THE JOURNAL OF FINANCE • VOL. LV, NO. 3 • JUNE 2000

# Conditional Skewness in Asset Pricing Tests

CAMPBELL R. HARVEY and AKHTAR SIDDIQUE*

**ABSTRACT**

If asset returns have systematic skewness, expected returns should include rewards for accepting this risk. We formalize this intuition with an asset pricing model that incorporates conditional skewness. Our results show that conditional skewness helps explain the cross-sectional variation of expected returns across assets and is significant even when factors based on size and book-to-market are included. Systematic skewness is economically important and commands a risk premium, on average, of 3.60 percent per year. Our results suggest that the momentum effect is related to systematic skewness. The low expected return momentum portfolios have higher skewness than high expected return portfolios.

THE SINGLE FACTOR CAPITAL ASSET PRICING MODEL (CAPM) of Sharpe (1964) and Lintner (1965) has come under recent scrutiny. Tests indicate that the cross-asset variation in expected returns cannot be explained by the market beta alone. For example, a growing number of studies show that "fundamental" variables such as size, book-to-market value, and price to earnings ratios account for a sizeable portion of the cross-sectional variation in expected returns (see, e.g., Chan, Hamao, and Lakonishok (1991) and Fama and French (1992)). Fama and French (1995) document the importance of SMB (the difference between the return on a portfolio of small size stocks and the return on a portfolio of large size stocks) and HML (the difference between the return on a portfolio of high book-to-market value stocks and the return on a portfolio of low book-to-market value stocks).

There are a number of responses to these empirical findings. First, the single-factor CAPM is rejected when the portfolio used to proxy for the market is inefficient (see Roll (1977) and Ross (1977)). Roll and Ross (1994) and Kandel and Stambaugh (1995) show that even very small deviations from efficiency can produce an insignificant relation between risk and expected returns. Second, Kothari, Shanken, and Sloan (1995) and Breen and Korajczyk (1993) argue that there is a survivorship bias in the data used to test these new asset pricing specifications. Third, there are several specification issues. Kim (1995) and Amihud, Christensen, and Mendelson (1993) argue that errors-in-variables impact the empirical research. Kan and Zhang (1997) focus on time-varying risk premia and the ability of insignificant factors to appear

* The authors are from Duke University and Georgetown University respectively. We appreciate the comments of Philip Dybvig, Stephen Brown, Alon Brav, S. Viswanathan, and seminar participants at Georgetown, Indiana University, the University of Toronto, the 1996 WFA (Oregon), and the AFA (New Orleans) meetings. We appreciate the helpful comments of an anonymous referee and the detailed suggestions of the editor.

significant as a result of low-powered tests. Jagannathan and Wang (1996) show that specifying a broader market portfolio can affect the results. Finally, Ferson and Harvey (1998) show that even these new multifactor specifications are rejected because they ignore conditioning information.

The goal of this paper is to examine the linkage between the empirical evidence on these additional factors and systematic coskewness. The following is our intuition for including skewness in the asset pricing framework. In the usual setup, investors have preferences over the mean and the variance of portfolio returns. The systematic risk of a security is measured as the contribution to the variance of a well-diversified portfolio. However, there is considerable evidence that the unconditional returns distributions cannot be adequately characterized by mean and variance alone.[1] This leads us to the next moment—skewness. Everything else being equal, investors should prefer portfolios that are right-skewed to portfolios that are left-skewed. This is consistent with the Arrow–Pratt notion of risk aversion. Hence, assets that decrease a portfolio's skewness (i.e., that make the portfolio returns more leftskewed) are less desirable and should command higher expected returns. Similarly, assets that increase a portfolio's skewness should have lower expected returns.

One clue that pushed us in the direction of skewness is the fact that some of the empirical shortcomings of the standard CAPM stem from failures in explaining the returns of specific securities or groups of securities such as the smallest market-capitalized deciles and returns from specific strategies such as ones based on momentum. These assets are also the ones with the most skewed returns. Skewness may be important in investment decisions because of induced asymmetries in ex post (realized) returns. At least two factors may induce asymmetries. First, the presence of limited liability in all equity investments may induce option-like asymmetries in returns (see Black (1972), Christie (1982), Nelson (1991), and Golec and Tamarkin (1998)). Second, the agency problem may induce asymmetries in portfolio returns (see Brennan (1993)). That is, a manager has a call option with respect to the outcome of his investment strategies. Managers may prefer portfolios with high positive skewness.

We present an asset pricing model where skewness is priced. Our formulation is related to the seminal work of Kraus and Litzenberger (1976) and to the nonlinear factor models presented more recently in Bansal and Viswanathan (1993) and Leland (1997). We use an asset pricing model incorporating conditional skewness to help understand the cross-sectional variation in several sets of asset returns.

Our work differs from Kraus and Litzenberger (1976) and Lim (1989) in our focus on conditional skewness rather than unconditional skewness as well as in our objective of explaining the cross-sectional variation in ex-

[1] Merton (1982) shows that if instantaneous returns are normal, then the price process is lognormal and, unless the measurement interval is very small, the simple returns are not normal.

pected returns. Conditional skewness also captures asymmetry in risk, especially downside risk, which has come to be viewed by practitioners as important in contexts such as value-at-risk (VaR). Our work focuses primarily on monthly U.S. equity returns from CRSP. We form portfolios of equities on various criteria such as industry, size, book-to-market ratios, coskewness with the market portfolio (where we define coskewness as the component of an asset's skewness related to the market portfolio's skewness), and momentum using both monthly holding periods as well as longer holding periods. Additionally, we also examine individual equity returns.

We analyze the ability of conditional coskewness to explain the cross-sectional variation of asset returns in comparison with other factors. We find that coskewness can explain some of the apparent nonsystematic components in cross-sectional variation in expected returns even for portfolios where previous studies have been unsuccessful. The pricing errors in portfolio returns using other asset pricing models can also be partly explained using skewness. Our results, however, show that the asset pricing puzzle is quite complex and the success of a given multifactor model depends substantially on the methodology and data used to empirically test the model. We also find that an important role is played by the degree of precision involved in computing the asset betas with respect to the factors—that is, what may be a proxy for estimation risk.

Our paper is organized as follows. In Section I, we use a general stochastic discount factor pricing framework to show how skewness can affect the expected excess asset returns. We also develop specific implications for the price of skewness risk based on utility theory. The data used in the paper and summary statistics are in Section II. Section III contains the econometric methodology and empirical results. Some concluding remarks are offered in Section IV.

## I. Skewness in Asset Pricing Theory

The first-order condition for an investor holding a risky asset (in a representative agent economy) for one period is

$$E[(1 + R_{i,t+1})m_{t+1}|\mathbf{\Omega}_t] = 1, \tag{1}$$

where $(1 + R_{i,t+1})$ is the total return on asset $i$, $m_{t+1}$ is the marginal rate of substitution of the investor between periods $t$ and $t + 1$, and $\mathbf{\Omega}_t$ is the information set available to the investor at time $t$. The marginal rate of substitution $m_{t+1}$ can be viewed as a pricing kernel or a stochastic discount factor that prices all risky asset payoffs.[2]

[2] See Harrison and Kreps (1979), Hansen and Richard (1987), Hansen and Jagannathan (1991), Cochrane (1994), Carhart et al. (1994), and Jagannathan and Wang (1996).

Under no arbitrage, the discount factor in equation (1), $m_{t+1}$, must be nonnegative (see Harrison and Kreps (1979)). The marginal rate of substitution is not observable. Hence, to obtain testable restrictions from this first-order condition, we need to define observable proxies for the marginal rate of substitution. Different asset pricing models differ primarily in the proxies they use for the marginal rate of substitution and the mechanisms they use to incorporate the proxies into the asset pricing model. The proxies can be either observed returns of financial assets such as equity portfolios or nonmarket variables such as growth rate in aggregate consumption as in Hansen and Singleton (1983). The form and specification of the marginal rate of substitution is determined jointly by the assumptions about preferences and distributions of the proxies. A specification for the marginal rate of substitution can also be viewed as a restriction on the set of trading strategies that the marginal investor can use to achieve the utility-maximizing portfolios. Thus, the standard capital asset pricing model implies that the optimal trading strategy for the marginal investor is to invest in the risk-free rate and the market portfolio.

## *A. A Three-Moment Conditional CAPM*

In the traditional CAPM, one of two routes is usually pursued. In a two-period world with homogeneous agents, the representative agent's derived utility function (in wealth) may be restricted to forms such as quadratic or logarithmic which guarantee that the discount factor is linear in the value-weighted portfolio of wealth. The other route involves making distributional assumptions on the asset returns, such as the elliptical class, which also guarantees that the discount factor is linear in the value-weighted portfolio of wealth. The empirical predictions (i.e., restrictions on the moments of the returns) are identical in either case. The assumption that the marginal rate of substitution is linear in the market return,

$$m_{t+1} = a_t + b_t R_{M,t+1}, \tag{2}$$

produces the *classic* CAPM with the weights $a_t$ and $b_t$ being functions of period-$t$ information set. To see this, expand the expectation in equation (1):

$$\mathrm{Cov}_t[m_{t+1},(1 + R_{i,t+1})] + E_t[1 + R_{i,t+1}]E_t[m_{t+1}] = 1, \tag{3}$$

which can also be written as

$$E_t[1 + R_{i,t+1}] = \frac{1}{E_t[m_{t+1}]} - \frac{\mathrm{Cov}_t[m_{t+1},(1 + R_{i,t+1})]}{E_t[m_{t+1}]}. \tag{4}$$

Assuming the existence of a conditionally risk-free asset and given equation (2), we get the standard CAPM

$$E_t[r_{i,t+1}] = \frac{\mathrm{Cov}_t[r_{i,t+1}, r_{M,t+1}]}{\mathrm{Var}_t[r_{M,t+1}]} E_t[r_{M,t+1}]$$

or

$$E_t[r_{i,t+1}] = \beta_{i,t} E_t[r_{M,t+1}], \tag{5}$$

where $r$ represents returns in excess of the conditionally risk-free return. This expression decomposes the expected excess return into the product of the asset's beta and the market risk premium. The econometric restriction such a model imposes is that in a time-series regression of the excess returns on the market excess return, the intercept should be zero, the betas should be significant, and the market risk premium estimate should be the same across all the assets. In a cross-sectional regression of the excess returns on the betas, the slope, the market risk premium, should be significantly different from zero.

An alternative to the linear specification is to assume that the marginal rate of substitution is nonlinear in its observed proxies. Here we are confronted with the large number of choices for nonlinear functions, each of which implies a different restriction on the marginal investor's trading strategies. We assume that the stochastic discount factor is quadratic in the market return; that is,

$$m_{t+1} = a_t + b_t R_{M,t+1} + c_t R_{M,t+1}^2. \tag{6}$$

We choose the quadratic form because we show later that the quadratic form can be linked to an important property that all admissible utility functions must have. Additionally, it also is one of the simplest types of nonlinearities.[3] The quadratic form for the marginal rate of substitution implies an asset pricing model where the expected excess return on an asset is determined by its conditional covariance with both the market return and the square of the market return (conditional coskewness).

Bansal and Viswanathan (1993) assume that the marginal rate of substitution is nonlinear in several factors and they directly test the first-order condition on the marginal rate of substitution; however, they do not have explicit expressions for premia for the risk factors in their model. In contrast, our approach involves a similar initial assumption that the marginal

[3] We can derive this expression for the marginal rate of substitution ab initio using several different models of preferences and return distributions or by using a second-order Taylor expansion of the marginal rate of substitution. Alternatively, a two-period model with asymmetric return distribution will also produce the same expression. For example, expected utility maximization in an infinite-horizon economy of representative agents with logarithmic preferences and an asymmetric return distribution will produce the expression for the marginal rate of substitution that includes $R_{M,t+1}^2$.

rate of substitution is a nonlinear function of market, SMB, and HML. However, with an explicit functional form for the marginal rate of substitution, we derive explicit expressions for risk premia. Additionally, our formulation permits us to accommodate nonincreasing absolute risk aversion. Nonincreasing absolute risk aversion (i.e., risk aversion should not increase if wealth increases) is a property that all utility functions should have. This property can be explicitly modeled as skewness in a two-period model.

Assuming the existence of a conditionally risk-free asset, we obtain

$$E_t[r_{i,t+1}] = \lambda_{1,t}\mathrm{Cov}_t[r_{i,t+1}, r_{M,t+1}] + \lambda_{2,t}\mathrm{Cov}_t[r_{i,t+1}, r_{M,t+1}^2] \tag{7a}$$

where

$$\lambda_{1,t} = \frac{\mathrm{Var}_t[r_{M,t+1}^2]E_t[r_{M,t+1}] - \mathrm{Skew}_t[r_{M,t+1}]E_t[r_{M,t+1}^2]}{\mathrm{Var}_t[r_{M,t+1}]\mathrm{Var}_t[r_{M,t+1}^2] - (\mathrm{Skew}_t[r_{M,t+1}])^2}, \tag{7b}$$

$$\lambda_{2,t} = \frac{\mathrm{Var}_t[r_{M,t+1}]E_t[r_{M,t+1}^2] - \mathrm{Skew}_t[r_{M,t+1}]E_t[r_{M,t+1}]}{\mathrm{Var}_t[r_{M,t+1}]\mathrm{Var}_t[r_{M,t+1}^2] - (\mathrm{Skew}_t[r_{M,t+1}])^2}. \tag{7c}$$

The restriction this model imposes on a cross-section of assets is that $\lambda_{1,t}$ and $\lambda_{2,t}$ are the same across all the assets and are statistically different from zero. This is the conditional version of the three-moment CAPM first proposed by Kraus and Litzenberger (1976) who use a utility function defined over the unconditional mean, standard deviation, and the third root of skewness.[4] Rewriting equation (7) as

$$E_t[r_{i,t+1}] = A_t E_t[r_{M,t+1}] + B_t E_t[r_{M,t+1}^2], \tag{8}$$

where $A_t$ and $B_t$ are functions of the market variance, skewness, covariance, and coskewness, illustrates the relation between our model and the Kraus and Litzenberger three-moment CAPM. $A_t$ and $B_t$ are analogous to the beta in the traditional CAPM.[5] Equation (8) is an empirically testable restriction imposed on the cross section of expected asset returns by the asset pricing model incorporating skewness, and as such it is an alternative to equation (5).

The empirical studies of asset pricing may be seen as attempts to find the best among these competing specifications of the pricing kernel. However, it is also possible that no one model solves the asset pricing puzzle and differ-

[4] Also see Friend and Westerfield (1980) and Ingersoll (1990). Alternative models with three moments are used by Sears and Wei (1985), Nummelin (1994), Lim (1989), and Waldron (1990). Coskewness could also be important for hedging the volatility shocks to the market portfolio as shown by Racine (1995).

[5] Another simple nonlinearity is to assume that the marginal investor's trading strategies are restricted to the risk-free asset and a call option on the market. This produces the Bawa and Lindenberg (1977) asset pricing model.

ent combinations of factors work for different settings. Therefore, we consider an asset pricing model that is a combination of the multifactor model along with a simple nonlinear component derived from skewness. Our choice is also consistent with the findings in Ghysels (1998) that nonlinear multifactor models are more successful empirically than linear beta models.

### B. How Skewness Enters Asset Pricing

The various asset pricing specifications can also be viewed as competing approximations for the discount factor or the intertemporal marginal rate of substitution. The nonmarket variables in equations (7) or (8) may also be viewed as proxies for the hedge portfolios (information about future returns) in a dynamic model such as that of Campbell (1993). If we relate the discount factor to the marginal rate of substitution between periods $t$ and $t+1$, in a two-period economy, a Taylor's series expansion allows us to make the following identification:

$$m_{t+1} = 1 + \frac{W_t U''(W_t)}{U'(W_t)} R_{M,t+1} + o(W_t), \tag{9}$$

where $o(W_t)$ is the remainder in the expansion and $W_t U''(W_t)/U'(W_t)$, which is $-b_t$ in equation (2), is relative risk aversion. Then $a_t = 1 + o(W_t)$ and $b_t < 0$. A negative $b_t$ implies that with an increase in next period's market return, the marginal rate of substitution declines. This decline in the marginal rate of substitution is consistent with decreasing marginal utility.

In a similar fashion we assume that the pricing kernel is quadratic in the market return, that is, $m_{t+1} = a_t + b_t R_{M,t+1} + c_t R^2_{M,t+1}$. Expanding, as before, the marginal rate of substitution in a power series gives

$$m_{t+1} = 1 + \frac{W_t U''(W_t)}{U'(W_t)} R_{M,t+1} + \frac{W_t^2 U'''(W_t)}{2U'(W_t)} R^2_{M,t+1} + o(W_t). \tag{10}$$

Then $b_t < 0$ and $c_t > 0$ since nonincreasing absolute risk aversion implies $U''' > 0$.[6] According to Arrow (1964), nonincreasing absolute risk aversion is one of the essential properties for a risk-averse individual.

Nonincreasing absolute risk aversion for a a risk-averse utility-maximizing agent can also be linked to prudence as defined by Kimball (1990). Prudence relates to the desire to avoid disappointment and is usually linked to the precautionary savings motive. Nonincreasing absolute risk aversion implies that in a portfolio, increases in total skewness are preferred. Since adding an asset with negative coskewness to a portfolio makes the resultant port-

[6] Nonincreasing absolute risk aversion implies that its derivative should be less than or equal to zero. $U''' \geq 0$ is a necessary condition to satisfy this. Also see Scott and Horvath (1980) for a discussion of the preference of moments beyond variance.

folio more negatively skewed (i.e., reduces the total skewness of the portfolio), assets with negative coskewness must have higher expected returns than assets with identical risk-characteristics but zero-coskewness. Thus, in a cross section of assets, the slope of the excess expected return on conditional coskewness with the market portfolio should be negative. Thus, the premium for skewness risk over the risk-free asset's return (assuming that the risk-free asset possesses zero betas with respect to all the factors being examined to explain the cross section of returns) should also be negative. In equation (7) we are able to decompose contributions of conditional covariance and coskewness with the market to the expected excess return of a specific asset. Alternative nonlinear frameworks such as Bansal and Viswanathan (1993) are unable to provide this decomposition.

### *C. The Geometry of Mean-Variance-Skewness Efficient Portfolios*

Figure 1, Panel A, presents a mean-variance-skewness surface. Slicing the surface at any level of skewness, we get the familiar positively sloping portion of the mean-variance frontier. Skewness adds the following possibility: at any level of variance, there is a negative trade-off of mean return and skewness. That is, to get investors to hold low or negatively skewed portfolios, the expected return needs to be higher. This is evident in the graph.

Panel B of Figure 1 introduces the risk-free rate. The capital market "line" starts out at zero variance–zero skewness. Think of a ray from the risk-free rate (at zero variance) that is tangent to the surface at a particular variance-skewness combination. For that level of variance, there are many possible portfolios with different skewnesses. The tangency point is the one with the highest skewness. Now add another ray from the risk-free rate that is tangent to a different variance-skewness point.

In the usual mean variance analysis, there is a single efficient risky-asset portfolio. In the mean-variance-skewness analysis, however, there are multiple efficient portfolios. The optimal portfolio for the investor is chosen as the tangency of the investor's indifference surface to the capital market plane.

## II. Does Skewness Exist in the Returns Data?— Portfolio Formation and Summary Statistics

For the empirical work, we use monthly U.S. equity returns from CRSP NYSE/AMEX and Nasdaq files. We form portfolios from the equities as well as analyze individual equity returns. Most of our work focuses on the period July 1963 to December 1993. We use a longer sample to investigate the interactions of momentum and skewness. As factors capable of explaining cross-sectional variations in excess returns, we use the CRSP NYSE/AMEX value-weighted index as the market portfolio. To capture the effects of size and book-to-market value, we use the SMB and HML hedge portfolios formed by Fama and French. These portfolios are constructed to capture the

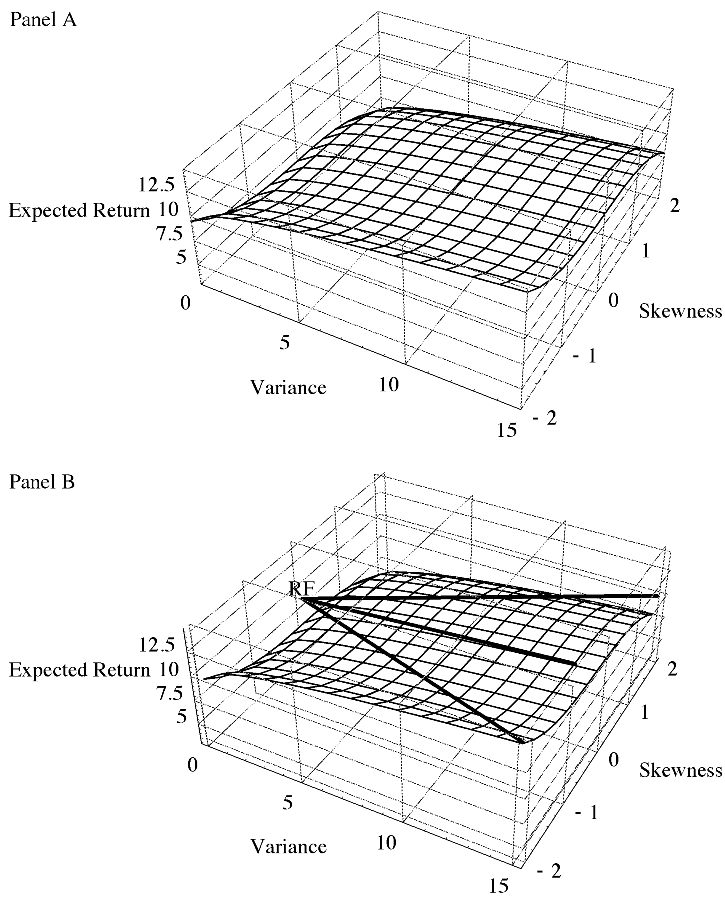


**Figure 1. A mean–variance–skewness surface.** The trade-offs between mean, variance, and skewness are illustrated. The surfaces are generated using a positive trade-off between mean and variance and a negative trade-off between mean and skewness. Panel A presents the surface without a risk-free rate. In Panel B, rays are drawn from the risk-free rate to be tangential to the surface. The tangent points represent efficient portfolios.

marketwide effect of size and book-to-market value. The average annualized returns on these portfolios from July 1963 to December 1993 are 3.5 percent and 5.6 percent respectively.

Table I presents some summary statistics that compare the different measures of coskewness across five portfolio groups. The first group represents 32 value-weighted industry portfolios.[7] The second set are the 25 portfolios sorted on size and book-to-market value used by Fama and French (1995,

[7] Of the 32 industry portfolios, we exclude five portfolios from the regression because they include fewer than 10 firms. Summary statistics for portfolios constructed on other criteria are available from the authors.

**Table I**

## Summary Statistics on Portfolios

This table summarizes four sets of portfolios formed from monthly U.S. equity returns. The market portfolio is the value-weighted NYSE/AMEX index. Standardized unconditional skewness is the third central moment about the mean. Standardized unconditional coskewness of the $i$th asset is defined as $E[\epsilon_{i,t}\epsilon_{M,t}^2]/\sqrt{E[\epsilon_{i,t}^2]E[\epsilon_{M,t}^2]}$, where $\epsilon_t$ are residuals from regressing the excess return of asset $i$ on the market return. The $\beta$s are computed from univariate regressions of the portfolio return on the risk factor. Time-variation in conditional coskewness is captured through the autoregression $E_t[\epsilon_{i,t+1}\epsilon_{M,t+1}^2] = \rho_0 + \rho_1\epsilon_{i,t}\epsilon_{M,t}^2 + \rho_2\epsilon_{i,t-1}\epsilon_{M,t-1}^2$ and whether it is significant at the 10 percent level. Cross-sectional correlations between the average excess returns and other portfolio-specific variables are also reported. S = smallest third in size, M = middle third in size, B = largest third in size. Significance levels for unconditional skewness and coskewness are computed by generating the statistic 10,000 times by simulating it under the null, specifically using a *Normal*(0,1) for skewness, and an ARMA(2,0) process using a bivariate *Normal* for coskewness. With 366 observations, skewness is significant at −0.253 and 0.254 at the 5 percent level and −0.210 and 0.212 at the 10 percent level. Coskewness is significant at ±0.095 at the 10 percent level and ±0.108 at the 5 percent level.

Panel A. Portfolios Formed on Industrial Classification, July 1963–December 1993

| Industry | Standardized Unconditional Skewness | Standardized Unconditional Coskewness | $\beta$ to $S^-$-$S^+$ | $\beta$ to $S^-$-$R_f$ | $\beta$ to $(R_M - R_f)^2$ | Time-Varying Coskewness | Average Excess Return | $\beta$ to $R_M - R_f$ | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|
| Extractive | −0.305** | −0.102** | −0.518** | 0.745** | −0.013* | Yes | 0.642 | 0.834** | 5.414 |
| Oil & gas | −0.016 | −0.032 | −0.738** | 0.746** | −0.010 | No | 0.499 | 0.879** | 5.276 |
| Building & construction | 0.037 | 0.103** | −0.058 | 1.223** | −0.010 | No | 0.467 | 1.278** | 6.257 |
| Chemicals | −0.258** | −0.003 | 0.043 | 0.953** | −0.010 | Yes | 0.402 | 0.976** | 4.723 |
| Computers, electrical & electronics & electronic equipment | −0.140 | 0.103** | 0.079 | 1.040** | −0.008 | Yes | 0.377 | 1.065** | 5.355 |
| Engineering—Primary metals, machining | −0.364** | −0.157** | −0.230* | 1.058** | −0.015* | Yes | 0.363 | 1.129** | 5.551 |
| Vehicles | 0.052 | −0.230* | 0.221 | 0.934** | −0.020** | No | 0.461 | 0.941** | 5.914 |
| Paper, pulp, & printing | 0.476** | 0.063 | −0.078 | 1.010** | −0.009 | Yes | 1.071 | 1.056** | 6.644 |
| Textiles & apparel | −0.284** | −0.204** | 0.127 | 1.130** | −0.019** | No | 0.570 | 1.148** | 6.162 |
| Food manufacturers | 0.081 | −0.012 | 0.167 | 0.854** | −0.009 | Yes | 0.641 | 0.848** | 4.638 |
| Beverages | −0.166 | 0.109** | 0.310** | 0.981** | −0.006 | No | 0.800 | 0.961** | 5.150 |
| Household goods | 0.075 | 0.118** | 0.145 | 1.095** | −0.005 | No | 0.574 | 1.111** | 5.865 |
| Healthcare | 0.112 | 0.136** | 0.198 | 1.469** | −0.005 | Yes | 0.952 | 1.465** | 9.597 |
| Pharmaceuticals | −0.129 | 0.009 | −0.269* | 1.033** | −0.010 | Yes | 0.486 | 1.099** | 5.756 |
| Tobacco | 0.018 | 0.077 | 0.128 | 0.888** | −0.006 | Yes | 0.994 | 0.891** | 5.683 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Distributors | −0.293** | −0.058 | −0.053 | 1.174** | −0.014* | No | 0.583 | 1.223** | 6.258 |
| Leisure and hotels | −0.436** | −0.212** | 0.482** | 1.381** | −0.023** | Yes | 0.889 | 1.351** | 7.163 |
| Media | −0.194 | −0.163** | 0.158 | 1.155** | −0.018** | Yes | 0.801 | 1.175** | 6.214 |
| Food retailers | 0.715** | −0.005 | 0.051 | 0.900** | −0.009 | No | 0.455 | 0.909** | 5.429 |
| General retailers | −0.143 | −0.139** | 0.382** | 1.174** | −0.018** | No | 0.593 | 1.151** | 6.060 |
| Support services | −0.014 | 0.078 | 0.078 | 1.273** | −0.012 | No | 0.609 | 1.309** | 6.628 |
| Transportation | −0.178 | −0.020 | −0.077 | 1.147** | −0.012 | No | 0.446 | 1.196** | 6.265 |
| Electric & water | 0.462 | 0.184** | 0.193** | 0.612** | 0.002 | No | 0.251 | 0.595** | 4.085 |
| Telecommunications | 0.068 | 0.002 | 0.110 | 0.572** | −0.004 | No | 0.356 | 0.573** | 4.087 |
| Depository financial institutions | 0.129 | 0.140** | 0.182 | 1.111** | −0.005 | Yes | 0.342 | 1.120** | 6.432 |
| Nondepository financial institutions & brokerages | 0.283** | 0.280** | 0.168 | 1.124** | −0.000 | Yes | 0.530 | 1.135** | 6.023 |
| Holding companies & investment companies | −0.171 | 0.089 | −0.200* | 0.888** | −0.006 | No | 0.486 | 0.952** | 4.512 |
| Property | 0.173 | −0.068 | 0.069 | 1.263** | −0.017 | Yes | 0.327 | 1.312** | 7.830 |
| Agriculture & forestry | 0.592** | −0.075 | 0.128 | 1.100** | −0.016 | No | 0.462 | 1.119** | 9.920 |
| Aerospace, aircraft | −0.118 | −0.087 | 0.069 | 1.196** | −0.015 | Yes | 0.562 | 1.241** | 6.421 |
| Oil & gas transportation | 0.118 | 0.045 | −0.488** | 0.725** | −0.006 | Yes | 0.408 | 0.824** | 4.605 |
| Auto & gas retailers | 0.363** | 0.074 | 0.089 | 1.195** | −0.012 | Yes | 1.430 | 1.226** | 9.205 |
| Correlation with $\bar{r}$ | −0.143 | −0.067 | 0.207 | 0.288 | −0.161 | | | 0.260 | 0.348 |

Panel B. Portfolios Formed on Size and Book/Market Value, July 1963–December 1993

| Size Quintile | Book/ Market Quintile | Standardized Unconditional Skewness | Standardized Unconditional Coskewness | $\beta$ to $S^{-}$-$S^{+}$ | $\beta$ to $S^{-}$-$R_f$ | $\beta$ to $(R_M - R_f)^2$ | Time-Varying Coskewness | Average Excess Return | $\beta$ to $R_M - R_f$ | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | −0.303** | −0.276** | 0.138 | 1.339** | −0.027** | Yes | 0.310 | 1.403** | 7.665 |
| | 2 | −0.274** | −0.330** | 0.166 | 1.216** | −0.026** | Yes | 0.698 | 1.263** | 6.744 |
| | 3 | −0.359** | −0.349** | 0.160 | 1.098** | −0.024** | Yes | 0.818 | 1.142** | 6.135 |
| | 4 | −0.135 | −0.350** | 0.196 | 1.023** | −0.024** | Yes | 0.949 | 1.054** | 5.842 |
| | 5 | 0.001 | −0.334** | 0.234* | 1.048** | −0.025** | No | 1.082 | 1.071** | 6.142 |
| 2 | 1 | −0.416** | −0.196** | −0.002 | 1.337** | −0.020** | No | 0.481 | 1.422** | 7.128 |
| | 2 | −0.445** | −0.332** | 0.085 | 1.185** | −0.022** | Yes | 0.720 | 1.246** | 6.250 |
| | 3 | −0.384** | −0.372** | 0.116 | 1.078** | −0.022** | Yes | 0.905 | 1.124** | 5.708 |
| | 4 | −0.268** | −0.257** | 0.131 | 0.995** | −0.017** | Yes | 0.921 | 1.030** | 5.231 |
| | 5 | −0.306** | −0.328** | 0.128 | 1.085** | −0.022** | Yes | 1.095 | 1.127** | 5.943 |

*Continued*

**Table I**—*Continued*

Panel B (*Continued*)

| Size Quintile | Book/ Market Quintile | Standardized Unconditional Skewness | Standardized Unconditional Coskewness | $\beta$ to $S^-$-$S^+$ | $\beta$ to $S^-$-$R_f$ | $\beta$ to $(R_M - R_f)^2$ | Time-Varying Coskewness | Average Excess Return | $\beta$ to $R_M - R_f$ | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | −0.353** | −0.181 | 0.098 | 1.277** | −0.018** | No | 0.439 | 1.344** | 6.512 |
| | 2 | −0.566** | −0.324** | 0.074 | 1.091** | −0.018** | Yes | 0.676 | 1.146** | 5.527 |
| | 3 | −0.552** | −0.341** | 0.059 | 0.986** | −0.018** | Yes | 0.746 | 1.036** | 5.111 |
| | 4 | −0.277** | −0.173** | 0.091 | 0.930** | −0.013** | Yes | 0.857 | 0.965** | 4.794 |
| | 5 | −0.377** | −0.261** | 0.122 | 1.020** | −0.018** | Yes | 1.055 | 1.060** | 5.484 |
| 4 | 1 | −0.244* | 0.053 | 0.020 | 1.174** | −0.010 | Yes | 0.511 | 1.241** | 5.857 |
| | 2 | −0.491** | −0.232** | −0.066 | 1.053** | −0.015** | Yes | 0.388 | 1.131** | 5.273 |
| | 3 | −0.314** | −0.158** | 0.039 | 0.991** | −0.013* | Yes | 0.638 | 1.043** | 4.975 |
| | 4 | 0.177 | 0.098 | 0.123 | 0.929** | −0.006 | Yes | 0.799 | 0.965** | 4.811 |
| | 5 | −0.066 | −0.070* | 0.154 | 1.074** | −0.012 | Yes | 1.039 | 1.112** | 5.664 |
| 5 | 1 | −0.069 | 0.214 | 0.049 | 0.974** | −0.005 | Yes | 0.366 | 1.025** | 4.842 |
| | 2 | −0.286** | 0.013 | −0.149 | 0.910** | −0.009 | Yes | 0.387 | 0.994** | 4.604 |
| | 3 | −0.104 | 0.039 | −0.206** | 0.801** | −0.007 | No | 0.370 | 0.883** | 4.277 |
| | 4 | 0.189 | 0.186** | −0.032 | 0.777** | −0.003 | Yes | 0.551 | 0.832** | 4.181 |
| | 5 | 0.131 | 0.020 | −0.034 | 0.839** | −0.007 | Yes | 0.715 | 0.889** | 4.901 |
| Correlation with $\bar{r}$ | | 0.067 | −0.498 | 0.648 | −0.021 | −0.319 | | | −0.092 | 0.122 |

Panel C. Portfolios Formed on Size Deciles, July 1963–December 1993

| Size | Portfolio No. | Standardized Unconditional Skewness | Standardized Unconditional Coskewness | $\beta$ to $S^-$-$S^+$ | $\beta$ to $S^-$-$R_f$ | $\beta$ to $(R_M - R_f)^2$ | Time-Varying Coskewness | Average Excess Return | $\beta$ to $R_M - R_f$ | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|---|
| Smallest | 1 | 1.156** | −0.181** | 0.410** | 1.011** | −0.022* | No | 2.122 | 1.005** | 7.994 |
| | 2 | 0.390** | −0.292** | 0.249 | 1.026** | −0.026** | Yes | 1.051 | 1.052** | 6.992 |
| | 3 | 0.144 | −0.302** | 0.212 | 1.061** | −0.025** | Yes | 0.702 | 1.096** | 6.639 |
| | 4 | 0.112 | −0.294** | 0.235 | 1.097** | −0.024** | Yes | 0.664 | 1.130** | 6.465 |
| | 5 | −0.154 | −0.330** | 0.162 | 1.107** | −0.024** | Yes | 0.526 | 1.151** | 6.253 |
| | 6 | −0.213* | −0.307** | 0.163 | 1.120** | −0.021** | Yes | 0.516 | 1.163** | 6.052 |
| | 7 | −0.386** | −0.341** | 0.082 | 1.115** | −0.021** | Yes | 0.505 | 1.173** | 5.841 |
| | 8 | −0.445** | −0.312** | 0.076 | 1.090** | −0.018** | Yes | 0.574 | 1.145** | 5.548 |
| | 9 | −0.513** | −0.313** | 0.024 | 1.050** | −0.016** | Yes | 0.549 | 1.111** | 5.198 |
| Largest | 10 | −0.232* | 0.236** | −0.011 | 0.984** | −0.007 | Yes | 0.437 | 1.044** | 4.683 |
| Correlation with $\bar{r}$ | | 0.697 | 0.370 | 0.794 | −0.358 | −0.636 | | | −0.527 | 0.818 |

Panel D. Twenty-Seven Portfolios Formed on Book/Market, Size, Momentum, July 1963–December 1993

| B/M | Size | Momentum | Portfolio No. | Standardized Unconditional Skewness | Standardized Unconditional Coskewness | $\beta$ to $S^{-}$-$S^{+}$ | $\beta$ to $S^{-}$-$R_f$ | $\beta$ to $(R_M - R_f)^2$ | Time-Varying Coskewness | Average Excess Return | $\beta$ to $R_M - R_f$ | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | S | Loser | 1 | −0.013 | −0.128* | 0.010 | 1.209** | −0.017* | Yes | −0.224 | 1.282** | 6.719 |
| | S | Middle | 2 | −0.465** | −0.264** | 0.035 | 1.165** | −0.019** | Yes | 0.388 | 1.231** | 6.146 |
| | S | Winner | 3 | −0.718** | −0.425** | −0.020 | 1.271** | −0.028** | No | 0.995 | 1.354** | 6.822 |
| Low | M | Loser | 4 | 0.086 | 0.164** | 0.024 | 1.153** | −0.006 | No | −0.074 | 1.216** | 6.017 |
| | M | Middle | 5 | −0.468** | −0.182** | −0.024 | 1.067** | −0.014** | Yes | 0.189 | 1.138** | 5.396 |
| | M | Winner | 6 | −0.569 | −0.291** | −0.060 | 1.173** | −0.019** | No | 0.962 | 1.258** | 6.087 |
| | B | Loser | 7 | 0.173 | 0.406* | −0.000 | 0.997** | 0.002 | Yes | 0.131 | 1.046** | 5.281 |
| | B | Middle | 8 | −0.159 | 0.135* | 0.028 | 0.913** | −0.006 | Yes | 0.271 | 0.960** | 4.547 |
| | B | Winner | 9 | −0.312** | −0.102 | −0.138 | 0.981** | −0.012* | No | 0.655 | 1.069** | 5.399 |
| | S | Loser | 10 | 0.454** | −0.058 | 0.180 | 1.057** | −0.012 | Yes | 0.361 | 1.093** | 6.058 |
| | S | Middle | 11 | −0.214* | −0.306** | 0.087 | 0.919** | −0.018** | Yes | 0.686 | 0.960** | 5.014 |
| | S | Winner | 12 | −0.736** | −0.532** | 0.111 | 1.149** | −0.029** | Yes | 1.206 | 1.201** | 6.158 |
| | M | Loser | 13 | 0.494** | 0.270** | 0.064 | 0.977** | 0.000 | No | 0.432 | 1.024** | 5.461 |
| M | M | Middle | 14 | −0.359** | −0.272** | 0.042 | 0.879** | −0.014** | Yes | 0.567 | 0.924** | 4.546 |
| | M | Winner | 15 | −0.905** | −0.592** | 0.052 | 1.041** | −0.025** | Yes | 0.825 | 1.095** | 5.384 |
| | B | Loser | 16 | 0.466** | 0.499** | −0.003 | 0.844** | 0.008 | No | 0.471 | 0.893** | 4.855 |
| | B | Middle | 17 | 0.160 | 0.141* | −0.019 | 0.793** | −0.004 | No | 0.401 | 0.851** | 4.240 |
| | B | Winner | 18 | −0.352** | −0.180** | −0.146 | 0.902** | −0.013** | No | 0.635 | 0.985** | 4.902 |
| | S | Loser | 19 | 0.961** | −0.109 | 0.294** | 1.043** | −0.015* | No | 0.596 | 1.057** | 6.519 |
| | S | Middle | 20 | 0.187 | −0.316** | 0.243** | 0.963** | −0.022** | No | 1.105 | 0.978** | 5.649 |
| | S | Winner | 21 | −0.302** | −0.455** | 0.225* | 1.099** | −0.030** | Yes | 1.396 | 1.124** | 6.320 |
| High | M | Loser | 22 | 0.461** | −0.043 | 0.174 | 1.071** | −0.012 | No | 0.617 | 1.105** | 6.158 |
| | M | Middle | 23 | −0.026 | −0.250** | 0.158 | 0.992** | −0.017** | No | 0.963 | 1.021** | 5.360 |
| | M | Winner | 24 | −0.832** | −0.538** | 0.094 | 1.092** | −0.029** | Yes | 1.371 | 1.142** | 5.928 |
| | B | Loser | 25 | 0.709** | 0.225** | 0.133 | 0.959** | −0.000 | No | 0.645 | 0.998** | 5.629 |
| | B | Middle | 26 | −0.058 | −0.106 | −0.045 | 0.881** | −0.011* | Yes | 0.645 | 0.944** | 4.782 |
| | B | Winner | 27 | −0.266** | −0.187** | −0.013 | 0.975** | −0.015** | Yes | 0.988 | 1.024** | 5.378 |
| Correlation with $\bar{r}$ | | | | −0.407 | −0.705 | 0.237 | 0.081 | −0.685 | | | 0.068 | 0.201 |

** and * denote $t$-statistics significant at the 5 percent and 10 percent levels, respectively.

1996). Third, we investigate 10 momentum portfolios formed by sorting on past return over $t-12$ to $t-2$ months and holding the stock for six months. The fourth group are size (market capitalization) deciles used in a number of empirical studies. Finally, we look at the three-way classification based on book-to-market value, size, and momentum detailed in Carhart (1997).[8] We describe four ways to compute coskewness. The first two are "direct" measures and the last two are based on sensitivities to coskewness hedge portfolios (much in the same way Fama and French construct factor loadings on SMB and HML). Figure 2 plots the density functions for the market risk premium, SMB portfolio, and the smallest and largest size deciles. The skewness in the smallest decile is prominent.

We first construct a direct measure of coskewness, $\beta_{SKD}$, which is defined as

$$\hat{\beta}_{\mathrm{SKD}_i} = \frac{E[\epsilon_{i,t+1}\epsilon_{M,t+1}^2]}{\sqrt{E[\epsilon_{i,t+1}^2]}E[\epsilon_{M,t+1}^2]}, \tag{11}$$

where $\epsilon_{i,t+1} = r_{i,t+1} - a_i - \beta_i(r_{M,t+1})$, the residual from the regression of the excess return on the contemporaneous market excess return. $\beta_{SKD}$ represents the contribution of a security to the coskewness of a broader portfolio. A negative measure means that the security is adding negative skewness. According to our utility assumptions, a stock with negative coskewness should have a higher expected return—that is, the premium should be negative.

Another approach to estimating coskewness is to regress the asset return on the square of the market return. Although we report in Table I the coefficient on the square term, we believe that there are two advantages to examining $\beta_{SKD}$. The first is that $\hat{\beta}_{\mathrm{SKD}_{i,t}}$ is constructed from residuals that are independent of the market return by construction. The second is that $\beta_i$ is similar to the traditional CAPM beta. As defined, standardized coskewness is unit free and analogous to a factor loading.[9]

We investigate two value-weighted hedge portfolios that capture the effect of coskewness. Using 60 months of returns, we compute the standardized direct coskewness for each of the stocks in the NYSE/AMEX and the Nasdaq universe. We then rank the stocks based on their past coskewness and form three value-weighted portfolios: 30 percent with the most negative coskewness, which we call $S^-$; the middle 40 percent, which we call $S^0$; and 30 percent with the most positive coskewness, which we call $S^+$. The 61st month

[8] We thank Mark Carhart for giving us these data used in Carhart (1997.) These portfolios are formed by dividing all stocks into thirds based on book/market values. These portfolios are then divided into three portfolios based on size. The second-level portfolios are then divided into "losers," "middle," and "winners" based on their past 12-month performance. Thus, there are 27 portfolios.

[9] $\beta_{SKD}$ is related to the coefficient obtained from regressing the excess return on the square of the market return, if the market return and squared market return are orthogonalized. The numerator of $\beta_{SKD}$ is also similar to $\mathrm{Cov}[r_{i,t+1}, r_{M,t+1}^2]$ in equation (7a).

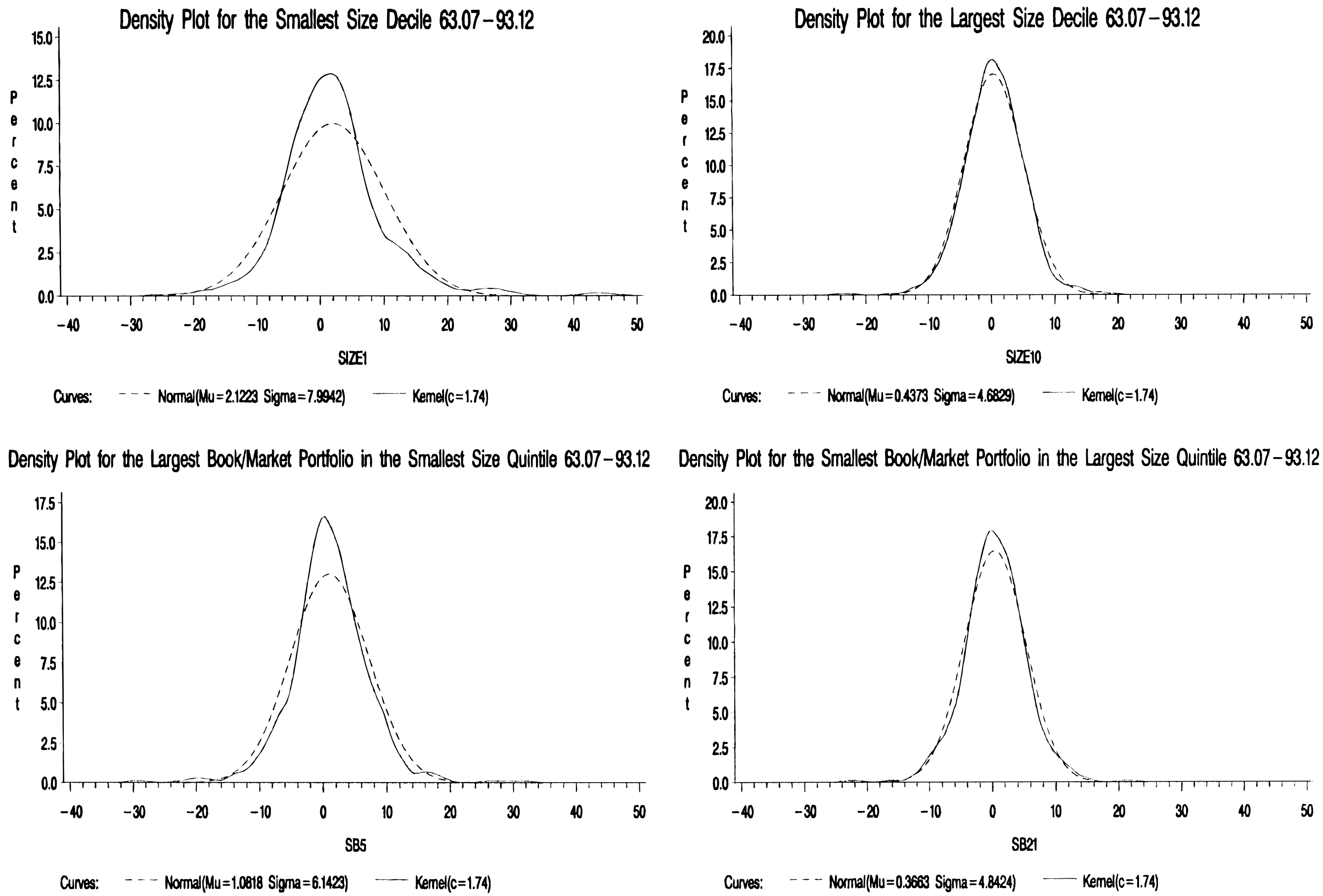


**Figure 2. Density plots for the extreme size and size and book/market portfolios.** The nonparametric density is computed using a quadratic kernel with the smoothing parameter selected by minimizing mean integrated squared error.

(i.e., post-ranking) excess returns on $S^-$ and $S^+$ are then used to proxy for systematic skewness. The average annualized spread between the returns on the $S^-$ and $S^+$ portfolios is 3.60 percent over the period July 1963 to December 1993 (this is greater than the return on the SMB portfolio over the same period.) We reject the hypothesis that the mean spread is zero at the 5 percent level of significance. We compute the coskewness for a risky asset from its beta with the spread between the returns on the $S^-$ and $S^+$ portfolios and call this measure $\beta_{\mathrm{SKS}}$. Another measure of coskewness for an asset is from its beta with the excess return on the $S^-$ portfolio. We call this measure $\beta_{S^-}$. For the hedge portfolios, a high factor loading should be associated with high expected returns. This is analogous to the factor loading on the SMB portfolio in the Fama–French model where SMB is defined as the return on the small-size stocks minus the return on the large-size stocks. This difference, that is, the risk premium for SMB factor loading—should be positive. Analogously, the risk premium for the skewness factor loading should be positive.

Table I also reports the unconditional skewness, a test of whether coskewness is time-varying, the beta implied by the CAPM, the average return, and the standard deviation. We also report the cross-sectional correlation between a number of these risk measures and the average portfolio returns. Our test of time-varying coskewsness involves the estimation of the first two autocorrelations for $(\epsilon_{i,t}\epsilon^2_{M,t})$. An alternative method for capturing the time-series variation in skewness is provided in Harvey and Siddique (1999). Their approach involves using the noncentral-*t* distribution. We also examine and document (but do not report) time-variation in the conditional moments of the market returns including skewness.

The results in Table I are intriguing. For the industry portfolios in Panel A there is a negative correlation between the direct measures of coskewness and the mean returns and a positive association between the hedge portfolio loadings and the average returns—both as expected. Additionally, the loadings on the hedge portfolio appear to contain as much information as the CAPM betas. Different industries possess very different standardized unconditional coskewness, with the Vehicles industry having the most negative coskewness of −0.230 and the Nondepository Financial Institutions industry having the most positive coskewness of 0.280.[10] We compute the standard errors for standardized unconditional coskewness using 10,000 simulations to generate a test statistic under the null hypothesis of zero coskewness.

The results get more interesting when we examine the portfolio groupings that pose the greatest challenges to asset pricing models. In the 25 size and book-to-market value-sorted portfolios in Panel B, the highest mean return

[10] We compute these statistics without September, October, and November of 1987 as well. For these two industries, coskewness without these three months becomes −0.190 and 0.245 respectively. We also examine equity indices from eight countries using the Morgan Stanley Capital International (MSCI) world index as the market portfolio. Most have negative standardized coskewness as well. The market portfolio, measured by the NYSE/AMEX index, displays negative skewness.

portfolios have the smallest direct coskewness measures. There is a $-0.50$ correlation between the mean returns and the direct skewness measure. There is an even stronger relation with the SKS hedge portfolio. The differences in the factor loadings have 0.65 correlation with the mean returns. In this case, there is little evidence of a relation between the $S^-$ portfolio and the mean returns.

The size deciles are presented in Panel C. There is some evidence that coskewness is important. The high expected return portfolio, decile one (small capitalization), has a negative direct coskewness and low expected return portfolio, decile 10 (large capitalization), has a positive coskewness. However, the results for the middle portfolios are ambiguous. The factor loadings on SKS are almost monotonically decreasing as size increases. The correlation between the SKS betas and the average returns is almost 0.80.

The fourth group is the three-way (size, book-to-market, and momentum) sorted portfolios presented in Panel D. There is a remarkably sharp relation between the direct measure of coskewness and the mean returns ($-0.71$ correlation). There is also information in the hedge portfolio SKS betas that is relevant for the cross section of mean returns.

We are concerned that our results may be highly sensitive to the October 1987 observation. We estimate each panel with and without the last three months of 1987 and find that although the measures of coskewness change, the inference from the table does not change. We even compute the average conditional coskewness, average of $\beta_{\mathrm{SKD}_t}$, for all stocks in the United States month by month, using 60 months of observations for the conditional coskewness of the 61st month. The impact of the crash of 1987 on conditional coskewness is striking. The average coskewness for all stocks in the United States increases from $-0.03$ in September to 0.11 in October. However, the crash also causes a substantial increase in the cross-sectional dispersion of coskewness across the stocks.

The summary statistics suggest that coskewness plays a role in explaining the cross section of asset returns. Next, we formally test the information in coskewness relative to alternative asset pricing models.

## III. Results

### A. Can Skewness Explain What Other Factors Do Not?

The failures of traditional asset pricing models often appear in specific groups of securities such as those formed on "momentum" and small size stocks. One method to understand how skewness enters asset pricing is to analyze the pricing errors from other asset pricing models.

Fama and French (1995) carry out time-series regressions of excess returns,

$$r_{i,t} = \alpha_i + \hat{\beta}_i r_{\mathrm{M,t}} + \hat{s}_i \mathrm{SMB}_t + \hat{h}_i \mathrm{HML}_t + e_{i,t} \quad \text{for } i = 1,\ldots,N,\ t = 1,\ldots,T, \quad (12)$$

and jointly test whether the intercepts, $\alpha_i$, are different from zero using the $F$-test of Gibbons, Ross, and Shanken (1989) where $F \sim (N, T - N - 1)$. We test the Fama–French model for the momentum portfolios described in Panel C of Table I. The inclusion of the $S^-$ portfolio reduces the $F$-statistic from 42.82 to 2.57. Similarly, when we form 25 portfolios sorted by coskewness over July 1963 to December 1993, inclusion of the $S^-$ portfolio reduces the $F$-statistic from 68.74 using three factors to 0.82 when the skewness factor is added. We find similar results for the 27 momentum portfolios formed by Carhart (1997).[11] We carry out these tests for several other sets of portfolios and the results are reported in Table II. In all cases, the inclusion of a skewness factor dramatically reduces the Gibbons–Ross–Shanken $F$-statistic.

Different kinds of pricing errors arise from the cross-sectional regressions

$$r_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + e_i \quad \text{for } i = 1,\ldots,N, \tag{13}$$

where the $\lambda$s are computed every month in a two-step estimation using time-series betas from a Fama–MacBeth procedure. We take the pricing errors (i.e., the intercepts or $\lambda_0$) from these cross-sectional regressions and compute correlations between these pricing errors and the ex post realizations on the $S^-$ portfolio.

For the 10 momentum portfolios formed on short-term performance (from $t - 12$ to $t - 2$), the correlation between pricing errors and the ex post realizations on the $S^-$ portfolio is 0.61 over the period July 1964 to December 1993. Using Fisher's logarithmic transformation ($\frac{1}{2}[(1 + r)/(1 - r)]$) for computing the standard error of correlation coefficients, the correlation with 366 observations is 0.11 at the 5 percent significance level. Therefore, a correlation of 0.61 is highly significant.

We also examine other portfolio sets and find significant correlations between the pricing errors and the $S^-$ portfolio. In the case of the momentum portfolios where 25 portfolios are formed using the past six months of returns and holding period returns are computed over the next six months, the pricing errors have a correlation with $S^-$ portfolio of 0.35. For the 25 Fama–French portfolios formed on book/market and size, the correlation is 0.31. The corresponding correlations for the 27 momentum and 27 industry portfolios are 0.33 and 0.33, respectively. For the individual equities, when the Fama–French factors are used as explanatory variables, the correlation between the pricing errors and return on $S^-$ portfolio is 0.53. When the Fama–French factors are replaced with firm-specific market/book ratio and market value of equity, the correlation is 0.41.

[11] In addition to the three Fama–French factors we use the fourth factor used by Carhart (1997) and find that it has an effect similar to that of the $S^-$ portfolio for the 27 portfolios formed by Carhart but not of the other portfolio sets. Our results are also invariant for other multivariate tests using the intercepts as well as different methods for estimating the variance-covariance matrix.

**Table II**

**Tests of Intercepts from the Fama–French Model**

We report the results from multivariate tests on intercepts from time-series regressions with the three Fama–French factors and four factors including skewness as defined by the excess return on $S^-$ portfolio. The test-statistic is the Gibbons–Ross–Shanken $F$-test statistic distributed as $F \sim (N, T - N - 1)$, where $N$ is the number of portfolios and $T$ is the number of observations. The significance levels are presented in parentheses. The correlation is the correlation of intercepts obtained from month-by-month cross-sectional Fama–French regressions on the three Fama–French factors with the ex post return on the $S^-$ portfolio for 366 months. A correlation above 0.11 is significantly different from zero at the 10 percent level using the Fisher transformation.

| Criterion | No. of Portfolios | Period | $F$-test for Three Factors | $F$-test for Four Factors (with $S^-$) | Correlation with $S^-$ |
|---|---|---|---|---|---|
| Industrial, | 27 | 1963.07–1993.12 | 8.56 | 1.40 | 0.330 |
| one-month holding | | | (0.000) | (0.093) | |
| Size and B/M sorted, | 25 | 1963.07–1993.12 | 1.92 | 1.43 | 0.340 |
| one-month holding | | | (0.006) | (0.086) | |
| Size, | 10 | 1963.07–1993.12 | 12.32 | 7.56 | 0.410 |
| one-month holding | | | (0.000) | (0.003) | |
| $t - 12$, $t - 2$ momentum, | 10 | 1964.07–1995.12 | 42.82 | 2.57 | 0.120 |
| six-month holding | | | (0.000) | (0.010) | |
| Book/market, size | 27 | 1963.07–1993.12 | 4.63 | 1.82 | 0.312 |
| $t - 12$, $t - 2$ momentum, | | | (0.000) | (0.011) | |
| one-month holding | | | | | |
| Other portfolios | | | | | |
| $t - 12$, $t - 2$ momentum, | 10 | 1964.07–1995.12 | 11.36 | 1.56 | 0.610 |
| one-month holding | | | (0.000) | (0.118) | |
| Coskewness, | 25 | 1963.07–1993.12 | 74.59 | 0.698 | 0.306 |
| one-month holding | | | (0.000) | (0.859) | |
| Coskewness, | 25 | 1963.07–1993.12 | 78.85 | 1.20 | 0.423 |
| six-month holding | | | (0.000) | (0.235) | |

These results show that conditional skewness can explain a significant part of the variation in returns even when factors based on size and book/market like SMB and HML are added to the asset pricing model. However, as we show later, conditional skewness is not successful in explaining all of the abnormal expected returns. Additionally, the impact of conditional skewness varies substantially by the econometric methodology used. Several reasons might explain these findings. The first is that we use conditional coskewness to explain the variation in next period's returns. However, our measurement of conditional coskewness is based on historical returns and, thus, is an imperfect proxy for true (ex ante) conditional coskewness. A second important reason is that the additional factors besides the *market*, namely SMB and HML, that we use in our asset pricing equation may capture the same economic risks that underlie conditional skewness. For example, book/market and size effects in asset returns may proxy for conditional skewness in asset returns. Partial evidence for this is found in our results for industry portfolios where adding conditional skewness alone or adding it along with SMB and HML produces very similar increases in $R^2$ from the single beta model.

### *B. Results of Cross-Sectional Regression Tests on Different Portfolio Sets*

We conduct tests on the first sets of portfolios in Table I using several econometric methods. These methods differ in how the betas vary through time as well as in how the standard errors are computed. In the traditional cross-sectional regression (CSR) approach pioneered by Fama and MacBeth (1973), a two-stage estimation is carried out period by period with betas estimated in the time-series and the risk premia estimated in the cross section. However, this approach ignores dependence across portfolios as well as the impact of heteroskedasticity and autocorrelation. These problems of CSR estimation are well known and have been analyzed in Shanken (1992) as well as more recently in Kim (1995), Kan and Zhang (1997), and Jagannathan and Wang (1996). These studies suggest that the Fama–MacBeth procedure, because the betas are assumed to be fixed over 60 months, does not capture the time-series variation in the betas. Therefore, we estimate risk premia for the various factors using the two-step Fama–MacBeth approach (CSR) as well as a full-information maximum likelihood (FIML) method that does not allow time-series variation in the betas. Indeed, the most important difference between the CSR and the FIML methods is that, in an FIML, we explicitly assume that the betas are constant over time.

The full-information maximum likelihood is a multivariate version of equation (13) and is similar to Shanken (1992). We assume that the residuals are distributed as $N(0,\mathbf{\Sigma})$ where $\mathbf{\Sigma}$ is an $N \times N$ heteroskedasticity and autocorrelation consistent variance and covariance matrix. This method permits the intercepts as well as the beta estimates to vary across the portfolios, though remaining constant in time. We maximize the likelihood function and use the beta estimates to run the cross-sectional regressions:

I Fama–French:

$$\hat{\mu}_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + e_i \tag{14a}$$

II Fama–French + $\hat{\beta}_{\mathrm{SKS}_i}$:

$$\hat{\mu}_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + \lambda_{\mathrm{S}}^{-}\hat{\beta}_{\mathrm{S}_i^-} + e_i, \tag{14b}$$

where $\hat{\mu}_i$ are $\sum_{t=1}^{T}(r_{i,t}/T_i)$, unconditional mean excess returns for every portfolio. This is a two-stage estimation procedure where we first estimate the mean excess returns as well as the betas using all the returns and then estimate the risk premia, from the mean excess returns and betas, permitting cross-sectional heteroskedasticity.

In constrast, the CSR method uses 60 time-series observations to estimate the betas and these betas are employed in cross-sectional regressions using the 61st period returns to estimate the risk premia, $\lambda$s. To alleviate the errors-in-variables (EIV) problem, following Shanken (1992), we compute the EIV-adjustments for the two-pass estimates for the the factor risk premia.

There are a number of interesting observations in Table III. First, the FIML with constant betas tends to present more explanatory power than the CSR with rolling betas. These results are consistent with Ghysels (1998). Second, the industry sort produces the lowest explanatory power, because as Berk (2000) emphasizes, this is the only sort that is *not* based on an attribute correlated with expected returns.

Contrary to the other tables, we contrast the Fama–French three-factor model with the CAPM. The three-factor model always does better than the one-factor CAPM. However, the addition of a skewness factor makes the single-factor model strikingly more competitive. For example, in the size and book-to-market value portfolios, the three-factor model produces an $R^2$ of 71.8 percent in the FIML estimation. The CAPM with the $S^-$ portfolio delivers a 68.1 percent $R^2$ (compared to only 11.4 percent in a one-factor model.) Similar results are found with the momentum portfolios. The CAPM with the $S^-$ skewness portfolio explains 61 percent of the cross-sectional variation in returns (compared to only 3.5 percent with the CAPM alone). The three-factor model explains 89.1 percent of the variation. When the $S^-$ portfolio is added to the three-factor model, the explanatory power increases to 95 percent.[12] The message of this analysis is that HML and SMB, to some extent, capture information similar to that captured by skewness.

Importantly, the addition of the skewness factor adds something over and above the three-factor model. Concentrating on the $S^-$ portfolio, the addition of skewness raises the explanatory power in all of the portfolio groups with the exception of the size portfolios. However, it should be noted that for the size portfolios, the explanatory power of the CAPM plus skewness is, to begin with, slightly higher than that of the three-factor model.

### C. Individual Equity Returns

Our last task is to analyze the individual securities in the FIML framework. We start with the 9,268 individual equities in the CRSP NYSE/AMEX and Nasdaq files. For the 9,268 stocks in CRSP files, the average betas to market, size, and book-to-market value as measured using the NYSE/AMEX value-weighted index, SMB, and HML portfolios are, respectively, 0.90, 1.16, and 0.47 over the period July 1963 to December 1993. The correlation between betas to SMB and HML and $\beta_{\mathrm{SKS}}$ are $-0.04$ and $-0.09$ respectively.

[12] We are concerned that both the size and momentum portfolios have too few cross-sectional observations (10). As a result, we calculate (but do not report) 25 size and 25 momentum portfolios. The results are similar.

**Table III**

**Results of Regressions for Portfolio Groups**

We form portfolios of U.S. NYSE/AMEX and Nasdaq equities over the period July 1963 to December 1993. We then estimate asset pricing models for each portfolio group using several factors as independent variables and report the adjusted $R^2$s for these models. We use two methods for estimation: (1) joint full-information maximum likelihood (with post-portfolio formation returns) assuming constant betas, and (2) month-by-month cross-sectional regressions after estimating the betas in rolling regressions using 60 months at a time. The equations we estimate in FIML are:

I 3 factors: $\hat{\mu}_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + e_i$

II 3 factors, $\hat{\beta}_{S_i^-}$: $\hat{\mu}_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + \lambda_{\mathrm{S}^-}\hat{\beta}_{S_i^-} + e_i$

III 3 factors + $\hat{\beta}_{SKS_i}$: $\hat{\mu}_i = \lambda_0 + \lambda_{\mathrm{M}}\hat{\beta}_i + \lambda_{\mathrm{SMB}}\hat{s}_i + \lambda_{\mathrm{HML}}\hat{h}_i + \lambda_{\mathrm{SKS}}\hat{\beta}_{SKS_i} + e_i$,

where $\hat{\beta}_{\mathrm{S}_i^-}$ and $\hat{\beta}_{SKS_i}$ are, respectively, the betas with respect to the excess return on the $S^-$ portfolio and the spread between returns on the $S^-$ and $S^+$ portfolios. In the CSRs the next period's excess return is used as the dependent variable. The $R^2$s reported are the adjusted $R^2$s for FIML and time-series average adjusted $R^2$s for CSRs. The CAPM, CAPM + $\mathrm{S}^-$, and CAPM + SKS regressions use the market beta and the respective coskewness betas as the regressors.

| | FIML (constant betas) | | | 2 pass CSR (rolling betas) | | |
|---|---|---|---|---|---|---|
| Sample | CAPM | CAPM, $\mathrm{S}^-$ | CAPM, SKS | CAPM | CAPM, $\mathrm{S}^-$ | CAPM, SKS |
| Industry | 1.53 | 13.2 | 9.2 | 9.6 | 17.9 | 18.2 |
| Size and book/market | 11.4 | 68.1 | 62.7 | 21.5 | 25.2 | 25.2 |
| 10 Momentum, six-month holding | 3.5 | 61.1 | 59.6 | 30.0 | 46.9 | 45.6 |
| 10 Size | 44.7 | 84.9 | 81.3 | 25.6 | 54.1 | 55.6 |
| 27 Size, B/M, momentum | −3.9 | 1.4 | 1.4 | 11.3 | 19.3 | 18.2 |

| | FIML (constant betas) | | | 2 pass CSR (rolling betas) | | |
|---|---|---|---|---|---|---|
| Sample | 3 Factor | 3 Factor, $\mathrm{S}^-$ | 3 Factor, SKS | 3 Factor | 3 Factor, $\mathrm{S}^-$ | 3 Factor, SKS |
| Industry | 25.3 | 30.1 | 28.5 | 18.3 | 28.1 | 29.5 |
| Size and book/market | 71.8 | 82.5 | 81.0 | 46.0 | 48.9 | 49.5 |
| 10 Momentum, six-month holding | 89.1 | 95.8 | 86.9 | 57.3 | 67.1 | 61.8 |
| 10 Size | 84.7 | 83.0 | 81.7 | 62.9 | 65.1 | 65.7 |
| 27 Size, B/M, momentum | 6.8 | 3.2 | 8.5 | 32.3 | 38.6 | 37.4 |

However, the averages and correlations vary substantially based on the length of return history. For example, for the 3,990 equities with fewer than 60 months of returns, the correlations between $\hat{\beta}_{\mathrm{SKS}}$ and betas to SMB and HML are 0.09 and $-0.29$. The average market, SMB, and HML betas for this sample are 0.81, 1.30, and 0.46. For the 5,278 stocks with greater than 90 months of returns the correlations between $\hat{\beta}_{\mathrm{SKS}}$ and betas to SMB and HML are 0.37 and 0.30. The average market, SMB, and HML betas for this sample are 0.95, 0.96, and 0.30.[13]

After estimating the betas in the time-series, we estimate the risk premia in cross-sectional regressions as in equation (16). In using individual equity returns, the large idiosyncratic variations are of concern and partly motivate portfolio formation. However, portfolio formation also causes information loss through the reduction in the number of cross-sectional observations. To control for these variations without losing observations, we weigh the securities using $1/\sigma(\hat{e}_i)$, where $\sigma(\hat{e}_i)$ is the standard deviation of the residuals from the beta estimation in equation (16). Such a weighting scheme allows us to use all the observations while controlling the idiosyncratic variations.[14]

We estimate equation (14) for all the stocks with both $\hat{\beta}_{\mathrm{SKS}}$ and $\hat{\beta}_{\mathrm{SKD}}$. The results are presented in Table IV. Premia estimates for both measures of coskewness are significant and the signs are generally as predicted. ($\beta_{\mathrm{SKD}}$ always has a negative premium and $\beta_{\mathrm{SKS}}$ has a positive premium for the whole sample.) Premia estimates for the other factors, market, SMB, and HML, are also significant.

Given that the correlations between the betas are different for stocks with different lengths of return histories, we permit the premia to vary by the length of return history available. We use all 9,268 stocks and estimate the models with four indicator variables that allow the slopes to differ for the following return histories: fewer than 24 months, 24 to 59 months, 60 to 89 months, and greater than or equal to 90 months. These results are also presented in Table IV.

The results show that the risk premium estimate for the market is positive for all return history lengths but the premia are inconsistent for SMB and HML. SMB (i.e., size) is significant only for stocks with fewer than 60 months of returns.

Thus, our results show that variation in size does not appear useful in explaining the variation in returns for stocks with more than 60 months of returns (i.e., 5,278 of the full sample of 9,268 stocks) though the length of return history and size are related. Additionally, premia for all four factors

[13] Empirical analyses of individual equities' returns are always suspect because of possible nonsynchronous returns. To allay such concerns, we also estimate the betas using the Scholes–Williams (1977) correction through instrumental estimation, using a moving average of previous, contemporaneous, and next-period excess return on the four-factor portfolios as instruments. Except for a decline in the magnitude of $\lambda_{\mathrm{M}}$, the results do not change substantially using heteroskedasticity and autocorrelation adjustments. These results are available from the authors.

[14] As a robustness check, we also consider alternative weighting schemes. The premia estimates change with the weights, however our inference about the significance of skewness does not.

**Table IV**

## Estimation of Risk Premia with Dependence on Return History Length for Individual Stocks

We estimate the risk premia for 9,268 stocks. We weight each of the stocks by $1/\sigma(\hat{e}_i)$, where $\sigma(\hat{e}_i)$ is the standard deviation of residuals from the beta estimation. The first line reports the estimated premium and the second line the WLS standard error in parentheses.

Panel A. Fama–French Model

| | No. of Stocks | $\lambda_{M}$ | $\lambda_{SMB}$ | $\lambda_{HML}$ |
|---|---|---|---|---|
| Full sample | 9268 | 0.290** | 0.013 | −0.060** |
| | | (0.023) | (0.012) | (0.012) |
| $T < 24$ | 1707 | 0.308** | 0.005 | −0.060** |
| | | (0.025) | (0.013) | (0.013) |
| $24 \leq T < 60$ | 2283 | 0.250** | 0.113** | −0.054 |
| | | (0.064) | (0.045) | (0.051) |
| $60 \leq T < 90$ | 1240 | 0.321** | −0.066 | −0.132 |
| | | (0.116) | (0.087) | (0.097) |
| $T \geq 90$ | 4038 | 0.088 | 0.072 | −0.269** |
| | | (0.077) | (0.062) | (0.088) |

Panel B. Fama–French Model with $\beta_{SKD}$, Directly Computed Coskewness

| | No. of Stocks | $\lambda_{M}$ | $\lambda_{SMB}$ | $\lambda_{HML}$ | $\lambda_{SKD}$ |
|---|---|---|---|---|---|
| Full sample | 9268 | 0.293** | 0.012 | −0.060** | −0.019** |
| | | (0.023) | (0.012) | (0.012) | (0.009) |
| $T < 24$ | 1707 | 0.308** | 0.006 | −0.059** | −0.022 |
| | | (0.025) | (0.013) | (0.013) | (0.019) |
| $24 \leq T < 60$ | 2283 | 0.255** | 0.114** | −0.055 | −0.011 |
| | | (0.064) | (0.045) | (0.051) | (0.019) |
| $60 \leq T < 90$ | 1240 | 0.334** | −0.067 | −0.137 | −0.018 |
| | | (0.117) | (0.087) | (0.097) | (0.023) |
| $T \geq 90$ | 4038 | 0.098 | 0.071 | −0.268** | −0.011 |
| | | (0.078) | (0.062) | (0.089) | (0.015) |

Panel C. Fama–French Model with $\beta_{SKS}$, $S^- - S^+$ Spread Portfolio

| | No. of Stocks | $\lambda_{M}$ | $\lambda_{SMB}$ | $\lambda_{HML}$ | $\lambda_{SKS} \times 100$ |
|---|---|---|---|---|---|
| Full Sample | 9268 | 0.278** | 0.022* | −0.051** | 0.058** |
| | | (0.023) | (0.012) | (0.012) | (0.011) |
| $T < 24$ | 1707 | 0.294** | 0.019* | −0.049** | 0.074** |
| | | (0.026) | (0.013) | (0.013) | (0.011) |
| $24 \leq T < 60$ | 2283 | 0.343** | 0.144** | −0.012 | −0.232** |
| | | (0.066) | (0.046) | (0.052) | (0.048) |
| $60 \leq T < 90$ | 1240 | 0.349** | −0.051 | −0.136 | −0.054 |
| | | (0.116) | (0.087) | (0.097) | (0.099) |
| $T \geq 90$ | 4038 | 0.122* | 0.075 | −0.271** | 0.027 |
| | | (0.077) | (0.062) | (0.088) | (0.173) |

** and * denote $t$-statistics significant at 5 percent and 10 percent levels, respectively.

decline as more return observations become available.[15] Book-to-market value differs substantially across return histories as well. One possible reason is that book-to-market proxies for differences in return history lengths. This is supported by the fact that the average, log, book-to-market ratio for stocks with more than 90 months of returns is 3.85 as compared to an average, log, book-to-market ratio of 6.39 for stocks with fewer than 60 months of returns. These findings suggest that the size effect is related to how long a stock is listed. The size factor appears important only for new firms. SMB and HML appear to be more important for firms where the number of returns available to estimate the market beta is small. This may also be an IPO effect, in other words, factors other than market (such as SMB or HML) may be more useful in predicting the returns on firms with a short return history.

The results show that substantial differences exist between the performance of asset pricing models in portfolios and individual equities. Therefore, to investigate the robustness of our results, we also examine a number of portfolios of the CRSP returns formed on a variety of criteria such as book-to-market value, return on assets, etc. These criteria are known to be correlated with expected returns and, hence, as Berk (2000) shows, are likely to reject the model in question. However, we find that although the premia estimates do vary substantially across the portfolio sorts, skewness is usually significant as a factor in explaining the cross-asset variation in returns.

### *D. Economic Significance*

Our results show that coskewness is usually statistically significant. An important question is whether it is economically significant. We undertake two additional evaluation exercises. The first focuses on the impact on the model's pricing errors and the second measures the expected return implied by a change in coskewness.

Consider the root mean squared in-sample pricing error (RMSE). Ghysels (1998) argues that this is an important metric for model evaluation.[16] We look at each model's RMSE relative to the CAPM's RMSE. This exercise is related to Table III. With the size and book-to-market value-sorted portfolios, the addition of the SMB and HML factors reduces the base RMSE by 48 percent. Adding the skewness spread, reduces the RMSE by another 10 percent. Similarly, for the 10 "momentum" portfolios, RMSE is reduced from the CAPM RMSE by 29 percent on adding SMB and HML. Inclusion of coskewness reduces the RMSE by a further 17 percent. When we examine the individual stocks, the general reductions in the RMSEs are smaller (because of the limited explanatory power for individual securities). As with the portfolio results, the introduction of the SMB and HML portfolios reduces the

[15] We also analyze the 4,603 stocks in COMPUSTAT over the January 1970 through December 1993 period alone. The results are similar to those for CRSP stocks and are available from the authors. There is also an issue of survivorship in the comparisons between CRSP and COMPUSTAT data, see Kothari et al. (1995) and Breen and Korajczyk (1993).

[16] Chen, Kan, and Zhang (1997) provide a critique of using $t$-ratios in model evaluation.

RMSE from the base case and there is a further reduction when the skewness spread (or the direct estimate of skewness) is added. However, in contrast to the portfolio results, introducing skewness alone as a second factor reduces the RMSE more than adding both HML and SMB as second and third factors. The RMSE analysis is suggestive that skewness plays an important role.

A second exercise is to assess the impact of ignoring skewness on expected returns. We calculate the cross-sectional distribution of the skewness spread betas for all individual firms. If we perturb the coskewness beta by one cross-sectional standard deviation, it translates into a 2.34 percent expected return on an annual basis. This assumes that all other factor loadings are held constant. The corresponding additional returns for the same perturbation for size and book-to-market value are 0.62 percent and 6.01 percent, respectively.

### *E. Momentum Strategies and Skewness*

Momentum or relative strength strategies have posed the great challenge for asset pricing models. Jegadeesh and Titman (1993) present evidence that the strategy of buying winners and selling losers where the losers and winners are defined on the basis of past returns can produce abnormal returns. Grinblatt, Titman, and Wermers (1995) and Carhart (1997) provide evidence that a large component of mutual fund abnormal returns can be generated by the momentum strategy of buying past winners and selling past losers. We examine the various momentum strategies with the objective of understanding how the abnormal returns from momentum strategies relate to skewness. The summary statistics on two widely used sets of momentum portfolios in Panels C and E of Table I show that, for a one-month holding period, winners have substantially lower skewness than losers.

Table V presents a representative subset of the momentum portfolios for the 1927 to 1997 period.[17] Some of these results are connected to Figure 3. For every momentum definition, the skewness of the loser portfolio is higher than that of the winner portfolio. For example in the 6,24 strategy, the loser portfolio has a mean excess return of −0.88 percent and a skewness of 0.36. The winner portfolio has a mean excess return of 7.06 percent and a skewness of −0.33. The higher mean strategy is associated with lower skewness.

We construct several other sets of portfolios using momentum defined for individual stocks over five different horizons: from 36 months to two months before portfolio formation, from 24 months to two months before, from 12 months to two months before, from six months to two months before, and from three months to two months before. We maintain a one-month gap between the portfolio formation and the computation of the return on the

[17] Full results for every momentum definition are available at http://www.duke.edu/charvey/Research/momentum.htm

**Table V**

**Summary Statistics on Momentum Strategies**

We form equally weighted portfolios of U.S. NYSE/AMEX and Nasdaq equities using several definitions of momentum that rely on past returns. The past returns are measured over 36 months to two months before portfolio formation. We maintain a one-month gap between portfolio formation and computation of return on the portfolio. For each $J$ month momentum we rank all the stocks in NYSE/Amex over the periods January 1927 to December 1997 into 10 equally weighted deciles and follow the post-portfolio formation returns over six different holding periods. These six holding periods are: one month, three months, six months, 12 months, 24 months, and 36 months. This table presents average excess return, volatility, skewness, and kurtosis for the loser portfolio (first decile) and winner portfolio (tenth decile) for a representative subset of these momentum strategies.

| Holding Period No. | Momentum | Decile Return | Average | Volatility | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| 1 | 6 | 1 | 2.47 | 34.43 | 2.60 | 19.62 |
| 1 | 6 | 10 | 20.26 | 29.26 | 1.05 | 8.49 |
| 3 | 6 | 1 | −5.09 | 20.91 | 1.50 | 12.67 |
| 3 | 6 | 10 | 6.25 | 18.57 | 0.50 | 3.84 |
| 1 | 12 | 1 | 2.47 | 35.81 | 2.45 | 18.54 |
| 1 | 12 | 10 | 23.64 | 27.83 | 0.67 | 6.74 |
| 3 | 12 | 1 | −6.38 | 21.45 | 1.30 | 10.55 |
| 3 | 12 | 10 | 10.40 | 18.00 | 0.71 | 6.25 |
| 6 | 12 | 1 | −6.06 | 14.28 | 0.07 | 1.93 |
| 6 | 12 | 10 | 6.48 | 12.94 | 0.13 | 2.37 |
| 1 | 24 | 1 | 12.55 | 39.85 | 3.25 | 25.36 |
| 1 | 24 | 10 | 19.81 | 24.84 | −0.22 | 3.24 |
| 3 | 24 | 1 | −0.05 | 23.01 | 1.72 | 12.34 |
| 3 | 24 | 10 | 10.04 | 15.91 | −0.24 | 2.47 |
| 6 | 24 | 1 | −0.88 | 14.91 | 0.36 | 2.49 |
| 6 | 24 | 10 | 7.06 | 11.58 | −0.33 | 1.47 |
| 12 | 24 | 1 | 0.86 | 10.21 | 0.14 | 2.17 |
| 12 | 24 | 10 | 4.11 | 8.67 | −0.63 | 0.72 |

portfolio. For each $J$-month momentum we rank all the stocks in the NYSE/AMEX over the periods January 1926 to December 1997 into 10 equally weighted and value-weighted deciles and follow the post-portfolio formation returns over six different holding periods. These six holding periods are: one month, three months, six months, 12 months, 24 months, and 36 months. Thus, we examine a total of 30 sets of 10-decile portfolios.

Figure 3 plots the trade-off between mean and skewness of returns for the losers and winners for several sets of momentum strategies over the period January 1927 to December 1997. The height of the line along the Y-axis is the mean annualized return from the strategy and the length of the line along the X-axis indicates the difference in skewness between the losers and

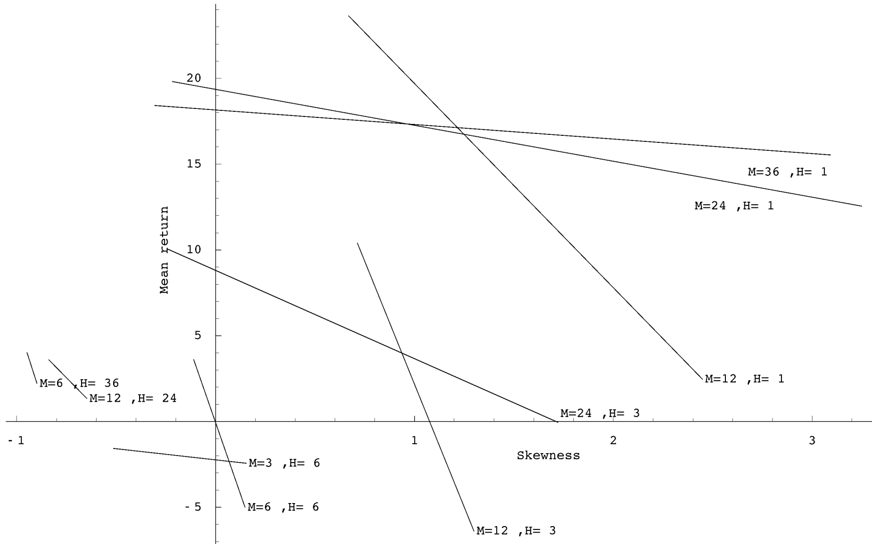


**Figure 3. Trade-off between mean and skewness of returns in momentum trading strategies.** This graph plots the trade-off in mean and skewness of annualized returns for several sets of momentum trading strategies using returns over the period January 1927 to December 1997. For each strategy M indicates the period over which the momentum is computed and H indicates the number of months over which the portfolio is held. The height of the line along the Y-axis indicates the mean annual return from the strategy and the length along the X-axis indicates the change in skewness.

winners as a result of the strategy. The negative slope of the line indicates that in a momentum-based trading strategy, buying the winner and selling the loser requires acceptance of substantial negative skewness.[18]

Next we examine how moments of the momentum-sorted deciles vary across the portfolios. We first test using all six holding periods across the five definitions of momentum jointly over the period January 1926 to December 1997. Using a multivariate $t$-test using the Bonferroni correction for dependence across deciles, we reject the hypothesis that the mean returns are equal across the deciles with a $p$-value of 0.001. Over this period, the average excess return for the loser decile is 1.15 percent per year in contrast to the average return in winner decile of 6.77 percent per year. When we exclude holding periods greater than 12 months, the respective average returns are 0.29 percent and 8.50 percent a year. Consistent with the results in Jegadeesh and Titman (1993), we find that the betas and volatilities of the momentum deciles are not significantly different from each other. ($p$-values are 1.000 for betas and 0.334 for volatilities.)[19] Similar to Jegadeesh and Titman, we also find that, for most of the deciles, the strategy of selling

[18] We have also examined a sample excluding the crash of 1929 and found similar results.

[19] Without a Bonferroni correction, the $p$-value is 0.083 for the equality for volatilities.

the loser and buying the winner has a negative beta. However, when we test if skewness is identical across the deciles, we reject with a $p$-value of 0.001. The average skewness for the losers is 0.494, whereas the average skewness for winners is $-0.380$. The results for kurtosis also reject the hypothesis that kurtosis is identical across the deciles.

We also examine various subsamples. For the period January 1951 to December 1997, we find that the average returns on the loser and winner deciles are 0.47 percent and 5.26 percent a year. As before, we fail to reject the hypothesis that volatilities are identical across the deciles. The average skewness for the loser decile is 0.031 and for the winner decile is $-0.673$, and the hypothesis that skewness is identical across the deciles is very strongly rejected. In contrast to the longer sample, for this subsample we fail to reject the hypothesis that average kurtosis across the deciles is identical.

The final exercise is to understand the fundamentals behind momentum. We select the second quarter of 1989 and 1993 and analyze the fundamental characteristics of stocks sorted on 12-month momentum in these periods. We compare these characteristics to those of stocks sorted on coskewness in the same periods. We choose 1989 and 1993 because they bracket the recession of 1990. We select only stocks that have remained in the same decile during the second quarter.

The results are presented in Table VI. For coskewness, our analysis suggests that characteristics of coskewed firms vary over time since 1989 and 1993 look different. In general, the most negatively coskewed and most positively coskewed firms are different from stocks in the middle. In both years, earnings growth for positively coskewed firms is substantially higher than for negatively coskewed firms. In 1993, positively coskewed firms are also larger in size than negatively coskewed firms and possess substantially higher P/E ratios. For momentum-sorted portfolios, both the "losers" and the "winners" are smaller than the middle deciles. Additionally, the "losers" have negative return on assets and much lower ratings on their debt.

Some caution should be used in interpreting these results. The regressions in Table IV indicate that coskewness is even more important for stocks with short histories and less important for stocks with longer histories. Nevertheless, these exercises suggest that skewness plays an important economic role.

### *F. Extensions*

The model and results presented in this paper have several interesting extensions. For both U.S. and world portfolios, the unconditional skewness is negative over the periods considered. Since the price of skewness should be negative, the implied risk premium from skewness should be positive. Hence, the implied market risk premium from variance and skewness should be higher than that from variance alone for a large number of the periods. This has the potential of helping to explain the equity market risk premium puzzle—the fact that unconditionally the equity market risk premium is

**Table VI**

**Fundamental Characteristics of Companies in Momentum and Coskewness Deciles**

We form value-weighted portfolios of U.S. NYSE/AMEX and Nasdaq equities in the second quarters of 1989 and 1993 based on coskewness and $t-12$ to $t-2$ momentum. We then select stocks that are in the same decile in all three months of the quarter. Next, we use COMPUSTAT to compute the mean fundamental characteristics of these stocks. All data except for growth are from the annual files. Return on assets is percentized. Leverage is defined as long-term debt over common equity (with the addition of balance sheet deferred taxes and investment tax credit to and subtraction of preferred stock value from common equity.) Book/market ratio is common equity over market capitalization. Rating is based on the 15-point scale used by Standard & Poor's with 10 representing BBB. Higher numbers mean lower ratings. Maturity is based on the S&P breakdown of debt into five categories representing how much debt is due in each year with all of the debt due in year five or later being combined. Sales growth, EPS (primary) growth, debt growth, and asset growth are percentized logarithmic growths from the first to the second quarter.

| Year | Decile | Return on Assets | Leverage | Book/ Market Ratio | Rating | Maturity | P/E Ratio | Size (in MM) | Sales Growth | EPS Growth | Debt Growth | Asset Growth |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A. Fundamental Characteristics of Coskewness-Sorted Deciles | | | | | | | | | | | | |
| 1989 | 1 | 2.77 | 2.38 | 0.75 | 10.19 | 2.98 | 19.18 | 1328.24 | 2.42 | −1.14 | 2.91 | 2.48 |
| | 5 | 3.44 | 0.58 | 0.74 | 10.62 | 2.79 | 14.30 | 1226.19 | 6.02 | 12.43 | 0.13 | 2.02 |
| | 6 | 4.06 | 0.91 | 0.82 | 9.64 | 2.83 | 25.69 | 1268.61 | 4.13 | 8.02 | 3.60 | 1.52 |
| | 10 | 4.27 | 1.04 | 0.78 | 10.91 | 2.93 | 17.88 | 1581.98 | 3.63 | 13.06 | 3.55 | 1.32 |
| 1993 | 1 | −2.16 | 0.68 | 0.81 | 13.13 | 2.64 | 16.99 | 345.73 | 3.56 | 4.51 | 1.59 | 4.41 |
| | 5 | 14.81 | 0.76 | 0.71 | 10.76 | 2.72 | 11.26 | 777.91 | 4.13 | 6.98 | 1.86 | 2.54 |
| | 6 | −6.61 | 0.74 | 0.59 | 11.07 | 2.56 | 1.54 | 896.91 | 8.02 | −3.08 | 0.17 | 5.08 |
| | 10 | −3.35 | 0.92 | 0.59 | 11.35 | 2.60 | 28.76 | 811.85 | 3.52 | 8.90 | −3.34 | 4.96 |
| Panel B. Fundamental Characteristics of $t-12$ to $t-2$ Momentum-Sorted Deciles | | | | | | | | | | | | |
| 1989 | 1 | −12.40 | 15.90 | 1.12 | 17.40 | 2.58 | 2.04 | 38.63 | 0.70 | 8.40 | 0.60 | −0.20 |
| | 5 | −13.30 | 0.55 | 0.76 | 11.00 | 2.74 | 5.13 | 287.05 | 11.30 | −5.90 | 5.80 | 2.10 |
| | 6 | 4.20 | 0.74 | 0.88 | 9.50 | 2.89 | 19.58 | 349.29 | 0.10 | −6.70 | 0.20 | 0.80 |
| | 10 | 9.30 | 0.73 | 0.34 | 13.60 | 2.44 | 17.13 | 865.20 | 14.50 | 5.20 | −8.80 | 7.30 |
| 1993 | 1 | −22.00 | 1.15 | 0.71 | 19.25 | 2.29 | 3.37 | 90.00 | 1.80 | 11.10 | 8.20 | 2.10 |
| | 5 | 0.80 | 0.57 | 0.78 | 11.44 | 2.68 | 21.08 | 2446.20 | −2.20 | 9.10 | −22.80 | 6.50 |
| | 6 | 6.00 | 0.70 | 0.79 | 10.00 | 2.69 | 20.69 | 999.74 | 2.00 | 32.40 | −8.50 | 2.30 |
| | 10 | 3.10 | 2.04 | 0.65 | 12.56 | 2.66 | 17.31 | 466.46 | 12.80 | 9.90 | −9.00 | 9.70 |

higher than what one should expect from variance alone. The model and results of this paper may also have implications for asset allocation and portfolio analysis. Instead of analyzing portfolios in a conditional mean-variance framework, a richer conditional mean-variance-skewness framework may be employed. The impact of skewness on option prices is also another possible extension of this paper. Finally, the asymmetric variance phenomenon (high conditional variance when returns are negative) found in many financial markets may also have an explanation based on conditional skewness.

## IV. Conclusions

Cross-sectional tests of the single factor asset pricing model have shown that systematic risk as measured by the covariance (or the beta) with the market and other factors does not satisfactorily explain the cross-sectional variation in expected excess returns. We provide a possible explanation for these failures. Our intuition is that if investors know that the asset returns have conditional coskewness at time $t$, expected returns should include a component attributable to conditional coskewness. Our asset pricing model formalizes this intuition by incorporating measures of conditional coskewness.

We estimate this model for several sets of equity returns, both individually as well as jointly, using portfolios formed using different criteria. Our results show that coskewness is important. In general, we find that a model incorporating coskewness is helpful in explaining the cross-sectional variation of equity returns. Coskewness also provides us with some insights as to why why variables such as size and book-to-market value are important in explaining the cross-sectional variation of asset returns. We also find that the momentum effect is related to systematic skewness. However, given the data limitations, measurement of ex ante skewness is difficult. Variables such as size and book-to-market value might capture information about ex ante skewness that we cannot measure from past returns alone.

## REFERENCES

Amihud, Yakov, Bent J. Christensen, and Haim Mendelson, 1993, Further evidence on the risk-return relationship, Working paper, New York University.

Arrow, Kenneth J., 1964, Aspects of the theory of risk-bearing; Lectures, Helsinki, Finland.

Bansal, Ravi, and S. Viswanathan, 1993, No arbitrage and arbitrage pricing, *Journal of Finance* 48, 1231–1262.

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

Berk, Jonathan B., 2000, Sorting out sorts, *Journal of Finance* 55, 407–427.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 44, 444–455.

Breen, William, and Robert A. Korajczyk, 1993, On selection biases in book-to-market based tests of asset pricing models, Working paper, Northwestern University.

Brennan, Michael, 1993, Agency and asset pricing, Unpublished manuscript, UCLA and London Business School.

Campbell, John, 1993, Intertemporal asset pricing without consumption data, *American Economic Review* 83, 487–511.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carhart, Mark M., Robert J. Krail, Robert I. Stevens, and Kelly E. Welch, 1994, Testing the conditional CAPM, Unpublished manuscript, University of Chicago.

Chan, Louis, Yasushi Hamao, and Josef Lakonishok, 1991, Fundamentals and stock returns in Japan, *Journal of Finance* 46, 1739–1764.

Chen, Nai-Fu, Raymond Kan, and Chu Zhang, 1997, A critique of the use of *t*-ratios in model selection, Working paper, University of California, Irvine.

Christie, Andrew A., 1982, The stochastic behavior of common stock variances: Value, leverage, and interest rate effects, *Journal of Financial Economics* 23, 407–432.

Cochrane, John, 1994, Discrete time empirical finance; Lecture notes, University of Chicago.

Fama, Eugene, and Kenneth French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 59, 23–49.

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

Fama, Eugene, and Kenneth French, 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* 50, 131–155.

Fama, Eugene, and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene, and Kenneth French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene, and James MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Ferson, Wayne E., and Campbell R. Harvey, 1998, Conditioning variables and the cross-section of stock returns, Working paper, Duke University.

Friend, Irwin, and Randolph Westerfield, 1980, Co-skewness and capital asset pricing, *Journal of Finance* 35, 897–914.

Ghysels, Eric, 1998, On stable factor structures in the pricing of risk: Do time-varying betas help or hurt?, *Journal of Finance* 53, 457–482.

Gibbons, Michael, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Golec, Joseph, and Maurry Tamarkin, 1998, Bettors love skewness, not risk, at the horse track, *Journal of Political Economy* 106, 205–225.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088–1105.

Handa, Puneet, S. P. Kothari, and Charles Wasley, 1989, The relation between the return interval and betas—Implications for the size effect, *Journal of Financial Economics* 23, 79–100.

Hansen, Lars, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars, and Scott Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.

Hansen, Lars, and Kenneth Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, *Journal of Political Economy* 91, 249–265.

Harrison, Michael, and David Kreps, 1979, Martingales and arbitrage in multi-period securities markets, *Journal of Economic Theory* 20, 381–408.

Harvey, Campbell, and Akhtar Siddique, 1999, Autoregressive conditional skewness, *Journal of Financial and Quantitative Analysis* 34, 465–487.

Ingersoll, Jonathan, Jr. 1990, *Theory of Financial Decision Making*, (Rowman and Littlefield, Totowa, New Jersey.)

Jagannathan, Ravi, and Z. Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 53, 3–53.

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

Kan, Raymond, and Chu Zhang, 1997, Tests of asset pricing models with useless factors, Working paper, University of Toronto.

Kandel, Shmuel, and Robert Stambaugh, 1995, Bayesian inference and portfolio efficiency, *Review of Financial Studies* 8, 1–53.

Kim, Dongcheol, 1995, The errors in the variables problem in the cross-section of expected stock returns, *Journal of Finance* 50, 1605–1634.

Kimball, Miles, 1990, Precautionary saving in the small and in the large, *Econometrica* 58, 53–73.

Kothari, S. P., Jay Shanken, and Richard Sloan, 1995, Another look at the cross-section of expected stock returns, *Journal of Finance* 50, 185–224.

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

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Leland, Hayne, 1997, Beyond mean-variance: Performance measurement of portfolios using options or dynamic strategies, Working paper, University of California, Berkeley.

Lim, Kian-Guan, 1989, A new test of the three-moment capital asset pricing model, *Journal of Financial and Quantitative Analysis* 24, 205–216.

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

Merton, Robert, 1982, On the mathematics and economics assumptions of continuous-time models, in William F. Sharpe and Cathryn M. Cootner, eds.: *Financial Economics: Essays in honor of Paul Cootner*, Prentice-Hall, Englewood Cliffs, NJ.

Nelson, Daniel, 1991, Conditional heteroskedasticity in asset returns, *Econometrica* 59, 347–370.

Nummelin, Kim, 1994, Expected asset returns and financial risks, Dissertation, Swedish School of Economics and Business Administration (Svenska Handelshögskolan), Helsinki, Finland.

Racine, Marie, 1995, Volatility shocks, conditional coskewness, conditional beta and asset pricing, Unpublished manuscript, Wilfrid Laurier University.

Roll, Richard, and Stephen Ross, 1994, On the cross-sectional relation between expected returns and betas, *Journal of Finance* 49, 101–121.

Roll, Richard, 1977, A critique of the asset pricing theory's tests, Part I: On past and potential testability of theory, *Journal of Financial Economics* 4, 129–176.

Ross, Stephen, 1977, The capital asset pricing model (CAPM), short-sale restrictions and related issues, *Journal of Finance* 32, 177–183.

Scholes, Myron, and Joseph Williams, 1977, Estimating betas from nonsynchronous data, *Journal of Financial Economics* 14, 327–348.

Scott, Robert, and Philip Horvath, 1980, On the direction of preference for moments of higher order than the variance, *Journal of Finance* 35, 915–919.

Sears, R. Stephen, and K-C. John Wei, 1985, Asset pricing, higher moments, and the market risk premium: A note, *Journal of Finance* 40, 1251–1253.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *Review of Financial Studies* 5, 1–33.

Sharpe, William, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Waldron, Paddy, 1990, Three-moment and three-fund results in portfolio theory, Working paper, University of Pennsylvania.