Journal of Financial Economics 18 (1987) 91-110. North-Holland

MULTIVARIATE PROXIES AND ASSET PRICING RELATIONS
Living with the Roll Critique

Jay SHANKEN*

University of Rochester, Rochester, NY 14627, USA

Received August 1984, final version received August 1986

A framework is developed in which inferences can be made about the validity of an equilibrium asset pricing relation, even though the central aggregate in this relation is unobservable. A multivariate proxy for the true market portfolio, consisting of an equal-weighted stock index and a long-term government bond index, is employed in an investigation of the Sharpe-Lintner CAPM. The empirical evidence suggests that we can reject the joint hypothesis that (a) CAPM is valid, and (b) multiple correlation between the true market portfolio and proxy assets exceeds 0.7. Connections to the equilibrium factor pricing literature are also explored.

1. Introduction

A feature common to many models in modern financial theory is that expected return is linear in the covariance of an asset's return with some fundamental economic aggregate; e.g., (the marginal utility of) aggregate wealth or consumption.[1] This covariance measures the asset's 'systematic risk'. Since the theoretical aggregates are not directly observable, however, proxies are employed in empirical tests. Consequently, the results of such investigations are open to various interpretations. In particular, empirical rejection of the implied risk-return relation may indicate a violation of the underlying theory; alternatively, it may simply reflect the misspecification of the proxy.

This point has been made emphatically by Roll (1977) in the context of the capital asset pricing model (CAPM), where the validity of the risk-return relation is equivalent to the mean-variance efficiency of the market portfolio. Noting that the true market portfolio may be efficient when a proxy is not

*Earlier versions of this paper were presented at the 1983 and 1985 AFA meetings and in seminars at the Universities of British Columbia, Wisconsin, and Rochester, Wharton, Columbia, Berkeley, UCLA, Cornell, Northwestern, Ohio State, and Oxford University. I am grateful to D. Breeden, S. Brown, G. Chamberlain, M. Grinblatt, S. Hodges, S. Kandel, B. Korajczyk, C. MacKinlay, J. Ohlson, R. Roll, G. Rubio, M. Sperling, A. Tajirian, S. Titman and, especially, the editor John Long for helpful comments and suggestions.

[1]See, for example, Sharpe (1964), Lintner (1965), Black (1972), Rubinstein (1976) and Breeden (1979).

0304-405X/87/$3.50

(and conversely), Roll concludes that 'the theory is not testable unless the exact composition of the true market portfolio is known and used in the tests'. He argues further that 'most reasonable proxies will be very highly correlated with each other and with the true market whether or not they are mean–variance efficient', and that 'this high correlation will make it seem that the exact composition is unimportant, whereas it can cause quite different inferences'.[2]

In this paper, an empirical framework is developed in which a prior belief about the correlation between a proxy and the true market portfolio can be explicitly incorporated. The usual notion of a proxy is expanded to accommodate a *vector* of variables which, together, account for much of the variation in the market portfolio return. In this context, the focus is on the *multiple* correlation between the proxy and the market portfolio. Particular attention is given to the case in which the proxy variables are asset returns. Here, we find that if the statistical evidence of the proxy's inefficiency is sufficiently strong, then the inefficiency of the true market may indeed be correctly inferred and the CAPM rejected. The strength of such an inference increases with the presumed correlation and, of course, is conditional on the correctness of this prior belief or joint hypothesis.

The usefulness of additional information of various sorts in testing mean–variance efficiency has been considered previously by Kandel (1984) and Shanken (1986a). The present paper differs from these analyses in its emphasis on the role of correlation and the development of a multivariate proxy perspective. Furthermore, our results are applicable to a variety of equilibrium models and are not limited to investigations of the mean–variance theory. Some of the issues addressed in this paper are also considered in recent work by Kandel and Stambaugh (1987), Green (1984) and Mayers (1972).

The pricing restrictions derived here also extend earlier work on equilibrium factor pricing models, with the proxy components serving as the 'factors'. In particular, a bound on an individual asset's deviation from exact multibeta pricing is obtained which does not require that the market portfolio is *exactly* well-diversified [as in the Connor (1986) model] or that the factor model disturbances are independent [as in the models of Dybvig (1983) and Grinblatt and Titman (1983)].

The paper is organized as follows. In section 2, some general pricing restrictions are developed and compared to previous results in the asset pricing literature. Additional restrictions are derived and interpreted in section 3, for the case where the proxy components are portfolio returns. In section 4, econometric procedures for testing these relations are introduced and applied in an investigation of the Sharpe–Lintner CAPM. Some extensions and topics

[2] Stambaugh (1982) obtains similar inferences about the CAPM using several different market proxies.

for future research are explored in section 5, and section 6 summarizes the main conclusions of the paper.

2. Restrictions with a general proxy

Let m be an unobservable scalar random variable. A K-dimensional random proxy, P, and an N-vector of security returns, R, are observable and the corresponding set of $N+K$ components is assumed to be linearly independent. Before imposing any additional structure, we derive a useful lemma.

Lemma 1. Consider the linear regression of m on P and the multivariate linear regression of R on P:

$$m = a_m + b_m P + e_m, \tag{1}$$

and

$$R = a + BP + e. \tag{2}$$

Then

$$\operatorname{cov}(e, e_m)' \Sigma_e^{-1} \operatorname{cov}(e, e_m) \leq \sigma^2(m)(1-\rho^2), \tag{3}$$

where ρ *is the multiple correlation between P and m and* Σ_e *is the* $N \times N$ *covariance matrix of e. Furthermore, equality holds if and only if* e_m *is an exact linear function of the N components of e.*

Proof. The N-vector of coefficients from the regression of e_m on e is $\Sigma_e^{-1}\operatorname{cov}(e, e_m)$ and $\operatorname{cov}(e, e_m)'\Sigma_e^{-1}\operatorname{cov}(e, e_m)$ is the corresponding 'explained variance'. The regression of m on P ensures that $\sigma^2(e_m) = \sigma^2(m)(1-\rho^2)$. Thus, the lemma is just the statement that the explained variance in the regression of e_m on e is bounded above by the total variance of e_m; i.e., the r-squared in this regression is at most one. Q.E.D.

Lemma 1 is quite general and makes no use of our intended interpretation of m as an economic aggregate. If, however, expected excess return is given by security covariance with m then $\operatorname{cov}(e, e_m)$ may be interpreted as a vector of deviations from an exact multibeta expected return relation. In this case, the lemma provides an upper bound on a weighted sum of squared deviations from the relation. The formal statement is given in:

Proposition 1. Assume

$$\mathrm{E}(R) = r1_N + \operatorname{cov}(R, m) \tag{4}$$

for some r. Then there exists a K-vector, γ, of proxy 'prices of risk' that satisfies

$$d'\Sigma_e^{-1}d \leq \sigma^2(m)(1-\rho^2), \tag{5}$$

where

$$d \equiv \mathrm{E}(R) - r1_N - B\gamma. \tag{6}$$

In particular, $\gamma = \mathrm{cov}(P, m)$ *satisfies* (5).

Proof. Substituting the regression of R on P into (4) gives $\mathrm{E}(R) = r1_N + B\,\mathrm{cov}(P, m) + \mathrm{cov}(e, m)$. By construction, $\mathrm{cov}(e, P) = 0$, hence $\mathrm{cov}(e, m) = \mathrm{cov}(e, e_m)$. Lemma 1 then implies that $\gamma = \mathrm{cov}(P, m)$ satisfies inequality (5). Q.E.D.

A case of particular interest is that in which m is proportional to the return, R_m, on a portfolio that is efficient with respect to a set of security returns which includes R. With the constant of proportionality equal to $[\mathrm{E}(R_m) - r]/\sigma^2(R_m)$, the risk–return relation (4) holds and $\sigma(m)$ equals the portfolio's 'Sharpe measure of performance', defined as[3]

$$\theta_m \equiv [\mathrm{E}(R_m) - r]/\sigma(R_m).$$

Henceforth, we refer to this simply as the 'efficient portfolio case'.

Proposition 1 takes on positive economic significance only when an asset pricing theory is postulated which specifies the identity of m. If $\rho = 1$, it follows that $d = 0$; i.e., $\mathrm{E}(R)$ is exactly linear in the columns of B, the $N \times K$ matrix of multivariate regression coefficients. In particular, if m is proportional to the return on the value-weighted market portfolio of all assets, then we have a 'multibeta interpretation of the CAPM', as in earlier papers by Rosenberg and Guy (1976), Ross (1976), and Sharpe (1977). Connor (1984) derives a similar exact result which does not assume mean–variance preferences but does require perfect (multiple) correlation between the proxy and the true market return.[4]

When ρ is merely 'close' to one, a bound on an individual asset's pricing deviation may be of interest. If security i is an element of R, then Proposition 1 yields

$$|\mathrm{E}(R_i) - r - b_i\gamma| \leq \sigma(e_i)\sigma(m)(1-\rho^2)^{1/2}. \tag{7}$$

[3] This term is usually reserved for the case in which r is a risk-free rate.

[4] Connor's assumption that the market portfolio is 'well-diversified' is identical to this condition.

Here, b_i is a row vector of slope coefficients from the regression of R_i on P. In addition to the multiple correlation, ρ, the bound depends on the aggregate parameter, $\sigma(m)$, and i's residual risk.[5] It approaches zero as either ρ approaches one or $\sigma(e_i)$ approaches zero. Related bounds are derived by Dybvig (1983) and Grinblatt–Titman (1983) in the context of a strict factor model (with independent residuals). A potential advantage of (7) is that a few well chosen factors may suffice to ensure that $\rho \approx 1$, while the empirical evidence of Dhrymes, Friend and Gultekin (1984) and others indicates that a rather large number of factors is needed to satisfy the strict factor model assumption.

3. Restrictions with a proxy return vector

In this section we assume that the proxy P is a vector of portfolio returns that satisfy the underlying pricing relation; i.e.,

$$\mathrm{E}(P) = r1_K + \mathrm{cov}(P, m). \tag{8}$$

As we shall see below, this entails additional restrictions that greatly simplify the task of testing the pricing theory. In situations where non-return factors are of interest, it may be feasible to form portfolios which roughly 'mimic' the factors.[6] Thus, the limitations imposed by the portfolio assumption may not be very great.

A few definitions are needed before we state our next result. Let Σ_p be the $K \times K$ non-singular covariance matrix of P. For a given r, Jobson and Korkie (1983) show that the maximum Sharpe performance measure over all portfolios of the components of P is θ_p, where

$$\theta_p^2 = [\mathrm{E}(P) - r1_K]'\Sigma_p^{-1}[\mathrm{E}(P) - r1_K]. \tag{9}$$

This parameter plays a fundamental role in:

Proposition 2. *If* $\mathrm{E}(P) = r1_K + \mathrm{cov}(P, m)$ *then*

$$\rho^2 = \theta_p^2/\sigma^2(m). \tag{10}$$

In this case, the conclusion of Proposition 1 reduces to

$$d'\Sigma_e^{-1}d \leq \theta_p^2(\rho^{-2} - 1), \tag{11}$$

[5] Recall that $\sigma(m)$ equals θ_m in the efficient portfolio case. To prove (7), define a 'new' R which consists of security i alone and apply Proposition 1, with $N = 1$.

[6] See footnote 7 of Breeden (1979) for a related observation.

where

$$d = [\mathrm{E}(R) - r1_K] - B[\mathrm{E}(P) - r1_K].$$

Proof. Consider the regression of m on P and a constant. The K-vector of slope coefficients is

$$\Sigma_p^{-1}\operatorname{cov}(P, m).$$

Since $\operatorname{cov}(P, m) = \mathrm{E}(P) - r1_K$ by assumption, the explained variance in this regression is

$$[\Sigma_p^{-1}\operatorname{cov}(P, m)]'\Sigma_p[\Sigma_p^{-1}\operatorname{cov}(P, m)]$$

$$= [\mathrm{E}(P) - r1_K]'\Sigma_p^{-1}[\mathrm{E}(P) - r1_K],$$

which equals θ_p^2. Hence, the r-squared of the regression is

$$\theta_p^2/\sigma^2(m),$$

as in (10). (11) follows easily. Q.E.D.

The significance of Proposition 2 lies in the fact that the zero-beta rate r and the multiple correlation ρ are the only parameters in the deviation restriction that cannot be estimated directly from observable data on (P, R). In particular, if r is an observable riskless rate of return, then only ρ need be specified as a joint hypothesis. The elimination of $\sigma(m)$ is especially important since, in general, this parameter may be a complicated function of aggregate marginal utilities which are difficult to assess. A corollary of Proposition 2 yields an interesting interpretation of our restriction in terms of the familiar mean–variance portfolio geometry.[7]

Corollary 1. Let P be a K-vector of returns and t the tangency portfolio determined by (r, P, R); *then*

$$d'\Sigma_e^{-1}d = \theta_p^2(\rho_t^{-2} - 1), \tag{12}$$

where ρ_t, *the multiple correlation between P and t, equals* θ_p/θ_t.[8]

[7]From now on, we assume that θ_p and ρ are both positive.

[8]Kandel and Stambaugh (1987) independently show that $\rho_t = \theta_p/\theta_t$ when $K = 1$. They also consider procedures for testing hypotheses about ρ_t.

Proof. The conclusions in Proposition 2 must hold for any m which satisfies the assumed pricing relations for P and R. In particular, these pricing relations must hold in the efficient portfolio case where m is proportional to the return on the tangency portfolio t; hence,

$$d'\Sigma_e^{-1}d \leq \theta_p^2\left(\rho_t^{-2} - 1\right).$$

The conclusion (12) follows by noting that the condition for equality in Lemma 1 is satisfied, as m is a linear combination of the components of P and R. Finally, recall that $\sigma(m) = \theta_t$ in the efficient portfolio case. Thus, $\theta_p/\theta_t = \theta_p/\sigma(m)$, which equals ρ_t by Proposition 2. Q.E.D.

It follows immediately from Corollary 1 that the deviation restriction in Proposition 2 holds if and only if the multiple correlation between P and the tangency portfolio, t, exceeds ρ. As a ratio of Sharpe measures, ρ_t is naturally interpreted as a measure of the relative potential efficiency of P when r is a riskless rate of return. Thus, our equilibrium restriction amounts to a lower bound on the relative efficiency of the proxy, in this case. In particular, if the proxy is perfectly correlated with the relevant economic aggregate, then equilibrium requires that $\theta_p/\theta_t = 1$ or $\theta_p = \theta_t$; i.e., some portfolio of the proxy components must equal the tangency portfolio.

4. Empirical analysis with a riskless asset and a proxy return vector

In this section we assume that the components of P are portfolio returns and r is an observable riskless rate. Several econometric difficulties are thereby avoided, although some remain. The test we propose makes use of the observation that the deviation vector, d, in Proposition 2, is the N-vector of intercepts in the multivariate linear regression of $R - r1_N$ on $P - r1_K$.

Assume the regression parameters d, B, and Σ_e are constant over T periods and the conditional distribution of R, given P, is multivariate normal. Let $\hat{d}$ and $\hat{\Sigma}_e$ be the usual unbiased estimators of the intercept vector and residual covariance matrix, respectively, obtained from N separate OLS excess return time-series regressions. Likewise, let $\hat{\theta}_p$ be the maximum likelihood estimator of the proxy performance measure, computed from the sample mean and covariance matrix of $P - r1_N$.

Gibbons, Ross and Shanken (1985) show that the conditional distribution of $[N^{-1}(T-N-K)/(T-K-1)]Q$, given P, is non-central F with degrees of freedom N and $T-N-K$ and non-centrality parameter[9]

$$\lambda \equiv Td'\Sigma_e^{-1}d/\left(1+\hat{\theta}_p^2\right), \tag{13}$$

[9]Also, see MacKinlay (1987).

where

$$Q \equiv T\hat{d}'\hat{\Sigma}_e^{-1}\hat{d}/\left(1+\hat{\theta}_p^2\right). \tag{14}$$

By Proposition 2, a necessary condition for the underlying pricing theory to hold is that

$$d'\Sigma_e^{-1}d \leq \theta_p^2\left(\rho^{-2}-1\right);$$

equivalently,

$$\mathrm{H}_0: \quad \lambda \leq T\theta_p^2\left(\rho^{-2}-1\right)/\left(1+\hat{\theta}_p^2\right). \tag{15}$$

Thus, our pricing restriction may be viewed as a constraint on the magnitude of the non-centrality parameter in the distribution of Q.[10] We exploit this observation below.

4.1. A test of the CAPM with a stock index proxy

As an initial exploration of this framework, we consider a test of the Sharpe–Lintner CAPM with the return on the CRSP equal-weighted stock index as P. This proxy is used in the early CAPM tests criticized by Roll (1977) and in many recent empirical investigations. The tests are carried out over five subperiods of equal length from February 1953 through November 1983. January returns have been deleted from the tests in light of much puzzling evidence which indicates that the return generating process may differ in January from that in the rest of the year.[11] As a result, each test period contains $T = 68$ months of data. Securities with complete data on the CRSP monthly return file, for a given subperiod, are stratified into $N = 20$ equal-weighted portfolios based on the market value of equity at the beginning of each subperiod. Returns on these portfolios constitute the vector R. Excess returns are computed using the monthly T-bill return series constructed by Ibbotson and Sinquefield.

Suppose the CRSP index is a perfect proxy ($\rho = 1$) for the true market portfolio. In this case the non-centrality parameter, λ, equals zero by (15), and our test statistic has a central F distribution; the null hypothesis requires that the index be the tangency portfolio. The five subperiod statistics and associated p-values are reported in table 1. An aggregate p-value of 0.02 for the overall period is obtained using a procedure suggested in Shanken (1985a).[12]

[10] The potential usefulness of noncentral distributions in testing approximate asset pricing relations was first noted in Shanken (1982a).

[11] See Keim (1983).

[12] For each subperiod, the standard normal z corresponding to the given p-value is determined. These z's are added and divided by $\sqrt{5}$ to obtain another standard normal variate from which the aggregate p-value is determined.

Table 1

Tests of the hypothesis that the CRSP equal-weighted stock index is equal to the tangency portfolio determined by the one-month T-bill rate, the index, and twenty equal-weighted portfolios stratified by firm size at the beginning of each subperiod. Tests are conducted over five subperiods from February 1953 through November 1983. January returns are deleted.

Subperiod	F-statistics[a]	p-value[b]
2/53 - 3/59	2.09	0.02
4/59 - 5/65	1.78	0.05
6/65 - 7/71	1.57	0.10
8/71 - 9/77	0.84	0.65
10/77 - 11/83	1.01	0.47

[a]All F-statistics have degrees of freedom 20 in the numerator and 47 in the denominator.

[b]An aggregate p-value for the overall period is 0.02. This is obtained by applying the inverse normal transformation to each of the subperiod p-values and aggregating the resulting five independent standard normal variates.

Thus, if the CAPM is true, we can infer that the CRSP index is not a perfect proxy.[13]

Now consider a test of the hypothesis that ρ exceeds, say, 0.7. Unfortunately, in order to evaluate the implied bound on λ, we need the true value of θ_p^2. Estimates of this 'nuisance parameter' are presented in table 2, under the assumption that excess returns on the CRSP index are independent and indentically normally distributed over each subperiod. Since the maximum likelihood estimates are biased upwards, unbiased estimates are also computed as in Jobson and Korkie (1980). The average of these rather volatile estimates is 0.023. The corresponding annualized value of θ_p is 0.52, implying a risk premium of 10.4% on a standard deviation of 20%.

If the true (monthly) value of θ_p^2 were known to be 0.023, we would proceed as follows. From table 2, $\hat{\theta}_p^2$ equals 0.121 in the first subperiod. With $\rho = 0.7$, the inequality in (15) is

$$\lambda \leq 68(0.023)(0.7^{-2} - 1)/1.121 = 1.45.$$

The test statistic for the first subperiod in table 1 is 2.09. The corresponding p-value, based on a non-central F distribution with degrees of freedom (20, 47) and non-centrality parameter 1.45, is 0.03; p-values for the other subperiods are computed similarly and aggregated as before.

The results of this experiment, for various values of ρ, are reported in the second column of table 3. The inferences are fairly strong; in particular, the hypothesis that ρ exceeds 0.7 is rejected at the 0.05 level. This implies that the CRSP index accounts for less than half ($0.7^2 = 0.49$) of the variation in the

[13]When January returns are included, the aggregate p-value is 0.07.

Table 2

Maximum likelihood (MLE) and unbiased estimates of θ_p^2 and corresponding annualized values of θ_p (in parentheses) for the CRSP equal-weighted stock index, over five subperiods from February 1953 through November 1983. Each subperiod contains 68 monthly observations. January returns are deleted. θ_p is the ratio of mean excess return to standard deviation of return for the index. Annualized θ_p in parentheses equals $(12 \times \text{monthly } \theta_p^2)^{1/2}$.

Estimator	Subperiod				
	2/53–3/59	4/59–5/65	6/65–7/71	8/71–9/77	10/77–11/83
MLE	0.121	0.017	0.003	0.010	0.045
	(1.20)	(0.45)	(0.19)	(0.35)	(0.73)
Unbiased	0.101	0.002	− 0.012	− 0.005	0.028
	(1.10)	(0.15)	(—)	(—)	(0.58)

Table 3

Aggregate p-values for tests of the hypothesis that the correlation betwen the CRSP equal-weighted index and the tangency portfolio exceeds ρ.[a] The tangency portfolio is that determined by the one-month T-bill rate, the index, and twenty equal-weighted portfolios stratified by firm size at the beginning of each subperiod. Tests are conducted over five subperiods from February 1953 through November 1983. January returns are deleted.

Correlation	Assumption about the value of θ_p[b]			
ρ	$\theta_p = 0.52$	$\theta_p = 1.0$	Posterior[c]	Bound[d]
1.0	0.02	0.02	0.02	0.02
0.9	0.02	0.04	0.02	0.03
0.8	0.03	0.11	0.04	0.04
0.7	0.05	0.28	0.06	0.06
0.6	0.09	0.62	0.12	0.13
0.5	0.19	0.94	0.24	0.29

[a]Aggregate p-values are obtained by applying the inverse normal transformation to each of the subperiod p-values and aggregating the resulting five independent normal variates.

[b]θ_p is the ratio of mean excess return to standard deviation of return for the CRSP equal-weighted index. θ_p values above are annualized.

[c]True value of monthly θ_p^2 is assumed to be randomly drawn from the posterior distribution in table 4.

[d]Upper bound on p-value for any posterior with first two moments equal to those of the posterior in table 4.

true market return, if the CAPM is true. Note that the p-values increase as ρ decreases and the corresponding non-centrality parameter increases.

If the true value of θ_p^2 is greater (less) than 0.023, the inferences above are biased toward rejection (acceptance) of the null hypothesis.[14] A more conservative approach would be to specify a value of θ_p which is greater than any

[14]If the value of θ_p^2 is too small, so is the non-centrality parameter, λ; hence, the p-value is too small.

conceivable true value. The value used, as an example, in the third column of table 3 corresponds to an annualized θ_p of one – an ex ante risk *premium* of 20% on a standard deviation of 20%. Interestingly, the hypothesis that ρ exceeds 0.9 is still rejected at the 0.05 level. The p-values rise sharply as ρ decreases, however.

Presumably, few individuals would assign much subjective probability to the notion that the annualized θ_p is as high as one. A more satisfactory pragmatic approach is to consider a range of possible values for θ_p^2 and compute conditional aggregate p-values for each one. A weighted average (unconditional) p-value can then be computed, with weights that vary according to the relative 'likelihood' of the different values of θ_p^2. The question, then, is how to devise this weighting scheme.

Bayesian analysis provides a useful starting point. Suppose the subperiod estimates of θ_p^2 were independent and identically normally distributed with known variance. Given a 'non-informative prior' distribution on the true value of θ_p^2, the posterior distribution would be normal with mean equal to the sample mean estimate and variance equal to the variance of the mean.[15] In fact, the unbiased estimator of θ_p^2 is not normally distributed, nor is its variance known exactly. Nonetheless, this theorem suggests a reasonable heuristic, if not an 'objectively correct' procedure, for assigning weights to the possible values of θ_p^2.[16]

Using results from Jobson and Korkie (1980), and assuming that θ_p^2 is constant over our five subperiods, we arrive at 0.019 as an estimate of the standard deviation of the sample mean value, 0.023, used earlier.[17] Motivated by the Bayesian theorem above, a simple discrete posterior distribution with the required mean and standard deviation has been constructed. The distribution, displayed in table 4, assigns mass to the mean, as well as values of θ_p^2 one standard deviation below or two standard deviations above the mean. The fourth column of table 3 contains results derived using this posterior distribution. The p-values lie between those obtained by our previous 'best guess' and 'worst case' methods, but are much closer to the former.

An example should clarify the nature of the posterior approach. Consider the case $\rho = 0.7$. First, conditional aggregate p-values are computed for each of the three possible values of θ_p^2 in table 4. The p-values are 0.02, 0.05, and

[15] See DeGroot (1970).

[16] A few colleagues have expressed horror at the notion that the inferences derived in this manner are Bayesian inferences. No such claim is made here. The posterior analysis is simply a pragmatic approach to the nuisance parameter problem. See Shanken (1986c) for a truly Bayesian approach to testing portfolio efficiency hypotheses.

[17] The variance of the unbiased estimator of θ_p^2 depends on the true value of θ_p^2. We replace this parameter by the sample mean estimate to get an estimate of the variance. The result is divided by five to obtain an estimate of the variance of the sample mean.

Table 4

Three-point posterior distribution for the ratio of mean excess return to standard deviation of return, θ_p. Parameters are based on data for the CRSP equal-weighted stock index over the period from February 1953 through November 1983. January returns are deleted. Annualized θ_p equals $(12 \times \text{monthly } \theta_p^2)^{1/2}$.

	Possible values			Mean	σ
Monthly θ_p^2	0.003	0.023	0.062	0.023	0.019
Annualized θ_p	0.20	0.52	0.86	0.47	0.23
Probability	1/3	1/2	1/6	—	—

0.18, for $\theta_p^2 = 0.003$, 0.023, and 0.062, respectively.[18] The final weighted-average p-value is just

$$0.06 = (1/3)(0.02) + (1/2)(0.05) + (1/6)(0.18).$$

Suppose we had considered some other posterior with the same mean and variance as the one in table 4. Is it possible that materially different conclusions would be reached? As demonstrated in the appendix, an uper bound on the aggregate p-value, over *all* possible posteriors with the given mean and variance, can easily be computed. The bounds reported in the last column of table 3 are, for $\rho \geq 0.6$, almost identical to the p-values obtained using the simple three-point posterior. Therefore, the particular form of the distribution is not an issue.

Table 5 presents the results of one final 'sensitivity analysis'. Increases of 25 percent in the mean and/or standard deviation of the posterior distribution for θ_p^2 are considered. The p-values barely change for $\rho = 0.8$ and $\rho = 0.9$, and we continue to reject $\rho = 0.7$ at the 10 percent significance level. Thus, our rejection of the joint hypothesis that (a) the CAPM is valid, and (b) ρ exceeds 0.7, appears to be quite robust.[19]

4.2. A test of the CAPM with a stock index – bond index proxy

The extent to which we interpret our previous empirical analysis as a rejection of the CAPM depends on our belief about the magnitude of the variation in the true market return left unexplained by the CRSP index. It

[18] Note that the p-value corresponding to $\theta_p^2 = 0.023$ is (necessarily) the same as that in column 2 of table 3, for $\rho = 0.7$.

[19] While we use the CAPM as a 'concrete' example, we could just as well refer to many other equilibria for which the joint hypothesis about ρ seems reasonable. In essence, we are testing a class of equilibrium models and our approach does not permit us to distinguish between models in this class.

Table 5

Aggregate p-values indicating sensitivity of tests of the relative efficiency of the CRSP equal-weighted stock index to increases of 25 percent in the mean, μ, and/or standard deviation, σ, of the posterior distribution for θ_p^2. θ_p is the ratio of mean excess return to standard deviation of return for the index. All p-values are upper bounds (see appendix).

Correlation		Parameter(s) increased 25%		
ρ	Bound[a]	μ	σ	(μ, σ)
1.0	0.02	0.02	0.02	0.02
0.9	0.03	0.03	0.03	0.03
0.8	0.04	0.04	0.04	0.04
0.7	0.06	0.08	0.07	0.08
0.6	0.13	0.16	0.15	0.18
0.5	0.29	0.37	0.34	0.43

[a]Upper bound on p-value for any posterior with first two moments equal to those of the posterior in table 4. Source, last column of table 3.

Table 6

Maximum likelihood (MLE) and unbiased estimates of θ_p^2 and corresponding annualized values of θ_p (in parentheses) for the proxy consisting of the CRSP equal-weighted stock index and the Ibbotson–Sinquefield long-term U.S. government bond index. Five subperiods from February 1953 through November 1983 are considered, each containing 68 monthly observations. January returns are deleted. θ_p^2 is the maximum squared ratio of mean excess return to standard deviation of return over all portfolios of the stock and bond indexes. Annualized θ_p in parentheses equals $(12 \times \text{monthly } \theta_p^2)^{1/2}$.

	Subperiod				
Estimator	2/53–3/59	4/59–5/65	6/65–7/71	8/71–9/77	10/77–11/83
MLE	0.121	0.031	0.036	0.036	0.068
	(1.20)	(0.61)	(0.66)	(0.66)	(0.90)
Unbiased	0.084	0.000	0.004	0.004	0.035
	(1.00)	(0.00)	(0.22)	(0.22)	(0.65)

seems reasonable to suppose that this variation is somewhat related to movements in interest rates. Therefore, we now extend our proxy to include the Ibbotson–Sinquefield long-term U.S. government bond index as well as the CRSP stock index.[20] Of course, including the bond index can only increase the multiple correlation between the proxy and the market. Hence, for any given value of ρ, our confidence in the validity of the joint hypothesis should increase as well.

In the present context, θ_p is the maximum Sharpe measure over all portfolios of the stock and bond indexes. Estimates of θ_p^2 for each of our five subperiods

[20]Chen, Roll and Ross (1984) also use this bond index to capture term-structure effects.

Table 7

Tests of the hypothesis that some portfolio of the CRSP equal-weighted stock index and the Ibbotson–Sinquefield long-term U.S. government bond index is equal to the tangency portfolio determined by the one-month T-bill rate, the indexes, and twenty equal-weighted portfolios stratified by firm size at the beginning of each subperiod. Tests are conducted over five subperiods from February 1953 through November 1983. January returns are deleted.

Subperiod	F-statistic[a]	p-value[b]
2/53 – 3/59	2.50	0.01
4/59 – 5/65	1.72	0.06
6/65 – 7/71	1.52	0.12
8/71 – 9/77	0.85	0.65
10/77 –11/83	0.91	0.57

[a]All F-statistics have degrees of freedom 20 in the numerator and 46 in the denominator.
[b]An aggregate p-value for the overall period is 0.02. This is obtained by applying the inverse normal transformation to each of the subperiod p-values and aggregating the resulting five independent standard normal variates.

Table 8

Aggregate p-values for tests of the hypothesis that the multiple correlation between the tangency portfolio and the proxy exceeds ρ. The proxy consists of the CRSP equal-weighted stock index and the Ibbotson–Sinquefield long-term U.S. government bond index. The tangency portfolio is that determined by the one-month T-bill rate, the indexes and twenty equal-weighted portfolios stratified by firm size at the beginning of each subperiod. The initial posterior parameters, used in column 2, are $\mu = 0.026$ and $\sigma = 0.023$. Increases of 25 percent in the mean, μ, and/or standard deviation, σ, of the posterior distribution for θ_p^2 are considered. θ_p is the maximum ratio of mean excess return to standard deviation of return over all portfolios of the stock and bond indexes. Symbol(s) over a given column indicates which parameter(s) has been increased. All p-values are upper bounds (see appendix).

Correlation	Initial	Parameters increased 25%		
ρ	posterior	μ	σ	(μ, σ)
1.0	0.02	0.02	0.02	0.02
0.9	0.02	0.03	0.02	0.03
0.8	0.04	0.04	0.04	0.05
0.7	0.07	0.08	0.08	0.09
0.6	0.14	0.18	0.17	0.20
0.5	0.34	0.43	0.41	0.50

are presented in table 6. The mean unbiased value, 0.026, and the standard deviation of the mean, 0.023, are both a bit higher than the values obtained earlier with the stock index proxy.[21]

Tests of the hypothesis that our two-dimensional proxy captures all of the variation in the true market return are reported in table 7. The restriction, in

[21] Note that, unlike the maximum likelihood estimates, some of the unbiased estimates in table 6 are lower than the corresponding estimates in table 2. This is due to the degrees of freedom correction in the unbiased estimator.

this case, is that some portfolio of the stock and bond indexes is equal to the tangency portfolio. The subperiod results are similar to those for the stock index alone, and the aggregate p-value is the same – 0.02. Thus, we conclude that ρ is less than one, provided the CAPM is true.

We proceed, therefore, to test whether some portfolio of the two indexes is 'near' the tangency portfolio. By Corollary 1 of section 3, the performance ratio, θ_p/θ_t, must exceed ρ if the CAPM holds. Recall that θ_t is the Sharpe measure of the tangency portfolio, t, while the performance ratio equals the multiple correlation between the proxy and t. The empirical results in table 8 differ very little from those reported in table 5 for the stock index proxy. The hypothesis that θ_p/θ_t exceeds 0.7 is rejected at the 10 percent significance level; either the CAPM is false or our proxy captures a relatively small portion of the movement in aggregate wealth.[22,23]

5. Extensions

In this section two additional results are presented which may be useful in extending the empirical methodology developed earlier. The first issue we address is that of an unknown zero-beta rate. This is not a problem when $\rho = 1$, since r can then be consistently estimated using the usual cross-sectional regression or maximum likelihood methods. When ρ is less than one, however, it would appear that r is not statistically identifiable. The following proposition suggests a way of dealing with this situation.

Proposition 3. Assume P is a portfolio return ($K = 1$). Let expected excess return on P and R be given by covariance with m. Consider all portfolios of (P, R) that have the same expected return as P and let P^* be the member of this class with the smallest variance of return. If ρ^* is the multiple correlation between P and P^*, then $\rho^* = \sigma(P^*)/\sigma(P)$ and $\rho^* \geq \rho$.

Proof. Chamberlain and Rothschild (1983) show that P^* equals the orthogonal projection (regression) of P on the minimum-variance frontier for (P, R), which is spanned by any two minimum-variance portfolios. Let r be the (unknown) zero-beta rate *for m*, and t the corresponding tangency portfolio

[22] Of course, a third alternative is that neither of these conclusions is true and we are simply observing a 'rare' event.

[23] For simplicity, we have assumed that θ_p^2 is constant across subperiods in constructing our posterior distribution. Alternatively, we can think of the posterior as the distribution of the maximum value of θ_p^2 over the five subperiods; the p-values computed are upper bounds in this interpretation. Since non-trivial increases in the parameters of the posterior distribution do not alter our basic conclusions, there is room for θ_p^2 to wander around a bit. In work not reported here, a 'new' true value of θ_p^2 is permitted to be independently drawn each subperiod. The results obtained under this scenario are slightly stronger (p-values smaller) than those discussed in the paper.

for (r, P, R). By the general properties of multiple regression, P^* must be the minimum-variance portfolio that is maximally correlated with P.[24] Thus, ρ^* cannot be less than the correlation, ρ_t, between P and t. Furthermore, as discussed below Corollary 1 of section 3, $\rho_t \geq \rho$, under our assumptions; hence $\rho^* \geq \rho$. Since $\mathrm{E}(P) = \mathrm{E}(P^*)$ and P^* is minimum-variance, the beta of P on P^* must be one; therefore, $\mathrm{cov}(P, P^*) = \sigma^2(P^*)$ and $\rho^* = \sigma(P^*)/\sigma(P)$. Q.E.D.

Proposition 4 tells us that the correlation between P and P^* cannot be less than the correlation between P and m; thus, a joint hypothesis bounding ρ below translates directly into a hypothesis about ρ^*. The empirical advantage of shifting the focus from m to P^* is that the zero-beta rate for P^* is a function of $\mathrm{E}(P)$ and the 'efficient set parameters' for (P, R); therefore, it can be consistently estimated.[25] With this rate statistically identified, we are one step closer to the empirical framework of section 4, where the zero-beta rate is known. Error in estimating this parameter is an added complication, however, which will have to be addressed in future work.[26]

Before concluding, we consider one other issue. The pricing restriction that has served as the focus of this paper involves Σ_e^{-1}, the inverse of the residual covariance matrix from the multivariate regression of R on P. This is somewhat limiting, in that estimation of Σ_e^{-1} requires that the dimension of R be less than the time-series length T. The following simple consequence of Proposition 1 does not involve Σ_e^{-1}, and may eventually provide the basis for a more powerful empirical test in which N is permitted to be very large.

Corollary 2. Under the conditions of Proposition 1,

$$d'd \leq \bar{c}\sigma^2(m)(1-\rho^2),$$

where

$$d = \mathrm{E}(R) - r1_N - B\gamma,$$

and $\bar{c}$ is the maximum eigenvalue of the residual covariance matrix, Σ_e.[27]

[24] When $K > 1$, P^* may be defined as the minimum-variance portfolio that is maximally correlated with some portfolio of the components of P. This is an example of the statistical problem of canonical correlations.

[25] As noted earlier, this is not the case for r, the zero-beta rate associated with m.

[26] Shanken (1986) derives an upper bound on the power function of the likelihood ratio test of portfolio efficiency when the unknown zero-beta rate is constrained to a given interval. This result may be useful in the context above.

[27] A similar sum of squared deviations appears in the original Ross (1976) arbitrage pricing theory (the prices of risk which define his deviation vector may differ from those defined in Proposition 1 of this paper, however). Ross establishes the existence of a finite bound that holds as N, the dimension of d, approaches infinity. Our general restriction, which involves Σ_e^{-1}, should be compared to Ingersoll's (1984) general APT result. See Shanken (1982b, 1985b) and Dybvig and Ross (1985) for analyses of the testability of the APT.

Proof. The proof is based on an observation in Chamberlain and Rothschild (1983). Since $\bar{c}$ is the maximum eigenvalue of Σ_e, $\bar{c}^{-1}$ is the minimum eigenvalue of Σ_e^{-1} and $\Sigma_e^{-1} - \bar{c}^{-1}I_N$ is positive semidefinite. Therefore, $d'[\Sigma_e^{-1} - \bar{c}^{-1}I_N]d \geq 0$ and the desired conclusion follows directly from Proposition 1. Q.E.D.

The usefulness of Corollary 2 will depend on the behavior of the eigenvalues of the residual covariance matrix as N increases. The scenario we have in mind is one in which the eigenvalues remain bounded (and 'small') – essentially the approximate factor structure assumption of Chamberlain and Rothschild (1983). This might be a reasonable assumption for some proxies but not for others.

It deserves emphasis that, even if the eigenvalues are 'well-behaved', the correlation between the proxy and the economic aggregate need not be close to one. Thus, it would be inappropriate to presume that an exact multibeta pricing relation necessarily provides a good approximation in this case. For example, suppose an important economic factor, orthogonal to P, affects the returns on a relatively small number of assets in the empiricist's subset. Omission of this factor implies that ρ is less than one, yet the components of P might well serve as the factors in an approximate factor model for the subset. Corollary 2 demonstrates that while the individual pricing deviations may not all be negligible, the sum of squared deviations must be appropriately bounded in equilibrium.[28]

6. Conclusions

Empirical evidence has been presented which suggests that either the Sharpe–Lintner CAPM is invalid or our proxies account for at most two-thirds (rejected at the 0.05 level), or perhaps only one-half (rejected at the 0.10 level), of the variation in the true market return. The results are essentially the same whether we use the CRSP equal-weighted stock index alone, or together with the Ibbotson–Sinquefield long-term U.S. government bond index, in a multivariate proxy. While an unambiguous inference about the validity of the CAPM is probably unattainable, our analysis demonstrates that it *is* possible to test the theory conditional on a prior belief about the proxy–true market correlation.

There is an important sense in which the results obtained here are stronger than has been suggested thus far. Consider a hypothetical regression of the true market portfolio return on the asset and proxy returns employed in our

[28] Connor and Korajczyk (1986) develop an exact pricing model which combines an approximate factor structure with the *assumption* that the multiple correlation betwen the factors and the market portfolio is one.

test of the CAPM. This regression yields a decomposition of the market return into a 'maximally correlated portfolio' (MCP) and an orthogonal component that is uncorrelated with the test returns.[29] In general, the orthogonal component may be substantial, perhaps reflecting a portion of the return on non-traded or international assets. It is easily shown, however, that under the assumption that the CAPM is true, the linear relation between expected return and 'beta' must hold with respect to the MCP as well as the market portfolio. Thus, our entire analysis could have been presented in terms of this portfolio rather than the market. For example, we can reject (at the 0.10 level) the joint hypothesis that (a) the CAPM is valid, and (b) the multiple correlation between our stock–bond indexes and the MCP exceeds 0.7.

The MCP perspective has important implications for the subjective evaluation of the adequacy of a given proxy and the associated interpretation of the rejection of our joint hypothesis. A good proxy need not 'capture' the orthogonal component of the market return and thus need not be very highly correlated with the market portfolio; it suffices that the proxy be highly correlated with the MCP. The main concern, therefore, should be with the extent to which a proxy fails to capture variation in the market return that is correlated with the assets used in the test.

It is instructive to consider the case in which the market portfolio's residual variation, i.e., the variation 'not explained' by the proxy, is unrelated to the given asset returns. Under this assumption, the multiple correlation between the proxy and the MCP is one. Thus, the proxy is 'perfect' and a test for exact (multibeta) pricing with respect to the proxy may be viewed as a test of the CAPM. A strong assumption of this sort is implicit in most of the existing asset pricing empirical literature. In this paper, a more flexible framework has been developed to accomodate the likely imperfection of commonly used proxies or factors. While our discussion has focused on the Sharpe–Lintner model, similar principles clearly apply in the testing of other equilibrium models such as the consumption CAPM and variants thereof.

Appendix

In this section an upper bound on the p-value for testing the relative efficiency of the stock index is derived. The same procedure, with different inputs, was used in the stock index–bond index proxy tests. The resulting bound relates the p-value to the mean and variance of the posterior distribution for θ_p^2.

Recall that the p-value of interest is actually an expected value of the conditional aggregate p-value, where the expectation is taken with respect to

[29]See Breeden (1979) or Breeden, Gibbons and Litzenberger (1986) for a discussion in the context of the consumption CAPM.

the posterior distribution. By (15), the conditional aggregate p-value, which we call y, is uniquely determined by the subperiod test statistics and estimates of θ_p^2, as well as the value of

$$x \equiv T\theta_p^2(\rho^{-2}-1). \tag{A.1}$$

Holding all else constant, y is a function of x.

Consider a cubic approximation to this functional relation:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + e(x), \tag{A.2}$$

where $e(x)$ is the approximation error for each value of x. Taking expected values, we have

$$\mathrm{E}(y) = a_0 + a_1\mathrm{E}(x) + a_2\mathrm{E}(x^2) + a_3\mathrm{E}(x^3) + \mathrm{E}(e(x)). \tag{A.3}$$

Since an upper bound on $\mathrm{E}(y)$ is desired, it is helpful to bound $e(x)$ above. Through a series of ad hoc numerical curve fitting (regression) experiments, it was found that with

$$(a_0, a_1, a_2, a_3) = (0.0203, 0.0202, 0.0242, -0.0003),$$

$e(x)$ is bounded above by 0.0006.[30] By Jensen's inequality, $\mathrm{E}(x^3) \geq [\mathrm{E}(x)]^3$, for $x \geq 0$; thus,

$$\mathrm{E}(y) \leq a_0 + a_1\mathrm{E}(x) + a_2\mathrm{E}(x^2) + a_3[\mathrm{E}(x)]^3 + 0.0006. \tag{A.4}$$

This bound only depends on the first two moments of the distribution of x. For given values of T and ρ, these moments are just constant multiples of the corresponding moments of θ_p^2, as is clear for (A.1).

[30] For very large values of x, $e(x)$ is negative and decreasing. Thus, this bound appears to be a global maximum.

References

Black, F., 1972, Capital market equilibrium with restricted borrowing, Journal of Business 45, 444–454.

Breeden, D., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265–296.

Breeden, D., M. Gibbons and R. Litzenberger, 1986, Empirical tests of the consumption-oriented CAPM, Working paper (Stanford University, Stanford, CA).

Chamberlain, G. and M. Rothschild, 1983, Arbitrage, factor structure, and mean–variance analysis on large asset markets, Econometrica 51, 1281–1304.

Chen, N., R. Roll and S. Ross, 1984, Economic forces and the stock market, Working paper (University of Chicago, Chicago, IL).

Connor, G., 1984, A unified beta pricing theory, Journal of Economic Theory 34, 13–31.

Connor, G. and R. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, Journal of Financial Economics 15, 373–394.

DeGroot, M., 1970, Optimal statistical decisions (McGraw-Hill, New York).
Dhrymes, P., I. Friend and N. Gultekin, 1984, A critical reexamination of the empirical evidence on the arbitrage pricing theory, Journal of Finance 39, 323–346.
Dybvig, P., 1983, An explicit bound on deviations from APT pricing in a finite economy, Journal of Financial Economics 12, 483–496.
Dybvig, P. and S. Ross, 1985, Yes, the APT is testable, Journal of Finance 40, 1173–1188.
Green, R., 1984, Benchmark portfolio efficiency and deviations from the security market line, Working paper (GSIA Carnegie-Mellon University, Pittsburgh, PA).
Gibbons, M., S. Ross and J. Shanken, 1985, A test of the efficiency of a given portfolio, Working paper (Graduate School of Business, Stanford University, Stanford, CA).
Grinblatt, M. and S. Titman, 1983, Factor pricing in a finite economy, Journal of Financial Economics 12, 497–507.
Ingersoll, J., 1984, Some results in the theory of arbitrage pricing, Journal of Finance 39, 1021–1039.
Jobson, J. and B. Korkie, 1980, Estimation for Markowitz efficient portfolios, Journal of the American Statistical Association 75, 544–554.
Jobson, J. and B. Korkie, 1982, Potential performance and tests of portfolio efficiency, Journal of Financial Economics 10, 433–466.
Kandel, S., 1984, On the exclusion of assets from tests of the mean variance efficiency of the market portfolio, Journal of Finance 39, 63–75.
Kandel, S. and R. Stambaugh, 1987, On correlations and the sensitivity of inferences about mean–variance efficiency, Journal of Financial Economics, this issue.
Keim, D., 1983, Size related anomalies and stock return seasonality: Further empirical evidence, Journal of Financial Economics 12, 13–32.
Lintner, J., 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, C., 1987, On multivariate tests of the CAPM, Journal of Financial Economics, forthcoming.
Mayers, D., 1972, Nonmarketable assets and capital market equilibrium under uncertainty, in: M. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).
Roll, R., 1977, A critique of the asset pricing theory's tests, Journal of Financial Economics 4, 129–176.
Rosenberg, B. and J. Guy, 1976, Prediction of beta from investment fundamentals, Financial Analysts Journal 32, 60–72.
Ross, S., 1976, Return, risk, and arbitrage, in: I. Friend and J. Bicksler, eds., Risk and return in finance, Vol. 1 (Ballinger, Cambridge, MA) 189–219.
Rubinstein, M., 1976, The valuation of uncertain income streams and the pricing of options, Bell Journal of Economics 7, 407–425.
Shanken, J., 1982a, An analysis of the traditional risk–return model, Ph.D. dissertation (Carnegie-Mellon University, Pittsburgh, PA).
Shanken, J., 1982b, The arbitrage pricing theory: Is it testable?, Journal of Finance 37, 1129–1140.
Shanken, J., 1985a, Multivariate tests of the zero-beta CAPM, Journal of Financial Economics 14, 327–348.
Shanken, J., 1985b, Multibeta CAPM or Equilibrium-APT? A reply to Dybvig and Ross, Journal of Finance 40, 1189–1196.
Shanken, J., 1986a, On the exclusion of assets from tests of the mean variance efficiency of the market portfolio: An extension, Journal of Finance 41, forthcoming.
Shanken, J., 1986b, Testing portfolio efficiency when the zero-beta rate is unknown: A note, Journal of Finance 41, 269–276.
Shanken, J., 1986c, A posterior-odds ratio approach to testing portfolio efficiency, Working paper (Graduate School of Management, University of Rochester, Rochester, NY).
Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.
Sharpe, W. 1977, The capital asset pricing model: A 'multibeta' interpretation, in: H. Levy and M. Sarnat, eds., Financial decision making under uncertainty (Academic Press, New York).
Stambaugh, R., 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, Journal of Financial Economics 10, 237–268.