skew is measured in even months (February, April, etc.). For each of these two subsamples there is no overlap in the dependent variable. Table XIII

Table XIII

Predicting CRASH Factor Returns With The Skew in S&P500 Index Put Option Implied Volatilities

Skew is measured as the implied volatility of the option that is one standard deviation out-of-the-money minus the implied volatility of the at-the-money option, averaged over each day in the month using the option series expiring after the end of the subsequent month. The CRASH factor is the return over the subsequent month or the average return over the subsequent two months. For the regressions with subsequent two months, the sample is also split into even and odd months so that the dependent variable does not overlap. Returns are monthly from February 1991 until December 2004. In Panel B. the regressions follow Amihud and Hurvich (2004) by estimating ϕ, the coefficient on the contemporaneous error in an AR(1) specification for the skew to reduce bias resulting from any Stambaugh (1999) problem.

Panel A. OLS estimates

Sample	Dep. Var	Int.	Skew_{-1}	Skew_{-2}	$\bar{R}^2$
Full Sample	1 month	-0.40	0.28		0.01
		(-0.60)	(1.80)		
Full Sample	1 month	-1.42		0.52	0.06
		(-2.17)		(3.46)	
Full Sample	2 month	-0.93	0.40		0.09
		(-2.22)	(4.18)		
Odd Months	2 month	-1.12	0.44		0.12
		(-1.98)	(3.46)		
Even Months	2 month	-0.70	0.35		0.05
		(-1.10)	(2.37)		

Panel B. Corrected estimates

Sample	Dep. Var	Int.	Skew_{-1}	ϕ	$\bar{R}^2$
Full Sample	1 month	-0.40	0.25	-0.96	0.08
		(-0.62)	(1.66)	(-3.63)	
Full Sample	2 month	-0.93	0.39	-0.73	0.18
		(-2.35)	(3.99)	(-4.46)	
Odd Months	2 month	-1.12	0.42	-0.52	0.22
		(-2.10)	(3.28)	(-3.40)	
Even Months	2 month	-0.70	0.32	-0.62	0.14
		(-1.15)	(2.17)	(-3.02)	

shows that the coefficients for the two subsamples are statistically indistinguishable from the full sample regression, although the smaller sample size does reduce the t-stats but not to the point of insignificance. The subsample regressions also point to the robustness of the predictive relationship, at least within the 1991-2004 sample[18].

To address the possibility of the results being driven by a Stambaugh problem, we use the Amihud and Hurvich (2004) bias-reduction method where the predictive regression is augmented with the errors from the AR(1) specification[19] for the predictive variable. The bias-reduced estimate is then the OLS coefficient from the augmented regression, with a standard error adjusted for the uncertainty about the autoregressive parameter.

As Amihud and Hurvich (2004) point out, the coefficient on the contemporaneous errors from the AR(1) specification, ϕ, is informative as it captures the response to the unanticipated component of the predictive variable. The hypothesis so far has been that when the skew is high then subsequent returns to the CRASH strategy should be high. Such a hypothesis is incomplete because forecasts of higher future returns should result in lower current prices. So if the skew increases unexpectedly over the subsequent period, then the factor return should be *lower* as current prices move to reflect the additional risk. So we expect the coefficient on the AR(1) errors, ϕ, to be negative.

Panel B of Table XIII presents the coefficients from the corrected regressions. In each of the various specifications, the point estimate of the coefficient on the past skew is marginally lower but the conclusions drawn remain the same. The ϕ coefficient is

[18]The sample can also be split into the 1991-1995 Figlewski sample and the 1996-2004 OptionMetrics sample without affecting the conclusions.

[19]Stambaugh (1999) shows that the bias in the OLS estimate of the predictive coefficient is proportional to the bias in the autoregressive parameter for the predictive variable process. As such, the Amihud and Hurvich (2004) method uses a corrected estimate of the autoregressive parameter when obtaining the AR(1) errors, as the standard OLS estimate of the autoregressive parameter is also biased in finite samples.

highly significant and negative, confirming the expected economic relationship.

Clearly, the skew in option prices is informative about the future returns of the CRASH strategy. Both the appearance of skew in implied volatilities and the increase in returns to the CRASH both occur following the 1987 crash, so the predictive relationship is strong support for crash-o-phobia as well as for the risk premia interpretation of the large returns to the strategy isolating crash risk.

6 Conclusions

Tail dependence is certainly an interesting lens through which to view the risk and return tradeoff. Strategies isolating crash risk and boom risk each earn economically large and statistically significant risk premia that cannot be attributed to market, size, book-to-market or momentum risk. The premia exhibit a large increase following the 1987 crash, demonstrating a lasting effect of that important market event. Finally, the returns to bearing crash risk can be predicted using a measure of the skew in implied volatilities of S&P 500 index put options, which is both support for the idea that "crash-o-phobia" is driving the skew in implied volatilities and confirmation that the returns to the crash risk strategy are indeed reward for bearing crash risk.

Appendix A. The form of extremal dependence

It is important to determine what is the correct form of dependence or, equivalently, does $\bar{\chi} = 1$? For this purpose, we can use the entire return history of each stock as we are simply interested in characterizing behavior.

To carry out estimation, we need to construct the z variable so that we can use equation (5). Transforming the observations to Fréchet margins is done by transforming the empirical distribution, since the true marginal distribution is unknown. Note that such a transformation discards all information about the distributions except for the ranking of the observations; any location and scale information is lost. The lesser of each pair of observations from the two transformed series is then taken. Estimation follows from assuming that (5) holds exactly above some high threshold, with the slowly varying function equal to a constant.

In order to test whether $\bar{\chi} = 1$, we need an estimator. Since $\bar{\chi}$ can be obtained from the tail index of the transformed series, we can use standard methods from univariate extreme value theory. While there are a number of methods of estimating positive tail indices, the workhorse for estimating the tail index of fat-tailed variables is due to Hill (1975):

$$\widehat{\eta^{H}(k)} = \frac{1}{k} \sum_{j=1}^{k} \left(\log z_{(n-j+1)} - \log z_{(n-k)} \right), \tag{A-1}$$

where k is the number of observations defined to be the tail and $z_{(i)}$ denotes the ith observation in the ordered sample of n observations. The Hill estimator follows from assuming that each of the k observations are following the tail distribution. The k observations then follow the Pareto Distribution for which the Hill estimator is the maximum likelihood estimator of the shape parameter. By the functional invariance of maximum likelihood estimators:

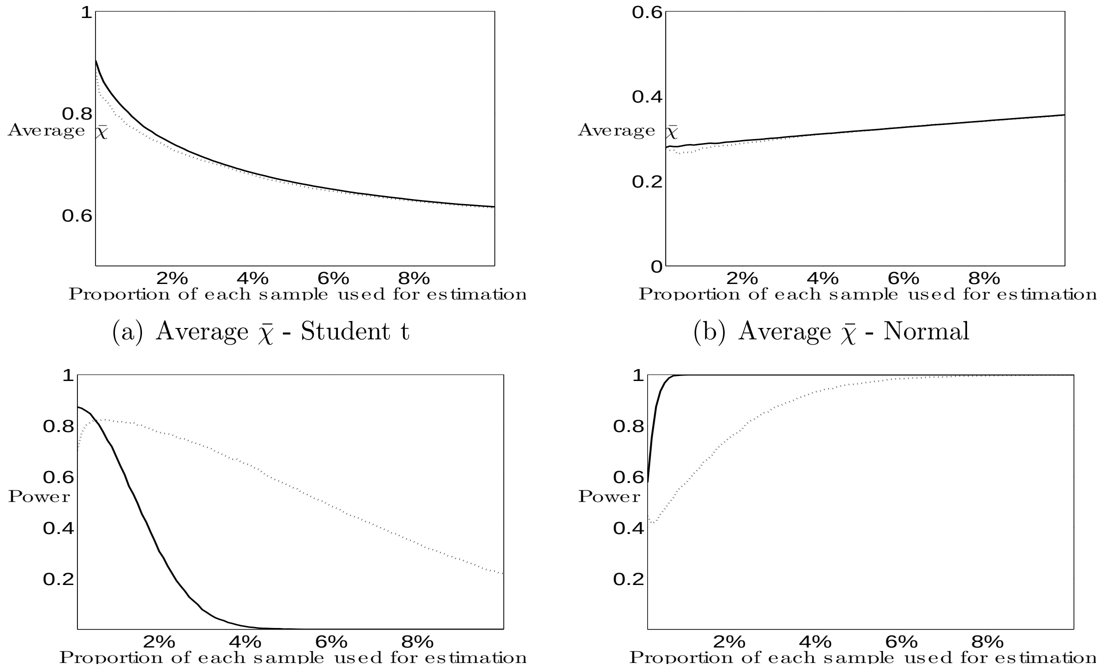
$$\widehat{\bar{\chi}(k)} = \frac{2}{k}\sum_{j=1}^{k}\left(\log x_{(n-j+1)} - \log x_{(n-k)}\right) - 1, \tag{A-2}$$

The estimator $\widehat{\bar{\chi}(k)}$ is aymptotically normal with variance $\frac{\left(\widehat{\bar{\chi}(k)}+1\right)^2}{k}$. This can be used to test $\bar{\chi} = 1$. While interest lies in estimating $\bar{\chi} \in (-1, 1]$, the upper constraint is not imposed[20] as we would lose asymptotic normality of the the estimator for the particularly relevant case when $\bar{\chi} = 1$.

One problem with this estimation procedure is with the standard errors for the estimates obtained from the maximum likelihood estimation. The standard errors are exact for cases where the data are i.i.d. and the tail approximation is exact. For financial time-series data, evidence supporting time-varying volatility for financial time-series data violates the i.i.d. assumption, leaving the estimates consistent but with the standard errors too low. The second problem is that the transformation to Fréchet marginals is not exact and so there is resulting uncertainty about the underlying data which will not be reflected in the standard errors. Finally, there is bias due to the slowly varying function being non-constant. Given that the standard test procedure for delineating asymptotic dependence relies on these standard errors, inference from these tests should be made with caution.

One particularly perilous situation can be encountered. If the two variables are asymptotically dependent, but the dependence structure is such that $\bar{\chi}$ is biased downwards at lower thresholds, then any researcher using a finite sample will be more likely to reject the true dependence structure. As an example of the problem, consider the bivariate t-distribution which exhibits asymptotic dependence or, equivalently,

[20]The lower bound is imposed implicitly in the functional form of the likelihood.



(a) Average $\bar{\chi}$ - Student t

(b) Average $\bar{\chi}$ - Normal

(c) Empirical Power - Student t

(d) Empirical Power - Gaussian

Figure A.1. Simulated extremal dependence diagnostic performance Five thousand samples of bivariate normal and bivariate Student t are simulated, with sample sizes of 1,000 and 10,000 observations. $\bar{\chi}$ is estimated for every possible threshold up to ten percent of the full sample. The degrees of freedom for the Student t are set to 4 and correlation is set to 0.3 for both distributions. Any bivariate t is asymptotically dependent with $\bar{\chi} = 1$, while this bivariate normal is asymptotically independent with $\bar{\chi} = \rho = 0.3$. The empirical power functions show the proportion of bivariate t replications which reject asymptotic independence and the proportion of bivariate normal replications which accept asymptotic independence. The solid lines represent the 10,000 observation series and the dashed line represent the 1,000 observation series.

has $\bar{\chi} = 1$. We simulate ten thousand replications of a bivariate t-distribution with common degrees of freedom of four and correlation of 0.3, for a sample size of one thousand observations. For each replication, estimate a series of $\bar{\chi}$ using k ranging from 1 to 100. Then, for each k, take the average over all replications. The process is repeated for a larger sample size of ten thousand observations, but using a maximum value of k equal to ten percent of the sample size.

Figure 6(a) plots the series of average $\bar{\chi}$ for each sample size for the bivariate t as well as for bivariate normal. The problem with detecting asymptotic dependence can be clearly seen in that the average $\bar{\chi}$ only approaches the true value of one for the very highest thresholds. For the average of the estimates using more observations, there is severe downward bias, which can lead the econometrician to falsely reject asymptotic dependence with too high a frequency. This is particularly evident when the empirical power is plotted in Figure 6(b).

The simulation exercise demonstrates that the test for diagnosing asymptotic dependence can be severely biased against detection in finite samples. The quandary is difficult to resolve since determining appropriate critical values for the test is impossible without an understanding of the dependence structure under examination. However, the simulation exercise suggests a simple strategy to avoid this problem. Figure 6(b) suggests that power is maximized at very high thresholds. In many situations, the researcher is concerned with the dependence structure between just two variables, so estimation using very high thresholds results in noisy estimates. In the current setting, however, we are trying to characterize the extremal dependence of many, many market and stock pairs so we can exploit the cross-section. Just as a clearer picture of the convergence was obtained in Figure 6 by averaging over many replications, we can average over many different stocks to get a better understanding of the extremal dependence for stocks. This technique requires making one weak assumption, namely that the form of extremal dependence of individual stocks is the same for all stocks.

So rather than test each stock individually, we test whether the average $\bar{\chi}$ is equal to one. The major advantage of this approach is that the cross-section can be exploited to bring precision to the overall estimate, rather than having to use more observations which introduce bias. The strategy is to estimate $\bar{\chi}$ for each stock using very few estimates, say $k = 1, ..., 5$. These estimates are far too noisy to be used in isolation, but have the advantage of being subject to far less bias. If we average over all stocks for each k, then the sample mean and standard deviation can be used to test for asymptotic dependence, which has the added benefit of avoiding using the MLE standard errors. Another benefit is that the series of $\bar{\chi}$ averages afford insight into the nature of the bias, in particular the sign.

Estimates with value of k from one to five as well as k such that the observations number 1%, 2.5%, 5% and 10% of the total sample size are presented in Table A-I. The stocks are grouped by the number of observations so that the impact of sample size can also be observed, with stocks with less than one thousand observations excluded. This also proxies for the size of the firm, which is difficult to do directly since both the firm size itself and the size breakpoints will vary over the trading history of each stock. The reported numbers are the cross-sectional averages of the $\bar{\chi}$.

The results for the crash dependence suggest that bias can have a major impact on inference. For the stocks with five thousand or more observations, the average $\bar{\chi}$ are one for the thresholds based on 1 to 5 most jointly extreme observations, leading to the conclusion that the stock and market are asymptotically dependent. Yet the averages for the higher thresholds are markedly lower, so inference based on these thresholds would conclude that the stocks were asymptotically independent of the market. The importance of sample size is also evident, since the smaller samples appear to be too small to reliably detect asymptotic dependence. The results for boom dependence are broadly consistent.

Since the estimates of $\bar{\chi}$ are limiting towards one, we conclude that individual stocks and the market are asymptotically dependent.

Table A-I

Diagnosing Extremal Dependence for Individual Stocks

Sample is all NYSE, Amex and Nasdaq stocks with at least one thousand valid daily returns included. Estimation of $\bar{\chi}$ is done at a number of different thresholds. Reported numbers are cross-sectional averages of the estimated $\bar{\chi}$.

Crash Sub-Asymptotic Dependence

Obs.	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k/n=1\%$	$k/n=2.5\%$	$k/n=5\%$	$k/n=10\%$
$1,000-2,499$	0.45	0.47	0.45	0.45	0.44	0.40	0.41	0.44	0.44
$2,500-4,999$	0.55	0.56	0.52	0.51	0.50	0.41	0.42	0.43	0.46
$5,000-7,499$	0.91	0.88	0.80	0.76	0.73	0.48	0.45	0.45	0.48
$7,500-9,999$	1.55	1.32	1.16	1.05	1.00	0.53	0.50	0.49	0.51
$10,000-10,698$	2.22	1.67	1.37	1.29	1.19	0.62	0.58	0.56	0.55

Boom Sub-Asymptotic Dependence

Obs.	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k/n=1\%$	$k/n=2.5\%$	$k/n=5\%$	$k/n=10\%$
$1,000-2,499$	0.29	0.30	0.30	0.30	0.30	0.31	0.32	0.35	0.34
$2,500-4,999$	0.45	0.42	0.42	0.40	0.40	0.33	0.32	0.34	0.37
$5,000-7,499$	0.58	0.58	0.55	0.52	0.51	0.38	0.35	0.34	0.40
$7,500-9,999$	0.76	0.65	0.64	0.61	0.59	0.43	0.40	0.39	0.42
$10,000-10,698$	1.12	0.89	0.78	0.77	0.73	0.53	0.49	0.47	0.48

References

Amihud, Yakov, and Clifford M. Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis* 39, 813–841.

Ang, Andrew, Joseph Chen, and Yuhang Xing, 2005, Downside risk, Working Paper, Columbia University.

Bai, Jushan, and Pierre Perron, 1998, Estimating and testing linear models with multiple structural changes, *Econometrica* 66, 47–78.

Bai, Jushan, and Peirre Perron, 2003, Computation and analysis of multiple structural change models, *Journal of Applied Econometrics* 18, 1–22.

Barro, Robert J., 2005, Rare events and the equity premium, Working Paper, Harvard University.

Bawa, Vijay S., and Eric B. Lindenberg, 1977, Capital market equilibrium in a mean-lower partial moment framework, *Journal of Financial Economics* 5, 189–200.

Black, Fisher, and Myron Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–654.

Boudoukh, Jacob, Matthew Richardson, and Robert Whitelaw, 2005, The myth of long-horizon predictibility, Working Paper, New York University.

Boyer, Brian H., Michael S. Gibson, and Mico Loretan, 1999, Pitfalls in tests for changes in correlations, *Federal Reserve Board* IFS Discussion Paper No. 597R.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, 1995, Survival, *Journal of Finance* 50, 853–873.

Bruun, John T., and Jonathan A. Tawn, 1998, Comparison of approaches for estimating the probability of coastal flooding, *Applied Statistics* 47, 405–423.

Chung, Y. Peter, Herb Johnson, and Michael J. Schill, 2006, Asset pricing when returns are nonnormal: Fama-french factors and higher order systematic co-moments, *Journal of Business* 79.

Coles, Stuart, 2001, *An Introduction to Statistical Modeling of Extreme Values* (Springer-Verlag: London).

———, Janet Heffernan, and Jonathan Tawn, 1999, Dependence measures for extreme value analyses, *Extremes* 2, 339–365.

Demarta, Stefano, and Alexander J. McNeil, 2005, The t copula and related copulas, *International Statistical Review* 73, 111–129.

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

Fama, Eugene F., and Kenneth R. French, 1992, The cross section of expected returns, *Journal of Finance* 47, 427–465.

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

Figlewski, Stephen, 2002, Assessing the incremental value of option pricing theory relative to an 'informationally passive' benchmark, *Journal of Derivatives* 10, 80–96.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, 2003, A theory of power-law distributions in financial market fluctuations, *Nature* 423, 267–270.

Harvey, Campbell R., and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *Journal of Finance* 55, 1263–1295.

Hill, Bruce M., 1975, A simple general approach to inference about the tail of a distribution, *The Annals of Statistics* 3, 1163–1174.

Ingersoll, Jr., Jonathan E., 1987, *Theory of Financial Decision Making* (Rowman and Littlefield: Maryland, USA).

Jansen, Dennis W., and Casper G. de Vries, 1991, On the frequency of large stock returns: Putting booms and busts into perspective, *The Review of Economics and Statistics* 73, 18–24.

Kraus, Alan, and Robert Litzenberger, 1976, Skewness preference and the valuation of risk assets, *Journal of Finance* 31, 1085–1100.

Ledford, Anthony W., and Jonathan A. Tawn, 1996, Statistics for near independence in multivariate extreme values, *Biometrika* 83, 169–187.

Lo, Andrew W., 2002, The statistics of Sharpe ratios, *Financial Analysts Journal* 58, 36–52.

——— , and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreactions?, *Review of Financial Studies* 3, 175–205.

Longin, François, and Bruno Solnik, 2001, Extreme correlation of international equity markets, *Journal of Finance* 56, 649–676.

Longin, Francois M., 1996, The asymptotic distribution of extreme stock market returns, *The Journal of Business* 69, 383–408.

Markowitz, Harry M., 1959, *Portfolio Selection* (Blackwell Publishers: Massachusetts, USA).

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Poon, Ser-Huang, Michael Rockinger, and Jonathan Tawn, 2004, Extreme value dependence in financial markets: Diagnostics, models and financial implications, *Review of Financial Studies* 17, 581–610.

Reitz, Thomas A., 1988, The equity risk premium: A solution, *Journal of Monetary Economics* 22, 117–131.

Rubinstein, Mark, 1994, Implied binomial trees, *Journal of Finance* 49, 771–818.

Rubinstein, Mark E., 1973, The fundamental theorem of parameter-preference security valuation, *Journal of Financial and Quantitative Analysis* 8, 61–69.

Sibuya, Masaaki, 1960, Bivariate extreme statistics, *Annals of the Institute of Statistical Mathematics* 11, 195–210.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.