Journal of Financial Economics 4 (1977) 129–176. © North-Holland Publishing Company

A CRITIQUE OF THE ASSET PRICING THEORY'S TESTS

Part I: On Past and Potential Testability of the Theory*

Richard ROLL**

University of California, Los Angeles, CA 90024, U.S.A.

Received June 1976, revised version received October 1976

Testing the two-parameter asset pricing theory is difficult (and currently infeasible). Due to a mathematical equivalence between the individual return/'beta' linearity relation and the market portfolio's mean-variance efficiency, any valid test presupposes complete knowledge of the true market portfolio's composition. This implies, inter alia, that every individual asset must be included in a correct test. Errors of inference inducible by incomplete tests are discussed and some ambiguities in published tests are explained.

> If the horn honks and the mechanic concludes that the whole electrical system is working, he is in deep trouble . . .
>
> Pirsig (1974)

1. Introduction and summary

The two-parameter asset pricing theory is testable *in principle*; but arguments are given here that: (a) No correct and unambiguous test of the theory has appeared in the literature, and (b) there is practically no possibility that such a

*This is Part I of a three-part study. Parts II and III are summarized in the introduction here, but will appear in later issues. A copy of the complete paper can be obtained by writing the author at: Graduate School of Management, University of California, Los Angeles, CA 90024, USA.

**This paper was written while the author was at the Centre d'Enseignement Supérieur des Affaires, France. Eugene Fama, Michael C. Jensen, John B. Long, Jr., Stephen Ross and Bruno H. Solnik provided many useful comments and Patricia Porter provided excellent secretarial service. While the paper was being written, Fama pointed out that his new book (1976) contains some of the same analysis and conclusions. New papers by Stephen Ross (forthcoming) and John B. Long (1976) contain results emphasized, and formerly believed to have been discovered, here. The reader will be able to verify, however, that most of this material is non-redundant.

To the authors criticised here: these papers were singled out because they are the best and most widely read on the subject. I have written some papers in this area too and have taught the subject to a number of unsuspecting students. So, the absence of detailed self-criticism should be attributed to the greater importance of the other papers and does not imply any personal prescience. None was present.

test can be accomplished in the future. This broad indictment of one of the three fundamental paradigms of modern finance will undoubtedly be greeted by my colleagues, as it was by me, with scepticism and consternation. The purpose of this paper is to eliminate the scepticism. (No relief is offered for the consternation.)

Here are the paper's conclusions:

(1) There is only a single testable hypothesis associated with the generalized two-parameter asset pricing model of Black (1972). This hypothesis is: 'the market portfolio is mean-variance efficient'.
(2) All other so-called implications of the model, the best known being the linearity relation between expected return and 'beta', follow from the market portfolio's efficiency and are not independently testable. There is an 'if and only if' relation between return/beta linearity and market portfolio mean-variance efficiency.
(3) In *any* sample of observations on individual returns, regardless of the generating process, there will always be an infinite number of ex-post mean-variance efficient portfolios. For each one, the sample 'betas' calculated between it and individual assets will be exactly linearly related to the individual sample mean returns. In other words, if the betas are calculated against such a portfolio, they will satisfy the linearity relation *exactly* whether or not the true market portfolio is mean-variance efficient. (The same properties also hold ex ante, of course). These results are implied in earlier literature [e.g., Ross (1972)], but I do not believe that their full consequences have been adequately explored previously. Some of these consequences are:
(4) The theory is not testable unless the exact composition of the true market portfolio is known and used in the tests. This implies that the theory is not testable unless *all* individual assets are included in the sample.
(5) Using a proxy for the market portfolio is subject to two difficulties. First the proxy itself might be mean-variance efficient even when the true market portfolio is not. This is a real danger since *every* sample will display efficient portfolios that satisfy perfectly all of the theory's implications. For example, suppose there exist 1000 assets but only 500 are used in the sample. For the sample, there will exist well-diversified portfolios of the 500 assets that seem to be reasonable proxies for the market and for which observed returns are exactly linearily related cross-sectionally to observed betas. On the other hand, the chosen proxy may turn out to be inefficient; but obviously, this alone implies nothing about the true market portfolio's efficiency. Furthermore, most reasonable proxies will be very highly correlated with each other and with the true market whether or not they are mean-variance efficient. This high correlation will make it seem that the exact composition is unimportant, whereas it can cause quite different inferences.

(6) As a case in point, a detailed discussion is provided of the papers by Fama and MacBeth (1973), Black, Jensen, and Scholes (1972) and Blume and Friend (1973), in the context of their rejection of the Sharpe–Lintner model. It is shown that their tests results are fully compatible with the Sharpe–Lintner model and a specification error in the measured 'market' portfolio. A misspecification would have created bias *and* non-stationarity in the fitted cross-sectional risk/return lines even if there were a constant riskless return. For the Black, Jensen and Scholes data, for example, there was a mean-variance efficient 'market' proxy that supported the Sharpe–Lintner model *perfectly* and that had a correlation of 0.895 with the market proxy actually employed. However, it cannot be ascertained without further analysis whether this other portfolio satisfied all the requirements of a good market proxy (such as positive proportions invested in all assets).

The market portfolio identification problem constitutes a severe limitation to the testability of the two-parameter theory. No two investigators who disagree on the market's measured composition can be made to agree on the theory's test results. However, suppose that advances in electronic monitoring of human capital and other non-traded assets make the market portfolio's true composition knowable; or more realistically, suppose a given composition is just agreed upon by everyone relevant. How should the mean-variance efficiency of this known composition portfolio be tested? Part II of the paper (to appear in a later issue) investigates the peculiar econometric problems associated with such testing, viz.:

(7) A direct test of the proxy's mean-variance efficiency is difficult computationally because the full sample covariance matrix of individual returns must be inverted and statistically because the sampling distribution of the efficient set is generally unknown. Some possible solutions to the statistical problems are presented. They include tests based on the fact that the market portfolio must have positive proportions invested in all assets; large sample distribution-free tests; and tests based on the sampling distribution of the efficient set assuming Gaussian returns.
(8) Testing for the proxy's efficiency by using the return/beta linearity relation also poses empirical difficulties:
    (a) The two-parameter theory does not make a prediction about parameter values but only about the form (linear) of the cross-sectional relation. Thus, econometric procedures designed to obtain accurate parameter estimates are not very useful.
    (b) Specifically, the widely-used portfolio grouping procedure can support the theory even when it is false. This is because individual asset deviations from exact linearity can cancel out in the formation of

portfolios. (Such deviations are not necessarily related to betas.) Some simulated data given by Miller and Scholes (1972) were used as an example of such an occurrence. Deviations in these data were known to be related to generating process asymmetry which would not have been detectable in grouped observations.

(9) Several others tests are proposed for the linearity relations. These include:
 (a) An Aitken-type procedure that gives unbiased cross-sectional tests with individual assets, and
 (b) a procedure that exploits asymptotic exact linearity by measuring the rate of decrease of cross-sectional residual variance with respect to increasing time-series sample size.

In Part III of the paper (to appear in a future issue), some of the common uses of the two-parameter theory are called into question:

(10) Deviations from the return/beta linearity relation are frequently linked with some other phenomenon. The validity of such linkages is criticised using the Jensen measure of portfolio performance as an example. If the 'market' proxy used in the calculations is exactly (not significantly different from) ex-post efficient, *all* of the individual Jensen performance measures gross of expenses will be identically (not significantly different from) zero. They can be (significantly) non-zero only if the proxy market portfolio is (significantly) not efficient. But if the proxy market portfolio is not efficient, what is the justification for using it as a benchmark in performance evaluation?

(11) The beta itself is criticised as a risk measure on two grounds: first, that it will always be (significantly) positively related to observed average individual returns if the market index is on (not significantly off) the positively sloped section of the ex-post efficient frontier, *regardless of investors' attitudes toward risk*; and second, that it depends, non-monotonically, on the particular market proxy used. About the second point: if two investors happen to choose two different 'market' portfolios, both of which are mean-variance efficient, the same security might have a beta of 1.5 for the first investor and 0.5 for the second. This is intuitively obvious since beta is supposed to be a *relative* measure of risk. But less obvious is the fact that if both investors increase the proportions this security represents in their 'markets', its beta will change and it can increase for one investor and decrease for another.

An appendix to this part (I) contains a compact analytic derivation of the efficient set propositions and includes a few original results (e.g., identity of the efficient portfolio that maximizes cross-sectional variation in beta).

2. The testable feature of asset pricing theory and the features that have been 'tested'

2.1. *Efficient set mathematics*

We should begin any quantitative enquiry by setting forth the relationships that are mutually and logically equivalent. The mathematics of the mean-variance efficient set serves just such a purpose, for it exposes several logically-equivalent relations among mean returns and covariances (which are the building blocks of the asset pricing theory). The mathematics is mostly available elsewhere [see, e.g., Sharpe (1970), Merton (1972), Black (1972), Szegö (1975), Fama (1976), Long (1976)] and a compact statement of all the familiar results plus some new ones is provided in the appendix.

The efficient set mathematics has been discussed most usually in terms of *ex-ante* returns and covariances. To emphasize the purely mathematical nature of the results, however, I should like to state it in terms of an observed sample of returns on N assets. No presumption is made about the population that generated this sample. It can be any probability law imaginable. Furthermore, no mention need be made about equilibrium, risk aversion, homogeneous anticipations, or anything else like that. There are only two assumptions:

(A.1) The sample product–moment covariance matrix, V, is non-singular.

(A.2) At least one asset had a different sample mean return from others.

These are very weak assumptions. (A.1) simply rules out assets whose returns were constant during every period in the sample and it excludes any pair of linear combinations of assets that were perfectly correlated during the sample period. (A.2) merely requires some sample variation in the critical variable of interest. After all, it is cross-sectional variation in the mean return which asset pricing theory strives to explain.

Given the sample covariance matrix and the arithmetic sample mean returns (expressed as an $N \times 1$ column vector R), the sample frontier of efficient ex-post portfolios can be easily obtained. This frontier enumerates all the portfolios that had minimal sample variance for each given level of mean sample return. Suppose we choose one of these portfolios, say portfolio m, with sample return r_m, which lies on the positively-sloped part of the efficient frontier. (That is, there is no other portfolio with the same sample variance that had a higher mean return.) Then the following statements are true:

(S.1) There exists a unique portfolio, denoted z, that had a correlation of zero with m during the sample period and that lies on the negatively-sloped segment of the sample efficient frontier; this implies that the sample return of m was greater than that of z, $r_m > r_z$. (For a formal proof, see the appendix, Corollary 3.)

(S.2) For any arbitrary asset or portfolio, say j, the sample mean return is equal to a weighted average of r_z and r_m where the weight of m is *exactly* the sample linear regression slope coefficient of j on m, i.e.,

$$r_j \equiv (1-\beta_j)r_z + \beta_j r_m, \qquad \text{for all } j, \tag{1}$$

where

$$\beta_j \equiv \frac{\text{sample covariance of } j \text{ and } m}{\text{sample variance of } m}$$

(Proof: Appendix, Corollary 6).

Statement (S.1) is related to the following facts:

(S.3) Every portfolio on the positively-sloped segment of the sample efficient set was positively correlated with every other one (Corollary 4).

(S.4) Every sample efficient portfolio except the global minimum sample variance portfolio has an orthogonal portfolio with finite mean return (Corollary 3).

It is easy to see that (S.3) and (S.4) imply that $r_m > r_z$ because we have chosen m to lie in the positively-sloped segment of the sample efficient frontier.

Proposition S.2, on the other hand follows from:

(S.5) The investment proportions of any sample efficient portfolio can be expressed as a weighted average of the proportions in any other two sample efficient portfolios whose means are different (Corollary 5).

Given (S.5), it is a simple matter to prove (1); see the appendix or, e.g., Black (1972, p. 450). In fact, a more general proposition than (1) follows readily from (S.5). Let A and B be *any* two arbitrary sample efficient portfolios, ex-post correlated or not, but with different sample mean returns. Then:

(S.6) The mean return on any arbitrary asset, j, is given exactly by

$$r_j \equiv (1-\beta'_j)r_A + \beta'_j r_B, \qquad \text{for all } j \tag{2}$$

(Corollary 6.A).

In eq. (2) β'_j is the multivariate *sample* slope coefficient for B from the regression of r_j on r_A and r_B. Furthermore, this regression coefficient has a simple form,

$$\beta'_j = (\sigma_{jB} - \sigma_{AB})/(\sigma_{BB} - \sigma_{AB}),$$

where σ_{ik} is the sample covariance of i and k. It is easy to see that (1) is merely a special case of (2) that obtains when A is chosen to be B's orthogonal sample portfolio.

Expression (2) can be considered a logical equivalent to assumptions (A.1) and (A.2). In other words, given an observed non-singular sample covariance matrix and at least two different sample mean returns, every observed mean return has exactly the relation shown in (2). Equivalently, every observed sample 'beta' conforms exactly to the rearrangement of (1),

$$\beta_j \equiv (r_j - r_z)/(r_m - r_z). \tag{3}$$

A converse statement is also true:

(S.7) Let β be the $(N \times 1)$ column vector of simple regression slope coefficients computed between individual assets and some portfolio m. Then the vector of mean returns R is an exact linear function of the vector β only if m is a sample efficient portfolio; i.e., in general,

$$R \equiv r_z \iota + (r_m - r_z)\beta, \tag{4}$$

if and only if r_m is ex-post efficient [r_z is the mean return on m's corresponding efficient orthogonal portfolio and ι is the unit vector, see Ross (1972, 1973)].

It follows that mean returns are *not* exact linear functions of betas when m is not efficient. This does not imply that mean returns are necessarily related to non-linear functions of beta. They are just not *exactly* linear. For example, the relation

$$R = \alpha + g\beta$$

is a possibility if m is inefficient; where α is a vector whose elements are non-constant but are unrelated to the elements of β, and g is a scalar constant.

Before going on to the theory of asset pricing, it is well to emphasize the nature of these mathematical relations. Identity symbols have been used in (1) through (4) because they really are identities. Given the choice of m as ex-post efficient, these expressions hold exactly. They do not, therefore, provide any information about the state of nature or about the process that generated the sample. The underlying probability law might be anything and the relations above would always be observed ex-post. This has relevant implications for testing the asset pricing theory, as we shall see.

2.2. *A review of some asset pricing theory tests*

Three widely-quoted empirical papers on asset pricing theory are Black, Jensen and Scholes (1972), Blume and Friend (1973), and Fama and MacBeth (1973).[1] Let us examine what they said they were testing: The statement in Fama and MacBeth is very clear. They refer to a portfolio m which is on the ex-ante efficient frontier as seen by a single investor. This leads to the derivation of an equation identical to (1) but with investors' subjective parameters instead of sample parameters. The resulting equation [Fama–MacBeth (1973, p. 610)]

> '... has three testable implications: (C.1) the relationship between the expected returns on a security and its risk in any efficient portfolio m is linear. (C.2) β_i is a complete measure of the risk of security i in the efficient portfolio m; no other measure of the risk of i appears in (6) [eq. (1) here]. (C.3) in a market of risk-averse investors, higher risk should be associated with higher return; that is $E(\tilde{R}_m)-E(\tilde{R}_0) > 0$.' [$R_0$ is the same as r_z here.]

Given that the word 'risk' has replaced the parameter β, we have already seen that Fama and MacBeth's (C.1), (C.2), and (C.3) are simply implications of the fact that m is assumed ex-ante efficient.[2] If m is known to be efficient, these relations are not independently testable. They are tautological. When m is efficient, the expected return *must* be linear in β and $E(\tilde{R}_m)$ must exceed $E(\tilde{R}_0)$. Incidentally, given the assumption that m is efficient, their last inequality has nothing to do with risk aversion. It is purely the mathematical implication of the assumption about m and the definition of β. It is totally independent of investor preferences since it follows from the mathematical property (S.1). Conversely, if Fama and MacBeth's (C.1) is true, and ex-ante β is an exact linear function of ex-ante expected return, then m must be ex-ante mean-variance efficient.

It is clear from the authors' discussion that they are aware of these internal

[1]There are other interesting papers containing similar tests, e.g., Petit and Westerfield (1974) and Modigliani, Pogue, Scholes and Solnik (1972). Petit and Westerfield's test of the asset pricing theory is actually identical to Black, Jensen and Scholes' although Petit and Westerfield seem to deny this. Modigliani, Pogue, Scholes and Solnik carry out a similar test for eight *different* European stock markets. Palacios (1973) and Rosa (1975) present detailed investigations for Spain and France respectively. Roll (1973) gives a comparative test of the asset pricing theory and the optimal growth model using the same methodology. See also Fama and MacBeth (1974a). Fama and MacBeth (1974b) investigate the extension of asset pricing theory into a multi-period context. Roll and Solnik (1975) apply the methodology to exchange rates.

[2](C.2), the statement that no other risk measure except β is important, presupposes that β measures risk. Whether it measures risk or not, however, it is the only variable on the right side of (1). (r_z and r_m are constant cross-sectionally.) Thus, it is the only cross-sectional explanatory variable of any kind.

relations. For example, on page 609 they state, '... there are conditions on expected returns that are implied by the fact that in a two-parameter world investors hold efficient portfolios.' But on page 610 they make a statement inconsistent with the facts and with their own knowledge of the mathematics: 'To test conditions (C.1)–(C.3) we must identify some efficient portfolio m.' Of course, if m is identified as efficient, there is no need to test (C.1)–(C.3). (See also the self-contradictory second paragraph on p. 614.)

But there are testable hypotheses in the Fama–MacBeth paper. The hypotheses really are:

(H.1) Investors regard as *optimal* those particular investment portfolios that are mean-variance efficient.

Assuming identical probability assessments by all investors, this hypothesis leads to:

(H.2) The 'market portfolio' is ex-ante efficient.

The 'market portfolio' is defined as a value-weighted combination of all assets (p. 611). Fama and MacBeth credit Black (1972) with deducing (H.2) given (H.1) and given homogeneous investor expectations. The Black proof is quite simple: Since all investors have identical beliefs and hold efficient portfolios, every investor holds a linear combination of two arbitrary efficient portfolios. Since the market portfolio is by construction a linear combination of the portfolios of individual investors, it is also a linear combination of these two efficient portfolios and is therefore also efficient [because the linear combination of any two efficient portfolios is also efficient by the basic mathematical property of the efficient set, (S.5)]. Interestingly, Black states that Lintner (1969) '... has shown that removing [the] assumption [of homogeneous anticipations] does not change the structure of capital asset prices in any significant way' (p. 445). Nevertheless, Black's proof of the market portfolio's efficiency does require homogeneity. This might be relaxed in a more general (and as yet unknown) proof; but Fama (1976, ch. 7) has argued that, in fact, no equilibrium model with non-homogeneous anticipations is testable.[3]

[3]On page 447 at the beginning of his discussion of efficient portfolios, Black makes a statement that seems to be in conflict with the results here. He claims that Cass and Stiglitz (1970) have shown

> '... that if the returns on securities are not assumed to be joint normal, but are allowed to be arbitrary, then the set of efficient portfolios can be written as a weighted combination of two basic portfolios *only* for a special class of utility functions' (italics added).

It is clear from his subsequent discussion that Black was referring to *mean-variance* efficient portfolios. Thus, his statement is false. Efficient mean-variance portfolios can *always* be constructed as a 'weighted combination of two basic portfolios'. Furthermore Cass and Stiglitz never claimed the contrary. What they did was to enumerate the set of utility functions for

In both the Black paper and the Fama–MacBeth paper there exists a bit of unfortunate wording about the efficient set mathematics and about optimal investment choices. At first, it might seem that the resulting confusion would be only minor. But when it comes to empirical testing and to specifying exactly those relations that are empirically rejectable and are valid scientific hypotheses, this possible confusion is of great significance.

The only viable (i.e., rejectable) hypothesis that we have so far been able to uncover is (H.2), the market portfolio is mean-variance efficient.[4] The assumptions which are sufficient for this result are rather strong: Perfect capital markets, homogeneous anticipations, two-parameter probability distributions of returns. But there is also another assumption that has received little attention in the literature: namely, the market portfolio must be identifiable.

This last assumption is very important when we consider that there will always be *some* portfolio which is ex-post efficient and will bring about exact observed linearity among ex-post sample mean returns and ex-post sample betas. If we do not know the composition of the market portfolio, we might by chance select a proxy that is close to mean-variance efficient. In fact, it may be hard to find a highly-diversified portfolio that is sufficiently far inside the ex-post efficient frontier to permit the detection of statistically significant departures from mean return/beta exact linearity. We will return to this point later. First, let us see what some of the other papers have been testing.

The widely-quoted paper by Black, Jensen, and Scholes (1972) makes no mention of the possible efficiency of the market portfolio and its importance for the linear relation between return and 'beta'. In fact, however, the authors modestly claim that their '. . . main emphasis has been to test the strict traditional form of the asset pricing model' (p. 113), by which they mean the original Sharpe (1964), Lintner (1965) model similar to (1), but with r_z replaced by a 'riskless' return. [This model results from the asset pricing theory assumptions listed above, and used by Black to derive (1), plus the extra assumption that investors can borrow and lend as much as they like at a riskless interest rate.] Black, Jensen, and Scholes explicitly deny that they have provided tests of any other hypothesis. However, the Black model is clearly in the backs of their minds and on page 81 they even go so far as to provide a historical glimpse of Black's theoretical progress by asserting that he 'was able' to derive his model 'after we had observed this phenomena' (that mean returns are linearly related to calculated beta coefficients but supposedly with a different slope and intercept than those implied by the Sharpe–Lintner theory). The graphs plotted by Black, Jensen, and Scholes appear to portray a very linear mean return/beta relation over long sample periods. Unlike Fama and MacBeth, however, no formal test of linearity is provided. (It must not have seemed necessary given the authors' goal.) Thus, no formal information is given on the possible efficiency of the measured market portfolio, nor about the hypothesis (H.2).

A direct test of linearity was provided by one of the authors [Jensen (1972a, 1972b)], who, using the same data as those used by Black, Jensen and Scholes, presented results from a regression similar to the one later computed by Fama and MacBeth. In fact, the fitted equations are nearly identical in form but the measurement methods were somewhat different and the 'market' portfolios used in calculating betas were different.[5] This was evidently sufficient to create some disparity between the two sets of results. We cannot ascertain the exact extent of the disparity because the sub-periods reported in the two papers were not identical. At least the signs of coefficients on squared beta terms were in agreement, being negative during the longer sample periods. The statistical significance of these negative signs is less clear. For example, the coefficient given by Fama–MacBeth for the squared beta term during 1946–55 was -0.0076 (p. 623). The same coefficient given by Jensen for an overlapping period, July 1948 to March 1957, was -0.0055. The associated t-statistics were far apart, however, Fama–MacBeth's was -2.16 whereas Jensen's was only -0.524.[6] This difference may very well be due to Fama and MacBeth's presumably more powerful test but there is no way to be sure without a complete replication.

It might be worthwhile carrying out such a replication because the linearity is directly related to the market portfolio's efficiency. We can already be sure that the 'market' portfolios used by Jensen and by Fama and MacBeth did not lie exactly on the sample efficient frontier. If they had been exactly efficient, the relation between the mean return vector and the vector of sample betas would have been exactly linear and it was not.[7] But it is not necessary for the basic hypothesis (H.2) (the market portfolio is *ex-ante* efficient) that the observed market portfolio be exactly *ex-post* efficient in every period. It only needs to be efficient over 'sufficiently' long periods. Now both Jensen and Fama–MacBeth find no significant non-linearity over the longest sample period nor do they find

which all investors would construct their *optimal* portfolios as a weighted average of two basic portfolios. Under a restrictive set of preferences, each investor would regard a mean-variance efficient portfolio as optimal; but as Cass and Stiglitz show, there are other investor preferences which would lead to 'separation' (or the choice of an optimal portfolio which is a weighted combination of two others), under which the optimal portfolio is *not* mean-variance efficient. [See also Hakansson (1969), Jacob (1970) and Ross (1976). The last reference gives separation results for probability distributions instead of utility functions.]

[4]Fama and MacBeth also provide an ingenious time-series test of market competition, given the hypothesis (H.2). However, this part of their paper is about a different set of hypotheses than our subject here.

[5]Cf. Jensen (1972b, pp. 385–388) with Fama and MacBeth (1973, pp. 615–617).

[6]Jensen reported a t-statistic with respect to a non-zero theoretical value which he derived from Merton's (1973) continuous time model. The number above is his t-statistic for a hypothesized coefficient of zero.

[7]Actually, this is not entirely true for Fama–MacBeth since they did not use *concurrent* sample mean returns and betas. Thus, some deviation from linearity might have been observed in their results, even if the 'market' portfolio had been exactly sample efficient, because the sample betas might not have been stationary. Part II of this paper will examine the importance for testing the basic hypothesis of attempting to purge measurement errors from sample betas. (This was the reason Fama and MacBeth did not use concurrent observations.)

any significance for non-beta 'measures of risk' such as standard deviation. Although this is consistent with their market portfolio proxies having been sample efficient over the long term, it is also consistent with their proxies having been inefficient (as we will soon see).

Interestingly, Fama and MacBeth offer a possible explanation for the significance of non-linear beta terms during some sample periods: viz., they suggest that there are omitted variables from the theory for which the non-linear terms act as proxies. Of course, their results are also consistent with the simpler explanation that the Fama–MacBeth 'market' portfolio was not exactly ex-post efficient in every sub-period. This alone implies that non-linear terms could be significant. It is also true that the non-exact ex-post efficiency of the market might induce significance in individual standard deviations (i.e., in non-portfolio risk measures).

2.3. *Tests of the Sharpe–Lintner model*

Let us now turn to an ancillary examination of the evidence offered by Black, Jensen and Scholes and by others against the original Sharpe–Lintner theory. It will be useful to have the following supplementary results from the efficient set mathematics.

Given the following additional assumption:

(A.3) There exists an asset whose return was a constant, r_F, during the sample period.

Then:

(S.7) The sample efficient set (in the mean-variance space) is a parabola with a tangent on the return axis at r_F.

(S.8) Suppose we denote the 'risky efficient set' as the ensemble of portfolios with minimum variance *excluding* asset F. Then results (S.1) through (S.6) still hold for the portfolios composing this 'risky efficient set'.

In particular, for any ex post portfolio composed entirely of risky assets and lying on the positively-sloped segment of the 'risky efficient set', sample mean returns on all assets are exact linear functions of sample betas as portrayed by eq. (1); sample mean r_z in (1) is the return on a portfolio lying on the negatively-sloped segment of the risky efficient set whose return was uncorrelated with the return on m during the sample period.

In other words, we have the familiar diagram shown in fig. 1, where m, m^* and z are all portfolios composed of risky assets only and are all on the sample risky efficient boundary. The portfolio m^* is the sample 'tangent' portfolio

whose return, according to Corollary 3.A of the appendix, is determined by the riskless return, r_F, and of some simple functions of the mean return vector of individual assets and of the sample covariance matrix. Portfolio z has been chosen to have zero sample correlation with portfolio m, a feat that is always possible for any position of m.

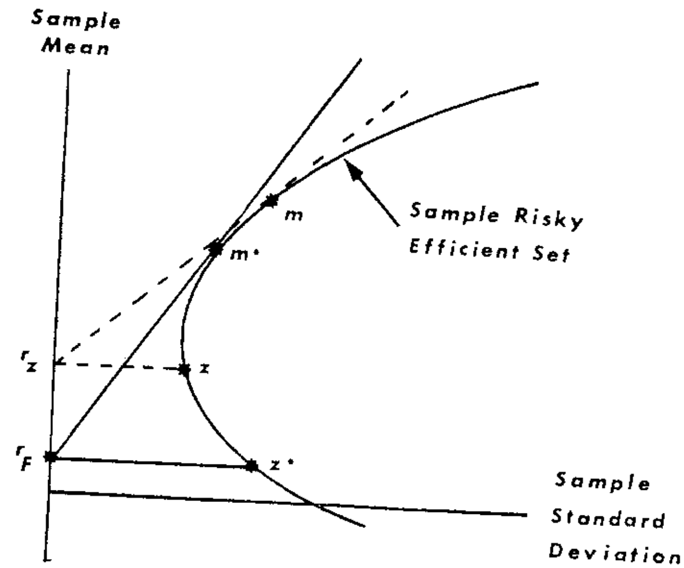

Fig. 1. The sample risky efficient set, sample market proxy and sample zero-beta portfolio.

Now let us consider the sample linearity property between mean return and beta. First, if portfolio m is used to compute beta, we must have the mathematical result already found,

$$r_j = r_z + (r_m - r_z)\beta_j, \qquad \text{for all } j. \tag{5a}$$

On the other hand we might choose portfolio m^* to compute the betas. This will produce a different set of sample betas because m and m^* are not perfectly correlated. Denoting these second betas by β_j^*, we must have also another linearity relation,

$$r_j = r_F + (r_{m^*} - r_F)\beta_j^*, \qquad \text{for all } j. \tag{5b}$$

What about z^*, the *risky* efficient portfolio that is uncorrelated with m^*? Since it too must be usable in yet another linearity relation with the β^*'s, it must have the same mean return as r_F. In fact, it is quite easy to prove that this is so.[8] Furthermore, since there is an infinite number of efficient risky portfolios along the positively-sloped boundary, there is an infinite number of these linearity relations, all equally satisfied exactly (but all with different beta vectors). In particular, r_z and r_m would have their own β_z^* and β_m^* in (5b) and would satisfy the second linearity relation above. *Note that β_z^* must be non-zero* because efficient orthogonal portfolios are unique. Thus, even though m and z are uncorrelated, m^* and z must be correlated. Furthermore, although m^*'s orthogonal portfolio is constrained to have the same sample return as the riskless return, there is no such restriction on portfolio z. Depending on the relative positions of m and m^*, r_z can be greater or less than r_F.[9]

Armed with these purely logical results which are true for any sample satisfying assumptions (A.1), (A.2), and (A.3), let us turn to the published tests of the original Sharpe–Lintner theory. First, what are the principal hypotheses of this theory? They are:

(H.3) Investors can borrow or lend at the riskless rate, r_F.

(H.1) (Same as before.) They consider that mean-variance efficient portfolios are optimal.

Thus, each individual would compose his portfolio of the riskless asset F and his subjective tangent portfolio m^*. If investors had homogeneous probability assessments, they would all have the same tangent portfolio. Thus:

(H.4) The ex-ante efficient tangent portfolio is the market portfolio of all assets.

Of course, since there seems to be little possibility of rejecting (H.3) or even (H.1) with direct information, we are left with (H.4) as the testable hypothesis.

Black, Jensen, and Scholes rejected the Sharpe–Lintner theory as a result of the following 'test': First, a 'market' portfolio was chosen and sample betas were calculated via a procedure designed carefully to remove measurement error. Then, the cross-sectional mean return/beta linearity relation was estimated in the form

$$r_j - r_F = \hat{\gamma}_0 + \hat{\gamma}_1 \beta_j + \hat{\varepsilon}_j, \tag{6}$$

where $\hat{\varepsilon}_j$ is the estimated residual.

The basic results were that $\hat{\gamma}_0$ exceeded zero, that $\hat{\gamma}_1$ was less than $r_m - r_F$, and that $\hat{\gamma}_0$ was highly variable from one sub-period to another. This led them to reject hypothesis (H.4).

Given our preceding analysis of the efficient set mathematics, we are entitled to be suspicious of their conclusion. Unless Black, Jensen, and Scholes were successful in choosing m^* (in fig. 1) for their market portfolio, their results are fully compatible with the original Sharpe–Lintner model. This is readily seen in the two ex-post equations (5a) and (5b). Suppose, for example, that their 'market' portfolio was really m rather than m^*. Then solve (5b) for $j = z$ and use this to replace r_z in (5a). The result is

$$r_j - r_F = \beta_z^*(r_{m^*} - r_F) + [r_m - r_F - \beta_z^*(r_{m^*} - r_F)]\beta_j.$$

Now we have already seen that β_z^* *must* be non-zero[10] and that the Sharpe–Lintner theory implies $r_{m^*} > r_F$ ex-ante (and ex-post asymptotically with increasing time-series sample size and with stationarity). Thus, the estimated coefficient $\hat{\gamma}_0$ in (6) is seen to be equal to $\beta_z^*(r_{m^*} - r_F)$ given validity of the Sharpe–Lintner theory and efficiency of the measured 'market' portfolio m. Furthermore, since Black, Jensen, and Scholes' constant term, $\hat{\gamma}_0$, is a function of the true tangent portfolio, m^*, whose return is a random variable, we should *expect* to see an intertemporal variation in their constant term even when r_F is a fixed number. There will be an offsetting variation in $\hat{\gamma}_1$, the slope of (6).

No calculations were made by Black, Jensen, and Scholes to ascertain whether their market portfolio was in fact close (statistically) to the ex-post tangent portfolio over long periods. But we can be absolutely certain that it was not! Why? Because the pure mathematics of the efficient set tell us that the relation (5b) is *exactly* satisfied in every ex-post sample for which assumptions (A.1), (A.2), and (A.3) were true. Assumption (A.3), a constant return existed, was indeed approximately satisfied during all their sample periods. Thus, we can be sure that for each sample period there was a portfolio m^* whose associated sample beta vector was a linear function of the mean return vector and for which the coefficients of (6) satisfied $\hat{\gamma}_0 = 0$. Since the sample beta vector calculated by Black, Jensen, and Scholes differed significantly from the vector that satisfied (5b) and did not approach that vector as the time series sample size increased,

[8] By Corollary 3.A of the appendix, a tangent drawn to any point p on the efficient frontier intersects the return axis at the level of the mean return on p's orthogonal portfolio. Since m^* is, by definition, located at the tangency drawn from r_F, we must have $r_F = r_{z^*}$. In the mean-*variance* space, there is a little-known analogous property: a line from any point p on the efficient frontier that passes through the global minimum variance position also intersects the return axis at the level of p's orthogonal portfolio. In general, if p is efficient, every portfolio orthogonal to p will have return $r_z = (a - br_p)/(b - cr_p)$ where a, b, and c are the 'efficient set constants' (see appendix, Definition A.9).

[9] r_z will exceed r_F if and only if $r_m > r_{m^*}$.

[10] As a reminder, $\beta_z{}^*$ is defined the sample analog of $\mathrm{Cov}(r_z, r_{m^*})/\mathrm{Var}(r_{m^*})$; i.e., the beta for the *proxy* zero-beta portfolio (z) computed against the *true* market portfolio (m^*). N.B.: This 'zero-beta' portfolio has a beta of zero only against m. It has a non-zero beta against *all* other efficient portfolios.

we know that their 'market' portfolio was not statistically close to the tangent portfolio.[11]

On the other hand, one should note also that an ex-post verification of (5b) would not have implied that (A.3) was valid. In other words, the purely mathematical proposition (5b) can be observed even if investors are totally prohibited from access to a riskless asset. Consider the following scenario as an example: Investors are totally excluded from riskless borrowing and lending. Nevertheless, the government publishes each period a number called the riskless rate of interest. It follows that each period there will exist some portfolio m^* whose associated betas along with the published number exactly satisfy (5b). This observed m^* will not necessarily be the market portfolio, of course. How can we distinguish empirically this scenario from the Sharpe–Lintner model where riskless borrowing and lending is fully permissible? We cannot do so from the linearity relation (6) *alone*. We must have independent information on the true market portfolio's identity. Only then can we determine whether this particular portfolio is or is not the tangent portfolio and thereby distinguish between the two scenarios.

In summary, even if Black, Jensen, and Scholes had been unable to reject the hypothesis that $\hat{\gamma}_0$ equals zero and that there is a linear beta/mean return tradeoff, they would not have been entitled to support the Sharpe–Lintner theory. They shouldn't have rejected the theory either upon not finding $\hat{\gamma}_0 = 0$. Their test is simply without rejecting power for hypothesis (H.4).

Black, Jensen, and Scholes realized that using a misspecified 'market' portfolio would result in a measured $\hat{\gamma}_0$ from (6) not equal to zero. However, they thought mistakenly that the $\hat{\gamma}_0$ would have to be constant even with the misspecification (cf. their page 115). This was a critical oversight, for it led to a professional consensus that the Sharpe–Lintner theory was false. It seems probable (at least to me) that such an opinion would have been held less widely if the market index' composition had been correctly perceived as *the* critical variable in understanding the test results; that is, if we had realized that a readjustment of the market portfolio's proportions might have reconciled the test results as well to Sharpe's and Lintner's theory as to Black's.

It may occur to the reader that the Black, Jensen, and Scholes paper tested a joint hypothesis: the Sharpe–Lintner theory *and* the hypothesis that the portfolio they used as the 'market' proxy was the true market portfolio. This joint hypothesis was indeed tested and it was rejected. We can conclude therefrom that either

(a) the Sharpe–Lintner theory is false, or
(b) the portfolio used by Black, Jensen, and Scholes was not the true market portfolio, or
(c) both (a) and (b).

[11]In the next section their results are used to actually calculate the mean and variance of this sample tangent portfolio (see table 1).

There lies the trouble with joint hypotheses. One never knows what to conclude. Indeed, it would be possible to construct a joint hypothesis to reconcile *any* individual hypothesis to *any* empirical observation. In the present case, fortunately, there is at least the information that (b) is false. The portfolio used by Black, Jensen, and Scholes was certainly not the true market portfolio; but whether it was statistically close to the true market portfolio [thus leading to conclusion (a)] or whether it was closer than the Sharpe–Lintner assumptions are to reality is beyond our capacity to know.

As for the other papers, Fama and MacBeth present tests of the Sharpe–Lintner theory which are similar in spirit, form, and conclusion to those of Black, Jensen, and Scholes (see their section VI, pages 630-633). The explicit stated hypothesis is that $\hat{\gamma}_0$ from (6) be insignificantly different from zero.[12] Their conclusion is that '. . . the most efficient tests of the S–L (Sharpe–Lintner) hypothesis . . . support the negative conclusions of others' (p. 633), (because $\hat{\gamma}_0$ was found to be significantly different from zero). Probably because of the nature of their methodology, Fama and MacBeth, unlike Black, Jensen, and Scholes, did not consider the variability of $\hat{\gamma}_0$ as an additional piece of condemning evidence against the Sharpe–Lintner hypothesis. Thus, they did not draw Black, Jensen, and Scholes' second erroneous inference.[13]

Blume and Friend (1973) provide an equivalent set of empirical results but interpret them quite differently. They begin by explicitly stating the Black model [essentially eq. (1)], and they take a similar tack in asserting that the observed zero-beta return, r_z, must equal the riskless rate, r_F,[14] in order for the Sharpe–Lintner hypothesis to be supported. They also find that the observed estimate of r_z is significantly different from r_F and thus reject the Sharpe–Lintner hypothesis.

They are clearly bothered by this conclusion, however, because they are convinced that a nearly riskless interest rate *did* exist. They state that: 'If returns are measured in real terms, the only risk in holding governments of appropriate maturities would stem from unexpected changes in the price level . . . [and] . . . this risk . . . has been very small' (p. 20). The second step in the argument leading to their inquietude is the conclusion that if a riskless asset *does* exist, the intertemporal variance in the zero-beta portfolio's return (i.e., in r_z) must be

[12]On page 630, they state: 'In the Sharpe–Lintner two-parameter model of equilibrium one has, in addition to conditions (C.1)–(C.3), the hypothesis that $E(\hat{\gamma}_0) = R_{ft}$.' (This is equivalent to Black, Jensen, and Scholes' model because Fama and MacBeth did not subtract R_{ft} from both sides of the linearity relation.)

[13]The first inference was $\hat{\gamma}_0$'s significant positivity. BJS stated clearly that misspecification in the market proxy portfolio could cause this. The second inference was intertemporal variation in $\hat{\gamma}_0$. They incorrectly thought that misspecification could not cause this. It was really this second crucial inference which induced them to state that the Sharpe–Lintner model was rejected by the data.

[14]This is, of course, equivalent to $\hat{\gamma}_0$ being zero in eq. (6).

zero (see their discussion on pages 22–23),[15] and they claim to have demonstrated the falsity of this empirical implication in their earlier article (1970).

This leads to an interesting conclusion: namely, the 'return generating process' corresponding to Black's model '. . . cannot explain the observed returns of all financial assets . . .. Nonetheless . . . it may be . . . adequate . . . for a subset of all financial assets, such as common stocks on the NYSE . . .. If this be so, the minimum variance zero-beta portfolio consisting only of common stocks would not be *the* zero-beta portfolio of the capital asset pricing model. However, . . . the expected return on all zero-beta assets and in particular a zero-beta portfolio consisting only of common stocks must be the same, namely the risk-free rate if such an asset exists' (pp. 22–23, italics theirs).

Blume and Friend have been quoted here at some length because their article illustrates the confusion that can arise from an insufficient understanding of efficient set mathematics. Some of their statements might very well be true; for example, that a riskless asset exists and that 'the zero-beta portfolio consisting only of common stocks would not be *the* zero-beta portfolio [of the global market]'. This last phrase might have led them to a correct understanding, for they seemed to be considering two 'market' portfolios, one consisting only of equities and one consisting of all assets in existence. Their mistake was brought about by concluding that two such distinct 'market' proxy portfolios would be associated with zero-beta (or orthogonal) portfolios having the *same* mean return and that this return must be equal to the riskless rate of interest. That conclusion is false. For example, suppose we consider the possibility that both the equities-only portfolio *and* the global all-assets portfolio are both mean-variance efficient. If these two portfolios had different mean returns and are not perfectly correlated, then the mean returns of their associated zero-beta portfolios *must* differ. Of course, if the Sharpe–Lintner hypothesis is valid, the global market portfolio's associated zero-beta portfolio would have an expected return equal to the riskless interest rate. This would imply nothing whatever about the equities-only zero-beta portfolio, the one actually used by Blume and Friend in their tests.

Blume and Friend conclude with some statements that illustrate the dangers of ad hoc theorizing. Their results supposedly (1) 'indicate a negative differential between the required rates of return on high-grade corporate bonds and on stock on a risk-adjusted basis', and (2) indicate that the supposed differential '. . . is consistent either with segmentation of markets, inadequacies of the return generating model used in this paper,[16] or a deficient short sales mechanism' (p. 32). Since corporate bonds were not included in the empirical work,[17] the first statement must be due to the observation that $\hat{\gamma}_0$ was not zero, i.e., that the

[15]Their argument is a bit clouded by being couched in the framework of the 'return generating process', but the inference above is indeed there.

[16]The generating model corresponds to Black's theory.

[17]A brief mention of bonds was contained in their note 24, page 31.

measured zero-beta return exceeded significantly the measured riskless return. This observation is perfectly consistent with *non*-segmented markets, with the Black model *or* the Sharpe–Lintner model, and with *perfect* short selling opportunities; in other words, with the precisely opposite set of circumstances to those postulated in their second statement.

One page earlier, Blume and Friend assert that '. . . the observed risk-return tradeoff would certainly have been highly non-linear in all periods' if corporate bonds had been '. . . included in the analysis' (p. 31). The evidence offered to support this is that corporate bonds indexes have measured betas close to zero, a fact that has no relevance for linearity.

If bonds had been included in the analysis, they might have been included in a new 'market' proxy and a new observed efficient set would have been obtained. The resulting linearity, or lack thereof, would have been completely dependent on the ex-post efficiency of this new market portfolio. If bonds were not made part of the market portfolio proxy, the risk-return tradeoff would still have been linear, including the bonds' returns and betas, unless the market proxy was significantly not ex-post mean-variance efficient. If it was significantly not efficient, there was no justification for its use as a proxy.

Blume and Friend conclude from their analysis: '. . . Even without allowing for the tax advantages of debt financing, the cost of bond financing may have been substantially smaller than the risk-adjusted cost of stock financing and probably smaller than the risk-adjusted cost of internal financing' (pp. 31–32). We suddenly encounter a conclusion about an important economic quantity (internal financing) upon literally its first and only mention in the entire paper, and we are told that is dearer than bond financing on a risk-adjusted basis. Still reeling, we come to the final paragraph and its assessment that '. . . in the current state of testing of the capital asset theory, the evidence points to segmentation of markets as between stocks and bonds, even though there are few legal restrictions which would have this effect' (p. 32)!

In summarizing all these empirical exercises about the Sharpe–Lintner theory, one is obliged to conclude that not a single paper contains a valid test of the theory. In fact, as Fama (1976, ch. 9) has recently concluded, there has been *no* unambiguous test of this theory in the published literature. Furthermore, it is easy to see that the prospect is dim for the ultimate achievement of such a test. We can well imagine that the critical issue of contention will always be the *identity* of the true market portfolio. Some portfolio will always occupy the Sharpe–Lintner tangency position; but whether the position will be occupied by a value-weighted average of all the assets in existence seems to be a difficult question.

In summarizing the three major papers in a broader context, two of them contained a formal test of efficiency for the market portfolio proxy. This test was the explicit inclusion of non-linear beta terms in the cross-sectional risk-return relation. Both Fama and MacBeth and Blume and Friend concluded that the non-linear terms were insignificantly different from zero. What does this tell us about the major hypothesis (H.2) of generalized asset pricing theory? In so far as we are ignorant of how close their proxy market portfolios were to the real thing, it tells us nothing at all. On the other hand, if we are willing to *assume* a close approximation between real and proxy markets, then the test results do not reject the basic hypothesis that the true market portfolio is efficient. (I shall argue in the next section, however, that such a 'good' approximation should be confronted with a strong dose of scepticism.)

Black, Jensen, and Scholes did not present a formal test of the linearity relation and thus gave no formal evidence about *their* proxy's efficiency. [Jensen (1972a, 1972b) did do this with the same data, however.] Their other stated test, of the Sharpe–Lintner theory, is certainly open to question since no information was provided about the proxy market's relation to the Sharpe–Lintner tangency. (In fact, we know there was a difference between the ex-post Sharpe–Lintner tangency portfolio and Black, Jensen, and Scholes' 'market'. See above.) Therefore, for the Black, Jensen, Scholes paper taken in isolation from Jensen's addition, no hypothesis *whatever* was tested unambiguously.

3. Measuring the market and testing the theories

3.1. The Sharpe–Lintner case

As mentioned earlier in connection with Black, Jensen, and Scholes' conclusions, there has been in the literature some consideration of mis-measuring the market portfolio. Black, Jensen, and Scholes thought that a mis-specified market would cause a bias in the cross-sectional risk-return intercept from the Sharpe–Lintner prediction, but that the intercept would be intertemporally constant. But as we have seen, an incorrect market portfolio can cause *both* a bias *and* variation of the intercept over time, even when the Sharpe–Lintner theory is the true state of nature.

Mayers (1973) also considered the question of omitted (and non-marketable) assets and reached a similar conclusion with respect to the empirical implications: '. . . the primary testable propositions of the extended [Mayers] model are the linearity of the risk-expected return relationships . . . and the implication that no other variables . . . should be systematically related to expected return' [Mayers (1973, p. 266)].

These conclusions about the empirical implications of Mayers' model are very interesting for the following reason (among others): his derived risk coefficient, though denoted by the symbol 'β', is not the simple regression slope coefficient of the other models. For a given marketable asset, Mayers' beta depends on that marketable asset's return covariance with aggregate non-marketable assets' returns. This implies that a mean-variance efficient marketable portfolio for one investor need not necessarily be mean-variance efficient for another; and thus, there is no longer a mathematical equivalence between mean-variance efficiency and beta/expected return linearity.

It is not clear whether this makes the Mayers model more or less easily testable. The problem of non-identifiability of the market portfolio (in this case, of the marketable market portfolio) is still present since its return also appears in Mayers' linearity relation. In addition, there is a new problem in measuring the return to aggregate nonmarketable capital. On the other hand, if these measurement problems were resolved, the Mayers model may be more easily testable because the linearity relation is more structured – it requires a particular relation between marketable and non-marketable aggregate portfolio returns. Furthermore, the testability of this structure does not seem to be hindered by a mathematical equivalence to the mean-variance efficiency of either portfolio.

Returning now to the simpler Sharpe–Lintner theory, despite the overwhelming importance for testing of measuring the market return properly, references to the consequences of doing it improperly are rather rare. In a typical reference, Petit and Westerfield simply say that the market '. . . is commonly measured by a stock market index. such as the Fisher Link Relative Index or the Standard and Poor's 500 . . .' (p. 581), and they pick yet a third proxy for their own calculations (the Fisher Combination Investment Performance Index). Blume and Friend also use this latter index and make no mention of its being only a proxy. Curiously, they do mention that the all-equities 'zero-beta' portfolio may be only an approximation (p. 23), but, as already noted, they draw an incorrect inference from this fact and they make no reference to the one-to-one relation between an error in the market proxy and an error in the zero-beta proxy.

Fama and MacBeth used 'Fisher's Arithmetic Index, an equally weighted average of all stocks listed on the New York Stock Exchange' (p. 614). This index is not even close to a value-weighted index and should never be suggested as a market proxy. But Fama and MacBeth make no mention of possible error in the proxy's measurement, despite the fact that their paper comes closest to a systematic exploitation of the efficient set mathematics and its implications. Given Fama's more recent statements (1976), it is safe to say that he would not choose this index again.

One analysis of mis-measurement of the market portfolio was presented by Miller and Scholes (1972, pp. 63–66). They report an experiment which had an important influence on the research of others. It is often mentioned in conversations and sometimes in print. For example, in his review article, Jensen (1972a) states that Miller and Scholes '. . . conclude that the improper measurement of the market portfolio returns does not seem to be causing substantial problems' (p. 365).

Miller and Scholes studied the following problem: Suppose that individual returns are generated by a process containing a 'true' market index, m^*,

$$\tilde{R}_i = \beta_i \tilde{R}_{m^*} + \tilde{\eta}_i, \tag{7}$$

where β_i is constant and $\tilde{\eta}_i$ is a random variable with zero mean.

Suppose also that only a proxy index, m, is identifiable and that its returns satisfy the same equation,

$$\tilde{R}_m = \beta_m \tilde{R}_{m^*} + \tilde{\eta}_m. \tag{8}$$

Miller and Scholes then ask the question, what would be the large sample value of $\hat{\gamma}_1$ in a cross-sectional model of the form

$$\bar{R}_i = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{b}_i + \hat{\varepsilon}_j, \tag{9}$$

where $\bar{R}_i$ is the time-series sample mean of $\tilde{R}_i$, $\hat{b}_i$ is the simple least-squares time-series regression slope coefficient of the individual return $\tilde{R}_i$ on the *proxy* market, $\tilde{R}_m$ and $\hat{\varepsilon}_j$ is the estimated residual. They show under quite general conditions that $\hat{\gamma}_1$ will be asymptotic to

$$\beta_m \bar{R}_{m^*}[r^2(\hat{b}_i, \beta_i)/r^2(R_m, R_{m^*})].$$

The term in brackets contains two squared correlation coefficients, a cross-sectional one between true and estimated beta and a time-series one between true and proxy market return.[18]

Miller and Scholes went on to an empirical analysis. Having first estimated the cross-sectional model (9) using an all-equities proxy for the market, they re-estimated (9) with a 25% bond index and then with a 50% bond index. The coefficients $\hat{\gamma}_1$ '. . . were virtually unchanged . . .' (p. 66) in the three cases. They state that empirically '. . . the correlation between the old and new indexes was very close to one' (p. 66), i.e., that $r^2(R_m, R_{m^*}) \approx 1$ if R_{m^*} is taken as the 'old' index. Also, the old and new 'coefficients of risk' were almost perfectly correlated, $r^2(\hat{b}_i, \beta_i) \approx 1$. This implied that the old and new estimates of γ_1 were proportional by the factor β_m which is the beta of the new proxy index with respect to the old proxy.

Conclusion: if the market proxy is perfectly correlated with the true market, the resulting cross-sectional model would yield a $\hat{\gamma}_1$ exactly proportional to the γ_1 computed by using the true market. It is easy to see, therefore, that the Sharpe–Lintner basic hypothesis (H.4) would be supported by the data, and by this test procedure, if it were true.[19]

The key to understanding the nature and significance of this conclusion is the perfect correlation between the proxy and the true market. Of course, if such a perfect correlation were the state of nature (and everyone knew it), the mean-standard deviation efficient frontier would be a line composed of various combinations of the proxy and true markets. This alone implies the existence of a riskless return, one particular linear combination, and it also implies an infinity of Sharpe–Lintner tangent portfolios, any one of which would support (H.4) in the cross-sectional tests.

Since the mere presence of perfect correlation between the true and proxy markets implies the Sharpe–Lintner result, how are the Miller–Scholes results to be reconciled with the results of Black, Jensen, and Scholes, Blume and Friend and Fama and MacBeth, all of whom rejected the Sharpe–Lintner theory. Miller and Scholes actually anticipated an econometric reconciliation which will be discussed in detail in the next section. There exist other explanations and one very simple possibility will be discussed next.

Actually, Miller and Scholes (and others)[20] only found *almost* perfect correlation between two *proxy* market portfolios. The demonstration of such a correlation for the *true* market was beyond their (and is beyond our) econometric ingenuity for the simple reason that the true market portfolio is unknown. This suggests a reconciliation of the body of empirical results based on either (a) the true market is not *perfectly* correlated with the measured proxies, or (b) perfect correlation only exists among inefficient portfolios. Explanation (b) is inconsistent with equilibrium unless there are restrictions on short-selling. Even if there were such restrictions, however, the computation of sample betas with an inefficient portfolio would give an asymptotically (time-series-wise) not exactly-linear mean return beta relation. It would therefore seem unlikely that this particular explanation has much validity.

To understand explanation (a), we need to know the effect of market proxy correlation on the deviation between the Sharpe–Lintner implications and the observed results. For example, referring again to fig. 1, where m^* is the true market portfolio and m is the proxy, what is the relation between the distance $r_z - r_F$ on the one hand and the correlation between $\tilde{r}_{m^*}$ and $\tilde{r}_m$ on the other hand? From the geometry alone, we observe that this must depend upon the curvature of the risky efficient set and on its distance from the return axis. It also must depend on the absolute and relative positions of m and m^*. If both are located far out on the positive segment of the efficient frontier, they might be nearly perfectly correlated and yet imply a large and significant difference between the returns r_F and r_z on their orthogonal portfolios.

Some simple numerical examples may serve to illustrate the possible magnitudes involved. There are two hypothetical states of nature contained in the two examples in table 1. The numbers are not just made up, however. Those

[18]Note that the γ_0 would be intertemporally constant in the Miller–Scholes framework. Thus, their model is consistent with the Black, Jensen, Scholes interpretation of $\hat{\gamma}_0$, which is misleading in the case of a mis-measured market proxy portfolio.

[19]A simple way to see this is as follows: Suppose the returns in (7) and (8) are *excess* returns, that $\eta_m \equiv 0$, and that the Sharpe–Lintner (H.4) is valid. Then the cross-sectional model (9) would yield the asymptotic result, $\hat{\gamma}_0 = 0$ and $\hat{\gamma}_1 = \bar{R}_m$, where $\bar{R}_m$ is the market proxy excess mean return. Then it would appear from the data that the market proxy is efficient and equal to the Sharpe–Lintner tangent portfolio.

[20]See, for example, Fisher (1966). Table 4.5 (p. 81) of Lorie and Brealey (1972), gives correlation coefficients for five commonly-used indexes, for data from the mid-20's to the mid-60's, ranging between 0.906 and 0.985.

in the 'Given' panel come directly from Black, Jensen, and Scholes' tables 5 and 7 (1972) – for Example 1 – and from Morgan's table 3 (1975) – for Example 2. Example 1 contains ex-post results calculated from monthly returns for 1931–65. Example 2 also contains ex-post numbers but for 5-day intervals from July 1962 through December 1972.[21] Only the 3.0 riskless interest rate in Example 2 is a

Table 1
Examples of the ex-post risky efficient set and of sample correlation among efficient portfolios.

Quantity	Symbol	Numerical values (%/annum) Example 1 (BJS results)	Example 2 (Morgan results)
Given:			
Riskless return	r_F	1.920	3.0
Market proxy return	r_m		
Mean		18.96	15.54
Standard deviation		106.9	101.4
Return on market proxy's efficient orthogonal portfolio	r_z		
Mean		5.976	4.067
Standard deviation		51.10	52.14
Implied by the above:			
Global minimum variance portfolio	r_0		
Mean		8.392	6.469
Standard deviation		46.10	33.12
Sharpe–Lintner tangent portfolio	r_{m^*}		
Mean		12.34	12.75
Standard deviation		58.49	77.74
Correlation coefficient between market proxy's return and return on S–L tangent portfolio	ρ	0.8952	0.9860

pure assumption. The source paper provided no measure of the riskless return and 3 percent was chosen as a reasonable but conservative figure for the period.

Given one additional and strong assumption, estimates for the global minimum variance and Sharpe–Lintner tangent portfolios are implied by the riskless return, the market proxy and the market proxy's orthogonal (zero-beta) efficient

[21]The BJS and Morgan numbers have been made comparable by using suitable annualisation multipliers. The BJS numbers were multiplied by 1200 and the Morgan numbers by 36500/7. Morgan gave several different measures of m and z. I used the numbers in the first column of his table 3, p. 371, and the first of each pair. To obtain the standard deviation of the market proxy, I assumed that his 'risk premium: mean/std. error' was computed as

$$[(r_m - r_z)\sqrt{522}]/\sqrt{(\sigma_z^2+\sigma_m^2)}, \qquad 522 = \text{sample size.}$$

I am indebted to him for private correspondence that certified the validity of this assumption.

portfolio that were provided in the source papers. The crucial assumption is that the market proxy and its associated zero-beta portfolio are actually located on the ex-post efficient frontier.[22] If m and z are both efficient, the variance of each one is related to its mean by the efficient set quadratic equation, (A.11) of the appendix, which contains the three 'efficient set constants', a, b and c. In addition, since m and z were orthogonal by construction, their means are related by a third expression (A.15) which is the general equation relating the mean returns of orthogonal efficient portfolios. Since this expression also contains the efficient set constants, there results a non-linear system of three equations in the three unknowns, a, b and c. Usually, as in the case of our examples here, the system has a unique solution. Once the three efficient set constants are determined, all the other information of table 1 is computable in a straightforward way. The Sharpe–Lintner tangent portfolio's return requires additionally that the riskless interest rate be assigned a value.

The examples' relevance derives from the tangent portfolio and its correlation with the proxy market portfolio. In Example 1, Black, Jensen, and Scholes' data indicate that the Sharpe–Lintner tangent portfolio had an average monthly return of 12.3 (percent per annum) from 1931–65 and that its ex-post correlation with their market proxy was on the order of 90 percent. Notice that the tangent return was only 65 percent of the market proxy return, despite the significant correlation. Also note that the Black, Jensen, and Scholes zero-beta proxy returned 5.976 (percent per annum) on average. As mentioned previously, this finding was used by them to deny the validity of Sharpe–Lintner theory (because 5.976 was significantly greater than 1.920, the estimated riskless return).

There is a possible way to examine the validity of their conclusion. Using the same data, a different consistency check of the Sharpe–Lintner model would involve the individual asset investment proportions in the observed tangency portfolio. If any of these were significantly negative, the tangency portfolio would not satisfy the qualities of a market portfolio, which must have positive investments in all assets.[23] The suggested exercise (it has not yet been done by anyone, to my knowledge), has been termed a 'consistency check' rather than a 'test' of the Sharpe–Lintner theory because of the many assets omitted from the Black, Jensen, and Scholes sample. The omission of even a single asset can in principle cause an observed tangency portfolio to alter in composition

[22]There are several reasons why this is a strong assumption and why the results of table 1 should only be considered as examples. In both the Black, Jensen, Scholes and the Morgan papers, the samples consisted only of equities. Thus, it is very unlikely that the market proxy was exactly mean-variance efficient. Even if the samples had included all assets, the zero-beta measured portfolios were probably not precisely on the ex-post mean-variance boundary because the full covariance matrix was never inverted to find the efficient set constants. For example, Morgan estimated the efficient set by using a sample of 89 portfolios of 6 securities each (p. 365). This estimate of the efficient set would have differed, although perhaps in only a minor way, from the efficient set computed using the 534 (89 × 6) individual stocks.

[23]I am indebted to Michael C. Jensen for suggesting this procedure.

from totally positive to some negative proportions. (The alteration is not merely an allocation of the former weight of the omitted asset to the remaining assets because the entire efficient set can change.) Nevertheless, the calculation would be worthwhile because it would at least provide an insight into the possibility of incorrect inferences arising from market proxy portfolio misspecification within the Black, Jensen, and Scholes universe of securities. Unfortunately, the calculation cannot be reported here because it requires the full sample covariance matrix and the sample mean return vector of individual assets. These are not in my possession.

The Morgan data, which cover a later period than the Black, Jensen, and Scholes data, imply an even higher correlation between the market proxy and the Sharpe–Lintner ex-post tangency portfolios. This is due partly to a lower mean return of the market proxy and partly to a larger *assumed* riskless return which has caused the tangent portfolio to lie closer to the proxy.[24] However, the same qualitative conclusions obtain: Efficient portfolios are highly correlated and the Sharpe–Lintner theory is consistent with the data and a mis-specified market index.

Recall that *all* efficient portfolios on the positively-sloped segment are positively correlated. It is also true that the correlation increases with increasing mean returns of the two portfolios in question (holding constant the difference between their means). In the two numerical examples of table 1, for instance, all efficient portfolios with returns between 14 and 49 percent for Example 1 and with returns between 12 and 36 percent for Example 2 had *squared* correlations with the market proxy greater than 90 percent.

The implications of this are clear: Any hypothesis, such as Sharpe–Lintner, that makes a specific prediction about the position of the market portfolio, is likely to be highly susceptible to a type II error – being rejected when it is true. Heuristically, a small error in measuring the market's composition can cause an error in testing the theory. The market proxy may be almost perfectly correlated with the true market and yet a significant difference can emerge between the proxy zero-beta return and the true zero-beta return (or the riskless return).

3.2. *The generalized asset pricing theory case*

For testing the Sharpe–Lintner hypothesis (H.4), the identifiability of the market portfolio is a serious problem. For the more general asset pricing hypothesis (H.2), it is perhaps even a more serious problem. For (H.4), we can at least get an idea of the consequences of a mis-specified market portfolio by *assuming* that the proxy market is efficient. Then, the ex-post tangent portfolio can be calculated as in the above examples and its return and reasonableness can be judged against external criteria. For example, it might have been expected that the true market portfolio had less variance and less return than the Black, Jensen, and Scholes and the Morgan all-equities proxies, perhaps because equities are more variable than the average asset. Or, if the all-equities proxies had been combined with bonds, human capital, and real estate in reasonable proportions, the resulting mixture might have been closer to the observed tangent portfolio. Naturally, such possibilities are mere conjectures: They are not testable hypotheses, again for the simple reason that the true market portfolio has an unknown composition.

[24]The effect of the riskless rate assumption is easy to assess. For an assumption of 2% rather than 3%, the correlation between market proxy and tangent portfolio would have been 0.9587. For 4%, the correlation would have been 0.9999. Thus a considerable range of assumptions for the riskless rate would have given the same general impression.

For the more general hypothesis (H.2), such judgements based on common-sense interpretation of the results are likely to be unavailable. (H.2) merely requires that the true market be somewhere on the positively-sloped segment of the mean-variance efficient frontier. For relatively small (time-series) sample sizes, this hypothesis is highly susceptible to a type I error, being acceptable when it is false; but as the number of time-series observations increases, the hypothesis will almost surely be rejected, even when it is true. To see why, first consider the fact that the true market portfolio has a positive proportion invested in every individual asset. This implies that every reasonable candidate for the market proxy must have totally positive investment proportions. In many cases, in fact, the investment proportions are either the positive constant $1/N$ for the included assets (and zero for excluded assets), or the proportions display little cross-sectional variation. We know, therefore, that all such candidates for the proxy market portfolio must lie in a relatively small region of the mean-variance space.

Suppose, for example, that the true efficient set is given by the curve labeled 'I' in fig. 2. In this particular example, efficient portfolios between A and B are assumed to have totally positive investment proportions. As shown in the appendix, Theorem 3, an efficient set like I, whose global minimum variance portfolio has totally positive investment proportions, will occur if the variance of every individual asset exceeds its covariances with all other individual assets (if $\sigma_j^2 > \sigma_{ij}$ for all $j \neq i$). Above the point A and below the point B, at least one asset has negative investment proportions. Suppose the asset that leaves the efficient set at point A is indexed j. Then the curve I_{-j} would be the efficient set if j did not exist. I and I_{-j} are tangent at a single point at most. (There might have been no finite tangency because j might have had a positive or a negative weight in all portfolios on I.)[25] But since the weight is assumed negative above A and positive below A, it must be zero at A. Since it is zero everywhere on I_{-j}, A must be a tangency. (I and I_{-j} obviously cannot cross since they are minima.)

[25]The curve I_{-j}, with one omitted asset, is offered as an expositional example. In general, there will be more than one omitted asset from any empirical sample and so there will be no common point between the true efficient frontier and the sample efficient frontier, except in an unusual circumstance (the unusual circumstance being that *all* omitted assets have their zero investment proportions at a common point on the true efficient set).

If (H.2) is true, the true market portfolio must lie on the boundary I between A and B, say at m^*. If m is chosen as the market proxy, all empirical tests will support (H.2) because the proxy m does indeed lie on the reduced efficient boundary I_{-j} which lacks asset j. Although proxy portfolio m is inefficient globally (since it lies below I), this fact will not be detectable by any test using the reduced subset without j.

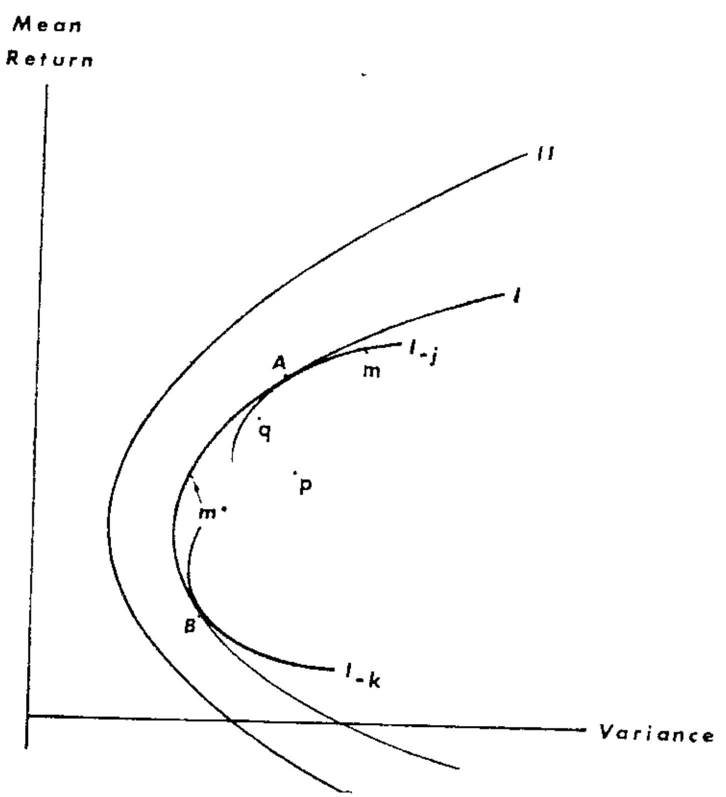

Fig. 2. Totally positive portfolios and the efficient boundary.

Thus (H.2) will be supported correctly, but for the wrong reason. On the other hand, suppose that the true market is really inefficient and lies within I at the point labelled p. Then the same exact test with market proxy m will support (H.2) incorrectly. In fact, it seems that this is the greatest danger. For any subset of assets, there exists an 'efficient frontier' whose constituent portfolios will satisfy all empirical tests of (H.2). As long as the investment proportions in the subset of assets are totally positive somewhere along this reduced boundary, a 'reasonable' market proxy will be available and it will support (H.2) since it will be subset efficient.

On the other hand, the geometry makes evident that a type II error is also quite possible. Suppose that portfolio q has been chosen as the market proxy and that (H.2) is true and that m^* is truly efficient and lies on I. In this case, a large sample will almost surely reject (H.2) since the proxy does not even lie on the reduced efficient set I_{-j}. Note that this will occur even when the proxy is highly correlated with the true market m^* and is also highly correlated with a proxy which is subset efficient.

The empirical situation is aggravated by this likelihood of a high correlation between the true market and the proxy, whether (H.2) is true or false. Such a high correlation is bound to make it seem that the exact identity of the market is of relatively minor significance and the temptation will be great to modify the market proxy slightly to obtain the *desired* result, whether it be rejection or acceptance of (H.2). In fact, if the identity of the true market is a matter of dispute among different researchers, there may be no way to settle the validity of (H.2) with any size of sample. The only exception would seem to be when (H.2) is false, the true state of nature is depicted by efficient set II in fig. 2 and the sample contains every asset. In this case, *no* totally-positive portfolio is efficient and large (time-series) samples would reject (H.2) unambiguously.

In principle, such a test is easy to construct. As shown in the appendix, the investment proportions for portfolios that lie on the ex-post efficient boundary are given by the $N \times 1$ vector

$$X = B\begin{pmatrix} r \\ 1 \end{pmatrix},$$

where B is an $(N \times 2)$ matrix of constants that depend only on the mean returns and sample covariances of the N individual assets, and r is the mean return of the ex-post efficient portfolio whose investment proportions are given by X. Since B contains constants, the test simply involves computing X at two points – at the minimum and maximum observed individual returns, $r_{\min}$ and $r_{\max}$, to obtain

$$X_{\min} \equiv B\begin{pmatrix} r_{\min} \\ 1 \end{pmatrix} \quad \text{and} \quad X_{\max} \equiv B\begin{pmatrix} r_{\max} \\ 1 \end{pmatrix}.$$

Then if there is a single element that is significantly negative in both $X_{\min}$ and $X_{\max}$, (H.2) is false. This follows because every totally positive portfolio lies between $r_{\min}$ and $r_{\max}$. Thus, if all efficient portfolios in this interval have one or more significantly negative investment proportions, there is no totally positive mean-variance efficient portfolio. Unfortunately, this 'test' is only valid in principle. The full covariance matrix of all individual assets is required to compute the matrix B. Furthermore, the sampling variation of B would generally be unknown. Finally, the cost of collecting data *for every* existing asset would be prohibitive.

In summary of this section, the sole viable hypothesis of generalized asset pricing theory is (H.2) – the true market portfolio is ex-ante efficient. This hypothesis offers a non-trivial challenge to our econometric ingenuity and the challenge has not yet been satisfactorily met. The problem can be summarised by noting that in a given sample there are always portfolios which do not reject (H.2) and that little external information is available on the true market portfolio's exact composition. Furthermore, even a small mis-specification of the proxy's composition can lead to the wrong conclusion. What might seem a trivial mis-specification in an ordinary statistical application can be of crucial importance for testing (H.2).

Appendix

The efficient set mathematics[26]

The efficient set (or efficient portfolio frontier) is composed of portfolios with minimum variance at each possible level of mean return. Ex-ante variances and mean returns of individual assets must be estimated or subjectively determined. The ex-post efficient set is a sample statistic, the set of minimum sample variance portfolios.

Given the characteristics of individual assets, a portfolio is completely characterized by the proportions invested in its constituent securities,

$$X_p = \|x_{ip}\|,$$

where x_{ip} is the proportion of portfolio p invested in asset i, and X_p is a vector subject to the constraint

$$X_p'\iota = 1, \tag{A.1}$$

ι denoting the unit vector. When speaking of a single portfolio, we will suppress the subscript p on the vector X.

The parameters of the efficient set problem are the mean return vector of individual assets,

$$R = \|r_i\|,$$

[26]As mentioned in the text, most of the results in this appendix have appeared previously. See particularly Merton (1972), which is the original full analytic treatment of the efficient set mathematics. Szegö (1975) seems to be the only other relatively complete treatment but Sharpe (1970), Black (1972) and Fama (1976) also contain many of the same results. I have made no attempt to ascribe originality.

and the covariance matrix of individual returns,

$$V = \|\sigma_{ij}\|.$$

These can be either population values or they can be sample product moments. The mathematics that follows does not require a distinction, merely a set of numbers R and V. Any portfolio's mean and variance are given by

$$r_p = X'R, \tag{A.2}$$

$$\sigma_p^2 = X'VX. \tag{A.3}$$

Similarly, the covariance of any two arbitrary portfolios (say p_1 and p_2) is given by

$$\sigma_{p_1p_2} = X_{p_1}'VX_{p_2}. \tag{A.4}$$

Necessary and sufficient condition for a portfolio to be efficient

The efficient set is found by minimizing σ_p^2 subject to the two constraints (A.1) and (A.2). The Lagrangian is

$$L = X'VX - \lambda_1(X'R - r_p) - \lambda_2(X'\iota - 1),$$

where λ_1 and λ_2 are undetermined multipliers. The first extremum conditions are the vector

$$VX = \tfrac{1}{2}(\lambda_1 R + \lambda_2 \iota), \tag{A.5}$$

plus the constraints (A.1) and (A.2).

If the joint distribution of individual returns is non-degenerate, the covariance matrix is positive definite (and non-singular), and all efficient portfolios satisfy

$$X = \tfrac{1}{2}V^{-1}(R\ \iota)\begin{pmatrix}\lambda_1\\ \lambda_2\end{pmatrix}. \tag{A.6}$$

Non-degeneracy simply implies that no two distinct linear combinations of assets are perfectly correlated and that no asset has zero variance. We shall see later how to find the efficient set when this condition is not satisfied. When the probability distribution *is* non-degenerate, the second-order conditions for a minimum are satisfied because the covariance matrix is positive definite.

The equation of the efficient set

Different efficient portfolios are determined by different values of the multipliers in eq. (A.6). The result can be written in a more intuitive way, however, as shown in the following:

Theorem 1. If no linear combination of assets has zero variance and at least two assets have different mean returns, the investment proportions of a mean-variance efficient portfolio whose mean return is r_p are given by the vector

$$X = V^{-1}(R\,\iota)A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix}, \tag{A.7}$$

where the (2×2) matrix A is defined as

$$A \equiv (R\,\iota)'\,V^{-1}(R\,\iota). \tag{A.8}$$

Proof. The assumptions in the 'if' clause guarantee that V is positive definite and that $(R\,\iota)$ has rank two. This implies that A is non-singular and positive definite. Then pre-multiplying eq. (A.6) by $(R\,\iota)'$ gives

$$\tfrac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = A^{-1}\,(R\,\iota)'\,X,$$

and by using constraints (A.1) and (A.2), this simplifies to

$$\tfrac{1}{2}\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix}.$$

Substitution for $\tfrac{1}{2}\binom{\lambda_1}{\lambda_2}$ in eq. (A.6) then gives eq. (A.7). Q.E.D.

The matrix A is the fundamental matrix of information about the basic data contained in the means and covariances of individual assets. As we shall see, the elements of A contain sufficient information to prove all the important results of the efficient set mathematics. Since A is 2×2 and symmetric, it contains only three distinct constants.

Definition

$$a = R'V^{-1}R, \quad b = R'V^{-1}\iota, \quad c = \iota'V^{-1}\iota, \tag{A.9}$$

are the 'efficient set constants' contained in the matrix

$$A \equiv \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

For example:

Corollary 1. The variance of any mean-variance efficient portfolio is related to its mean by the parabola

$$\sigma_p^2 = (r_p\ 1)A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix}, \tag{A.10}$$

which can be written in scalar notation as

$$\sigma_p^2 = (a - 2br_p + cr_p^2)/(ac - b^2). \tag{A.11}$$

Proof. The efficient investment proportions given by eq. (A.7) can be used in the general formula for a portfolio's variance, (A.3), to give

$$\begin{aligned}\sigma_p^2 &= (r_p\ 1)A^{-1}(R\,\iota)'V^{-1}\,VV^{-1}(R\,\iota)A^{-1}(r_p\ 1)' \\ &= (r_p\ 1)A^{-1}\,AA^{-1}(r_p\ 1)'.\end{aligned}$$

Q.E.D.

The minimum variance portfolio

One portfolio of special interest is the global minimum variance portfolio. Its mean and variance are found easily from the minimum of eq. (A.10).

Corollary 2. The global minimum variance is

$$\sigma_0^2 = 1/c. \tag{A.12}$$

The portfolio with this variance has mean return

$$r_0 = b/c, \tag{A.13}$$

and its investment proportions are given by

$$X_0 = V^{-1}\iota/c. \tag{A.14}$$

It is positively correlated with all portfolios and assets and its covariance with all individual assets and all portfolios is a fixed constant, σ_0^2, which is its own variance.

Proof. Eq. (A.13) is obtained from the zero of the first derivative of eq. (A.11),

$$0 = -b - cr_0.$$

This gives a minimum if the second derivative of eq. (A.11),

$$2c/(ac - b^2),$$

is positive. From eq. (A.9), we note that c must be positive since it is a quadratic form of the positive definite matrix V^{-1}. The denominator is positive because it is the determinant of A which is also positive definite.

Eq. (A.12) is obtained by substituting eq. (A.13) in the general formula for the variance of efficient portfolios, (A.11). Similarly, substituting eq. (A.13) into eq. (A.7) gives

$$\begin{aligned} X_0 &= V^{-1}(R\ \iota)\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}\begin{pmatrix} b/c \\ 1 \end{pmatrix}\Big/(ac-b^2) \\ &= V^{-1}(R\ \iota)\begin{pmatrix} 0 \\ 1/c \end{pmatrix} \\ &= V^{-1}\iota/c. \end{aligned}$$

which is eq. (A.14).

The last statement about the covariance of the minimum-variance portfolio is obtained as follows: Let X_j be any arbitrary vector of investment proportions. Then the covariance between this portfolio, j, and the global minimum variance portfolio is

$$\sigma_{j0} = X_j'VX_0 = X_j'\iota/c = 1/c. \qquad \text{Q.E.D.}$$

Heuristically, the minimum variance portfolio *must* be positively correlated with all other portfolios. If it were not, a further combination would result in another portfolio whose variance was even smaller. A heuristic reason for the constancy of the covariance is more subtle; but consider the following: Find the minimum variance portfolio that could be obtained from any arbitrary pair of assets. It can be verified readily that this minimum coincides with one or the other assets if and only if their covariance equals one of their variances. The same thing is true when one of this arbitrary pair is the global minimum variance portfolio. Since it is indeed the minimum, *its* variance must equal the arbitrary covariance (which is thus seen to be constant and independent of the second asset).

Efficient portfolios and correlation

Corollary 3. For every efficient portfolio except the global minimum variance portfolio there exists a unique orthogonal efficient portfolio with finite mean. If the first efficient portfolio has mean r_p, its orthogonal portfolio has mean r_z given by

$$r_z = (a-br_p)/(b-cr_p). \tag{A.15}$$

Furthermore, portfolio z always lies on the opposite-sloped segment of the efficient set from portfolio p.

Proof. By using eq. (A.7) in eq. (A.4), we can obtain the covariance between any arbitrary pair of efficient portfolios, say between p and q, as

$$\sigma_{qp} = (r_q 1)\, A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix}. \tag{A.16}$$

If p and q are orthogonal, this covariance is zero. Thus, putting $q = z$ gives the equation

$$(r_z\ 1)\begin{pmatrix} c & -b \\ -b & a \end{pmatrix}\begin{pmatrix} r_p \\ 1 \end{pmatrix} = 0,$$

from which eq. (A.15) follows directly. [Note that the determinant of A has been eliminated from eq. (A.16).] Uniqueness is obvious from eqs. (A.7) and (A.15).

The second part of the corollary follows by noting that if r_z is on the negatively sloped efficient set segment, its return must be smaller than the return on the global minimum variance portfolio, i.e., $r_z < b/c$. Then from eq. (A.15), $(a-br_p)/(b-cr_p) < b/c$. If the return on r_p also satisfied $r_p < b/c$ then $b-cr_p$ would be positive and we would have

$$a-br_p < (b/c)(b-cr_p),$$

or $a-b^2/c < 0$. But since c is positive, this is a contradiction because $ac-b^2$ is recognized as the determinant of A, which is positive. Thus, r_p must exceed b/c and be on the positively-sloped segment. Q.E.D.

The geometry of these orthogonal portfolios is useful, as can be seen in the following:

Corollary 3.A. In the return-variance space, the line passing between the efficient portfolio p and the global minimum variance portfolio intersects the return axis at r_z. In the return-standard deviation space, the tangent to the efficient set at r_p intersects the return axis at r_z.

Proof. In the mean-variance space, the slope of the line connecting portfolios p and 0 is

$$(r_p-r_0)/(\sigma_p^2-\sigma_0^2),$$

and its intercept on the return axis is

$$r_p-\sigma_p^2\,\{(r_p-r_0)/(\sigma_p^2-\sigma_0^2)\}.$$

Substituting for r_0 and σ_0^2 from eqs. (A.13) and (A.12) and simplifying, this expression reduces to $(br_p - a)/(cr_p - b)$ which is r_z in eq. (A.15). Thus the first statement is proven. To prove the second, it is easiest to use the first derivative of eq. (A.11). This provides an expression for the tangent to the efficient set in the mean-standard derivation space. Multiplying this tangent by σ_p and subtracting the result from r_p gives eq. (A.15). Q.E.D.

Corollary 4. All portfolios on the positively-sloped segment of the efficient set are positively correlated.

Proof. From eq. (A.16) we can see that any two efficient portfolios, p and q, will be negatively correlated if and only if

$$a - b(r_p + r_q) + c\, r_p r_q < 0,$$

which implies that

$$a - br_p < r_q(b - cr_p).$$

If p is on the positively-sloped efficient segment, $b - cr_p < 0$, and thus $r_q < r_z$, where z is p's orthogonal portfolio. But z must lie on the negatively-sloped segment, from Corollary 3. Thus, q is also on the negatively-sloped segment since its return is smaller than r_z. Q.E.D.

It is easy to see by the same argument that all portfolios on the negatively-sloped segment are positively correlated too. Only portfolios that lie sufficiently far apart, and on opposite sides of the efficient set, are negatively correlated. Heuristically, if investors do wish to minimize variance and maximize expected returns, all investors who agree on the probability distribution would hold positively correlated portfolios.

The separation or 'two-fund' theorem

Corollary 5 ('Two-Fund Theorem'). The investment proportions vector of every mean-variance efficient portfolio is a linear combination of the proportions vectors of two other efficient portfolios whose means are different.

Proof. From eq. (A.7), the investment proportions are seen to be linear in the mean return. This is so because the $(N \times 2)$ matrix, $B \equiv V^{-1}(R\,\iota)A^{-1}$, contains only constants (N is the number of individual assets). Thus, if p_1, p_2 and p_3 are efficient portfolios and α is a constant given by $\alpha \equiv (r_3 - r_2)/(r_1 - r_2)$,

then

$$X_{p_3} = B\begin{pmatrix} r_3 \\ 1 \end{pmatrix} = B\begin{pmatrix} \alpha r_1 + (1-\alpha) r_2 \\ 1 \end{pmatrix}$$
$$= \alpha X_{p_1} + (1-\alpha) X_{p_2}, \qquad \text{Q.E.D.}$$

According to Corollary 5, if we identify two efficient portfolios, all others can be constructed as a linear combination of these two. We might as well pick two portfolios whose means and variances are easy to compute. One would certainly be the global minimum variance portfolio with mean and variance given by eqs. (A.13) and (A.12). Its investment proportions, eq. (A.14), are simply the sums of the rows of the inverse covariance matrix, V^{-1} (normalized by the sum of all elements of V^{-1}). A second easily-computable efficient portfolio is the one with mean return $r_1 \equiv a/b$. Its variance is $\sigma_1^2 = a/b^2$ and its investment proportions are

$$X_1 = V^{-1}R/b. \tag{A.17}$$

From Corollary 3, eq. (A.15), we observe that this portfolio's orthogonal portfolio has a mean of zero.[27] Thus, its associated covariance vector is *proportional* to the mean return vector. Its investment proportions are very easy to compute as can be seen in eq. (A.17).

Relations among individual asset parameters

Corollary 6. The covariance vector of individual assets with any portfolio can be expressed as an exact linear function of the individual mean returns vector if and only if the portfolio is efficient.

Proof of Sufficiency. The vector of covariances between individual assets and a particular efficient portfolio is given by VX, where X is the portfolio's investment proportions vector. From eq. (A.7), this implies

$$VX = (R\,\iota)A^{-1}\begin{pmatrix} r_p \\ 1 \end{pmatrix}. \tag{A.18}$$

Since $A^{-1}\binom{r_p}{1}$ contains two constants, VX is linear in R. Q.E.D.

Although eq. (A.18) is perfectly acceptable as a result, it can be rewritten in a more traditional and perhaps more recognizable form. Note that the (2×1)

[27] I am indebted to Eugene Fama for pointing out this fact.

vector $A^{-1}\binom{r_p}{1}$ can be simplified as

$$\begin{pmatrix} cr_p-b \\ -br_p+a \end{pmatrix}/(ac-b^2) = \frac{cr_p-b}{ac-b^2}\begin{pmatrix} 1 \\ -r_z \end{pmatrix} = \frac{\sigma_p^2}{r_p-r_z}\begin{pmatrix} 1 \\ -r_z \end{pmatrix},$$

where z is p's orthogonal portfolio. Substitution back into eq. (A.18) gives

$$R = r_z\iota+(r_p-r_z)\beta, \tag{A.19}$$

where $\beta \equiv VX/\sigma_p^2$ is the vector of simple regression slope coefficients of individual assets on efficient portfolio p (the 'betas'). Since the covariances are linear in the mean return, of course the 'betas' are too.

Proof of Necessity. Let the vector of covariances with an arbitrary portfolio m be an exact linear function of the mean returns,

$$VX_m = \alpha_1 R+\alpha_2\iota,$$

where α_1 and α_2 are arbitrary constants. Then the vector X_m is equal to

$$\begin{aligned} X_m &= \alpha_1 V^{-1}R+\alpha_2 V^{-1}\iota \\ &= \alpha_1 bX_1+\alpha_2 cX_0, \end{aligned}$$

where X_1 is the vector of investment proportions for special efficient portfolio 1 [see eq. (A.17)], and X_0 is the vector of proportions for the global minimum variance portfolio (b and c are two of the three efficient set constants).

Since the constraint (A.1) must apply to all such vectors, including the vector X_m, we must have

$$\alpha_1 b+\alpha_2 c = 1.$$

Thus X_m is a weighted average of two efficient vectors X_1 and X_0 and, by Corollary 5, X_m also is efficient. Q.E.D.

A slightly more general form of the linearity property can be obtained. Let us suppose that portfolios A and B are chosen arbitrarily and that the *multivariate* regression coefficients are computed between individual asset returns and these two portfolios' returns. Then:

Corollary 6.A.[28] *Let vectors of multivariate regression coefficients be calculated between individual asset returns (as dependent variables) and the returns o*

[28] I am indebted to John B. Long, Jr. for pointing out a previous error in this corollary.

two imperfectly-correlated portfolios A and B (as explanatory variables). Then both multivariate regression coefficient vectors are exact linear functions of the mean return vector if and only if A and B are efficient. Furthermore,

(a) *the coefficient vector for asset j on portfolio A is given by*

$$(VX_A-\sigma_{AB}\iota)/(\sigma_A^2-\sigma_{AB}) = (R-r_B\iota)/(r_A-r_B), \tag{A.20}$$

and similarly for portfolio B with the A and B subscripts interchanged; and

(b) *the two coefficient vectors sum to the unit vector.*

Proof of Sufficiency. If statement (a) is true, the proof of exact linearity is identical to the proof of Corollary 6 because (A.20) shows that the multivariate coefficient vector is linear in the covariance vector VX_A. Since a linear function of a linear function is also linear, the result follows immediately.

To prove parts (a) and (b), first note that the usual multivariate coefficient of A obtained by regressing j on A and B would have been

$$\beta_j' \equiv \frac{\sigma_{jA}\sigma_B^2-\sigma_{jB}\sigma_{AB}}{\sigma_A^2\sigma_B^2-\sigma_{AB}^2}. \tag{A.21}$$

But since A and B are mean-variance efficient, this multivariate coefficient can be simplified. Noting from eq. (A.19) that $\sigma_{jB} = \sigma_B^2(r_j-r_{z_B})/(r_B-r_{z_B})$, where z_B is B's orthogonal portfolio, and noting analogous expressions for σ_{jA} and σ_{AB}, eq. (A.21) can be rewritten as

$$\beta_j' = \frac{(r_j-r_{z_A})(r_B-r_{z_B})-(r_j-r_{z_B})(r_B-r_{z_A})}{(r_A-r_{z_A})(r_B-r_{z_B})-(r_A-r_{z_B})(r_B-r_{z_A})},$$

or

$$\beta_j' = \frac{r_j-r_B}{r_A-r_B} = \frac{\sigma_{jA}-\sigma_{AB}}{\sigma_A^2-\sigma_{AB}}. \tag{A.22}$$

This proves part (a). Since the same development can be made in behalf of portfolio B, the multivariate coefficient against B must be $(r_j-r_A)/(r_B-r_A) = 1-\beta_j'$. This proves part (b). Q.E.D.

Proof of Necessity. Let both multivariate coefficient vectors be exact linear functions of the mean return vector. Using the general definition (A.21) of multivariate coefficients and the same procedure as that contained in the necessity proof of Corollary 6, it can be shown that the vectors W_A and W_B, defined by

$$W_A \equiv (X_A-\beta_{AB}X_B)/(1-\beta_{AB}),$$

and

$$W_B \equiv (X_B - \beta_{BA} X_A)/(1-\beta_{BA}),$$

are efficient. (Here, $\beta_{AB} = \sigma_{AB}/\sigma_B^2$.)

Adding the two equations in order to eliminate X_B gives

$$W_A(1-\beta_{AB}) + W_B\beta_{AB}(1-\beta_{BA}) = X_A(1-\beta_{BA}\beta_{AB}).$$

Since the coefficients of W_A and W_B sum to $1-\beta_{BA}\beta_{AB}$, X_A is a linear combination of two efficient vectors and is therefore also efficient (by Corollary 5). An identical development proves the efficiency of vector X_B. Since A and B are efficient, statements (a) and (b) are true. Q.E.D.

Suppose that some non-efficient portfolio q has been used to estimate a vector of betas. There exists an efficient portfolio p that has the same mean return, $r_p = r_q$, and for which eq. (A.19) is satisfied. But the betas of eq. (A.19) are VX_p/σ_p^2 whereas the betas estimated with portfolio q are $\hat{\beta} \equiv VX_q/\sigma_q^2$. Using the estimates $\hat{\beta}$, (A.19) can be rewritten as

$$R \equiv r_z + (r_q - r_z)\hat{\beta} + \frac{r_q - r_z}{\sigma_q^2} V(X_p/k - X_q), \tag{A.23}$$

where $k = \sigma_p^2/\sigma_q^2$. The expression (A.23) is linear in the vector $\hat{\beta}$ if and only if the last term is a constant vector. But if q is not efficient, then $k < 1$. This implies that the last term can be constant only if $r_p \neq r_q$, which is a contradiction. (This is actually an equivalent proof to that used for Corollary 6; but it has the advantage that the deviations from a linear return-beta relation caused by using an inefficient base portfolio are given explicit display.)

One feature of eq. (A.23) is of particular interest: It would be quite possible that a subset of the elements of X_p and X_q be equal; and if the returns and betas for this particular subset were calculated, they would be exactly linearly related when taken alone. Thus the full set of data is not *continuously* non-linear even when the base portfolio is *in*efficient.

To reiterate, an efficient set calculated from a subset of assets will pass through portfolios that are inefficient globally. But since such portfolios are efficient for the subset, their associated subset of betas will be exactly linear in the subset of mean returns.

Corollary 7. The proportion invested in a given individual asset changes monotonically along the efficient frontier.

Proof. By inspection from eq. (A.7). (The gradient vector of X with respect to r_p is a constant vector.)

This corollary implies that if an individual asset represents a non-zero proportion in any efficient portfolio, it is held in *all* efficient portfolios except at most one. It is either held in positive amount or sold short in all others. N.B. It is mathematically possible that some assets are positively (or negatively) represented in all efficient portfolios and even that the individual proportion is a constant.

The gradient vector of eq. (A.7) can be written as

$$\partial X/\partial r_p = \sigma_0^2 V^{-1}(R/r_0 - \imath)/(r_1 - r_0), \tag{A.24}$$

where r_1 and r_0 are means of, respectively, the special efficient portfolio 1 (see Corollary 5), and of the global minimum variance portfolio. σ_0^2 is the global minimum variance. The gradient vector of the covariance between individual assets and efficient portfolios is $\partial(VX)/\partial r$ and thus it is also monotonic. Since $r_1 > r_0$ if r_0 is positive, we see that as the efficient portfolio's return increases, the covariance between a given asset and the portfolio increases (is zero, or decreases) when the asset's mean return is larger (equal to, or smaller) than the return on the global minimum variance portfolio.

In contrast, the *betas* computed with efficient portfolios are *not* monotonic. In fact:

Corollary 7.A. To every individual asset, there corresponds a unique pair of orthogonal portfolios which provides the maximum and minimum betas for that asset. These portfolios have returns

$$r = r_{z_j} \pm \sigma_0 \sigma_{z_j} \sqrt{|A|}, \tag{A.25}$$

where z_j is the efficient portfolio whose return is orthogonal to the return of asset j, σ_0^2 is the global minimum variance, and A is the efficient set information matrix (A.8).

Proof. The beta of any individual asset with efficient portfolio p is given by

$$\beta_j = (r_j\ 1)A^{-1}(r_p\ 1)'/\sigma_p^2. \tag{A.26}$$

Differentiating with respect to r_p gives a function whose zero is a quadratic equation in r_p. Some tedious algebra simplifies the result to (A.25). (Note that σ_p^2 is a function of r_p.)

To prove that the maximum and minimum beta efficient portfolios are orthogonal, simply use (A.4). Their covariance is

$$(r_z + K\,1)A^{-1}(r_z - K\,1)',$$

where

$$K = \sqrt{(\sigma_{z_j}^2 \sigma_0^2 |A|)}.$$

In scalar notation, this reduces to

$$(a-2br_{z_j}+cr_{z_j}^2-cK^2)/|A|,$$

and since

$$\sigma_{z_j}^2 \equiv (a-2br_{z_j}+cr_{z_j}^2)/|A| \doteq cK^2/|A|,$$

the covariance is zero. Q.E.D.

An asset whose return is greater than the return on the minimum variance portfolio will be positively correlated with all efficient portfolios on the positive segment. This implies that its maximum-beta efficient portfolio will have return $r_{z_j}+\sigma_0\sigma_{z_j}\sqrt{|A|}$, and its minimum-beta efficient portfolio will have return equal to the smaller root of eq. (A.25).[29]

A straightforward application of l'Hôpitals' rule to eq. (A.26) shows that every individual beta converges to zero as the efficient portfolio's return grows indefinitely large (or small). (This does not violate the fact that the weighted average beta is always one because the investment proportions grow indefinitely large in absolute value.) As an implication, the cross-sectional variance in beta is determined by the particular efficient portfolio used in the computation. When the global minimum variance portfolio is used, all betas are equal to 1.0 and their cross-sectional variance is zero. At an infinite return, all betas converge to zero and again their cross-sectional variance is zero. Naturally, there must exist an efficient portfolio whose associated betas have a maximum cross-sectional variation:

Corollary 7.B. The maximum cross-sectional variance in beta is given by either one of the two efficient portfolios whose returns are

$$r_p = r_0 \pm \sigma_0^2 \sqrt{|A|}, \tag{A.27}$$

where r_0 and σ_0^2 are the mean and variance of the global minimum variance portfolio and A is the efficient set information matrix (A.8).

[29]As an example, for the Black, Jensen, and Scholes data of table 1 in the paper, an asset with the same return as the proxy market, $r_j = 18.96$ percent, would have had a maximum beta of 1.66 and a minimum of -0.659 over all efficient portfolios. (Assuming that the BJS market proxy and *its* zero beta portfolio were in fact efficient.) The two efficient portfolios that would have provided these maximum and minimum betas had returns of 11.6 and 0.375 percent, respectively.

Proof. It is possible to prove this result directly from the definition of beta, but the algebra is tedious. A shorter proof uses the fact that all beta vectors computed for efficient portfolios are exactly linear in the mean return vector (see Corollary 6). From the linearity relation (A.19), we must have

$$\beta = (R - r_z\iota)/(r_p - r_z), \tag{A.28}$$

and the cross-sectional variance in beta is therefore

$$\text{Var}(\beta) = \text{Var}(R)/(r_p - r_z)^2. \tag{A.29}$$

The cross-sectional variance of individual returns, Var (R), is a constant with respect to movements along the efficient set. Thus, Var (β) is maximized when $|r_p - r_z|$ is minimized. From Corollary 3, we note that r_z is a simple function of r_p [eq. (A.15)]. The first derivative of $r_p - r_z$ with respect to r_p gives a quadratic equation in r_p whose solution is eq. (A.27). Q.E.D.

It is evident that the two portfolios which satisfy eq. (A.27) lie symmetrically on oppositely-sloped segments of the efficient frontier. The elements of their two associated beta vectors must therefore be equal and of opposite sign.[30] They are also orthogonal. These are the only orthogonal efficient portfolios whose variances are equal and by Corollary 3.A, their variance is $2\sigma_0^2$.

The efficient set when the minimum variance is zero

A special problem arises when the least-risky portfolio has zero variance.[31] This could be caused by the existence of a riskless asset but the possibilities for its occurrence are much broader. In general, it occurs if the covariance matrix of returns is only positive *semi-definite*; that is, if there exists a vector of investment proportions X_F such that $X_F' V X_F = 0$.

Since V is no longer of full rank, its singularity precludes the direct calculation of the investment proportions vector. The simplest and most intuitive way to find them is to proceed in two steps as follows. First, find the efficient set parabola for the restricted group composed of all mutually non-degenerate assets. This means that any non-risky assets and one member of each pair of perfectly-correlated assets must first be discarded. A 'risky efficient set' is constructed from the remaining assets. For example, if the original N assets had covariance

[30]Continuing the numerical example with the Black, Jensen and Scholes data (table 1), the efficient portfolios with returns 13.44 and 3.44 percent would have maximized the cross-sectional beta variance.

[31]Mathematically speaking, there is nothing to preclude several zero-variance portfolios with different mean returns. Of course rational asset pricing would preclude such an event since in the absence of restrictions on short-selling, an infinite return could be obtained without a risk.

matrix V with rank $L < N$, $N-L$ rows and columns of V would be discarded to obtain an $L \times L$ covariance matrix, V_L, with full rank. *It is obvious that all the results found previously for V when it had rank N must apply to the restricted set of L assets whose covariance matrix is V_L.* In particular, Theorem 1 and Corollaries 1–7 all apply to the reduced space of L assets. A reduced 'risky efficient set' can be derived; the betas calculated against portfolios that lie on this 'risky efficient set' are linear in the reduced vector of mean returns. The minimum variance among portfolios of the L assets is $1/\iota' V_L^{-1}\iota$; and so on.

This implies:

Theorem 2. Let an efficient set be computed from the L assets whose covariance matrix is non-singular. Then the global efficient set is composed of

$$\alpha X_L = \alpha\sigma_0^2 V_L^{-1}[(R_L - r_F\iota)/(r_0 - r_F)], \tag{A.30}$$

as investment proportions in the L assets and $1-\alpha\iota' X_L$ in the zero-variance asset F. The scalar α is positive for efficient portfolios on the positive (negative) segment if r_F is less (greater) than the return r_0 on the minimum variance portfolio of the L assets. σ_0^2 is this minimum variance.

Proof. Since the global minimum variance is zero, the efficient set is composed of line segments in the mean-standard deviation space. This means that there is some tangent to the reduced (L asset) mean-standard deviation efficient set which passes through r_F and gives the global efficient set. From Corollary 3.A, we see that the return at the tangency point must be the solution to

$$r_F = (br_\tau - a)/(cr_\tau - b),$$

where a, b, and c are the three 'efficient set constants' for the L assets only. This equation is reversible, which implies

$$r_\tau = (br_F - a)/(cr_F - b).$$

Substituting r_τ in (A.7) gives the X_L investment proportions vector used in (A.30).

The last part of the theorem is established by noting that

$$r_F - r_\tau = (a - 2br_F + cr_F^2)/(cr_F - b).$$

The numerator is positive since $|A_L| = ac - b^2 > 0$.

Thus $r_\tau \gtreqless r_F$ implies $r_F \lesseqgtr b/c = r_0$. Q.E.D.

An alternative proof proceeds directly from the variance of the reduced set

of efficient portfolios. In other words, we can find the minimum with respect to r_p of

$$\alpha^2\sigma_p^2,$$

subject to

$$r = \alpha r_p + (1-\alpha) r_F.$$

Using the efficient set variance formula for σ_p^2, the first derivative of $\alpha^2\sigma_p^2$ is zero when

$$(r_p - r_F)(cr_p - b) - (a - 2br_p + cr_p^2) = 0,$$

and this reduces to

$$r_p = (br_F - a)/(cr_F - b).$$

Corollary. When V is singular, all positive-variance efficient portfolios are perfectly correlated.

Proof. Theorem 2 shows that the vector αX_L of investments in the reduced L asset space is proportional for all positive variance efficient portfolios. The remaining investment, $1-\alpha\iota' X_L$, is placed in a zero-variance asset. Thus, the returns on all efficient portfolios are exactly linearly related. Q.E.D.

Qualitative results for investment proportions vectors

The sign pattern of the investment proportions vector of an efficient portfolio is an important datum in several contexts. For example, the interdiction of short-selling would leave a region of the efficient frontier unchanged if and only if that region were characterized by totally non-negative investment proportions, $X \geqq 0$.

Perhaps more importantly, there may exist no efficient vector such that $X > 0$. This means that *no* efficient portfolio has positive investments in all individual assets; which would imply, in turn, that the 'market' portfolio (composed of all assets) could not be efficient. If the reader thinks there must be *some* efficient portfolio whose X vector is positive, the counter-example in the footnote is offered.[32] There seems to be no necessary economic reason to con-

[32] Let three assets have mean return vector and covariance matrix,

$$R' = (1\ 2\ 3) \quad \text{and} \quad V = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 5 & 6 \\ 2 & 6 & 10 \end{pmatrix}.$$

Using eq. (A.7), it can be verified readily that every efficient portfolio contains a *zero* investment in the second asset.

Jacob, N., 1970, The measurement of market similarity for securities under uncertainty, Journal of Business 43, 328–340.
Jensen, M.C., 1972a, Capital markets: Theory and evidence, Bell Journal of Economics and Management Science 3, 357–398.
Jensen, M.C., 1972b, The foundations and current state of capital market theory, in: M.C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).
Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, Review of Economics and Statistics 47, 13–47.
Lintner, J., 1969, The aggregation of investor's diverse judgements and preferences in purely competitive security markets, Journal of Financial and Quantitative Analysis 4, 347–400.
Long, J.B., 1976, Mean-variance efficiency and portfolio choice in the context of differential taxation of dividends and capital gains, unpublished manuscript (Graduate School of Management, University of Rochester, NY).
Lorie, J.F. and R. Brealey, 1972, Modern developments in investment management (Praeger, New York).
Mayers, D., 1973, Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset, Journal of Business 46, 258–267.
Merton, R.C., 1972, An analytic derivation of the efficient portfolio frontier, Journal of Financial and Quantitative Analysis 7, 1851–1872.
Merton, R.C., 1973, An inter-temporal capital asset pricing model, Econometrica 41, 867–887.
Miller, M.H. and M. Scholes, 1972, Rates of return in relation to risk: A re-examination of some recent findings, in: M.C. Jensen, ed., Studies in the theory of capital markets (Praeger, New York).
Modigliani, F., G. Pogue, M. Scholes and B. Solnik, 1972, Efficiency of European capital markets and comparison with the American market, Proceedings of the First International Conference on Stock Exchanges, CISMEC.
Morgan, I.G., 1975, Prediction of return with the minimum variance zero-beta portfolio, Journal of Financial Economics 2, 361–376.
Palacios, J., 1973, The stock market in Spain: Tests of efficiency and capital market theory, unpublished dissertation (Stanford University, CA).
Petit, R.R. and R. Westerfield, 1974, Using the capital asset pricing model and the market model to predict security returns, Journal of Financial and Quantitative Analysis 9, 579–605.
Pirsig, R.M., 1974, Zen and the art of motorcycle maintenance (William Morrow, New York).
Quirk, J. and R. Saposnik, 1968, Introduction to general equilibrium theory and welfare economics (McGraw-Hill, New York).
Roll, R., 1973, Evidence on the 'growth-optimum' model, Journal of Finance 28, 551–566.
Roll, R. and B.S. Solnik, 1975, A pure foreign exchange asset pricing model, Journal of International Economics (forthcoming).
Rosa, J.J., 1975, Rentabilité, risque et equilibre à la bourse de Paris, Working Paper 7508 (Institut Orléanais de Finance).
Ross, S.A., 1972, The arbitrage theory of capital asset pricing, Working Paper 2–73 (Rodney L. White Center for Financial Research, University of Pennsylvania, Philadelphia, PA), forthcoming in Econometrica.
Ross, S.A., 1973, Return, risk and arbitrage, Working Paper 17–73a (Rodney L. White Center for Financial Research, University of Pennsylvania, Philadelphia, PA).
Ross, S.A., 1976, Mutual fund separation in financial theory – The separating distributions, Working Paper 1–76 (Rodney L. White Center for Financial Research, University of Pennsylvania, Philadelphia, PA).
Ross, S.A., forthcoming, A note on the capital asset pricing model (CAPM), short-selling restrictions, and related issues, Journal of Finance.
Sharpe, W.F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425–442.
Sharpe, W.F., 1970, Portfolio theory and capital markets, (McGraw-Hill, New York).
Szegö, G.P., 1975, Nuovi risultati analitici nella teoria della selezione del portafoglio (C.N.R., Comitato Scienze Economiche, Sociologiche e Statistiche).