THE JOURNAL OF FINANCE • VOL. XLIX, NO. 1 • MARCH 1994

On the Cross-sectional Relation between Expected Returns and Betas

RICHARD ROLL and STEPHEN A. ROSS*

ABSTRACT

There is an *exact* linear relation between expected returns and true "betas" when the market portfolio is on the ex ante mean-variance efficient frontier, but empirical research has found little relation between sample mean returns and estimated betas. A possible explanation is that market portfolio proxies are mean-variance inefficient. We categorize proxies that produce particular relations between expected returns and true betas. For the special case of a zero relation, a market portfolio proxy must lie inside the efficient frontier, but it may be close to the frontier.

CONTRARY TO THE PREDICTIONS of the Sharpe, Lintner, and Black Capital Asset Pricing Model (hereafter the SLB CAPM or SLB Model; see Sharpe (1964), Lintner (1965), and Black (1972)), a decade of empirical studies has reported little evidence of a significant cross-sectional relation between average returns and betas. Yet it is well known (Roll (1977), Ross (1977)) that a positive and *exact* cross-sectional relation between ex ante expected returns and betas *must* hold if the market index against which betas are computed lies on the positively sloped segment of the mean-variance efficient frontier. Not finding a positive cross-sectional relation suggests that the index proxies used in empirical testing are not ex ante mean-variance efficient.

Some of the empirical studies have uncovered variables other than beta that have power in explaining the sample cross-sectional variation in mean returns. But the true cross-sectional expected return-beta relation is exact when the index is efficient, so *no* variable other than beta can explain any part of the true cross-section of expected returns. Conversely, if the index is not efficient, the ex ante cross-sectional relation does not hold exactly and other variables can have explanatory power. Indeed, any variable that happens to be cross-sectionally related to expected returns could have discernible empirical power when the index proxy is ex ante *in*efficient. Again, the empirical evidence supports an inference that market index proxies used in testing are not on the ex ante efficient frontier.

But the puzzle in the empirical work is not so much that the cross-sectional mean return-beta relation is imperfect nor that other variables have empiri-

* Roll is from the Anderson School of Management, University of California, Los Angeles, and Ross is from the Yale School of Management, Yale University. We are grateful for comments from T. Daniel Coggin, Mark Grinblatt, John E. Hunter, Chi-Cheng Hsia, Andrew Lo, Simon Wheatley, three referees, the coeditor of the *Journal*, David Mayers, and the editor, René Stulz.

cal power. This is to be expected given that direct tests reject mean-variance efficiency for many market index proxies.[1] Instead, the surprising thing is that the cross-sectional mean-beta relation appears to be virtually zero. Intuitively, it would seem that there should be *some* nonzero cross-sectional relation if the index is not too far inside the ex ante efficient frontier, even if it is statistically reliably inside. Why should we not anticipate at least a modest connection between expected returns and betas even on indices that are unmistakably inefficient?[2]

Yet the recent paper by Fama and French (1992) forcefully resurrects an old finding that there is virtually no detectable cross-sectional beta-mean return relation. They state, "...the relation between market β and average return is flat, even when β is the only explanatory variable" (Abstract). Earlier papers report the same result. For instance, Reinganum (1981), using two different indices, concludes, "...cross-sectional differences in portfolio betas estimated with common market indices are not reliably related to differences in average portfolio returns" (p. 460). Lakonishok and Shapiro (1986), after an extensive series of empirical tests, conclude, "...neither the traditional measure of risk (beta) nor the alternative measures (variance or residual standard deviation), can explain—again, at standard levels of significance—the cross-sectional variation in returns; only size appears to matter" (p. 131).[3]

Fama and French find no cross-sectional mean-beta relation after controlling for size and the ratio of book-to-market value, variables which *do* play statistically significant roles. Similar findings are reported by others, for a variety of different explanatory variables. For instance, Chen, Roll, and Ross (1986) conclude, "Although stock market indices 'explain' much of the intertemporal movements in other stock portfolios, their estimated exposures (their betas) do not explain cross-sectional differences in average returns after the betas of the economic state variables have been included" (p. 399).[4]

[1] Among the papers that reject efficiency for various market index proxies are Ross (1980), Gibbons (1982), Jobson and Korkie (1982), Shanken (1985), Kandel and Stambaugh (1987) and (1989, pp. 134, 135), Gibbons, Ross, and Shanken (1989), Zhou (1991), and MacKinlay and Richardson (1991).

[2] Note that the puzzle has no bearing on *market* efficiency. It is purely a mathematical and statistical problem. Whatever the distribution of returns, however well or poorly the market is operating, there exists an ex ante efficient frontier of portfolios. Any market index is located somewhere, either on the frontier or inside. The cross-sectional relation between expected return and beta, whether it is exact, imperfect, or zero, is completely determined by the position of the index.

[3] Coggin and Hunter (1985) find a negative relation between beta and mean return for large firms.

[4] Unlike Fama and French (1992), however, Chen, Roll, and Ross (1986) *do* find a nonzero cross-sectional mean return-beta relation in a univariate test. They use the value-weighted and the equally weighted New York Stock Exchange–listed indices. Similarly, Lakonishok and Shapiro find that "the coefficient of beta generally has the correct sign" (p. 131) across various subperiods, though it is not statistically significant.

The Fama and French paper made us wonder where an index would have to be located to produce a set of *true* betas that had no relation whatever to true expected returns. We soon discovered that such indices exist and that they lie within a set whose boundaries can be directly calculated from basic parameters (expected returns and covariances of returns). More generally, for *any* arbitrary cross-sectional linear slope coefficient between betas and expected returns, there is a bounded set of possible indices.

In Section I of this paper, we derive the analytic characterization of indices that produce an arbitrary cross-sectional relation between expected return and beta. Section II presents some "back-of-the-envelope" calculations of plausible locations for widely used market index proxies, i.e., how far inside the ex ante efficient frontier do such proxies lie? This section also discusses the implications of the empirical findings for the CAPM both as a scientific theory and as a practical tool for financial analysis. Sampling error, the other major possible explanation of the empirical findings, is analyzed briefly. Section III provides a summary and conclusion.

I. Indices That Produce a Given Ordinary Least Squares Slope Coefficient in the Cross-sectional Relation between Expected Return and Beta

To characterize market index proxies that produce particular cross-sectional mean-beta relations, we derive the boundary of the set of possible indices by finding members of the set with minimum return variance. This involves minimizing portfolio return variance subject to three constraints: (1) that the index portfolio's expected return is a given value, (2) that the index portfolio's investment proportions (weights) sum to unity, and (3) that a cross-sectional regression of expected returns on betas computed against the resulting index portfolio has a particular slope. Our derivation applies to any universe or subuniverse of assets provided that the index portfolio is composed only of stocks in the same group.

We employ the following notation:[5]

$\mathbf{R}$ = Expected returns vector for N individual assets in the universe,
$\mathbf{V}$ = $N \times N$ Covariance matrix of returns,
$\mathbf{1}$ = Unit vector,
$\mathbf{q}$ = Portfolio weights vector,
r = Scalar expected portfolio return, $\mathbf{q}'\mathbf{R}$,
σ^2 = Scalar portfolio return variance, $\mathbf{q}'\mathbf{V}\mathbf{q}$,
σ_j^2 = Cross-sectional or time series variance of j,
μ = Cross-sectional mean of expected returns, $\mathbf{R}'\mathbf{1}/N$,
$\boldsymbol{\pi}$ = Vector of scaled expected return deviations from the cross-sectional mean, $(\mathbf{R} - \mu\mathbf{1})/N$,
$\boldsymbol{\beta}$ = Beta vector, $\boldsymbol{\beta} \equiv \mathbf{V}\mathbf{q}/\mathbf{q}'\mathbf{V}\mathbf{q}$,

[5] Vectors and matrices are denoted in boldface.

k = The cross-sectional covariance of $\mathbf{R}$ and $\boldsymbol{\beta}$; i.e., the numerator of the ordinary least squares (OLS) slope from regressing individual expected returns on betas computed with an index-portfolio having weights $\mathbf{q}$.[6]

The appendix proves that any portfolio that is a solution to this problem must lie within a mean-variance region whose boundary is given by the equation

$$B\sigma^4 + Cr\sigma^2 + Dr^2 + F\sigma^2 + Gr + H = 0, \tag{1}$$

where the upper case constants are, $B = k^2(ac - b^2)$, $C = -2dkc$, $D = gc$, $F = 2dkb - g(ac - b^2) + cd^2$, $G = -2gb$, and $H = ag - d^2$, and where the lower case constants and parameters are as follows: three of these scalar elements, $a \equiv \mathbf{R}'\mathbf{V}^{-1}\mathbf{R}$, $b \equiv \mathbf{R}'\mathbf{V}^{-1}\mathbf{1}$, $c \equiv \mathbf{1}'\mathbf{V}^{-1}\mathbf{1}$, are the efficient frontier information constants (cf. Roll (1977), appendix). The two elements new in this paper are, $d \equiv \mathbf{R}'\mathbf{R}/N - \mu^2$, which is the *cross-sectional variance of expected returns*, ($d \equiv \sigma_R^2$), and $g \equiv \mu^2\sigma_{R-1}^2$, where σ_{R-1}^2 denotes the time series variance of the *difference* in returns between two portfolios, one weighted proportionately to the vector of expected returns and the second one equally weighted.

Equation (1) is the general form of a second-degree equation in r/σ^2 space. It is a parabola, a circle, an ellipse, or a hyperbola, depending on the value of $C^2 - 4BD$. The Appendix shows that $C^2 - 4BD$ is either zero (for $k = 0$) or negative. For $k \neq 0$, equation (1) is an *ellipse* in r/σ^2 space. The axes of the ellipse are oblique, i.e., not parallel to the r/σ^2 axes. In the special case $k = 0$ (a zero cross-sectional slope between expected returns and betas), equation (1) describes a parabola with an axis parallel to the σ^2 axis. Figures 1 and 2 illustrate these two cases, Figure 1 for $k = 0$ and Figure 2 for $k \neq 0$.

Portfolios that produce a zero cross-sectional slope, $\mathrm{Cov}(\mathbf{R}, \boldsymbol{\beta}) = k = 0$, lie within a parabola that is tangent to the efficient frontier at the global minimum variance point. It has long been known that the global minimum variance portfolio used as an index produces $\beta = 1$ for *every* asset, and, of course, $\mathrm{Cov}(\mathbf{R}, \mathbf{1}) = 0$. No other mean-variance efficient portfolio produces $k = 0$.

The minimum distance between the efficient frontier and a market index proxy with $\mathrm{Cov}(\mathbf{R}, \boldsymbol{\beta}) = 0$, measured along the return dimension at a given portfolio variance σ^2, is

$$\begin{aligned} M &\equiv r^* - r \\ &= \left\{[(c\sigma^2 - 1)(ac - b^2)]^{1/2} - [(c\sigma^2 - 1)(ac - b^2 - cd^2/g)]^{1/2}\right\}/c, \end{aligned} \tag{2}$$

where r is the expected return on the market proxy and r^* is the return on an efficient portfolio with the same variance as the proxy. In Figure 1, M is

[6] The parameter k is one measure of the relation between expected returns and β's. In the cross-sectional OLS regression, $\mathbf{R} = \gamma_0 + \gamma_1\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, (with $\boldsymbol{\varepsilon}$ the residual), the slope coefficient is $\gamma_1 = k/\sigma_\beta^2$, where σ_β^2 is the cross-sectional variance of $\boldsymbol{\beta}$.

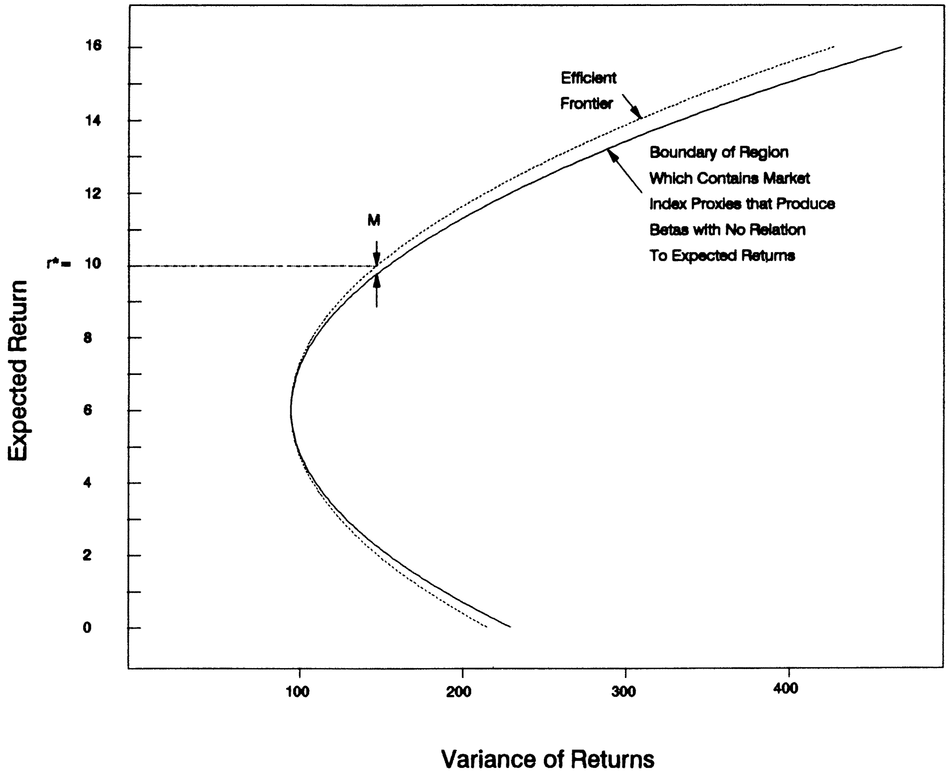

Figure 1. Market index proxies that produce betas having no relation to expected returns. These proxies are located within a restricted region of the mean-variance space, a region bounded by a parabola that lies inside the efficient frontier except for a tangency at the global minimum variance point. The distance, M, between the bounded region and the efficient frontier is proportional to the cross-sectional standard deviation of expected returns, σ_R. The M depicted is for $\sigma_R = 3\%$/annum and a market index proxy with expected return 9.78%/annum. The proxy is located on the boundary at a distance of $M = 22$ basis points below the efficient frontier. While betas against this market proxy have zero cross-sectional correlation with expected returns, a market proxy on the efficient frontier just 22 basis points above it would produce betas that are perfectly positively collinear with expected returns.

plotted for the case $\sigma_R = 3$, $\sigma_{R-1} = 5$, $\mu = 10$, and a proxy corresponding to an efficient portfolio with $r^* = 10\%$.

A useful and particularly tractable variant of (2) can be obtained by dividing both sides by $r^* - r_0$ where $r_0 = b/c$ is the expected return of the global minimum variance portfolio. The result is

$$M = (r^* - r_0)\left[1 - \left[1 - \frac{cd^2}{g(ac - b^2)}\right]^{1/2}\right], \tag{3}$$

i.e., the return distance of the proxy from the efficient frontier is a *constant* multiple (the term in large brackets) of the excess return $r^* - r_0$ of the efficient portfolio over the global minimum variance portfolio return, r_0. The

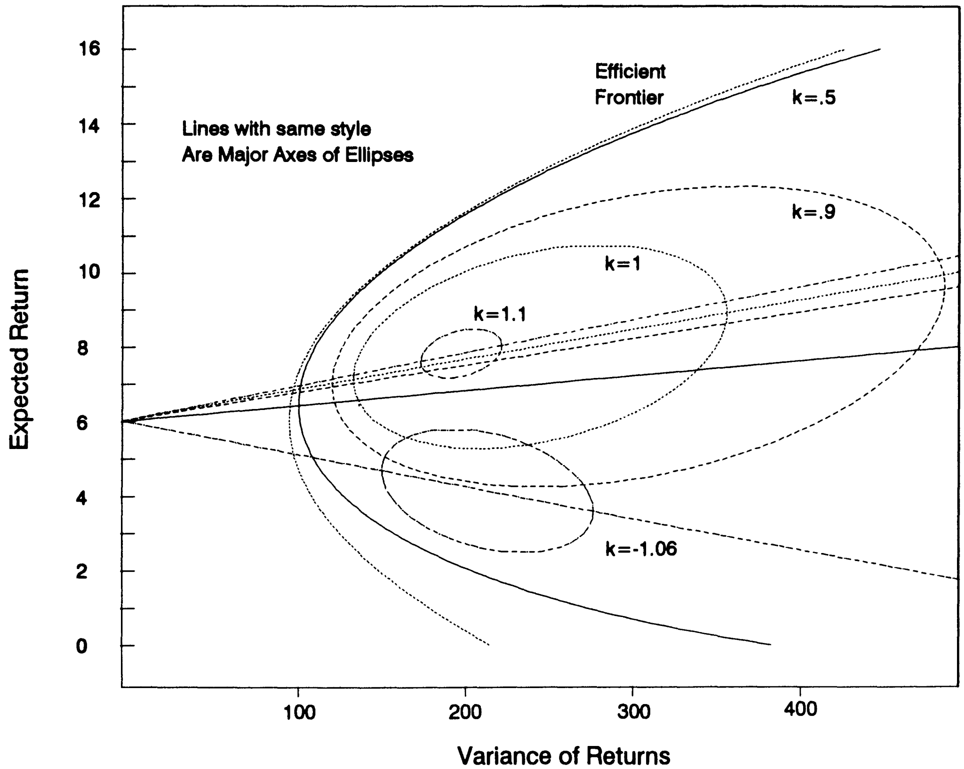

Figure 2. Market index proxies that produce betas having particular cross-sectional relations with expected returns. To produce a particular nonzero cross-sectional relation between betas and expected returns, a market index proxy must lie within a closed region of the mean-variance space. The regions are bounded by ellipses that may or may not have a tangency with the efficient frontier. If there is no tangency, then no mean-variance efficient market proxy can produce that particular relation. The major axes of the ellipses have positive (or negative) slopes when the resulting betas are positively (or negatively) related to expected returns. Ellipses are depicted for several values of k, the cross-sectional covariance between beta and expected return. The bounded region becomes smaller as this covariance increases. There is a maximum value of k beyond which the region vanishes; i.e., *no* market index proxy can produce a larger k.

term in large brackets in (3) is invariant with respect to the cross-sectional dispersion of expected returns.[7]

Index proxies that happen to lie within the sliver of space between the upper branch of the efficient frontier and the upper branch of the parabola,

[7] To see this, use the concept of a "mean-preserving spread" in the cross-sectional distribution of expected returns; i.e., $\mathbf{R} \equiv \sigma_R \mathbf{Z} + \mu \mathbf{1}$, where $\mathbf{Z}$ is a standardized vector of expected returns (mean zero and cross-sectional standard deviation of unity). Define standardized counterparts to the efficient set parameters (a and b) as $a^* \equiv \mathbf{Z}'\mathbf{V}^{-1}\mathbf{Z}$ and $b^* \equiv \mathbf{Z}'\mathbf{V}^{-1}\mathbf{1}$. It is straightforward to show that $ac - b^2 = \sigma_R^2(a^*c - b^{*2})$, $d = (\sigma_R^2/N)\mathbf{Z}'\mathbf{Z}$, and $g = (\sigma_R^2/N^2)\mathbf{Z}'\mathbf{V}\mathbf{Z}$. Thus, the expression in (3), $cd^2/[g(ac - b^2)] = c(\mathbf{Z}'\mathbf{Z})^2/[\mathbf{Z}'\mathbf{V}\mathbf{Z}(a^*c - b^{*2})]$, which is independent of σ_R. A similar development shows that $M = r^* - r$ in (2) is proportional to σ_R; thus, the standardized difference, $(r^* - r)/\sigma_R$, between the efficient frontier and the inner $k = 0$ parabola is invariant with respect to the cross-sectional dispersion of expected returns.

produce positive cross-sectional slopes. To prove this, note that if some index within the upper sliver had a negative slope, then by choosing appropriate weights the index could be combined with the corresponding efficient portfolio having the same mean such that the resulting combination had a zero slope. But, such a combined portfolio must lie under the $k = 0$ parabola of minimum variance portfolios with zero cross-sectional slopes, a contradiction.

The situation of $k \neq 0$ is more complex. The Appendix shows that the set of indices producing $\mathrm{Cov}(\mathbf{R}, \boldsymbol{\beta}) = k$, is bounded by an ellipse which may or may not be tangent to the efficient frontier. For any k greater in absolute value than formula (A9) in the Appendix, there is no tangency between the efficient frontier and the ellipse bounding the set of all index proxies that produce a cross-sectional covariance of k.

In Figure 2, ellipses have been plotted for several choices of the cross-sectional covariance k. The major axes of the ellipses have slopes in r/σ^2 space with the same sign as their associated k and they all intersect the return axis at r_0, the expected return of the global minimum variance portfolio. Notice that as k becomes larger, the ellipse becomes more concentrated about its center (which, incidentally, lies at the point $\sigma^2 = \frac{1}{2}g/k^2$, $r = r_0 + \frac{1}{2}d/k$). The collapse becomes complete at $k = \pm\frac{1}{2}\sqrt{cg}$. For larger absolute values of k, the ellipse becomes imaginary; i.e., there are *no* market index proxies that produce a larger cross-sectional covariance between $\mathbf{R}$ and $\boldsymbol{\beta}$.

Our results are reminiscent of those in two papers by Kandel and Stambaugh (1987, 1989) and in a paper by Shanken (1987). In their first paper, Kandel and Stambaugh derive the correlation between an arbitrary portfolio and a portfolio on the efficient frontier. They prove that this correlation is maximized when the two portfolios have the same expected return and they use this result to derive tests for the efficiency of an unknown market proxy that has a given correlation with the observed proxy. The idea is that an observed proxy may not be the true market index whose mean-variance efficiency is required by CAPM theory, but if one is willing to assume that the unobservable true market index has a given level of correlation with the observable proxy, an unambiguous test of the CAPM can still be conducted (conditional on the assumed correlation).[8]

A section of their paper deduces the boundary of the set of all portfolios that possess a particular minimum correlation with any given index. These sets may be closed. As the minimum correlation approaches 1.0, the set collapses to the single point coincidental with the index. At low correlations, however, the sets may be unbounded. For instance, when the index is inefficient, zero-beta portfolios (portfolios possessing zero correlation with the index) exist at all levels of expected return, a result derived by Roll (1980). Kandel and Stambaugh show that intermediate correlations can produce

[8] Using a similar approach, Shanken (1987) presents evidence that the SLB Model is invalid unless each of the several market proxies he employs is only weakly correlated (multiple correlation less than 0.7) with the true market portfolio.

bounded but open sets, e.g., with a minimum or maximum expected return but no limit on variance.

These Kandel-Stambaugh (1987) sets contain portfolios with a given minimum correlation to the original index proxy, whereas the sets we derive here contain index proxies that produce a given cross-sectional relation between expected return and beta. Thus, they are formally distinct, but they do possess some common properties. Perhaps the most important to emphasize is that the sets (the regions are graphed in our Figures 1 and 2 and in Kandel and Stambaugh's Figure 1) are not *exclusive*. There are other portfolios lying within these regions which do *not* produce the same result. Within the Kandel-Stambaugh regions are portfolios with *higher* correlations to the index proxy than the specified minimum correlation. Within our regions are portfolios that produce other values of the cross-sectional mean-beta relation. For both types of regions, no portfolio lying outside can produce the given relation, but an infinite number of portfolios *inside* can produce some other relation.

Figure 2 provides an intuitive depiction of nonexclusivity. Notice that some ellipses plotted there fall entirely within others. Thus, within the $k = 1$ ellipse, $[k = \mathrm{Cov}(\mathbf{R},\boldsymbol{\beta})]$, are market proxies producing $k = 0.9$, $k = 0.5$, etc., although there are no market proxies producing $k = 1.1$ or $k = -1.06$ unless they lie also within their respective ellipses.

In contrast, the later paper by Kandel and Stambaugh (1989) derives exclusive regions of mean-variance space, but for a different purpose. Kandel and Stambaugh (1989) develop likelihood ratio tests for the ex ante mean-variance efficiency of a given index proxy. They show that the rejection region (or a given significance level) is bounded by a "critical hyperbola" in *sample* mean-variance space. Portfolios that lie away from the sample efficient frontier beyond this critical hyperbola should be judged inefficient. One only needs to plot the position of proxy being tested in order to conduct the test.

It is instructive to understand intuitively why a statistical test for proxy efficiency might lead to an exclusive rejection region while correlation sets and mean-beta relation sets would not be exclusive. In the first case, the further a proxy lies below the sample efficient frontier, the less likely it lies on the *true* ex ante frontier, provided that one is willing to assume stationarity of the expected return vector and the covariance matrix. However, there is only an indirect connection between the position of the proxy in mean-variance space and either its correlation with other portfolios or its cross-sectional mean-beta relation. For example, take correlation: if the covariance matrix is nonsingular and the number of assets is finite, there is no other portfolio *perfectly* positively correlated with the index proxy. Thus, a correlation of 1.0 implies a single position in mean-variance space. But if the index proxy is inefficient enough, there are other distinct portfolios *with the same mean and variance* having *zero* correlation with the proxy! Thus, two uncorrelated portfolios can lie at exactly the same point in mean-variance space. Clearly, there are an infinite number of portfolios, all lying at exactly the same mean-variance position, yet possessing an infinite number of different correlations with the index proxy.

The nonexclusivity of our sets makes it impossible to determine the cross-sectional mean-beta relation simply by plotting the position of the proxy in the mean-variance space. We wish this were possible. It is not. We know only that particular cross-sectional mean-beta relations *cannot* be produced by index proxies that lie outside the boundaries of the sets we derive here. Each set places an upper or a lower bound on the cross-sectional covariance between $\mathbf{R}$ and $\boldsymbol{\beta}$.

II. The Cross-sectional Return-Beta Relation and Tests of the CAPM

A. The Plausibility of Test Sensitivity to the Choice of a Market Proxy

The SLB Model implies mean-variance efficiency of the market index; this efficiency is equivalent to a perfect cross-sectional relation between expected returns and betas computed against the market index. But, when the market index is proxied by an *in*efficient portfolio, these two representations of the same theory are no longer strongly related. We have shown that the cross-sectional slope can have any absolute value below a certain maximum (including zero) depending on the index proxy's position inside the ex ante mean-variance efficient frontier. This implies that an index proxy can conceivably be substantially inefficient and still produce a strong cross-sectional regression between expected returns and betas or it can conceivably be close to the efficient frontier and yet produce a zero cross-sectional relation. What actually *is* produced in the empirical cross-sectional regression depends on the ensemble of expected returns, variances, and covariances.

This suggests that the slope of the cross-sectional return-beta relation may be of little direct use in assessing the distance of the index proxy from the ex ante efficient frontier and, therefore, it may not be useful for determining how inefficient is the true market index. An inefficient proxy with a zero cross-sectional slope may be quite close to the true market portfolio and the true market portfolio may be efficient.

The plausibility of such possibilities can be examined with back-of-the-envelope calculations using reasonable guesses of parameter values. For instance, given current levels of inflation, it seems reasonable to assume an expected return on the global minimum variance portfolio of 6 percent (per annum) and a minimum standard deviation of 10 percent; $r_0 = 6\%$, and $\sigma_0 = 10\%$. Similarly, an expected return of, say, 11 percent, seems reasonable for the efficient portfolio located where a ray from the origin through the global minimum variance position intercepts the efficient frontier, $r_1 = 11\%$. These values are sufficient to determine the equation of the efficient frontier. We also need to guess the values of three other parameters: μ, the average expected return on risky assets; σ_R, the cross-sectional dispersion of expected returns; and σ_{R-1}, the time series standard deviation of the difference between an expected return-weighted portfolio and an equally weighted portfolio. Reasonable values might be: $\mu = 10\%$, $\sigma_R = 3\%$, and $\sigma_{R-1} = 5\%$.

Notice that the last value is relatively small, but this is appropriate given that two well-diversified portfolios are likely to be significantly correlated.[9]

Using these parameter values in equation (3) gives $M = 0.055542(r^* - r_0)$ as the expected return distance of a market index proxy from the efficient frontier. If we happened to select a proxy whose corresponding mean-variance efficient portfolio with equal variance had the same mean as the global average mean, $r^* = 10\%$ and since $r_0 = 6\%$, $M = 0.2222\%$. Thus, given these parameter values, the mean return of an index proxy that produces a cross-sectional mean-beta relation of *zero* could lie only about 22 basis points below the efficient frontier; its expected return would be 9.78 percent while the efficient portfolio with the same variance would have an expected return of 10 percent. These positions are plotted in Figure 1; see the arrows below "M." Thus, the index proxy could produce a zero cross-sectional mean-beta slope while the corresponding efficient portfolio, if used as a proxy, would produce a *perfect* cross-sectional relation with a positive slope.

The presence of sampling error only strengths the caution with which we must approach cross-sectional empirical tests. If expected returns and betas could be measured with little or no error, then we could reject index mean-variance efficiency by finding a flat cross-sectional relation. But, with measurement error we can only say that we cannot reject a flat relation. For that matter, we probably also cannot reject that the slope is, say, 3 percent. With 60 years of observations on an index with an annual standard deviation of 20 percent, the standard error of the *sample* mean would be $20\%/\sqrt{60} = 2.6\%$.

With a standard error of, say, 3 percent in the measurement of index expected returns, the power of cross-sectional tests is suspect. If the true market portfolio is, in fact, efficient, index proxies that produce a flat *sample* cross-sectional relation may be positioned well within a 3 percent interval of the ex post efficient frontier. Thus, the probability of not rejecting a flat slope when the slope is actually not flat, may be quite high.[10]

It is perplexing, then, that some authors relate the absence of a detectable cross-sectional slope for a particular market index proxy to a general condemnation of the SLB CAPM model. Fama and French (1992) include a section entitled "Can the SLB Model be Saved?" (p. 459), where they state, "We are forced to conclude that the SLB model does not describe the last 50 years of average stock returns" (p. 464). We would add, "for this particular market index proxy."

[9] The assumed value of σ^2_{R-1} is one-half the standard deviation of the global minimum variance portfolio; larger values of σ^2_{R-1} would cause the index proxy to lie closer to the efficient frontier.

[10] Cross-sectional mean-beta tests are different from direct tests of the mean-variance efficiency of a given index (cf. Gibbons, Ross, and Shanken (1989)). The null hypothesis of cross-sectional tests is that the theory is *not* true. In contrast, the null hypothesis of direct tests is that the index is efficient. The power of cross-sectional tests is the probability of accepting a cross-sectional relation when there really is one. The power of direct tests is the probability of rejecting index efficiency when the index really is not on the efficient frontier. Thus, these two index efficiency tests have the null and alternative hypotheses reversed.

An alternative interpretation of their results is that the SLB Model may be of little use in explaining cross-sectional returns no matter how close the index is to the efficient frontier unless it is *exactly* on the frontier. Since such exactitude can never be verified empirically, we would endorse (again, as we have in the past when we first asserted the proposition; see, e.g., Roll (1977), and Chen, Roll, and Ross (1986)), that the SLB is of little practical use in explaining stock returns.

In a different section of their paper, Fama and French argue that

> different approaches to the tests are not likely to revive the Sharpe-Lintner-Black model. Resuscitation of the SLB model requires that a better proxy for the market portfolio (a) overturns our evidence that the simple relation between β and average stock returns is flat and (b) leaves β as the only variable relevant for explaining average returns. Such results seem unlikely, given Stambaugh's (1982) evidence that tests of the SLB model do not seem to be sensitive to the choice of a market proxy. Thus, if there is a role for β in average returns, it is likely to be found in a multi-factor model that transforms the flat simple relation between average return and β into a positively sloped conditional relation (p. 449).

This essentially alleges that *no* reasonable market proxy can produce a nonzero cross-sectional expected return/beta relation in which beta is the sole relevant explanatory variable.

But, viewed in the context of our analysis, such a statement seems at least questionable. It appears that a proxy can be quite close to the ex ante frontier and still produce a cross-sectional beta-return relation with a slope near zero, and a proxy that is far from the frontier can still have a significant cross-sectional relation. In particular, another proxy can be close to the ones used now and have a positive cross-sectional relation or a zero one. An empirical slope near zero tells us little, if anything, about whether the SLB Model describes "average stock returns," but it does tell us something about the market index proxies we are using. As for whether an inefficient proxy can be found with betas that alone explain average returns, there is no a priori reason to reject such a possibility.[11]

B. Plausibility and Short Positions

Several readers of a previous version of this paper speculated that the central results may be driven by short positions in market index proxies that produce a particular mean-beta cross-sectional slope. Indices with short positions have not been used in the empirical tests. Yet the indices we

[11] This can be true notwithstanding the observation that size, for example, appears to be a significant explanatory variable in cross-sectional studies. Given the hundreds of parameters that have been used in such studies, it would be astonishing if the best performing of them were not significant by chance alone.

characterize in Section I have no restrictions against short positions and thus may not be empirically relevant.

We have not yet been able to assess this objection in a completely general context, but a limited assessment is possible given a few more assumptions about the process generating asset returns. The objection is valid for some relatively simple asset return structures including the example represented by a limited version of the single-factor arbitrage pricing theory (APT) model. If there is just one priced APT factor *and every asset has positive sensitivity to that factor*, any well-diversified market index proxy without short positions will produce a positive cross-sectional expected return–CAPM beta relation if the market risk premium is positive.[12]

However, this simple example fails to generalize into a more complicated world. For instance, there need be no necessary relation between expected return and beta, even when there is only a single generating factor, when the APT is not true. Suppose there is cross-sectional variability in expected returns that is unrelated to the asset's factor sensitivity. Although this would admit the potential for arbitrage cash flows (with virtually no risk and no investment),[13] it permits any variety of cross-sectional relation between expected return and CAPM beta even when the market index proxy has nonnegative weights on all assets.

In the absence of arbitrage opportunities but with a multiple factor asset return structure, totally positive well-diversified market index proxies may produce an insignificant cross-sectional mean-beta relation. A simple numerical example is provided by the two-factor hypothetical economy described in Table I. In this economy, the APT holds exactly but some positively weighted portfolios produce betas that have no cross-sectional relation to expected returns; even an *equally* weighted market index proxy produces a slightly negative but statistically insignificant cross-sectional slope.[14] The hypothetical economy in Table I represents a counterexample to the objection that our results are driven by short positions. There are, of course, other possible asset

[12] In a single-factor APT model, every asset j has returns in time t given by $\rho_{jt} = r_j + b_j\delta_t + \varepsilon_{jt}$ where r_j is the asset's expected return, δ_t is the mean-zero single factor, b_j (> 0 by assumption) is the asset's factor sensitivity, and ε_{jt} is an idiosyncratic white noise disturbance. If the APT holds perfectly, there exist constants γ_0 and γ_1 such that $r_j = \gamma_0 + \gamma_1 b_j$. A well-diversified market proxy M is simply a portfolio with negligible idiosyncratic disturbance, i.e., $\rho_{Mt} \approx r_M + b_M\delta_t$. If M has nonnegative investment proportions in all individual assets, then since $b_j > 0\ \forall j$, $b_M > 0$. In this situation, the CAPM beta is approximately $\beta_j = b_j/b_M$. Thus, the cross-sectional slope coefficient between individual asset expected returns and CAPM betas is $\mathrm{Cov}(r_j, \beta_j)/\mathrm{Var}(\beta_j) = \gamma_1 b_M$, which is positive if the market price of risk, γ_1, is positive.

[13] Pure arbitrage cash flows, *zero* risk and no investment, would technically be feasible only with an infinite number of assets.

[14] Note that an equally weighted index is *not* likely to be on the boundary of one of our sets. The equally weighted index is 200 basis points below the frontier but there are positively weighted proxies closer to the efficient frontier that produce roughly the same cross-sectional mean return-beta relation.

Table I

A Simulated Two-Factor APT Economy

Number of Assets: 25

Every asset j has return ρ_{jt} in time t generated by a two-factor model,

$$\rho_{jt} = r_j + b_{1j}\delta_{1t} + b_{2j}\delta_{2t} + \varepsilon_{jt},$$

where r_j is j's expected return, the b's are factor sensitivities, the δ's are mean-zero factors, and ε is a disturbance independently distributed across assets and over time.

The APT holds exactly: $r_j = \gamma_0 + \gamma_1 b_{1j} + \gamma_2 b_{2j}$.

In the simulated economy, $\gamma_0 = 5\%$ and $\gamma_1 = \gamma_2 = 8\%$ per period. Each of the two factors is independently and normally distributed over time with a standard deviation of 13% per period. The 25 values of b_1 are randomly selected from a normal distribution with mean 1.0 and standard deviation 0.4. Twenty-three of the 25 values of b_2 are zero, but $b_{21} = -b_{22} = 3.34$. Finally, each asset's generating equation is fully specified by selecting a random R-square from a uniform distribution between 0.15 and 0.30 (this conforms roughly to actual stock returns).

Once the R-square is selected, the asset's total return variance is readily calculated from the generating equation. It is also possible to calculate the exact composition of the Markowitz efficient frontier, to determine the mean-variance position of any potential market index proxy, and to calculate true values of each asset's CAPM betas. Here are the resulting calculations when the market index proxy is an *equally* weighted portfolio.

	True Parameters	
	Mean Return (%)	Std. Dev. (%)
Equally weighted portfolio	13.0	14.8
Efficient portfolio, same mean	13.0	12.0
Efficient portfolio, same standard deviation	15.0	14.8
Global minimum variance portfolio	8.74	8.77

The cross-sectional OLS regression of true expected returns on CAPM betas computed with the equally weighted portfolio as a market index proxy is (t-statistics in parentheses):

$$\begin{aligned} r_j &= 13.1 - 0.215\beta_j \\ &\,(4.12)\quad(-0.0761) \end{aligned}$$

The adjusted R-square of the cross-sectional regression is -0.0432.

structures that would bring about our results, but one counterexample is sufficient to dispel the notion that the objection is valid in general.

C. The Potential Sensitivity of CAPM Tests to the Econometric Method

Although the superiority of generalized least squares (GLS) to OLS is well-recognized by finance empiricists, our results above depend on the cross-sectional regressions being OLS. Most of the existing literature relies on this technique. There are, however, some exceptions. A recent paper by Amihud, Christensen, and Mendelson (1992), for instance, replicates the Fama and French tests while employing the more advanced econometric techniques of GLS and pooled time series–cross-section analysis. Although Amihud *et al.* find the same results as Fama and French using OLS, their results are reversed when using either pooled time series–cross-section methods or when using GLS; the estimated impact of beta on expected return is particularly strong when both methods are employed.[15] They conclude that "beta is still alive and well" (p. 1).

One might be tempted to conclude that more powerful econometric techniques and better estimation reveal that the market index proxy is not too far from the efficient frontier after all. But our analysis above is based on true expected returns, variances, and covariances; estimation problems are assumed away. We show above that the OLS definition of the cross-sectional mean-beta coefficient can be *truly* zero if the market index is sufficiently mean-variance inefficient. This result does not depend on statistical misestimation of any relevant parameter, but it does assume that the cross-sectional mean-beta regression coefficient is calculated with the OLS formula.

Thanks to a private communication from Simon Wheatley in 1992, we learned that using a GLS calculation rather than OLS can have a significant effect on the resulting true cross-sectional coefficient. GLS produces a *positive* cross-sectional relation between true expected returns and true betas regardless of the inefficiency of the market index proxy so long as its expected return exceeds the expected return, r_0, of the global minimum variance portfolio![16]

Intuitively, the GLS method diagonalizes the covariance matrix of regression residuals. It is equivalent to using OLS when the covariance matrix of returns, $\mathbf{V}$, is proportional to the identity matrix. But if $\mathbf{V}$ is proportional to the identity matrix $\mathbf{I}$, $\boldsymbol{\beta} = \mathbf{Iq}/\mathbf{q}'\mathbf{q}$. Thus, to obtain a portfolio with expected return $r = \mathbf{R}'\mathbf{q}$ and with $\mathrm{Cov}(\mathbf{R}, \boldsymbol{\beta}) = 0$, we must have $\mathbf{R}'\boldsymbol{\beta} - \mu\mathbf{1}'\boldsymbol{\beta} = 0$, which implies $r = \mu$. There is no solution to the problem $k = 0$ unless the portfolio's expected return, r, is the cross-sectional mean of the expected returns of all assets, μ. And when $\mathbf{V} \propto \mathbf{I}$, the expected return, r_0, of the global minimum variance portfolio is also the cross-sectional mean expected return, μ.

The use of GLS is likely to overturn the Fama and French empirical result of a zero cross-sectional slope. Unless the index proxy is grossly inefficient, with expected return less than or equal to r_0, a GLS regression would almost certainly find a significant and positive mean-beta relation in large samples. But what would this really imply about the validity of the CAPM, about whether the true market portfolio of all assets is ex ante mean-variance efficient? If the mean return-beta relation is positive for *every possible* market proxy whose mean return exceeds r_0, what conceivable set of empirical results would cause us to *reject* the CAPM?

Kandel and Stambaugh (1993) derive a goodness-of-fit statistic, expressed as an R-square, for the true cross-sectional GLS relation between expected returns and betas. They show that R-square decreases (or increases) as the index proxy lies farther (or closer) to the efficient frontier. Thus, if the true

[15] However, the strength of beta as an explanatory factor is much greater in the 1953 to 1971 sample period than in the 1972 to 1990 sample period. In the later period, beta is *not* significant.

[16] A formal proof is in the GLS section of the Appendix. Kandel and Stambaugh (1993) derive and elaborate the same result.

parameters were known, the Kandel and Stambaugh R-square is a metric of the index proxy's degree of inefficiency. The problem is that the true parameters are *not* known; thus, any observed empirical GLS R-square consists both of sampling error and (possibly) true ex ante scatter. It is not immediately clear how an empirical investigator can tell the difference. Perhaps it will prove best to employ a direct test of the index proxy's efficiency, such as the Kandel and Stambaugh (1989) likelihood ratio test which depends only on the proxy's location relative to the *sample* efficient frontier.

We don't want to leave the impression that the Wheatley–Kandel and Stambaugh result fully explains the differences between the findings of Fama and French and of Amihud, Christensen, and Mendelson. The GLS proof assumes knowledge of all true parameters in the spirit of this paper. The empirical researchers have only estimates. Also, the GLS method used by Amihud *et al.* is somewhat different than that assumed by Wheatley and Kandel and Stambaugh. Nonetheless, we think it is appropriate to bring attention to the bizarre idea that the very range of *possible* findings can be affected by the econometric technique. Shanken (1992) provides a thoughtful analysis of the different inferences that might be obtained with various econometric techniques. He investigates not only OLS versus GLS but also the impact of errors in the variables on familiar two-pass tests of beta pricing models. In the context of factor models, he also shows that autocorrelation in the underlying factors can lead to problems of inference.

III. Summary and Conclusion

The empirical absence of a detectable relation between average returns and betas is an indictment of the SLB Model, at least for use with the most widely employed market index proxies. If the SLB Model cannot tell us about average returns, then it is not of practical value for a variety of applications including the computation of the cost of capital and the construction of investment portfolios.

As we have seen, though, the empirical findings are not by themselves sufficient cause for rejection of the theory. The cross-sectional OLS relation is very sensitive to the choice of an index and indices can be quite close to each other and to the mean-variance frontier and yet still produce significantly different cross-sectional slopes, positive, negative, or zero. The finding that a market index proxy does not explain cross-sectional returns is consistent with even a very close, but unobserved, true market index being efficient.

The almost pathological knife-edged nature of the expected return-beta OLS cross-sectional relation, even without measurement error, is a shaky base for modern finance. Surely the idea of a tradeoff between risk and expected return is valid and meaningful. Whatever model is eventually used to measure and apply that basic idea will have to be considerably more robust.

As proved by Wheatley (1992) and Kandel and Stambaugh (1993), using a GLS cross-sectional fit between true expected returns and betas renders the relation less subject to these knife-edged properties. The GLS slope is positive so long as the expected return on the index proxy exceeds the expected return of the global minimum variance portfolio. This implies that virtually *any* proxy for the market index that is not grossly inefficient will produce a positive cross-sectional relation between mean returns and betas in large samples. But since every conceivable proxy candidate produces a positive relation, an empirical finding of a positive slope by itself implies very little about whether the proxy is ex ante efficient. Such a finding must be abetted by other direct tests of efficiency.

Sampling error makes these problems all the more troublesome. Since estimates of the efficient frontier and of the index proxy's mean and variance are subject to serious sampling error, the proxy itself may have a true positive cross-sectional expected return-beta OLS relation that cannot be detected in the sample mean return–estimated beta relation. For the GLS version, one is obliged to detect the difference between sampling scatter and ex ante scatter about the true cross-sectional relation. Again, it seems likely that cross-sectional tests of the mean-beta relation will take a back seat to direct tests of portfolio efficiency.

Despite these problems with the SLB Model, market value weighted index proxies are of considerable interest in their own right because they reflect averages of investor holdings. Whether or not such indices produce betas that are cross-sectionally related to average returns, their own returns serve as a benchmark for investment comparisons. Beating or trailing a value-weighted index has become the most widely accepted criterion of investment performance. It is an appropriate criterion relative to the wealth-weighted average returns of other investors.

Appendix: Derivation of Index Proxies That Produce a Given Cross-sectional Slope between Expected Returns and Betas

Notation:[17]

$\mathbf{R}$ = Expected returns vector for N individual assets,
$\mathbf{V}$ = $N \times N$ Covariance matrix of returns,
$\mathbf{1}$ = Unit vector,
$\mathbf{q}$ = Portfolio weights vector,
r = Scalar expected portfolio return, $\mathbf{q}'\mathbf{R}$,
σ^2 = Scalar portfolio return variance, $\mathbf{q}'\mathbf{V}\mathbf{q}$,
σ_j^2 = Cross-sectional or time series variance of j,
μ = Cross-sectional mean of expected returns, $\mathbf{R}'\mathbf{1}/N$,
$\boldsymbol{\pi}$ = Vector of scaled expected return deviations from the cross-sectional mean, $(\mathbf{R} - \mu\mathbf{1})/N$,
k = Scalar slope from cross-sectionally regressing $\mathbf{R}$ on betas computed for individual assets against portfolio $\mathbf{q}$.

[17] Vectors and matrices are denoted in boldface.

The mathematical problem is to find a *minimum variance* portfolio-index proxy that satisfies three conditions: (1) that the portfolio's expected return is a fixed value r, (2) that its weights $\mathbf{q}$ sum to unity, and (3) that a cross-sectional regression of expected returns $\mathbf{R}$ on betas $\boldsymbol{\beta} \equiv \mathbf{Vq}/\mathbf{q}'\mathbf{Vq}$ has a given slope.

Formally,

$$\text{minimize } \mathbf{q}'\mathbf{Vq} \text{ with respect to } \mathbf{q},$$

subject to

$$\mathbf{q}'\mathbf{R} = r$$
$$\mathbf{q}'\mathbf{1} = 1$$
$$\mathbf{q}'\mathbf{V}\boldsymbol{\pi} = k\mathbf{q}'\mathbf{Vq}.$$

The parameter k in the last constraint fixes the cross-sectional relation between expected returns and β's. In the cross-sectional regression, $\mathbf{R} = \gamma_0 + \gamma_1\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, the slope coefficient is $\gamma_1 = k/\sigma_\beta^2$, where σ_β^2 is the cross-sectional variance of $\boldsymbol{\beta}$.[18]

The first-order condition for a minimum is

$$\mathbf{Vq} - \lambda_1\mathbf{R} - \lambda_2\mathbf{1} - \lambda_3(\mathbf{V}\boldsymbol{\pi} - 2k\mathbf{Vq}) = 0,$$

where the λ's are Lagrange multipliers.

To eliminate the Lagrange multipliers, define the 3×3 matrix

$$\mathbf{A} \equiv [\mathbf{R} \quad \mathbf{1} \quad \mathbf{V}\boldsymbol{\pi}]'\mathbf{V}^{-1}[\mathbf{R} \quad \mathbf{1} \quad \mathbf{V}\boldsymbol{\pi}], \tag{A1}$$

collect terms and simplify the first-order condition to

$$\mathbf{q} = \mathbf{V}^{-1}[\mathbf{R} \quad \mathbf{1} \quad \mathbf{V}\boldsymbol{\pi}]\mathbf{A}^{-1}[r \quad 1 \quad k\sigma^2]' \tag{A2}$$

The equation of the boundary of the set of permissible indices in the r/σ^2 space can be obtained by using $\mathbf{q}$ from (A2) in the definition $\sigma^2 = \mathbf{q}'\mathbf{Vq}$ and then simplifying to obtain,

$$\sigma^2 = [r \quad 1 \quad k\sigma^2]\mathbf{A}^{-1}[r \quad 1 \quad k\sigma^2]'. \tag{A3}$$

Note that (A3) is not yet a functional relation since σ^2 appears on both sides.

To reduce the solution further, we are obliged to pay some attention to the structure of $\mathbf{A}^{-1}$. From (A1), the matrix $\mathbf{A}$ is a quadratic form in $\mathbf{V}$ and thus positive definite if $\mathbf{V}$ is positive definite (which we will assume); thus $|\mathbf{A}| > 0$. However, since (A3) is nonlinear in σ^2, $\mathbf{A}$ being positive definite does not guarantee that every solution to the first-order conditions is a minimum. Inspection of the cross-sectional beta constraint,

$$\mathbf{q}'\mathbf{V}\boldsymbol{\pi} = k\mathbf{q}'\mathbf{Vq},$$

[18] The constraint may be slightly confusing because only the expected return is de-meaned (while beta is not de-meaned). But when calculating a covariance, it is necessary to de-mean only one of the two random variables; i.e., $\mathrm{Cov}(x, y) = E\{x[y - E(y)]\} = E(xy) - E(x)E(y)$.

reveals that $\mathbf{q}$ is bounded from above; this implies that the constraint will provide both a maximum and a minimum. For our problem the appropriate second-order condition is the definiteness of

$$(1 + 2k\lambda_3)\mathbf{V},$$

which depends on the sign of $(1 + 2k\lambda_3)$ since $\mathbf{V}$ is positive definite. The first-order equation (A3) is a quadratic and has two roots corresponding to the minimum when the above expression is positive and the maximum when it is negative.

$\mathbf{A}$ can be written

$$\mathbf{A} = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & g \end{bmatrix}$$

where three of the scalar elements, $a = \mathbf{R}'\mathbf{V}^{-1}\mathbf{R}$, $b = \mathbf{R}'\mathbf{V}^{-1}\mathbf{1}$, $c = \mathbf{1}'\mathbf{V}^{-1}\mathbf{1}$, are the familiar efficient frontier information constants (cf. Roll (1977), appendix). The other three elements can be expanded and interpreted as follows:

$$d = \mathbf{R}'\boldsymbol{\pi} = \mathbf{R}'(\mathbf{R} - \mu\mathbf{1})/N = \mathbf{R}'\mathbf{R}/N - \mu^2. \tag{A4}$$

Thus, d can be recognized as the *cross-sectional variance of expected returns*, $d = \sigma_R^2$. Similarly,

$$e = \mathbf{1}'\boldsymbol{\pi} = \mathbf{1}'(\mathbf{R} - \mu\mathbf{1})/N = 0.$$

Finally,

$$g = \boldsymbol{\pi}'\mathbf{V}\boldsymbol{\pi} = [\mathbf{R}'\mathbf{V}\mathbf{R} - 2\mu\mathbf{R}'\mathbf{V}\mathbf{1} + \mu^2\mathbf{1}'\mathbf{V}\mathbf{1}]/N^2, \tag{A5}$$

and since $\mu = \mathbf{R}'\mathbf{1}/N$,

$$g = \mu^2\sigma_{R-1}^2,$$

where σ_{R-1}^2 denotes the time series variance of the *difference* in returns between two portfolios, one weighted proportionately to the vector of expected returns and the second one equally weighted.

Since the scalar element e is zero, the matrix inversion is simplified slightly and

$$\mathbf{A}^{-1} = \begin{bmatrix} cg & -bg & -cd \\ -bg & ag - d^2 & bd \\ -cd & bd & ac - b^2 \end{bmatrix} \frac{1}{|\mathbf{A}|}$$

where $|\mathbf{A}| = g(ac - b^2) - cd^2$. Using this expression for $\mathbf{A}^{-1}$, the formula describing the boundary of possible indices, equation (A3), can be written as

$$B\sigma^4 + Cr\sigma^2 + Dr^2 + F\sigma^2 + Gr + H = 0 \tag{A6}$$

where

$$B = k^2(ac - b^2), \quad C = -2dkc, \quad D = gc,$$
$$F = 2dkb - g(ac - b^2) + cd^2, \quad G = -2gb, \quad \text{and} \quad H = ag - d^2.$$

Equation (A6) can be recognized as the general form of a second-degree equation in r/σ^2 space. From analytic geometry, it is a parabola, a circle, an ellipse, or a hyperbola, depending on the value of $C^2 - 4BD$. Examining this expression,

$$C^2 - 4BD = 4d^2k^2c^2 - 4k^2(ac - b^2)gc = -4k^2c|\mathbf{A}|,$$

and since c and $|\mathbf{A}|$ are positive, $C^2 - 4BD$ is either zero (for $k = 0$) or negative. For $k \neq 0$, equation (A6) is an *ellipse* in r/σ^2 space. The axes of the ellipse are oblique, i.e., not parallel to the r/σ^2 axes. In the special case $k = 0$, (a zero cross-sectional slope between expected returns and betas), equation (A6) describes a parabola with an axis parallel to the σ^2 axis.

The situation for $k \neq 0$ is complex; the set of k-slope-producing indices is bounded by an ellipse that may or may not have a tangency point to the efficient frontier, depending on the value of k. To prove this assertion, note that the cross-sectional slope between expected returns and betas computed against a mean-variance efficient portfolio has the value $\Delta \equiv r^* - r_z$, where r^* is the portfolio's expected return and r_z is the return on its companion "zero-beta" portfolio. It is straightforward to show[19] that $r_z = (br^* - a)/(cr^* - b)$. Thus,

$$\partial\Delta/\partial r^* = 1 - \left[(ac - b^2)/(cr^* - b)^2\right]$$
$$= 0 \Rightarrow r^* = r_0 \pm (ac - b^2)^{1/2}/c, \quad \text{(A7)}$$

where $r_0 = b/c$ is the return on the global minimum variance portfolio. Equation (A7) indicates the presence of two local extrema. Checking the second-order conditions,

$$\partial^2\Delta/\partial r^{*2} > 0 \Rightarrow r^* > r_0.$$

Thus, the positive root of (A7) is a local minimum above which $\Delta > 0$ while the negative root is a local maximum below which $\Delta < 0$. There is a discontinuity at r_0, at which point Δ is undefined. There is no efficient portfolio with a "risk premium," Δ, between the two extrema. By direct substitution, the values of Δ at the extrema are,

$$\Delta_{\max} = -\Delta_{\min} = 2(ac - b^2)^{1/2}/c. \quad \text{(A8)}$$

For a mean-variance efficient portfolio, there is an exact cross-sectional linear relation between expected returns and betas,

$$\mathbf{R} \equiv r_z\mathbf{1} + (r^* - r_z)\boldsymbol{\beta}.$$

Thus, $\sigma_R^2 = (r^* - r_z)^2\sigma_\beta^2$, and since $k = \sigma_\beta^2(r^* - r_z) \Rightarrow k = \sigma_R^2/(r^* - r_z)$. This implies that $|k|$ has a *maximum* determined by the two extrema in (A8),

$$|k| \leq \frac{1}{2}\frac{\sigma_R^2/\sigma_0^2}{(ac - b^2)^{1/2}} \quad \text{(A9)}$$

[19] Cf. Roll (1977), appendix.

where $\sigma_0^2 = 1/c$ is the global minimum variance. For any value of k greater in absolute value than the expression above, there is no tangency between the efficient frontier and the ellipse bounding the set of all index proxies that produce a cross-sectional slope of k.

Notice, too, that since σ_β^2 is endogenous to the problem, constraining k is not the same as constraining $\gamma_1 = k/\sigma_\beta^2$, the cross-sectional slope coefficient, in the case where $k \neq 0$. This more complex problem introduces nonlinearities that will change the shapes of our boundaries but will not alter the qualitative properties we report.

A. Using GLS in the Cross-sectional Mean-Beta Regression

Begin with the familiar cross-sectional model, $\mathbf{R} = \gamma_0 \mathbf{1} + \gamma_1 \boldsymbol{\beta} \equiv \mathbf{B}\Gamma$, where $\mathbf{B} \equiv [\mathbf{1} \quad \boldsymbol{\beta}]$ and $\Gamma = (\gamma_0 \quad \gamma_1)'$. Since $\mathbf{V}$ is the covariance matrix of returns, it is natural to consider a GLS estimator based on the sample mean returns and a consistent estimator of $\mathbf{V}$. In large samples, the maximum likelihood consistent GLS estimator of Γ will be

$$[\mathbf{B}'\mathbf{V}^{-1}\mathbf{B}]^{-1}\mathbf{B}'\mathbf{V}^{-1}\mathbf{R}.$$

By expanding this expression, it is straightforward to show that the sign of the resulting estimator of γ_1 depends on the sign of

$$(\boldsymbol{\beta}'\mathbf{V}^{-1}\mathbf{R})(\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}) - (\mathbf{1}'\mathbf{V}^{-1}\boldsymbol{\beta})(\mathbf{1}'\mathbf{V}^{-1}\mathbf{R}).$$

But since $\boldsymbol{\beta} = \mathbf{V}\mathbf{q}/\mathbf{q}'\mathbf{V}\mathbf{q}$, where $\mathbf{q}$ is the vector of investment proportions of the market index proxy, the above expression is proportional to

$$\mathbf{q}'\mathbf{R} - (\mathbf{1}'\mathbf{V}^{-1}\mathbf{R})/(\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}) = r - r_0.$$

Thus, regardless of the position of the market index proxy, as the sample size grows larger, the sign of this particular GLS estimator of γ_1 will converge to a positive (or negative) value when the proxy's expected return, r, is greater (or less) than the expected return, r_0, of the global minimum variance portfolio.[20]

REFERENCES

Amihud, Yakov, Bent Jesper Christensen, and Haim Mendelson, 1992, Further evidence on the risk-return relationship, Working paper, Graduate School of Business, Stanford University.

Black, Fischer, 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444–455.

Chen, Nai-fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–403.

Coggin, T. Daniel, and John E. Hunter, 1985, Are high-beta, large-capitalization stocks overpriced? *Financial Analysts Journal* 41, 70–71.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 67, 427–465.

[20] We are indebted to Simon Wheatley for pointing out these results.

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* 10, 3–27.

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

Jobson, J. D., and Bob Korkie, 1982, Potential performance and tests of portfolio efficiency, *Journal of Financial Economics* 10, 433–466.

Kandel, Shmuel, and Robert F. Stambaugh, 1987, On correlations and inferences about mean-variance efficiency, *Journal of Financial Economics* 18, 61–90.

———, 1989, A mean-variance framework for tests of asset pricing models, *Review of Financial Studies* 2, 125–156.

———, 1993, Portfolio inefficiency and the cross-section of mean returns, Working paper.

Lakonishok, Josef, and Alan C. Shapiro, 1986, Systematic risk, total risk and size as determinants of stock market returns, *Journal of Banking and Finance* 10, 115–132.

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.

MacKinlay, A. Craig, and Matthew P. Richardson, 1991, Using generalized method of moments to test mean-variance efficiency, *Journal of Finance* 46, 511–527.

Reinganum, Marc R., 1981, A new empirical perspective on the CAPM, *Journal of Financial and Quantitative Analysis* 16, 439–462.

Roll, Richard, 1977, A critique of the asset pricing theory's tests, *Journal of Financial Economics* 4, 129–176.

———, 1980, Orthogonal portfolios, *Journal of Financial and Quantitative Analysis* 15, 1005–1023.

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

Shanken, Jay, 1985, Multivariate tests of the zero-beta CAPM, *Journal of Financial Economics* 14, 327-348.

———, 1987, Multivariate proxies and asset pricing relations, *Journal of Financial Economics* 18, 91–110.

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

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

Stambaugh, Robert F., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.

Wheatly, Simon, 1992, Private communication.

Zhou, Guofu, 1991, Small sample tests of portfolio efficiency, *Journal of Financial Economics* 30, 165–191.