Downside Risk*

Andrew Ang  
Columbia University and NBER

Joseph Chen  
University of Southern California

Yuhang Xing  
Rice University

This Version: 18 December 2003

JEL Classification: C12, C15, C32, G12
Keywords: asymmetric risk, cross-sectional asset pricing, downside risk, first-order risk aversion, higher-order moments

*This paper is a substantial revision of an earlier paper titled “Downside Correlation and Expected Stock Returns”. The authors thank Brad Barber, Geert Bekaert, Alon Brav, John Cochrane, Randy Cohen, Kent Daniel, Bob Dittmar, Rob Engle, Wayne Ferson, David Hirschleifer, N. Jegadeesh, Jonathan Lewellen, Qing Li, Terence Lim, Toby Moskowitz, Ľuboš Pástor, Akhtar Siddique, Rob Stambaugh and Zhenyu Wang. We especially thank Cam Harvey (the editor) and Bob Hodrick for detailed comments. We thank seminar participants at Columbia University, Koc University, NYU, USC, the European Finance Association, the Five Star Conference, an NBER Asset Pricing Meeting, the Texas Finance Festival, and the Western Finance Association for helpful discussions. We thank two referees whose comments greatly improved the paper. The authors acknowledge funding from a Q-Group research grant. Andrew Ang: aa610@columbia.edu; Joe Chen: joe.chen@marshall.usc.edu; Yuhang Xing: yx35@columbia.edu.
Abstract

Agents that place greater weight to the risk of losses on the downside than they are attracted to upside gains demand greater compensation for holding stocks with high downside risk. We show that the cross-section of stock returns reflects a premium for downside risk. Stocks that covary stronger with market declines than with market rises have high average returns. Specifically, stocks with high downside risk command a risk premia of approximately 6% per annum per unit of downside beta. We show that this compensation for bearing downside risk is not simply compensation for market beta. Moreover, the reward for downside risk is not subsumed by coskewness or liquidity risk, and is robust to controlling for momentum and other cross-sectional effects.
1 Introduction

If an asset tends to move downward in a declining market more so than it moves upward in a rising market, it is an unattractive asset to hold because it tends to have very low payoffs precisely when investors would like to receive high returns. Investors who are particularly sensitive to downside losses, relative to upside gains, ought to demand higher premium to hold these assets. This causes assets with high sensitivities to downside movements in the market to have high average returns in equilibrium. In this article, we show that the cross-section of expected returns reflects a premium for bearing downside risk. In contrast, we fail to find a significant discount for stocks that have high covariation with upside movements of the market.

The reason that stocks with large amounts of downside risk have high average returns is intuitive. As early as Roy (1952), economists have recognized that investors care differently about downside risk, than they care about upside risk. For example, Markowitz (1959) advocates using semi-variance as a measure of risk, rather than variance, because semi-variance weights upside gains differently from downside losses. More recently, the behavioral framework of Kahneman and Tversky’s (1979) loss aversion and the axiomatic approach taken by Gul’s (1991) disappointment aversion preferences allow agents to place larger weights on losses relative to gains in their utility functions. The high average returns for stocks that covary stronger with market declines than with market rises is consistent with agents adverse to downside risk demanding greater compensation for bearing such risk.

According to the Capital Asset Pricing Model (CAPM), a stock’s expected excess return is proportional to its market beta, which is constant across down-markets and up-markets. A natural extension of the CAPM that takes into account the asymmetric treatment of risk is to specify separate downside and upside betas, as Bawa and Lindenberg (1977) suggest. Downside (upside) betas are computed over periods where the excess market return is below (above) its mean. However, despite the intuitive appeal of downside risk with the way individuals actually perceive risk, there has been a dearth of empirical research into how downside risk is priced in the cross-section of stocks returns.

The paucity of research on a downside risk premium may be due to weak improvements over the CAPM that early researchers found by allowing betas to differ across the downside and upside. For example, in testing a model with downside betas, Jahankhani (1976) fails to find any improvement over the traditional CAPM. Harlow and Rao (1989) find more support for downside and upside betas, but they only evaluate downside risk relative to the CAPM using a maximum likelihood framework (see Black, Jensen and Scholes, 1972) that tests whether the return on the zero-beta asset is the same across all assets. Harlow and Rao do not directly
demonstrate that assets that covary more with the market, conditional on market downturns, have higher average returns.\(^1\) Hence, it is not surprising that the more recent multi-factor or APT models developed to account for the failure of the CAPM, like the Fama and French (1993) three-factor model, emphasize the addition of different factors, rather than incorporating asymmetry in the factor loadings across down-markets and up-markets.

Our strategy for finding a premium for bearing downside risk in the cross-section is as follows. First, we directly show, at the individual stock level, that higher downside beta stocks have higher average returns. Second, we claim that downside beta is a risk attribute because stocks that have high covariation with the market when the market declines exhibit high average returns over the same period. This contemporaneous relationship between a beta and a risk premium is the main underpinning of a cross-sectional risk-return relationship, and has been exploited from the earliest tests of the CAPM (see, among others, Black, Jensen and Scholes, 1979; Gibbons, 1982). Fama and French (1992) also seek, but fail to find, a relationship between post-formation betas and realized average stock returns.

Third, we differentiate the reward for holding high downside risk stocks from other known cross-sectional effects. In particular, since the coskewness measure of Harvey and Siddique (2000) has the potential to capture downside covariation, we are especially careful to control for coskewness risk in assessing the premium for downside beta. Rubinstein (1973), Friend and Westerfield (1980), Kraus and Litzenberger (1976 and 1983), and Harvey and Siddique (2000) show that agents dislike stocks with negative coskewness, so that stocks with low coskewness tend to have high average returns. The risk from downside beta is different from coskewness risk because downside beta explicitly conditions for market downside movements in a non-linear fashion, while coskewness is an unconditional measure. We also control for the standard list of known cross-sectional effects, including size and book-to-market (Fama and French, 1993; Daniel and Titman, 1997), liquidity risk (Pástor and Stambaugh, 2003), and past returns (Jegadeesh and Titman, 1993). Controlling for these and other cross-sectional effects, we estimate that the cross-sectional premium for bearing one unit of downside beta risk is approximately 6% per annum.

Finally, we check if past downside betas have the ability to predict future expected returns. We find that, for the majority of the cross-section, high past downside beta predicts high future

---

\(^1\) In a related approach, Pettengill, Sundaram and Mathur (1995) and Isakov (1999) estimate the CAPM by splitting the full sample into two subsamples that consist of observations where the realized excess market return is positive or negative. Naturally, they estimate a positive (negative) market premium for the subsample with positive (negative) excess market returns. In contrast, our approach examines premiums for asymmetries in the factor loadings, rather than estimating factor models on different subsamples.
returns over the next month, similar to the contemporaneous relationship between realized downside beta and realized average returns. However, high past downside beta fails to predict high average returns for stocks with very high volatility. This is due to the anomalously low returns on stocks with high volatility (see Ang et al., 2003). Fortunately, the proportion of the market where the past downside beta predictive relationship fails is small, at less than 4% in terms of market capitalizations. We also find that the predictive power of past coskewness for future returns is not picking up exposure to downside risk. Hence, past downside beta and past coskewness are capturing different risk components.

The rest of this paper is organized as follows. In section 2, we present a simple setting to show how downside risk is priced in the cross-section. The equilibrium setting uses a representative agent with a kinked disappointment aversion utility function of Gul (1991) that places larger weight on outcomes on the downside. Section 3 demonstrates that stocks with higher downside betas have higher average returns over the same period that covary strongly with declining market returns. In Section 4 we show that past downside beta also cross-sectionally predicts returns for the vast majority of stocks. However, the predictive downside beta relation breaks down for stocks with very high volatility, which have very low returns. Section 5 concludes.

2 A Simple Model of Downside Risk

In this section, we show how downside risk may be priced cross-sectionally in an equilibrium setting. Specifically, we work with a rational disappointment aversion (DA) utility function that embeds downside risk following Gul (1991). Our goal is to provide a simple motivating example of how a representative agent with a larger aversion to losses, relative to his attraction to gains, gives rise to cross-sectional prices that embed compensation for downside risk.

Our simple approach does not rule out other possible avenues for how downside risk may be priced in the cross-section. For example, Shumway (1997) develops an equilibrium behavioral model based on loss averse investors. Barberis and Huang (2001) use a loss averse utility function, combined with mental accounting, to construct a cross-sectional equilibrium. However, they do not relate expected stock returns to direct measures of downside risk. Aversion to downside risk also arises in models with constraints that bind only in one direction, for example, binding short-sales constraints (Chen, Hong and Stein, 2002) or wealth constraints (Kyle and Xiong, 2001).

Rather than considering models with one-sided constraints or agents with behavioral biases,
we treat asymmetries in risk in a rational representative agent framework. The advantage of treating asymmetric risk in a rational framework is that the disappointment utility function is globally concave, whereas optimal finite portfolio allocations for loss aversion utility may not exist (see Ang, Bekaert and Liu, 2002). Our example with disappointment utility differs from previous studies, because existing work with Gul (1991)’s first order risk aversion utility concentrates on the equilibrium pricing of downside risk for only the aggregate market, usually in a consumption setting (see, for example, Bekaert, Hodrick and Marshall, 1997; Epstein and Zin, 1990 and 2001; Routledge and Zin, 2003).

Gul (1991)’s disappointment utility is implicitly defined by the following equation:

\[
U(\mu W) = \frac{1}{K} \left( \int_{-\infty}^{\mu W} U(W)dF(W) + A \int_{\mu W}^{\infty} U(W)dF(W) \right)
\]

where \(U(W)\) is the felicity function over end-of-period wealth \(W\), which we choose to be power utility. That is \(U(W) = W^{(1-\gamma)}/(1 - \gamma)\), where \(0 < A \leq 1\) is the coefficient of disappointment aversion. Furthermore, \(F(\cdot)\) is the cumulative distribution function for wealth, \(\mu_W\) is the certainty equivalent (the certain level of wealth that generates the same utility as the portfolio allocation determining \(W\)) and \(K\) is a scalar given by:

\[
K = Pr(W \leq \mu_W) + APr(W > \mu_W).
\]

Outcomes above (below) the certainty equivalent \(\mu_W\) are termed “elating” (“disappointing”) outcomes. If \(0 < A < 1\), then the utility function (1) down-weights elating outcomes relative to disappointing outcomes. Put another way, the disappointment averse investor cares more about downside versus upside risk. If \(A = 1\), disappointment utility reduces to the special case of standard CRRA utility, which is closely approximated by mean-variance utility.

To illustrate the effect of downside risk on the cross-section of stock returns, we work with two assets \(x\) and \(y\). Asset \(x\) has three possible payoffs \(u_x, m_x\) and \(d_x\), and asset \(y\) has two possible payoffs \(u_y\) and \(d_y\). These payoffs are in excess of the risk-free payoff. Our set-up has the minimum number of assets and states required to examine cross-sectional pricing (the expected returns of \(x\) and \(y\) relative to each other and to the market portfolio, which consists of \(x\) and \(y\)), and to incorporate higher moments (through the three states of \(x\)). The full set of payoffs and states are given by:
The optimal portfolio weight for a DA investor is given by the solution to:

$$\max_{w_x, w_y} U(\mu_W)$$

where the certainty equivalent is defined in equation (1), $w_x$ ($w_y$) is the portfolio weight of asset $x$ ($y$), and end of period wealth $W$ is given by:

$$W = R_f + w_x x + w_y y,$$

where $R_f$ is the gross risk-free rate. An equilibrium is characterized by a set of asset payoffs, corresponding probabilities, and a set of portfolio weights so that equation (3) is maximized and the representative agent holds the market portfolio ($w_x + w_y = 1$) with $0 < w_x < 1$ and $0 < w_y < 1$.

The equilibrium solution even for this simple case is computationally non-trivial because the solution to the asset allocation problem (3) entails solving both the certainty equivalent $\mu_W$ and the portfolio weights $w_x$ and $w_y$ simultaneously. We extend a solution algorithm for (3) developed by Ang, Bekaert and Liu (2002) to multiple assets. Appendix A describes our solution method and details the values used in the calibration. Computing the solution is particularly challenging because for certain parameter values, equilibrium cannot exist because non-participation may be optimal for low $A$ under DA utility. This is unlike the case of standard CRRA, where holding risky assets with strictly positive risk premia is always optimal.

In this simple model, the standard beta with respect to the market portfolio ($\beta$) is not a sufficient statistic to describe the risk-return relationship of an individual stock. In our calibration, an asset’s expected returns increase with $\beta$, but $\beta$ does not fully reflect all risk. This is because the representative agent cares in particular about downside risk, through $A < 1$. Hence, measures of downside risk have explanatory power for describing the cross-section of expected returns. One measure of downside risk introduced by Bawa and Lindenberg (1977) is the downside beta $\beta^-$:

$$\beta^- = \frac{\text{cov}(r_i, r_m | r_m < \mu_m)}{\text{var}(r_m | r_m < \mu_m)},$$

where $r_i$ and $r_m$ are the returns of the individual asset and the market, respectively, and $\mu_m$ is the expected return of the market.
where \( r_i \) (\( r_m \)) is security \( i \)'s (the market’s) excess return, and \( \mu_m \) is the average market excess return. We also compute a relative downside beta \((\beta^- - \beta)\), relative to the standard CAPM \( \beta \), where \( \beta = \text{cov}(r_i, r_m) / \text{var}(r_m) \).

Figure 1 shows various risk-return relationships holding in our DA cross-sectional equilibrium. The mean excess return increases with \( \beta^- \). We define the CAPM \( \alpha \) as the excess return of an asset not accounted for by the asset’s beta, \( \alpha = E(r_i) - \beta E(r_m) \). The CAPM \( \alpha \) is also increasing with \( \beta^- \) or \((\beta^- - \beta)\). Hence, higher downside risk is remunerated by higher expected returns.

The bottom right-hand panel of Figure 1 plots the CAPM \( \alpha \) versus coskewness, defined as:

\[
\text{coskew} = \frac{E[(r_i - \mu_i)(r_m - \mu_m)^2]}{\sqrt{\text{var}(r_i)\text{var}(r_m)}},
\]

where \( \mu_i \) is the average excess return of asset \( i \). Harvey and Siddique (2000) advocate the use of coskewness as a way to investigate the relationships of expected returns with higher moments and find that since agents dislike negative coskewness (see, among others, Rubinstein, 1973 and Kraus and Litzenberger, 1976). They predict that lower coskewness should be associated with higher expected return. The coskewness measure can be motivated by a third-order Taylor expansion of a general Euler equation:

\[
E_t \left[ \frac{U'(W_{t+1})}{U'(W_t)} r_{i,t+1} \right] = 0,
\]

where \( W \) is the total wealth of the representative agent, and \( U'(\cdot) \) can be approximated by:

\[
U' = 1 + WU''r_m + \frac{1}{2} W^2 U'''r_m^2 + \ldots .
\]

The Taylor expansion in (8) is necessarily only an approximation. In particular, since the DA utility function is kinked, polynomial expansions of \( U \), such as the expansions used by Bansal, Hsieh and Viswanathan (1993), may not be good global approximations if the kink is large (or \( A \) is very small).\(^2\) Nevertheless, measures like coskewness based on the Taylor approximation for the utility function should also have some explanatory power for returns.

Downside beta and coskewness may potentially capture different effects. Note that for DA utility, both downside beta and coskewness are approximations because the utility function does not have an explicit form (equation (1) implicitly defines DA utility). Since DA utility is kinked at an endogenous certainty equivalent, skewness and other centered moments may not effectively capture aversion to risk across upside and downside movements in all situations.

\(^2\)Taylor expansions have been used to account for potential skewness and kurtosis preferences in asset allocation problems by Guidolin and Timmerman (2002), Jondeau and Rockinger (2002), and Harvey et al., (2003).
This is because they are based on unconditional approximations to a non-smooth function. In contrast, the downside beta (5) conditions directly on a downside event, that the market return is less than its unconditional mean. In Figure 1, our calibration shows that lower coskewness is compensated by higher expected returns. However, the Appendix describes a case where CAPM \( \alpha \)'s may increase as coskewness increases, the opposite predicted by the Taylor expansion.

Our equilibrium example shows how compensation to downside risk may arise in equilibrium. Of course, the example, having only two assets, is necessarily simplistic. Nevertheless, the example provides motivation to ask if downside risk demands compensation in the cross-section of US stocks, and if such compensation is different in nature from compensation for risk based on measures of higher moments (such as the Harvey-Siddique, 2000, coskewness measure). As our example shows, this compensation for downside risk is in addition to the reward already incurred for standard, unconditional risk exposures, such as exposure to the market factor. In our empirical work, we investigate a premium for downside risk also controlling for other known cross-sectional effects such as the size and book-to-market effect of Fama-French (1992), the liquidity effect of Pástor and Stambaugh (2003), and the momentum effect of Jegadeesh and Titman (1993).

3 Downside Risk and Realized Returns

In this section, we document that stocks that covary strongly with the market, conditional on down moves of the market, have high average returns over the same period. We document this by looking at patterns of realized returns for portfolios sorted on downside risk in Section 3.1. In Section 3.2, we examine the reward to downside risk controlling for other cross-sectional effects by using Fama-MacBeth (1973) regressions. Section 3.3 conducts robustness tests, while we specifically disentangle the different effects of coskewness risk and downside beta exposure in Section 3.4.

In our empirical work, we concentrate on presenting the results of equal-weighted portfolios. While a relationship between factor sensitivities and returns should hold for both an average stock (equal-weighting) or an average dollar (value-weighting), we focus on computing equal-weighted portfolios because past work on examining non-linearities in the cross-section has found risk due to asymmetries to be bigger among smaller stocks. For example, the coskewness effect of Harvey and Siddique (2000) is strongest for equal-weighted portfolios.

Nevertheless, we also examine the robustness of our findings using value-weighted portfolios. We also concentrate only on NYSE stocks to minimize the illiquidity effects of small
firms, but also consider all stocks on NYSE, AMEX and NASDAQ in robustness tests.

3.1 Unconditional, Downside, and Upside Betas

If there is a cross-sectional relation between risk and return, then we should observe patterns between average realized returns and the factor loadings associated with exposure to risk. For example, a standard unconditional CAPM implies that stocks that covary strongly with the market should have contemporaneously high average returns over the same period. In particular, the CAPM predicts an increasing relationship between realized average returns and realized factor loadings, or contemporaneous expected returns and market betas. More generally, an unconditional multi-factor model implies that we should observe patterns between average returns and sensitivities to different sources of risk.

In Table 1, we investigate patterns between realized average returns and realized betas. We work in intervals of twelve months, from \( t \) to \( t + 12 \). At the end of the year, \( t + 12 \), we compute a stock’s beta \( \beta \), downside beta \( \beta^- \), and upside beta \( \beta^+ \). The downside beta \( \beta^- \) is described in equation (5), while the upside \( \beta^+ \) takes the same form as equation (5), except we condition on movements of the market excess return above its average value:

\[
\beta^+ = \frac{\text{cov}(r_i, r_m | r_m > \mu_m)}{\text{var}(r_m | r_m > \mu_m)}. \tag{9}
\]

At the beginning of the year, at time \( t \), we sort stocks into five quintiles based on their \( \beta \), \( \beta^- \) or \( \beta^+ \) over the next twelve months. In the column labeled 'Return,' Table 1 reports the average realized excess return from \( t \) to \( t + 12 \) in each equally-weighted quintile portfolio. The table also reports the cross-sectional average realized \( \beta \), \( \beta^- \) or \( \beta^+ \) of each quintile portfolio. These average returns and betas are computed over the same 12-month period. Hence, Table 1 shows relationships between contemporaneous factor loadings and returns. Although we use a 1-year horizon, we evaluate 12-month returns at a monthly frequency. This use of overlapping information is more efficient, but induces moving average effects. To adjust for this, we report t-statistics of differences in average excess returns between quintile portfolio 5 (high betas) and quintile portfolio 1 (low betas) using 12 Newey-West (1987) lags. The sample period is from July 1963 to December 2001, with our last twelve-month return period starting in January 2001. As part of our robustness checks, we also examine non-overlapping sample periods.

Panel A of Table 1 shows a monotonically increasing pattern between realized average returns and realized \( \beta \). Quintile 1 (5) has an average excess return of 3.5% (13.9%) per annum, and the spread in average excess returns between quintile portfolios 1 and 5 is 10.4% per annum, with a corresponding difference in contemporaneous betas of 1.36. Our results are consistent
with the earliest studies testing the CAPM, like Black, Jensen and Scholes (1972), who find a reward for holding higher beta stocks. However, this evidence per se does not mean that the CAPM holds, because the CAPM predicts that no other variable other than beta should explain a firm’s expected return. Nevertheless, it demonstrates that bearing high market risk is rewarded with high average returns. Panel A also reports the positive and negative components ($\beta^-$ and $\beta^+$) of beta. By construction, higher $\beta^-$ or higher $\beta^+$ must also mean higher unconditional $\beta$, so high average returns are accompanied by high $\beta^-$, $\beta^+$ and regular $\beta$. In the remainder of the panels in Table 1, we decompose the reward for market risk into downside and upside components.

Panel B shows that stocks with high contemporaneous $\beta^-$ have high average returns. Stocks in the quintile with the lowest (highest) $\beta^-$ earn 2.7% (14.5%) per annum above the risk-free rate. The average difference between quintile portfolio 1 and 5 is 11.8%, which is statistically significant at the 1% level. These results are consistent with agents disliking downside risk and avoiding stocks that covary strongly when the market dips. Hence, stocks with high $\beta^-$ must carry a premium in order to entice agents to hold them. However, a second explanation is that agents have no particular emphasis on downside risk versus upside risk, and high $\beta^-$ stocks earn high returns simply because, by construction, high $\beta^-$ stocks have high $\beta$. The average $\beta^-$ spread between quintile portfolios 1 and 5 is very large (0.19 to 1.92), but sorting on $\beta^-$ also produces variation in $\beta$ and $\beta^+$. However, the variation in $\beta$ and $\beta^+$ is not as disperse as the variation in $\beta^-$. In Panels C and D, we demonstrate that it is the reward for downside risk alone that is behind the pattern of high $\beta^-$ stocks earning high returns.

Panel C shows a smaller spread (relative to the spreads for $\beta$ and $\beta^-$ in Panels A and B) for average realized excess returns for stocks sorted on realized $\beta^+$. We find that low (high) $\beta^+$ stocks earn, on average, 5.7% (9.8%) per annum in excess of the risk-free rate. This pattern of high returns to high $\beta^+$ loadings is inconsistent with agents having strong preferences for upside risk. Since $\beta^+$ only measures exposure to a rising market, stocks that rise more when the market return increases should be more attractive and, on average, earn low returns. We do not observe this pattern. Instead, the increasing pattern of returns in Panel C may be due to the patterns of $\beta$ or $\beta^-$, which increase from quintile portfolios 1 to 5. From the CAPM, high $\beta$ implies high returns, and if agents dislike downside risk, high $\beta^-$ also implies high returns.

Finally, in Panel D, we sort stocks by realized relative downside beta, defined as:

$$\text{Relative } \beta^- = (\beta^- - \beta).$$

Relative downside beta controls for the effect of upside exposure. Panel D shows that stocks with high realized $\beta^-$ have high average returns. The difference in average excess returns
between portfolio 1 and 5 is 6.6% per annum and is highly significant. We can rule out that this pattern of returns is attributable to unconditional beta. The average $\beta$ of each quintile portfolio has a fairly flat pattern, with a slight U-shape. Hence, by considering relative $\beta^-$, we produce a series of portfolios with little spread in $\beta$. By construction, stocks with high relative downside betas must have either high $\beta^-$, or low $\beta^+$, or both. Low upside beta, $\beta^+$, by itself, causes stocks to have low returns from Panel C. Hence, the high realized returns from high relative $\beta^-$ is produced by the exposure to downside risk, measured by high $\beta^-$ loadings.

In summary, Table 1 demonstrates that downside risk is rewarded in the pattern of cross-sectional returns. Stocks with high $\beta^-$ loadings earn high average returns over the same period that is not mechanically driven by high unconditional betas. In contrast, stocks that covary strongly with the market conditional on positive moves of the market do not command significant discounts.

### 3.2 Fama-MacBeth Regressions

While Table 1 establishes a relationship between $\beta^-$ and average returns, it does not control for the effects of other known patterns in the cross-section of stock returns. A long literature from Banz (1981) onwards has shown that various firm characteristics also have explanatory power in the cross-section. The size effect (Banz, 1981), the book-to-market effect (Basu, 1982), the momentum effect (Jegadeesh and Titman, 1991), exposure to coskewness risk (Harvey and Siddique, 2000), exposure to cokurtosis risk (Dittmar, 2002), and exposure to aggregate liquidity risk (Pástor and Stambaugh, 2003), all imply different patterns for the cross-section of expected returns. We now demonstrate that downside risk is different from all of these effects by performing a series of cross-sectional Fama and MacBeth (1973) regressions, over the sample period from July 1963 to December 2001.

We run Fama-MacBeth regressions of 12-month excess returns on firm characteristics and realized betas with respect to various sources of risk. Since the regressions are run with a 12-month horizon, but at a monthly frequency, we compute the standard errors of the coefficients by using 12 Newey-West (1987) lags. Table 2 reports the results listed by various sets of independent variables in regressions I-VI. We regress realized firm returns at time $t$ to $t + 11$ on realized market beta, downside beta and upside beta, $(\beta, \beta^-, \text{and } \beta^+)$ computed over the same period. Hence, these regressions demonstrate a relationship between contemporaneous returns and betas. We control for the log-size, the book-to-market ratio, and the past 12-month excess returns of the firm at the beginning of the period $t$. We also include as control variables: the realized standard deviation of the firm excess returns, the coskewness (equation (6)), and the
cokurtosis. All of these are also computed over the period from \( t \) to \( t + 12 \). We define cokurtosis in a similar manner to the coskewness in equation (6) (see Dittmar, 2002):

\[
cokurt = \frac{E[(r_i - \mu_i)(r_m - \mu_m)^3]}{\sqrt{\text{var}(r_i) \text{var}(r_m)^{3/2}}},
\]

(11)

where \( r_i \) is the firm excess return, \( r_m \) is the market excess return, \( \mu_i \) is the average excess stock return and \( \mu_m \) is the average market excess returns. Scott and Horvath (1980) and Dittmar (2002) argue that stocks with high cokurtosis should have low returns because investors dislike even-numbered moments. Finally, we also include the Pástor and Stambaugh (2003) historical liquidity beta at time \( t \) to proxy for liquidity exposure.

In order to avoid putting too much weight on extreme observations, we Winsorize all independent variables at the 1% and 99% levels.\(^3\) Winsorization has been performed in cross-sectional regressions by Knez and Ready (1997), among others, and ensures that extreme outliers do not drive the results. It is particularly valuable for dealing with the book-to-market ratio, because extremely large book-to-market values are sometimes observed due to low prices, particularly before a firm delists.

We begin with Regression I in Table 2 to show the familiar, standard set of cross-sectional return patterns. While beta carries a positive coefficient, the CAPM is overwhelmingly rejected: beta is not a sufficient statistic to explain the cross-section of stock returns. Small stocks and stocks with high book-to-market ratios also cause a firm to have high average returns, confirming Fama and French (1992). The coefficient on past returns is positive (0.017) but only weakly significant. The very large and highly significant negative coefficient (-8.43) on the firm’s realized volatility of excess returns confirms the anomalous finding of Ang et al. (2003), who find that stocks with high return volatilities have low average returns. Consistent with Harvey and Siddique (2000), stocks with high coskewness have low returns. Cokurtosis is only weakly priced, with a small and insignificant coefficient of 0.015.

In Regressions II-VI, we examine the downside and upside components of beta separately and shows that downside risk is priced. We turn first to Regression II, which reveals that downside risk and upside risk are treated asymmetrically. The coefficient on downside risk is positive (0.069) and highly significant, confirming the portfolio sorts of Table 1. The coefficient on \( \beta^+ \) is negative (-0.029) but less in magnitude than the coefficient on \( \beta^- \). We did not obtain a

\(^3\) For example, if an observation for the firm’s book-to-market ratio is extremely large and above the 99th percentile of all the firms’ book-to-market ratios that month, we replace that firm’s book-to-market ratio with the book-to-market ratio corresponding to the 99th percentile. The same is done for firms whose book-to-market ratios lie below the 1%-tile of all firms’ book-to-market ratios that month.
negative reward for $\beta^+$ in Table 1, Panel C, because in Table 1, we did not control for the effect of downside and upside beta separately.

Regression III shows that the reward for both downside and upside risk is robust to controlling for size, book-to-market and momentum effects. The Fama-MacBeth coefficients for $\beta^-$ and $\beta^+$ remain almost unchanged from their Regression II estimates at 0.064 and -0.025, respectively. However, once coskewness risk is accounted for in Regression IV, the coefficient on $\beta^+$ becomes very small (0.003) and becomes statistically insignificant, with a t-statistic of 0.22. Controlling for coskewness also brings down the coefficient on $\beta^-$ from 0.064 in Regression III (without the coskewness risk) to 0.028. Nevertheless, the premium for downside risk remains positive and statistically significant. In Regression V, where we add controls for realized firm volatility and realized firm cokurtosis, the coefficient on downside risk remains consistently positive at 0.062, and also remains highly statistically significant.

Finally, Regression VI investigates the reward for downside and upside risk controlling for the full list of firm characteristics and realized factor loadings. We lose five years of data in constructing the Pástor-Stambaugh historical liquidity betas, so this regression is run from January 1967 to December 2001. The coefficient on $\beta^-$ is 0.056, with a robust t-statistic of 5.25. In contrast, the coefficient on $\beta^+$ is statistically insignificant, whereas the premium for coskewness is significantly negative, at -0.188. Importantly, while $\beta^-$ and coskewness risk measure downside risk, the coefficients on both $\beta^-$ and coskewness are statistically significant. This shows that the risk exposure of downside beta is different from coskewness risk.

The consistent message from these regressions is that the reward for downside risk ($\beta^-$ is always positive and statistically significant. High downside beta leads to high average returns, and this result is robust to controlling for other firm characteristics and risk characteristics. Downside beta risk remains significantly positive in the presence of coskewness risk controls. On the other hand, the reward for upside risk ($\beta^+$) is fragile. A priori, we expect the coefficient on $\beta^+$ to be negative, but it often flips sign and is insignificant controlling for other variables.

### 3.3 Robustness

In Table 3, we subject our results to a battery of robustness checks. We report the robustness checks for realized $\beta^-$ in Panel A and for realized relative $\beta^-$ in Panel B. In each panel, we report average 12-month excess returns of quintile portfolios sorted by realized $\beta^-$, or realized relative $\beta^-$, over the same 12-month period. The table reports the differences in excess returns
between quintile portfolios 1 and 5 with robust t-statistics.\textsuperscript{4}

The first column shows that value-weighting the portfolios preserves the large spreads in average returns for sorts by $\beta^-$ and relative $\beta^-$. In Table 1, the 5-1 spread in average returns from equal-weighting the $\beta^-$ portfolios is 11.8% per annum, which reduces to 7.1% per annum when the portfolios are value-weighted. Similarly, the 5-1 spread in relative $\beta^-$ portfolios in Panel B reduces from 6.6% to 4.0%. In both cases, the spreads remain highly significant at the 1% level. Thus, while the effect of value-weighting reduces the effect of downside risk, it does not remove it.

In the second column, we use all stocks listed on NYSE, AMEX and NASDAQ, rather than restricting ourselves to stocks listed on the NYSE. We form equal-weighted quintile portfolios at the beginning of the period, ranking on realized betas, using breakpoints calculated over NYSE stocks. Because we include all stocks, many of which are small and illiquid, we also control for non-synchronous trading using a Scholes-Williams (1977) correction to compute the betas. We describe this correction in Appendix B. Naturally, using all stocks increases the average excess returns, as many of the newly included stocks are small. The 5-1 spreads in average returns also increase substantially. For the quintile $\beta^-$ (relative $\beta^-$) portfolios, the 5-1 difference becomes 15.2% (8.6%). By limiting our universe to NYSE stocks, we deliberately understate our results to avoid confounding influences of illiquidity and non-synchronous trading.

One concern about the 12-month horizon of Tables 1 and 2 is that they use overlapping observations. While this is statistically efficient, we examine the effect of using non-overlapping 12-month periods in the last column of Table 3. Our 12-month periods are from the beginning of January to the end of December each calendar year. With non-overlapping samples, it is not necessary to control for the moving average errors with robust t-statistics, but we have fewer observations. Nevertheless, the results show that the point estimates of the 5-1 spreads are still statistically significant at the 1% level. Not surprisingly the point estimates remain roughly unchanged. Hence, our premium for downsize beta is robust to value-weighting, using the whole stock universe and using non-overlapping observations.

\textsuperscript{4} We also examined robustness of the horizon period, but do not report the results to save space. If we examine realized betas and realized returns over a 60-month horizon using monthly frequency returns, we find the same qualitative patterns as using a 12-month horizon and the 5-1 difference in average returns remain statistically significant at the 5% level.
3.4 Downside Beta Risk and Coskewness Risk

While the Fama-MacBeth (1973) regressions in Section 3.2 demonstrate that both downside beta and coskewness have predictive power for the cross-section. Since both $\beta^-$ and coskewness measure asymmetric higher moments, we now measure the magnitude of the reward for exposure to downside beta, while explicitly controlling for the effect of coskewness. Table 4 presents the results of this exercise.

To control for the effect of coskewness, we first form 5 portfolios sorted on coskewness. Then, within each coskewness quintile, we sort stocks into five portfolios based on $\beta^-$. These portfolios are equally-weighted and both coskewness and $\beta^-$ are computed over the same 12-month horizon for which we examine realized excess returns. After forming the $5 \times 5$ coskewness and $\beta^-$ portfolios, we average the realized excess returns of each $\beta^-$ quintile over the five coskewness portfolios. This characteristic control procedure creates a set of quintile $\beta^-$ portfolios with near-identical levels of coskewness risk. Thus, these quintile $\beta^-$ portfolios control for differences in coskewness.

Panel A of Table 4 reports average excess returns of the 25 coskewness $\times \beta^-$ portfolios. The column labeled “Average” reports the average excess returns of the $\beta^-$ quintiles controlling for coskewness risk. The row labeled “High-Low” reports the differences in average returns between the first and fifth quintile $\beta^-$ portfolios within each coskewness quintile. The last row reports the 5-1 quintile difference for the $\beta^-$ quintiles, which control for the effect of coskewness exposure. The average return of 7.6% per annum in the bottom right entry of Panel A is the 5-1 difference in average returns between stocks with low and high $\beta^-$, controlling for coskewness risk. This difference has a robust t-statistic of 4.16. Hence, coskewness risk cannot account for the reward for bearing downside risk.

In Panel A, the patterns within each coskewness quintile (reading down each column) are very interesting. As coskewness increases, the differences in excess returns due to $\beta^-$ decrease. Stocks with low coskewness risk have very high spreads in excess returns due to different $\beta^-$ loadings. In the first coskewness quintile, the 5-1 difference in the $\beta^-$ quintiles average returns is 14.6% per annum. The 5-1 $\beta^-$ difference decreases to 2.1% per annum in the quintile of stocks with the highest coskewness.

The reason for this pattern is as follows. As defined in equation (6), coskewness is effectively the covariance of a stock’s return with the square of the market return, or with the volatility of the market. A stock with negative coskewness tends to have low returns when market volatility is high. These are also usually, but not always, periods of low market returns. Volatility of the market treats upside and downside risk symmetrically, so both extreme upside
and extreme downside movements of the market have the same volatility. Hence, stocks with large negative coskewness tend to decrease when the market falls, but also decrease when the market rises. In contrast, downside beta concentrates only on the former effect by explicitly considering the downside case. When coskewness is low, there is a wide spread in $\beta^-$ because there is large scope for market volatility to represent both large negative and large positive changes. This explains the large spread in average returns across the $\beta^-$ quintiles for stocks with low coskewness.

The small 2.1% 5-1 spread for the $\beta^-$ quintiles for the highest coskewness stocks is because these stocks already exhibit little asymmetry. The distribution of coskewness across stocks is skewed towards the negative side and is negative on average. Across the low to high quintiles of coskewness in Panel A, the average coskewness ranges from -0.41 to 0.09. Hence, highest quintile coskewness stocks have low coskewness and their stock returns essentially do not change their behavior across periods of stable or volatile market returns. In this case, the range of $\beta^-$ is also smaller. The small range of $\beta^-$ for the highest coskewness stocks explains the low 2.1% spread for the $\beta^-$ quintiles in the second last column of Panel A.

Panel B of Table 4 repeats the same exercise as Panel A, except we examine the reward for coskewness controlling for different levels of $\beta^-$ . Panel B first sorts stocks on coskewness before sorting on $\beta^-$, and then averages across the $\beta^-$ quintiles. This exercise examines the coskewness premium controlling for downside exposure. Controlling for $\beta^-$, the 5-1 difference in average returns for coskewness portfolios is -6.2% and highly statistically significant. Moreover, there are large and highly statistically significant spreads for coskewness in every $\beta^-$ quintile. Coskewness is able to maintain a high range within each $\beta^-$ portfolio, unlike the diminishing range for $\beta^-$ within each coskewness quintile in Panel A.

4 Predicting Downside Risk

The previous section demonstrates a strong positive relationship between stocks that exhibit high downside risk and returns for holding such stocks over the same period. While this is the essence of the relationship implied by a risk story, a remaining question is whether downside risk is predictable. That is, can we use information about how stocks have covaried with market down movements in the past to predict returns in the future?

In this section, we investigate the predictive relationship between past downside beta and future returns. In particular, we examine if stocks sorted on past downside risk provide ongoing exposure to downside risk in the future and thus compensation for exposure to risk. Section
4.1 examines returns of portfolios of stocks sorted by past $\beta^-$ and past coskewness. Section 4.2 delves deeper into the analysis of stocks for which exposures to downside risk is persistent and maintains the risk-return tradeoff even when we use past risk measures. In Section 4.3, we conclude by conducting robustness tests.

4.1 Past Downside Risk and Future Returns

Table 5 investigates the relationship between past downside risk measures and future returns. We compute a stock’s downside beta and coskewness with respect to the market using daily returns over the previous year. The sample period is from July 1962 to January 2001, with our first twelve-month risk measurement period ending in June 1963 for the portfolio formation in July 1963. At the beginning of each month $t$, we sort stocks into five quintiles based on their past $\beta^-$ and coskewness. In the column labeled “Return,” we report the average realized excess return over the next month from $t$ to $t + 1$. We also report the differences in the excess returns of the quintile 1 and 5 portfolios. The table also reports the cross-sectional realized $\beta$, $\beta^-$ and $\beta^+$ of each quintile portfolios. These realized betas are computed over the following 12-month period. The last column of Table 5 shows the cross-sectional realized coskewness of each quintile portfolios.

Panel A reports a monotonically increasing pattern in the realized betas of stocks sorted on past betas. That is, stocks with low (high) $\beta^-$ in the past continue to have low (high) $\beta^-$ going forward. The difference in the realized $\beta^-$ for portfolio 1 and 5 sorted on past $\beta^-$ is 0.77. Hence, past $\beta^-$ is a good indicator of future $\beta^-$. The 12-month autocorrelation of $\beta^-$ among stocks listed on NYSE is 0.43, which implies a monthly autocorrelation of 0.93, so individual firm $\beta^-$ loadings are also persistent. Panel A also shows that the past variation of $\beta^-$ produces little variation in realized coskewness.

While the ex-post $\beta^-$ loadings maintain their monotonic pattern, Panel A disappointingly shows a very weak relationship between past $\beta^-$ and future returns. We observe a strictly increasing pattern moving from the first quintile $\beta^-$ portfolio to the fourth quintile $\beta^-$ portfolio from 0.59% to 0.84% per month. However, the highest $\beta^-$ quintile portfolio performs has much lower average excess return of 0.70%. The weak pattern for past $\beta^-$ and future returns is similar to the observation made by Fama and French (1992) and Jagannathan and Wang (1996), who find that sorting stocks into 100 portfolios based on past $\beta$ and size generates almost no patterns in future average returns along the past $\beta$ dimension.

In contrast, Panel B shows the relation between past coskewness, future returns, and future

---

5 Similarly, the 12-month autocorrelation of $\beta$ among stocks listed on NYSE is 0.68.
risk attributes. The portfolios of stocks sorted by past coskewness does not exhibit large ex-post variation in coskewness. The average pre-ranking spread in coskewness between portfolios 1 and 5 is 0.50, while the average post-ranking spread in realized coskewness is only 0.05. Nevertheless, we observe a strong, strictly decreasing pattern in the returns of stocks sorted by past coskewness. The portfolio of the most negative coskewness stocks (quintile 1) has an average excess return of 0.84%, whereas the portfolio of the least negative coskewness stocks (quintile 5) has an average excess return of 0.57%. This pattern is consistent with Harvey and Siddique (2000). The strong predictive pattern of past coskewness and future returns does not have a relationship with downside beta. The future $\beta^-$ loadings from past coskewness are almost flat. Hence, the predictive pattern for cross-sectional returns from past coskewness is not picking up $\beta^-$ exposure.

It is puzzling why there is such a strong contemporaneous pattern between $\beta^-$ and average returns but such a weak relationship between past $\beta^-$ and future returns, despite the fact that the pre-ranking $\beta^-$ sorts produce ex-post variation in $\beta^-$. We now investigate why the past $\beta^-$ portfolios seem to fail to produce predictive return variation.

4.2 Can we Predict Downside Risk?

The reason that past downside beta provides exposure to downside risk over the following period, but does not seem to provide investors with compensation for bearing this risk, is due to a strong pattern of decreasing future average returns with increasing stock volatility characteristics. Ang et al. (2003) show that stocks with high volatility (both total or idiosyncratic volatility) have abysmally low future returns. High volatility stocks tend to be high beta stocks because, holding correlation between the market and the stock return constant, a high individual stock volatility must mean a high $\beta$ (since $\beta = \rho_{im}\sigma_i/\sigma_m$). The same is also true for $\beta^-$, so high volatility stocks tend to be have high $\beta^-$ loadings. Ang et al. (2003) volatility effect confounds the expected return pattern of $\beta^-$, making it hard to predict downside risk for stocks with very high volatility.

We illustrate the confounding volatility effect with the downside risk expected return patterns in Figure 2. In Panel A, we show the relationship between contemporaneous $\beta^-$, volatility ($\sigma$), and excess returns. We plot realized $\beta^-$ on the $x$-axis and realized $\sigma$ on the $y$-axis. Realized average excess returns are shown as contour lines, with the light (dark) lines indicating low (high) returns. Each circle represents approximately 0.5% of the number of stocks listed on the NYSE. First, as we move down the $y$-axis, from high to low $\sigma$, the lines become darker. This is the Ang et al. (2003) volatility effect, where stocks with high volatility have low returns.
Second, moving from low $\beta^{-}$ to high $\beta^{-}$ along the $x$-axis, Panel A shows that the contour lines become darker. Hence, high $\beta^{-}$ implies contemporaneously high average returns. In fact, if we average out the $\sigma$ effect, this is just a pictorial representation of the portfolio sorts on realized $\beta^{-}$ in Table 1.

In Panel B of Figure 2, we draw a similar graph to Panel A, except we use past $\beta^{-}$, past $\sigma$, and the contour lines represent next month’s future average excess returns. Again, moving down the $y$-axis from high $\sigma$ to low $\sigma$, the lines become progressively darker, showing the volatility effect. There is a large light area in the upper right hand corner of stocks with high past $\sigma$, high past $\beta^{-}$, and low returns. If we average out the volatility effect (by averaging out $\sigma$ along vertical lines), the high past $\beta^{-}$ levels, from approximately 1.8 to 2.5, must include this region of low returns. It is this volatility effect that confounds the $\beta^{-}$ expected return pattern.

Panel B also contains a long diagonal line emanating from the origin labeled “$\rho^{-} = 1.0$,” which represents the limiting case for a past downside correlation of unity. We can represent lines of different downside correlation by changing the diagonal line, holding fixed the origin, from perfectly horizontal (the $y$-axis), representing zero downside correlation, to the diagonal $\rho^{-} = 1.0$ line. For example, the line labeled “$\rho^{-} = 0.5$” represents the case of a downside correlation of a half. As we sweep the lines through the origin from the $y$-axis to the $\rho^{-} = 1$ line, we pass through regions that become increasingly dark. Notice the lines very close to the $\rho^{-} = 1$ line pass through the the darkest regions, with the highest average future returns. Hence, stocks with high downside correlation have high returns, which is the effect that Ang, Chen and Xing (2002) find. Note that the downside correlation effect holds both predictively (Panel B) and also contemporaneously (Panel A).

While downside correlation is immune from the volatility effect, it is an unsatisfactory measure of risk. Any economic model giving rise to downside risk, such as our model in Section 2 has implications for both magnitude and direction through a beta factor loading. Correlation lacks any sense of magnitude. Second, while downside correlation may produce strong predictive patterns in cross-sectional returns, it is not a coherent risk measure in the sense of Artzner et al. (1999). In particular, the downside correlation of a portfolio of two assets may be greater or less than the the sum of the weighted downside correlations of the two individual assets. Coskewness is also an incoherent risk measure, because coskewness coefficients defined by equation (6) are not additive. The coskewness statistic can be made additive if the normalization in the denominator is changed to $\text{var}(r_{m})^{3/2}$, rather than normalizing using $\sqrt{\text{var}(r_{i})\text{var}(r_{m})}$. 

---

6 In this analysis, we assume that $\sigma = \sigma^{-}$, where $\sigma^{-} = \text{var}(r_{i}|r_{m} < \mu_{m})$. However, the results do not change significantly when we use $\sigma^{-}$.

7 Coskewness is also an incoherent risk measure, because coskewness coefficients defined by equation (6) are not additive. The coskewness statistic can be made additive if the normalization in the denominator is changed to $\text{var}(r_{m})^{3/2}$, rather than normalizing using $\sqrt{\text{var}(r_{i})\text{var}(r_{m})}$.
of its constituent assets, making it appropriate for a performance evaluation (see Pedersen and Satchell, 2000).

Finally, Panel B of Figure 2 enables us to find a large subset of stocks where a predictive $\beta^-$ relationship holds. If we draw a horizontal line around $\sigma = 0.6$ in Panel B and only consider stocks with past volatility below this cut-off, then moving from left to right along the $x$-axis produces increasingly darker contour lines. Hence, if we exclude stocks with the highest levels of past volatility, we can find a large subsample of stocks where past $\beta^-$ predicts returns. We examine the expected return patterns to past $\beta^-$ when we exclude stocks with very high past volatility in Table 6.

We begin by first sorting stocks based on their past total volatility $\sigma$, calculated using daily returns over the past year. We sort stocks into quintiles, octiles, deciles and demi-decile (5%-tile) groups according to past $\sigma$. Panel A of Table 6 excludes stocks that fall into the highest quintile, octile, decile or demi-decile of $\sigma$, and then re sorts the remaining stocks into quintiles according to past $\beta^-$. Panel A reports the average excess returns of these $\beta^-$ quintiles that exclude very volatile stocks. We also report the difference in returns between the lowest and the highest $\beta^-$ portfolios (quintiles 1 and 5), as well as the difference in returns between quintile 1 and 4.

The first column in Panel A excludes stocks in the highest $\sigma$ quintile and shows a clear monotone pattern in the returns of stocks sorted by past $\beta^-$. The 5-1 difference of 0.34% per month is statistically significant with a t-statistic of 2.31. The difference between Quintile 4 portfolio and Quintile 1 portfolio is slightly smaller at 0.25% per month, but has approximately the same statistical significance. If we exclude fewer stocks and only exclude stocks in the highest volatility octile or decile, the return difference between highest $\beta^-$ portfolio and the lowest $\beta^-$ portfolio is about the same order of magnitude (roughly 0.31% per month), but the statistical significance is somewhat weaker, with p-values of 0.051 and 0.063, respectively. If we exclude only the stocks in the highest volatility demi-decile, the return difference between the lowest $\beta^-$ portfolio and the high $\beta^-$ portfolio becomes quite weak (0.22% per month). However, he return difference between quintiles 1 and 4 is unaffected for all the $\sigma$ exclusions at 0.25% per month and is always statistically significant. This is consistent with Panel B of Figure 2, which shows that the volatility effect is strongest among stocks with the highest past $\beta^-$.

Panel B of Table 6 shows how highly volatile stocks confound the predictive relation between past $\beta^-$ and returns. For the overall sample and for each of the highest volatility groups, we report the average market capitalization, volatility, past $\beta^-$, and average returns adjusted for size and book-to-market using a characteristic control similar to Daniel et al. (1997). The 20%
of stocks with the highest volatility constitutes, on average, only 3.9% of the market. Hence, by excluding the highest quintile of volatile stocks, we exclude stocks that represent only a small fraction of the market. Stocks in the highest \( \sigma \) quintile have an average annualized volatility of 61%, which is significantly higher than average annualized volatility of 36% volatility for all stocks. The highest quintile \( \sigma \) stocks have an average past \( \beta^- \) of 1.44, but over the next 12-months, these stocks have an average \( \beta^- \) of only 1.25. In contrast, Table 1 shows stocks within the highest realized \( \beta^- \) quintile have an average \( \beta^- \) of 1.92. Hence, high \( \sigma \) stocks tend to have high past \( \beta^- \), but this does not imply that they continue to exhibit high \( \beta^- \) the following period. Moreover, the average 12-month autocorrelation of \( \beta^- \) for these volatile stocks is 26%, which is much lower than the 12-month autocorrelation for the whole sample, which is 44%. If we narrow our focus to stocks on the highest octile, decile or demi-decile volatility stocks, we find that these stocks tend to be even smaller in terms of market capitalization and exhibit even lower persistence of \( \beta^- \) across the formation period and the holding period.

The final row of Panel B confirms Ang et al. (2003) that high \( \sigma \) stocks have low returns. We report size and book-to-market adjusted returns of stocks in the highest volatility groups, similar to Daniel et al. (1997), and outlined in the Appendix. Stocks in the highest volatility demi-decile underperform their benchmark portfolios by 0.67% per month. The volatility effect that stocks with high volatility have very low returns confounds the predictive \( \beta^- \) relationship of cross-sectional returns. Fortunately, these very volatile stocks constitute a small fraction of the total market capitalizations (up to only 4%), and if we focus on stocks not in the last volatility quintile, we find a strong pattern between \( \beta^- \) and future holding period returns.

### 4.3 Robustness Checks of Predictive Downside Risk

While we establish a predictive relation between \( \beta^- \) and future returns for a large proportion of the market, there remains a concern that past \( \beta^- \) may be proxying for size, book-to-market, momentum, coskewness, or liquidity effects and we are incorrectly attributing the return patterns to downside beta. Our final table ensures that this is not the case. We focus our attention on size and book-to-market adjusted returns, but continue to exclude the most volatile quintile of stocks from our analysis.

The first column of Table 7 shows the average size and book-to-market adjusted returns of portfolios sorted by past \( \beta^- \) after having excluded stocks within the highest volatility quintile. We observe a monotonically increasing relation between past \( \beta^- \) and next month returns.

---

8 Since we use all stocks listed on NYSE, AMEX and NASDAQ in constructing size and book-to-market benchmark portfolios, the average adjusted returns of all stocks listed on NYSE does not sum up to 0.0%. 

20
Controlling for size and book-to-market increases the 5-1 $\beta^-$ difference in returns to 0.44% per month from 0.34% per month in Table 6. This difference is highly statistically significant with a t-statistic of 3.36.

In the next three columns, we control for additional past return characteristics: momentum, coskewness, and liquidity. To control for momentum, we use past 12-month returns. We exclude stocks within the highest volatility quintile and sort the remaining stocks by into quintiles ranked on past returns. Then, within each quintile, we further sort stocks into quintiles ranked on past $\beta^-$. Then, we average the $\beta^-$ quintiles across the momentum quintiles, and report book-to-market and characteristic-matched returns within each $\beta^-$ quintile. A similar procedure is repeated to control for coskewness and liquidity, which is measured using the historical liquidity betas of Pástor and Stambaugh (2003). The 5-1 $\beta^-$ quintile spread in adjusted returns remains significant at the 5% level, at over 0.30% per month controlling for momentum, coskewness, or liquidity. Hence, our predictive pattern of returns for past $\beta^-$ are not due to size, book-to-market, past return, coskewness, or liquidity effects.

### 5 Conclusion

The cross-section of stock returns reflects a premium for downside risk. Stocks that covary stronger with market declines than with market rises have high average returns over the same period. This risk-return relationship is consistent with agents placing greater weight on downside risk than they place on upside gains. Agents with aversion to downside risk require a premium to hold assets that have high sensitivities to market downturns. Hence, stocks with high downside risk exposure, or downside betas, have high average returns.

We find that contemporaneous high average returns earned by stocks with high downside betas are robust to controlling for various cross-sectional effects, including size and book-to-market, coskewness, liquidity risk, and past returns. Controlling for these and other cross-sectional effects, we estimate that the cross-sectional premium for bearing downside beta risk is approximately 6% per annum. In particular, we find that the premium captured by downside beta is quite different from the coskewness effect of Harvey and Siddique (2000). Downside risk measures risk conditional only on market declines, whereas coskewness captures covariation of a stock with both extreme upside and extreme downside movements of the market. We show that, in a simple model with a representative agent having a Gul (1991) kinked utility function that places greater weight on downside losses, both downside beta and coskewness risk may arise.
Past downside beta is a good predictor of future covariation with down market movements. For the majority of stocks, past downside beta cross-sectionally predicts future returns. However, for stocks with very high volatility, the past downside beta predictive relationship breaks down, because the returns of these stocks are confounded by the Ang et al. (2003) volatility effect, in which stocks with very high volatility have extremely low returns. Fortunately, high volatility stocks constitute only a small fraction of the total market, and a predictive downside beta relationship holds for the vast majority of stocks.
Appendix

A Solution of the Disappointment Aversion Asset Allocation Problem

Ang, Bekaert and Liu (2002) develop an algorithm for solving the portfolio allocation problem for DA utility that transforms (3) into a series of standard CRRA problems under a transformed measure that involves the degree of disappointment aversion $A$. The simplicity of their algorithm relies crucially on the assumption of a discretized state space that is ordered by wealth. However, their set-up is only for a single risky asset. We extend their algorithm to a multiple asset case, by considering all possible combinations of the 6 states.

Epstein and Zin (1989 and 2001) show that the First Order Conditions (FOC) for (3) are given by:

$$
E \left[ \frac{\partial U(W)}{\partial W} x 1_{\{W \leq \mu \}} \right] + A \cdot E \left[ \frac{\partial U(W)}{\partial W} y 1_{\{W > \mu \}} \right] = 0
$$

(A-1)

where $1$ is an indicator function.

Over a discrete-state space over states $(x_s, y_s)$ indexed by $s$, the definition of the certainty equivalent $\mu_W$ in (1) can be written as:

$$
\mu_{W}^{1-\gamma} = \frac{1}{K} \left( \sum_{s: W_s \leq \mu_W} p_s W_s^{1-\gamma} + \sum_{s: W_s > \mu_W} A p_s W_s^{1-\gamma} \right)
$$

(A-2)

where wealth in state $s$ is given by:

$$
W_s = R_f + w_x x_s + w_y y_s,
$$

and the FOC (A-1) take the form:

$$
\begin{align*}
\sum_{s: W_s \leq \mu_W} p_s W_s^{-\gamma} x_s + A \sum_{s: W_s > \mu_W} p_s W_s^{-\gamma} x_s &= 0 \\
\sum_{s: W_s \leq \mu_W} p_s W_s^{-\gamma} y_s + A \sum_{s: W_s > \mu_W} p_s W_s^{-\gamma} y_s &= 0
\end{align*}
$$

(A-3)

Ang, Bekaert and Liu (2002) note that (A-3) is a standard CRRA maximization problem with a changed probability measure, where the probabilities for wealth above the certainty equivalent are down-weighted. That is, defining the probabilities as:

$$
\pi_i \equiv \frac{(p_1, ..., p_i, A p_{i+1}, ..., A p_N)^i}{(p_1 + ... + p_i) + A (p_{i+1} + ... + p_N)},
$$

(A-4)

allows (A-3) to be re-written as:

$$
\begin{align*}
\sum_s \pi_s W_s^{-\gamma} x_s &= 0 \\
\sum_s \pi_s W_s^{-\gamma} y_s &= 0
\end{align*}
$$

(A-5)

The algorithm starts with a state $i$, solves the standard CRRA problem with probability distribution $\{\pi_i\}$ for the optimal portfolio weights $w_{x_i}^*$ and $w_{y_i}^*$, and then computes the certainty equivalent $\mu_{W_i}^*$, given by:

$$
\mu_{W_i}^{*} = \left( \sum_{s=1}^{N} (W_s^*)^{1-\gamma} \pi_{is} \right)^{\frac{1}{1-\gamma}}.
$$

We must find the state $i$ where:

$$
\mu_{W_i}^* \in [R_f + w_{x_i}^* x_i + w_{y_i}^* y_i, R_f + w_{x_{i+1}}^* x_{i+1} + w_{y_{i+1}}^* y_{i+1}].
$$

(A-6)
If this condition holds, then the optimal portfolio weights for \( x \) and \( y \), \( w^*_x \) and \( w^*_y \), have been found, so \( w^*_x = w^*_x \) and \( w^*_y = w^*_y \), and the optimal utility is given by \( \mu^* = \mu^* \).

The condition (A-6) relies on ordering the states in increasing wealth. To modify this algorithm, we take all possible \( M \) orderings of the states. Then, we find state \( i \) of ordering \( j \) where (A-6) is satisfied. This gives the solution to the DA asset allocation problem over the two assets \( x \) and \( y \).

For our calibrations, we set \( \gamma = 6 \), \( A = 0.8 \) and set the gross-risk free rate to be \( R_f = 1.05 \). For a base-line case, we take \( u_x = 0.25 + \mu, x_m = 0.16 + \mu, \) and \( d_x = -0.25 + \mu, y_m = 0.40 + \mu, d_y = -0.15 \). The 6 states have probabilities given by \( p_1 = 0.15 \), \( p_2 = 0.20 \), \( p_3 = 0.15 \), \( p_4 = 0.25 \) and \( p_5 = 0.20 \). In equilibrium, the value of \( \mu = 0.0021 \). This gives us equilibrium weights of \( w^*_x = 0.5543 \) and \( w^*_y = 0.4457 \). In this specification, the mean excess returns, standard deviations and betas of the two assets and the market are given by:

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>stdev</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1168</td>
<td>0.1863</td>
<td>0.6944</td>
</tr>
<tr>
<td></td>
<td>0.1250</td>
<td>0.2750</td>
<td>1.3800</td>
</tr>
<tr>
<td>mkt</td>
<td>0.1200</td>
<td>0.1375</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

To obtain the relations between the \( \beta \), downside beta \( \beta^- \) and coskewness and alpha’s, we alter \( x_d \) from \(-0.16 \) to \(-0.30 \). Figure 1 shows the risk-return relations for asset \( x \).

With an alternative set of parameters, CAPM \( \alpha \)’s increase with increasing \( \beta^- \) but also increasing co-skewness. Here, \( \gamma = 6 \), \( A = 0.7 \), \( R_f = 1.05 \), \( x_m = 0.50 + \mu \), \( y_m = 0.35 + \mu \), \( y_m = 0.035 \) and \( y_d = -0.7 \). The probabilities are given by \( p_1 = 0.10 \), \( p_2 = 0.20 \), \( p_3 = 0.20 \), \( p_4 = 0.20 \) and \( p_5 = 0.20 \). If \( p_3 \) is altered between 0.08 and 0.10 and \( \mu \) solved for each case to obtain equilibrium, then we have:

<table>
<thead>
<tr>
<th>CAPM ( \alpha )</th>
<th>( \beta^- )</th>
<th>co-skew</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0273</td>
<td>0.9567</td>
<td>-2.9883</td>
</tr>
<tr>
<td>0.0317</td>
<td>0.8473</td>
<td>-3.1752</td>
</tr>
<tr>
<td>0.0348</td>
<td>0.7326</td>
<td>-3.3873</td>
</tr>
</tbody>
</table>

However, this case is unrealistic because the values of the \( \beta^- \)’s are extremely high.

### B Data and Portfolio Construction

We use data from the Center for Research in Security Prices (CRSP) to construct portfolios of stocks sorted by various characteristics of returns. We confine our attention to ordinary common stocks listed on NYSE, AMEX and NASDAQ, omitting ADRs, REITs, closed-end funds, foreign firms and other securities which do not have a CRSP share type code of 10 or 11. We use daily and monthly returns from CRSP for the period covering July 3rd, 1962 to December 31st, 2001, including NASDAQ data which is only available after 1972. We use the one-month Treasury bill rate from Ibbotson Associates the risk-free rate and take CRSP’s value-weighted returns of all stocks as the market portfolio. All our returns are expressed as continuously compounded returns. We also use book value information found on COMPUSTAT.

For every twelve month period, we construct portfolios based on measures of risk between asset \( i \)’s excess return, \( r_{it} \), and the market’s excess return, \( r_{mt} \). We exclude stocks with more than five missing observations from our analysis. We first demean the returns within each period, and denote \( \tilde{r}_{it} \) as the demeaned excess return of asset \( i \) and \( \tilde{r}_{mt} \) as the demeaned market excess return. We calculate \( \beta \) and \( \sigma \) in the usual manner as,

\[
\beta = \frac{E[r_{it} \tilde{r}_{mt}]}{E[r^2_{it} \tilde{r}_{mt}]} \quad \text{and} \quad \sigma = \sqrt{E[r^2_{it} \tilde{r}_{mt}]},
\]

where \( E[.] \) denotes the sample mean of the observations in the brackets. We estimate downside beta and upside beta by conditioning the observations used for the calculation of the sample mean. That is, we calculate \( \beta^- \) and \( \beta^+ \) as,

\[
\beta^- = \frac{E[r_{it} \tilde{r}_{mt} | \tilde{r}_{mt} < 0]}{E[\tilde{r}_{mt} | \tilde{r}_{mt} < 0]} \quad \text{and} \quad \beta^+ = \frac{E[r_{it} \tilde{r}_{mt} | \tilde{r}_{mt} > 0]}{E[\tilde{r}_{mt} | \tilde{r}_{mt} > 0]}.
\]

When we calculate betas for using daily frequency data, we have to be particularly wary of biases induced by non-synchronous trading. This is especially true of small illiquid stocks listed on AMEX and NASDAQ. While most of our analysis focuses on stocks listed on NYSE, in part to avoid this problem, we control for non-synchronous trading when we include stocks listed on AMEX and NASDAQ. Scholes and Williams (1977) suggest...
incorporating additional leads and lags of estimates of second-order moments. That is, Scholes and Williams (1977) betas are estimated as,

\[
\beta_{SW} = \frac{E[\tilde{r}_{i,t}\tilde{r}_{m,t-1}] + E[\tilde{r}_{i,t}\tilde{r}_{m,t}]}{E[\tilde{r}_{m,t}\tilde{r}_{m,t-1}] + E[\tilde{r}_{m,t}^2]} \tag{B-9}
\]

We control for non-synchronous trading in the estimation of downside and upside betas in an analogous manner. That is, we also estimate \(\beta^-\) and \(\beta^+\) as,

\[
\beta^-_{SW} = \frac{E[\tilde{r}_{i,t}\tilde{r}_{m,t-1} | \tilde{r}_{m,t} < 0] + E[\tilde{r}_{i,t}\tilde{r}_{m,t} | \tilde{r}_{m,t} < 0] + E[\tilde{r}_{i,t}\tilde{r}_{m,t+1} | \tilde{r}_{m,t} < 0]}{E[\tilde{r}_{m,t}\tilde{r}_{m,t-1} | \tilde{r}_{m,t} < 0] + E[\tilde{r}_{m,t}^2 | \tilde{r}_{m,t} < 0] + E[\tilde{r}_{m,t}\tilde{r}_{m,t+1} | \tilde{r}_{m,t} < 0]}
\]

\[
\beta^+_{SW} = \frac{E[\tilde{r}_{i,t}\tilde{r}_{m,t-1} | \tilde{r}_{m,t} > 0] + E[\tilde{r}_{i,t}\tilde{r}_{m,t} | \tilde{r}_{m,t} > 0] + E[\tilde{r}_{i,t}\tilde{r}_{m,t+1} | \tilde{r}_{m,t} > 0]}{E[\tilde{r}_{m,t}\tilde{r}_{m,t-1} | \tilde{r}_{m,t} > 0] + E[\tilde{r}_{m,t}^2 | \tilde{r}_{m,t} > 0] + E[\tilde{r}_{m,t}\tilde{r}_{m,t+1} | \tilde{r}_{m,t} > 0]}. \tag{B-11}
\]

We calculate higher-order moments of stock returns using continuously compounded daily returns over each twelve month period. Coskewness and cokurtosis are estimated as,

\[
\text{coskew} = \frac{E[\tilde{r}_{i,t}\tilde{r}_{m,t}^2]}{\sigma_m^2}, \quad \text{and} \quad \text{cokurt} = \frac{E[\tilde{r}_{i,t}\tilde{r}_{m,t}^3]}{\sigma_m^{3/2}}. \tag{B-12}
\]

We also collect for each stock, market capitalizations, book-to-market ratio and past twelve month returns at the end of each twelve month period. Finally, we calculate a liquidity risk measure as described below.

Once portfolios are formed, we calculate the returns to holding these portfolios. Over every twelve month period, we also collect the cumulative returns of each stock in excess of the one-month Treasury bill rate over the period. We also collect the excess stock return over the next month, as well as stock returns in excess of size and book-to-market matched benchmarked portfolios. These size and book-to-market adjusted returns are calculated in a manner similar to Daniel et al. (1997). Each month, stocks listed on NYSE, AMEX and NASDAQ are sorted into quintiles using their beginning of period market capitalizations based on NYSE breakpoints. Then within each of these quintiles, stocks are further sorted into quintiles using their book-to-market ratios based on NYSE breakpoints. For each 5x5 grouping, we calculate the return on equal-weighted portfolios consisting of all stocks in the grouping. For each stock, size and book-to-market adjusted returns are defined as the return in excess of the portfolio return of the 5x5 grouping to which the stock belongs to. All of these returns are calculated with an adjustment for delisting by taking the delisting return available at the time the stock is delisted. If a return is missing, we take the next available return.

**Liquidity Betas**

We follow Pástor and Stambaugh (2001) to construct betas with respect to an aggregate liquidity measure, \(L\). Stock return and volume data are obtained from CRSP. NASDAQ stocks are excluded in the construction of the aggregate liquidity measure. The liquidity estimate, \(\gamma_{i,t}\), for an individual stock \(i\) in month \(t\) is the ordinary least squares (OLS) estimate of \(\gamma_{i,t}\) in the following regression:

\[
r_{i,d+1,t} = \theta_{i,t} + \phi_{i,t}\tilde{r}_{i,d,t} + \gamma_{i,t}\text{sign} (r_{i,d,t} \, v_{i,d,t} + \epsilon_{i,d+1,t}), \quad d = 1, \ldots, D. \tag{B-13}
\]

In equation (B-13), \(\tilde{r}_{i,d,t}\) is the raw return on stock \(i\) on day \(d\) of month \(t\), \(r_{i,d,t} = \tilde{r}_{i,d,t} - \tilde{r}_{m,d,t}\) is the stock return in excess of the market return, and \(v_{i,d,t}\) is the dollar volume for stock \(i\) on day \(d\) of month \(t\). The market return on day \(d\) of month \(t\), \(\tilde{r}_{m,d,t}\), is taken as the return on the CRSP value-weighted market portfolio. A stock’s liquidity estimate, \(\gamma_{i,t}\), is computed in a given month only if there are at least 15 consecutive observations, and if the stock has a month-end share prices of greater than $5 and less than $1000.

The aggregate liquidity measure, \(L\), is computed based on the liquidity estimates, \(\gamma_{i,t}\), of individual firms listed on NYSE and AMEX from August 1962 to December 2001. Only the individual liquidity estimates that meet the above criteria is used. To construct the innovations in aggregate liquidity, we follow Pástor and Stambaugh and first form the scaled monthly difference:

\[
\Delta \gamma_t = \left( \frac{m_t}{m_1} \right) \frac{1}{N} \sum_{i=1}^{N} (\gamma_{i,t} - \gamma_{i,t-1}), \tag{B-14}
\]
where $N$ is the number of available stocks at month $t$, $m_t$ is the total dollar value of the included stocks at the end of month $t - 1$, and $m_1$ is the total dollar value of the stocks at the end of July 1962. The innovations in liquidity are computed as the residuals in the following regression:

$$
\Delta \hat{\gamma}_t = a + b \Delta \hat{\gamma}_{t-1} + c \left( m_t / m_1 \right) \hat{\gamma}_{t-1} + u_t. \quad \text{(B-15)}
$$

Finally, the aggregate liquidity measure, $L_t$, is taken to be the fitted residuals, $L_t = \hat{u}_t$.

To calculate the liquidity betas for individual stocks, at the end of each month, we identify stocks listed on NYSE, AMEX and NASDAQ with at least five years of monthly returns. For each stock, we estimate a liquidity beta, $\beta^L_i$, by running the following regression using the most recent five years of monthly data:

$$
r_{i,t} = \beta^0_i + \beta^L_i L_t + \beta^H_i M K T_t + \beta^S_i S M B_t + \beta^H_i H M L_t + \epsilon_{i,t}, \quad \text{(B-16)}
$$

where $L_t$ is the innovation in aggregate liquidity.
References


[38] Markowitz, H., 1959, Portfolio Selection. New Haven, Yale University Press.


Table 1: Returns of Stocks Sorted by Realized $\beta$

Panel A: Stocks Sorted by Realized $\beta$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Return</th>
<th>$\beta$</th>
<th>$\beta^-$</th>
<th>$\beta^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>3.52%</td>
<td>0.28</td>
<td>0.36</td>
<td>0.19</td>
</tr>
<tr>
<td>2</td>
<td>6.07%</td>
<td>0.59</td>
<td>0.67</td>
<td>0.51</td>
</tr>
<tr>
<td>3</td>
<td>7.58%</td>
<td>0.82</td>
<td>0.90</td>
<td>0.77</td>
</tr>
<tr>
<td>4</td>
<td>9.48%</td>
<td>1.10</td>
<td>1.18</td>
<td>1.06</td>
</tr>
<tr>
<td>High</td>
<td>13.95%</td>
<td>1.64</td>
<td>1.72</td>
<td>1.63</td>
</tr>
<tr>
<td>High-Low</td>
<td>10.43%</td>
<td>1.36</td>
<td>1.36</td>
<td>1.44</td>
</tr>
<tr>
<td>t-stat</td>
<td>[4.98]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Stocks Sorted by Realized $\beta^-$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Return</th>
<th>$\beta$</th>
<th>$\beta^-$</th>
<th>$\beta^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $\beta^-$</td>
<td>2.71%</td>
<td>0.40</td>
<td>0.19</td>
<td>0.42</td>
</tr>
<tr>
<td>2</td>
<td>5.62%</td>
<td>0.63</td>
<td>0.61</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>7.63%</td>
<td>0.83</td>
<td>0.89</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td>10.16%</td>
<td>1.06</td>
<td>1.23</td>
<td>0.99</td>
</tr>
<tr>
<td>High $\beta^-$</td>
<td>14.49%</td>
<td>1.49</td>
<td>1.92</td>
<td>1.34</td>
</tr>
<tr>
<td>High-Low</td>
<td>11.78%</td>
<td>1.09</td>
<td>1.72</td>
<td>0.92</td>
</tr>
<tr>
<td>t-stat</td>
<td>[6.16]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Stocks Sorted by Realized $\beta^+$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Return</th>
<th>$\beta$</th>
<th>$\beta^-$</th>
<th>$\beta^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $\beta^+$</td>
<td>5.73%</td>
<td>0.44</td>
<td>0.63</td>
<td>-0.04</td>
</tr>
<tr>
<td>2</td>
<td>7.42%</td>
<td>0.62</td>
<td>0.73</td>
<td>0.45</td>
</tr>
<tr>
<td>3</td>
<td>8.29%</td>
<td>0.82</td>
<td>0.90</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>9.33%</td>
<td>1.05</td>
<td>1.10</td>
<td>1.12</td>
</tr>
<tr>
<td>High $\beta^+$</td>
<td>9.83%</td>
<td>1.49</td>
<td>1.46</td>
<td>1.85</td>
</tr>
<tr>
<td>High-Low</td>
<td>4.11%</td>
<td>1.05</td>
<td>0.83</td>
<td>1.89</td>
</tr>
<tr>
<td>t-stat</td>
<td>[2.62]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel D: Stocks Sorted by Realized Relative $\beta^-$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Return</th>
<th>$\beta$</th>
<th>$\beta^-$</th>
<th>$\beta^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Relative $\beta^-$</td>
<td>4.09%</td>
<td>0.98</td>
<td>0.56</td>
<td>1.12</td>
</tr>
<tr>
<td>2</td>
<td>7.69%</td>
<td>0.83</td>
<td>0.73</td>
<td>0.84</td>
</tr>
<tr>
<td>3</td>
<td>8.53%</td>
<td>0.80</td>
<td>0.86</td>
<td>0.75</td>
</tr>
<tr>
<td>4</td>
<td>9.56%</td>
<td>0.84</td>
<td>1.08</td>
<td>0.72</td>
</tr>
<tr>
<td>High Relative $\beta^+$</td>
<td>10.73%</td>
<td>0.98</td>
<td>1.60</td>
<td>0.71</td>
</tr>
<tr>
<td>High-Low</td>
<td>6.64%</td>
<td>0.00</td>
<td>1.04</td>
<td>0.41</td>
</tr>
<tr>
<td>t-stat</td>
<td>[7.70]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table lists the equal-weighted average returns and risk characteristics of stocks sorted by realized betas. For each month, we calculate $\beta$, $\beta^-$, $\beta^+$ and relative $\beta^-$ with respect to the market of all stocks listed on NYSE using daily continuously compounded returns over the next 12 months. For each risk characteristic, we rank stocks into quintiles (1–5) and form equal-weighted portfolios at the beginning of each 12 month period. The sample period is from July 1963 to December 2001, with the last 12-month period from January 2001 to December 2001, and observations are at a monthly frequency. The number of stocks in each portfolio varies across time from 216 to 317 stocks. The column labeled “Return” reports the average return in excess of the one-month T-bill rate over the next 12 months (which is the same period as the period used to compute $\beta$, $\beta^-$ and $\beta^+$). The row labeled “High-Low” reports the difference between the returns of portfolio 5 and portfolio 1. The entry labeled “t-stat” in square brackets is the t-statistic computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags. The columns labeled “$\beta$”, “$\beta^-$” and “$\beta^+$” report the time-series and cross-sectional average of equal-weighted individual stock betas over the holding period.
This table shows the results of Fama-MacBeth regressions of 12-month excess returns on firm characteristics and realized risk characteristics. The sample period is from July 1963 to December 2001, with the last 12-month period from January 2001 to December 2001, and observations are at a monthly frequency (451 months) for all stocks listed on NYSE. The number of stocks in each regression varies across time from 1080 to 1582 stocks. The t-statistics in square brackets are computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags. The firm characteristics are log of market capitalizations (“Log-Size”), book-to-market ratios (“Bk-Mkt”), and past 12-month excess returns (“Past Ret”), both computed at the beginning of each period. The realized risk characteristics are $\beta$, $\beta^-$, $\beta^+$, standard deviations (“Std Dev”), coskewness and cokurtosis calculated over the following 12-month period using daily continuously compounded returns. We also include the P´astor-Stambaugh (2003) liquidity beta, $\beta_L$, for January 1967 to January 2001. All independent variables are Winsorized at the 1% level and at the 99% within each month.
Table 3: Robustness Checks of Realized $\beta^-$ Portfolios

Panel A: Sorts by Realized $\beta^-$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Value-Weighted</th>
<th>All Stocks</th>
<th>Non-Overlapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>2.62%</td>
<td>4.14%</td>
<td>3.69%</td>
</tr>
<tr>
<td>2</td>
<td>4.23%</td>
<td>7.23%</td>
<td>5.74%</td>
</tr>
<tr>
<td>3</td>
<td>6.04%</td>
<td>8.84%</td>
<td>8.33%</td>
</tr>
<tr>
<td>4</td>
<td>9.32%</td>
<td>11.35%</td>
<td>10.68%</td>
</tr>
<tr>
<td>High</td>
<td>9.76%</td>
<td>19.37%</td>
<td>16.15%</td>
</tr>
<tr>
<td>High-Low</td>
<td>7.14%</td>
<td>15.24%</td>
<td>12.46%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[3.30]</td>
<td>[5.57]</td>
<td>[3.51]</td>
</tr>
</tbody>
</table>

Panel B: Sorts by Realized Relative $\beta^-$

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Value-Weighted</th>
<th>All Stocks</th>
<th>Non-Overlapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>3.19%</td>
<td>5.58%</td>
<td>5.67%</td>
</tr>
<tr>
<td>2</td>
<td>6.86%</td>
<td>8.66%</td>
<td>8.78%</td>
</tr>
<tr>
<td>3</td>
<td>7.07%</td>
<td>9.41%</td>
<td>8.42%</td>
</tr>
<tr>
<td>4</td>
<td>7.56%</td>
<td>10.86%</td>
<td>9.67%</td>
</tr>
<tr>
<td>High</td>
<td>7.18%</td>
<td>14.21%</td>
<td>12.05%</td>
</tr>
<tr>
<td>High-Low</td>
<td>3.99%</td>
<td>8.63%</td>
<td>6.38%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[3.06]</td>
<td>[7.02]</td>
<td>[3.87]</td>
</tr>
</tbody>
</table>

The table presents robustness checks of the results in Table 1. For each month, we calculate $\beta^-$ and relative $\beta^-$ with respect to the market using daily continuously compounded returns over the next 12 months. We report the results using realized $\beta^-$ in Panel A and the results using realized relative $\beta^-$ in Panel B. For each risk characteristic, we rank stocks into quintiles (1–5). In the first column of each panel, we form value-weighted portfolios using stocks listed on NYSE at the beginning of each 12 month period. In the second column, we use all stocks listed on NYSE, AMEX and NASDAQ and form equal-weighted portfolios at the beginning of each period, using quintile breakpoints based on NYSE stocks. For this specification, we control for non-synchronous trading in a manner analogous to Scholes and Williams (1977). In the last column, we compute the risk characteristics using stocks listed on NYSE and form equal-weighted portfolios at the beginning of each January. We report the average return in excess of the one-month T-bill rate over the next 12 months. The row labeled “High-Low” reports the difference between the returns of portfolio 5 and portfolio 1. For the columns labeled “Value-Weighted” and “All Stocks,” the sample period is from July 1963 to December 2001, with the last 12-month period from January 2001 to December 2001, and observations are at a monthly frequency. For the column labeled “Non-Overlapping,” the sample period is from January 1964 to December 2001, with the last 12-month period from January 2001 to December 2001. In the last column, observations are non-overlapping from each January to December. The number of stocks in each portfolio varies across time from 216 to 317 stocks, except for “All Stocks,” where it varies from 289 to 2330 stocks. The entry labeled “t-stat” in square brackets is the t-statistic computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags, except for the column labeled “Non-Overlapping,” where 1 lag is used because there are no overlapping observations.
Table 4: Returns of Stocks Sorted by Realized $\beta$-s, Controlling for Coskewness

### Panel A: $\beta^-$ Sorts Controlling for Coskewness

<table>
<thead>
<tr>
<th>Portfolio $\beta^-$ Sorts Controlling for Coskewness</th>
<th>Coskewness Quintiles</th>
<th></th>
<th></th>
<th></th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $\beta^-$</td>
<td>7.21%</td>
<td>5.74%</td>
<td>4.03%</td>
<td>3.40%</td>
<td>0.22%</td>
</tr>
<tr>
<td>2</td>
<td>10.55%</td>
<td>8.40%</td>
<td>6.94%</td>
<td>5.59%</td>
<td>2.61%</td>
</tr>
<tr>
<td>3</td>
<td>13.63%</td>
<td>11.30%</td>
<td>8.30%</td>
<td>6.08%</td>
<td>3.76%</td>
</tr>
<tr>
<td>4</td>
<td>15.63%</td>
<td>12.82%</td>
<td>9.35%</td>
<td>6.74%</td>
<td>2.56%</td>
</tr>
<tr>
<td>High $\beta^-$</td>
<td>21.84%</td>
<td>15.85%</td>
<td>11.51%</td>
<td>6.81%</td>
<td>2.32%</td>
</tr>
<tr>
<td>High-Low</td>
<td>14.64%</td>
<td>10.11%</td>
<td>7.48%</td>
<td>3.41%</td>
<td>2.10%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[5.62]</td>
<td>[5.22]</td>
<td>[3.91]</td>
<td>[1.87]</td>
<td>[1.32]</td>
</tr>
</tbody>
</table>

### Panel B: Coskewness Sorts Controlling for $\beta^-$

<table>
<thead>
<tr>
<th>$\beta^-$ Quintiles</th>
<th>Portfolio Low coskew</th>
<th></th>
<th></th>
<th></th>
<th>High</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.41%</td>
<td>4.69%</td>
<td>7.15%</td>
<td>9.30%</td>
<td>12.59%</td>
<td>17.61%</td>
<td>10.27%</td>
</tr>
<tr>
<td>2.74%</td>
<td>4.17%</td>
<td>6.19%</td>
<td>9.61%</td>
<td>12.33%</td>
<td>18.21%</td>
<td>10.10%</td>
</tr>
<tr>
<td>3.00%</td>
<td>1.50%</td>
<td>5.24%</td>
<td>6.68%</td>
<td>9.16%</td>
<td>12.83%</td>
<td>7.08%</td>
</tr>
<tr>
<td>High coskew</td>
<td>0.41%</td>
<td>2.96%</td>
<td>3.86%</td>
<td>5.37%</td>
<td>7.65%</td>
<td>4.05%</td>
</tr>
<tr>
<td>High-Low</td>
<td>-4.28%</td>
<td>-4.18%</td>
<td>-5.45%</td>
<td>-7.22%</td>
<td>-9.96%</td>
<td>-6.22%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[4.23]</td>
<td>[5.64]</td>
<td>[7.12]</td>
<td>[8.09]</td>
<td>[7.94]</td>
<td>[8.17]</td>
</tr>
</tbody>
</table>

This table examines the relation between $\beta^-$ and coskewness. For each month, we compute $\beta^-$ and coskewness with respect to the market of all stocks listed on NYSE using daily continuously compounded returns over the next 12 months. In Panel A, we first rank stocks into quintiles (1–5) at the beginning of each 12 month period based on coskewness over the next 12 months. Then, we rank stocks further within each first-sort quintile into additional quintiles according to $\beta^-$, computed over the next 12 months. For each 5-by-5 grouping, we form an equal-weighted portfolio. In Panel B, we reverse the order so that we first sort on $\beta^-$, and then on coskewness. The sample period is from July 1963 to December 2001 and the number of stocks in each portfolio varies across time from 43 to 64 stocks. We report the average return in excess of the one-month T-bill rate over the next 12 months. For the column labeled “Average”, we report the average return of stocks in each second sort quintile. This controls for coskewness ($\beta^-$) in Panel A (B). The row labeled “High-Low” reports the difference between the returns of portfolio 5 and portfolio 1. The entry labeled “t-stat” in square brackets is the t-statistic computed using Newey-West (1987) heteroskedastic-robust standard errors with 12 lags.
Table 5: Returns of Stocks Sorted by Past Asymmetry Measures

**Panel A: Stocks Sorted by Past $\beta^-$**

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Average Next Month Return</th>
<th>Realized Statistics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>$\beta^-$</td>
<td>$\beta^+$</td>
</tr>
<tr>
<td>Low $\beta^-$</td>
<td>0.59%</td>
<td>0.54</td>
<td>0.61</td>
<td>0.48</td>
</tr>
<tr>
<td>2</td>
<td>0.71%</td>
<td>0.70</td>
<td>0.77</td>
<td>0.65</td>
</tr>
<tr>
<td>3</td>
<td>0.77%</td>
<td>0.85</td>
<td>0.93</td>
<td>0.80</td>
</tr>
<tr>
<td>4</td>
<td>0.84%</td>
<td>1.02</td>
<td>1.11</td>
<td>0.96</td>
</tr>
<tr>
<td>High $\beta^-$</td>
<td>0.70%</td>
<td>1.31</td>
<td>1.41</td>
<td>1.25</td>
</tr>
<tr>
<td>High-Low</td>
<td>0.11%</td>
<td>0.77</td>
<td>0.80</td>
<td>0.77</td>
</tr>
</tbody>
</table>

High-Low t-stat [0.60]

**Panel B: Stocks Sorted by Past Coskewness**

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Average Next Month Return</th>
<th>Realized Statistics</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>$\beta^-$</td>
<td>$\beta^+$</td>
</tr>
<tr>
<td>Low coskew</td>
<td>0.84%</td>
<td>0.91</td>
<td>1.01</td>
<td>0.85</td>
</tr>
<tr>
<td>2</td>
<td>0.82%</td>
<td>0.90</td>
<td>0.99</td>
<td>0.84</td>
</tr>
<tr>
<td>3</td>
<td>0.76%</td>
<td>0.89</td>
<td>0.97</td>
<td>0.83</td>
</tr>
<tr>
<td>4</td>
<td>0.60%</td>
<td>0.87</td>
<td>0.95</td>
<td>0.82</td>
</tr>
<tr>
<td>High coskew</td>
<td>0.57%</td>
<td>0.85</td>
<td>0.91</td>
<td>0.81</td>
</tr>
<tr>
<td>High-Low</td>
<td>-0.28%</td>
<td>-0.06</td>
<td>-0.10</td>
<td>-0.04</td>
</tr>
</tbody>
</table>

High-Low t-stat [2.76]

The table reports the equal-weighted average returns and risk characteristics of stocks sorted by past $\beta^-$ and coskewness. For each month, we compute $\beta^-$ and coskewness with respect to the market of all stocks listed on NYSE using daily continuously compounded returns over the previous 12 months. For each risk characteristic, we rank stocks into quintiles (1–5) and form equal-weighted portfolios at the beginning of each month. The sample period is from July 1962 to January 2001. The number of stocks in each portfolio varies across time from 221 to 346 stocks. The column labeled “Return” reports the average return in excess of the one-month T-bill rate over the next one month. The row labeled “High-Low” reports the difference between the returns of portfolio 1 and portfolio 5. The entry labeled “t-stat” is the simple OLS t-statistic in square brackets. The columns labeled “$\beta$”, “$\beta^-$”, and “$\beta^+$” report the time-series averages of equal-weighted cross-sectional averages of individual stock betas over the next 12 month period. The column labeled “coskew” reports the time-series averages of equal-weighted cross-sectional averages of individual stock coskewness over the next 12 month period.
Table 6: Stocks Sorted by Past $\beta^-$ Excluding Most Volatile Stocks

### Panel A: Average Excess Returns

<table>
<thead>
<tr>
<th>Excluding Stocks of the Highest Quintile $\sigma$</th>
<th>Octile $\sigma$</th>
<th>Decile $\sigma$</th>
<th>5%-tile $\sigma$</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Low $\beta^-$</td>
<td>0.58%</td>
<td>0.58%</td>
<td>0.57%</td>
<td>0.58%</td>
</tr>
<tr>
<td>2</td>
<td>0.69%</td>
<td>0.71%</td>
<td>0.72%</td>
<td>0.72%</td>
</tr>
<tr>
<td>3</td>
<td>0.82%</td>
<td>0.80%</td>
<td>0.79%</td>
<td>0.77%</td>
</tr>
<tr>
<td>4</td>
<td>0.82%</td>
<td>0.82%</td>
<td>0.83%</td>
<td>0.84%</td>
</tr>
<tr>
<td>5 High $\beta^-$</td>
<td>0.92%</td>
<td>0.89%</td>
<td>0.88%</td>
<td>0.80%</td>
</tr>
<tr>
<td>High-Low</td>
<td>0.34%</td>
<td>0.31%</td>
<td>0.31%</td>
<td>0.22%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[2.31]</td>
<td>[1.95]</td>
<td>[1.86]</td>
<td>[1.26]</td>
</tr>
<tr>
<td>Q4–Low</td>
<td>0.25%</td>
<td>0.25%</td>
<td>0.26%</td>
<td>0.26%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[2.28]</td>
<td>[2.17]</td>
<td>[2.31]</td>
<td>[2.21]</td>
</tr>
</tbody>
</table>

### Panel B: Selected Characteristics of $\sigma$ Portfolios

<table>
<thead>
<tr>
<th>All Stocks</th>
<th>Quintile $\sigma$</th>
<th>Octile $\sigma$</th>
<th>Decile $\sigma$</th>
<th>5%-tile $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Cap</td>
<td>100.0%</td>
<td>3.9%</td>
<td>1.9%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Annualized Volatility</td>
<td>35.9%</td>
<td>61.0%</td>
<td>68.4%</td>
<td>72.5%</td>
</tr>
<tr>
<td>Past $\beta^-$</td>
<td>0.99</td>
<td>1.44</td>
<td>1.47</td>
<td>1.48</td>
</tr>
<tr>
<td>12-month Autocorrelation of $\beta^-$</td>
<td>43.5%</td>
<td>25.8%</td>
<td>21.8%</td>
<td>20.5%</td>
</tr>
<tr>
<td>Size/Book-to-Market Adjusted Returns</td>
<td>-0.08%</td>
<td>-0.38%</td>
<td>-0.50%</td>
<td>-0.56%</td>
</tr>
</tbody>
</table>

In Panel A, each month, we calculate $\sigma$ and $\beta^-$ with respect to the market of all stocks listed on the NYSE using daily continuously compounded returns over the previous 12 months. We first sort stocks according to $\sigma$ into quintiles, octiles, deciles and demi-deciles (5%-tiles). Then for each $\sigma$ sort, we exclude the stocks that fall into the highest quintile, octile, decile or demi-decile of $\sigma$. We rank the remaining stocks into equal-weighted quintiles (1–5) according to past $\beta^-$. We report the average excess return over the next 1-month. The row labeled “High-Low” (“Q4-Low”) reports the difference between the returns of portfolio 5 (portfolio 4) and portfolio 1. We report simple t-statistics in square brackets. The number of stocks in each portfolio varies across time and groupings from 177 to 346 stocks. Panel B reports selected average characteristics of stocks in each $\sigma$ group. The first column reports the characteristics over the entire sample. The other columns report the characteristics within the highest $\sigma$ groups. The row labeled “Market Cap” reports the time-series averages of cumulative market capitalization represented by the stocks in each group. The other rows report the annualized past volatility ($\sigma$), past $\beta^-$, autocorrelation of $\beta^-$ between the past 12-months and the following 12-months, and returns adjusted for size and book-to-market using a characteristic control similar to Daniel et al. (1997). For each characteristic, we report the time-series averages of equal-weighted cross-sectional averages. The sample period is from July 1962 to January 2001.
Table 7: Characteristic Controls on Stocks Sorted by Past $\beta^-$

<table>
<thead>
<tr>
<th></th>
<th>Size/Bk-Mkt Adjusted</th>
<th>Include Additional Controls for Momentum</th>
<th>Coskewness</th>
<th>Liquidity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Low $\beta^-$</td>
<td>-0.25%</td>
<td>-0.21%</td>
<td>-0.21%</td>
<td>-0.17%</td>
</tr>
<tr>
<td>2</td>
<td>-0.09%</td>
<td>-0.07%</td>
<td>-0.07%</td>
<td>-0.02%</td>
</tr>
<tr>
<td>3</td>
<td>0.05%</td>
<td>0.04%</td>
<td>0.07%</td>
<td>0.05%</td>
</tr>
<tr>
<td>4</td>
<td>0.07%</td>
<td>0.10%</td>
<td>0.04%</td>
<td>0.10%</td>
</tr>
<tr>
<td>5 High $\beta^-$</td>
<td>0.20%</td>
<td>0.12%</td>
<td>0.15%</td>
<td>0.13%</td>
</tr>
<tr>
<td>High-Low</td>
<td>0.44%</td>
<td>0.32%</td>
<td>0.36%</td>
<td>0.30%</td>
</tr>
<tr>
<td>t-stat</td>
<td>[3.36]</td>
<td>[2.71]</td>
<td>[2.69]</td>
<td>[2.15]</td>
</tr>
</tbody>
</table>

The table reports robustness checks of the results in Table 6. For each month, we compute $\sigma$ and $\beta^-$ with respect to the market of all stocks using daily continuously compounded returns over the previous 12 months. We first sort stocks according to $\sigma$ into quintiles and exclude stocks that fall within the highest $\sigma$ ranking. We rank the remaining stocks into quintiles (1–5) according to past $\beta^-$ and form equal-weighted portfolios at the beginning of each month. The table reports characteristic-adjusted holding period returns over the next month of the $\beta^-$ quintiles that exclude stocks in the highest $\sigma$ quintile. In column labeled “Size/Bk-Mkt Adjusted”, we report the average returns in excess of size and book-to-market matched benchmark portfolios. In the next three columns, we include additional controls for momentum (as measured by past 12 month returns), coskewness, and historical liquidity betas, computed following Pástor and Stambaugh (2003). For each additional control, we first perform a quintile sort based on the characteristic and then on past $\beta^-$ excluding the highest $\sigma$ quintile of stocks. Then, we average the $\beta^-$ quintiles across the characteristic quintiles, and report book-to-market and characteristic-matched returns within each $\beta^-$ quintile. The number of stocks in each portfolio varies across time from 177 to 277 stocks. The row labeled “High-Low” reports the difference between the returns of portfolio 5 and portfolio 1. The entry labeled “t-stat” in square brackets is the simple t-statistic. The sample period is from July 1962 to January 2001, except in the last column where the sample period is from January 1967 to January 2001.
This figure shows risk-return relations for an asset in the DA cross-sectional equilibrium. Clockwise from the top left-hand corner we have: (i) a plot of the asset’s mean excess return versus downside beta $\beta^-$, (ii) a plot of the asset’s CAPM $\alpha$ versus downside beta $\beta^-$, (iii) a plot of the asset’s CAPM alpha versus relative downside beta $(\beta^- - \beta)$, and (iv) a plot of the asset’s CAPM alpha versus coskewness.
Figure 2: Distribution of Stocks According to $\beta^-$ and $\sigma$ versus Returns

Panel A shows the average realized 12-month excess return of stocks in contour lines, according to realized $\beta^-$ and realized $\sigma$. The average realized returns, realized $\beta^-$ and realized $\sigma$ are computed using daily continuously compounded returns over the same 12 month period. Panel B shows the average future excess 1-month return, according to past $\beta^-$ and past $\sigma$, computed over the previous 12-month period. Each marker, ‘o’, represents approximately 0.5% of the number of stocks listed on the NYSE. The contours represent low (high) average excess returns in light (dark) lines. The sample period is from July 1963 to December 2001. The y-axis $\sigma$ numbers are annualized in both panels.