NBER TECHNICAL PAPER SERIES

ARBITRAGE AND MEAN-VARIANCE ANALYSIS ON LARGE ASSET MARKETS

Gary Chamberlain

Michael Rothschild

Technical Paper No. 15

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

July 1981

NBER Technical Working Paper #15
July 1981

ARBITRAGE AND MEAN-VARIANCE ANALYSIS ON LARGE ASSET MARKETS

ABSTRACT

We examine the implications of arbitrage in a market with many assets. The absence of arbitrage opportunities implies that the linear functionals that give the mean and cost of a portfolio are continuous; hence there exist unique portfolios that represent these functionals. The mean-variance efficient set is a cone generated by these portfolios.

Ross [16, 18] showed that if there is a factor structure, then the distance between the vector or mean returns and the space spanned by the factor loadings is bounded as the number of assets increases. We show that if the covariance matrix of asset returns has only K unbounded eigenvalues, then the corresponding K eigenvectors converge and play the role of factor loadings in Ross' result. Hence only a principal components analysis is needed to test the arbitrage pricing theory. Our eigenvalue conditional can hold even though conventional measures of the approximation error in a K-factor model are unbounded.

We also resolve the question of when a market with many assets permits so much diversification that risk-free investment opportunities are available.

Professors Gary Chamberlain
and Michael Rothschild
Economics Department
Social Science Building
1180 Observatory Drive
University of Wisconsin
Madison, Wisconsin 53706

(608) 262-7789 and (608) 263-3880

INTRODUCTION

The two most significant developments in the field of finance have been the development of the capital asset pricing model (CAPM) and the working out of the implications of arbitrage beginning with the Modigliani-Miller Theorem and culminating in the theory of option pricing. While the principle that competitive financial markets do not permit profitable arbitrage opportunities to remain unexploited seems unexceptional, the same cannot be said for the major assumptions of the CAPM. Few believe that asset returns are well described by their first two moments or that some well-defined set of marketable assets contains most of the investment opportunities available to individual investors. Casual observation is sufficient to refute one of the main implications of the CAPM -- that everyone holds the market portfolio. Nonetheless, the CAPM seems to do a good job of explaining relationships among asset prices. Ross has argued that the apparent empirical success of the CAPM is due to three assumptions, which are together much weaker than the assumptions needed to derive the CAPM. These assumptions are first, that asset returns have a factor structure; second, that there are many assets; and finally, that the market permits no arbitrage opportunities. Ross presents a heuristic argument which suggests that on a market with an infinite number of assets there are sufficiently many well-diversified portfolios -- portfolios with no risk -- that prices of assets are largely determined by the arbitrage requirement that riskless portfolios which require no net investment should not have a positive return. If there are factors, asset prices are linear functions of factor loadings, and if there are no factors and all randomness is the result of what are called idiosyncratic disturbances,[2] arbitrage considerations alone determine relative prices. Although Ross' heuristics cannot be made rigorous,[3] he does prove that lack of arbitrage implies that asset prices are approximately linear functions of factor loadings, and Connor [4] has given conditions under which the conclusions of Ross' heuristic argument are precisely true.[4] Nonetheless, all of Ross' investigations of the implications of the absence of arbitrage opportunities take place in the context of a factor structure. This paper examines the implications of the absence of arbitrage opportunities on a market with many assets which does not necessarily have a factor structure.

The results of this exercise are as follows: In Section 1, we introduce our basic model of the asset market and show how markets with a countable number of assets have a simple Hilbert space structure. In Section 2 we formally define the absence of arbitrage opportunities. Our definition, essentially that given by Ross, is that it should not be possible to form a portfolio which has no variance, no cost, and a positive return. Ross [16] and Huberman [7] justified this definition by appeal to expected utility theory. They showed that if there are arbitrage opportunities, some expected utility maximizing traders will want to take infinitely large positions; this is, of course, inconsistent with equilibrium. Our results provide additional justification for this definition. We show that if there is a riskless asset, this definition of the lack of arbitrage opportunities is equivalent to the requirement that all riskless assets have the same rate of return. Furthermore, this definition fits easily into the Hilbert space structure of the asset market set out in Section 1. We prove that the absence of arbitrage opportunities implies that the linear functionals that give the mean and the cost of a portfolio are continuous. The Reisz Representation Theorm implies that there exist unique portfolios which represent these functionals. These portfolios play a crucial role in our analysis of mean-variance efficiency in Section 3. We also show that the absence of arbitrage opportunities implies that the distance between the vector of mean returns and a constant vector must be bounded (in the norm given by the inverse of the variance-covariance matrix). We call this bound δ.

In Section 3 we analyze the relationship between arbitrage, many assets, and the CAPM. As Roll [14] and Ross [17] have emphasized, the CAPM is equivalent to the statement that the market portfolio is mean-variance efficient; on finite markets the mean-variance efficient set is a cone generated by two portfolios or funds (see Chamberlain [2] and Roll [14]). The mean-variance efficient set also has this structure on markets with many assets where there are no arbitrage opportunities. If there are arbitrage opportunities, there is no well-defined tradeoff between mean and variance. The absence of arbitrage opportunities implies that the mean-variance efficient set is a cone generated by the portfolios that represent the linear functionals whose continuity is ensured by the lack of arbitrage opportunities. The mean-variance efficiency frontier is a very small subset of the space of all portfolios. The bound δ is intimately related to the mean-variance efficiency frontier. It is the slope of the tradeoff between mean and risk (as measured by standard deviation) along the efficiency frontier. That is, δ is the market price of risk.

In Sections 4 and 5 we examine the relationship between lack of arbitrage opportunities and factor structure. Ross showed that if there is a factor structure, then the distance between the vector of mean returns and the space spanned by the factor loadings is bounded. We show that a sufficient condition for Ross' result is that the covariance matrix of the asset returns ($\underset{\sim}{\Sigma}_N$) has only K unbounded eigenvalues as the number (N) of assets increases. This condition can hold even though conventional measures of the error in approximating $\underset{\sim}{\Sigma}_N$ by a K-factor model are unbounded as $N \to \infty$. Furthermore, the corresponding K eigenvectors converge in a relevant sense, and they play the role of factor loadings. The distance between the vector of mean returns and these loadings is bounded by $\lambda_{K+1}\delta^2$, where λ_{K+1} is a bound on the $(K+1)^{th}$ largest eigenvalue. These results have a number of implications for empirical work and provide some perspective on Roll and Ross' recent attempt to test the arbitrage pricing theory [15]. We also prove a converse of Ross' theorem. Under mild regularity conditions, if there is a factor structure and if asset prices are approximately linear functions of factor loadings, then there are no arbitrage opportunities.

In Section 6 we examine the question of whether large numbers of assets necessarily permit arbitrage opportunities. We show that there always is a nontrivial riskless hedge portfolio; this is not so if there are only a finite number of assets. We give necessary and sufficient conditions for the existence of a riskless asset. These conditions -- which are testable -- resolve the question of when a market with many assets permits so much diversification that risk-free investment opportunities are available.

1. THE HILBERT SPACE SETTING

We examine a market on which there are an infinite number of random assets, x_i, $i = 1,2,\ldots$. Each x_i represents a random return which can be purchased for $1. The means and variances of these returns are uniformly bounded, so

$$(1.1) \qquad E(x_i) = m_i \leq B_1$$

$$(1.2) \qquad V(x_i) = \sigma_{ii} \leq M ,$$

which, of course, implies

$$(1.3) \qquad E[x_i^2] \leq B = M + B_1^2 .$$

For any finite N, let $\underset{\sim}{x}_N' = (x_1,\ldots,x_N)$, the vector of the first N assets, and let

$$(1.4) \qquad \underset{\sim}{\Sigma}_N = E\{[\underset{\sim}{x}_N - E(\underset{\sim}{x}_N)][\underset{\sim}{x}_N - E(\underset{\sim}{x}_N)]'\}$$

denote the variance-covariance matrix of the random vector $\underset{\sim}{x}_N$. We assume that $\underset{\sim}{\Sigma}_N$ is positive definite for any finite N.

A portfolio is just a finite linear combination of the assets: $\underset{\sim}{p} = \Sigma_1^N \alpha_i x_i$.[5] The portfolio $\underset{\sim}{p}$ has mean

$$(1.5) \qquad E(\underset{\sim}{p}) = \sum_1^N \alpha_i m_i$$

and variance

$$(1.6) \qquad V(\underset{\sim}{p}) = \sum_{i,j=1}^N \alpha_i \sigma_{ij} \alpha_j .$$

Furthermore, the cost of the assets used to purchase the portfolio is just

$$(1.7) \qquad C(\underset{\sim}{p}) = \sum_1^N \alpha_i = \underset{\sim}{\alpha}' \underset{\sim}{e}_N ,$$

where $\underset{\sim}{\alpha}' = (\alpha_1, \ldots, \alpha_N)$ and $\underset{\sim}{e}_N' = (1, \ldots, 1)$, a vector of N ones.

Let F be the space of portfolios. For technical reasons, it is convenient to work not with F, but with the closures (in topologies defined by norms which are introduced below) of this space. These spaces are Hilbert spaces, which the space of portfolios is not. Let $\underset{\sim}{p} = \Sigma_1^N \alpha_i x_i$ and $\underset{\sim}{q} = \Sigma_1^N \beta_i x_i$ be typical elements of F. We consider two inner products on F. Define

$$(1.8) \qquad (\underset{\sim}{p}, \underset{\sim}{q}) = \sum_{i,j=1}^N \alpha_i \sigma_{ij} \beta_j = \mathrm{Cov}(\underset{\sim}{p}, \underset{\sim}{q}) .$$

It is easy to check that $(\underset{\sim}{p},\underset{\sim}{q})$ is an inner product and so

(1.9) $$||\underset{\sim}{p}|| = (\underset{\sim}{p},\underset{\sim}{p})^{1/2} = [V(\underset{\sim}{p})]^{1/2}$$

is a norm. For many purposes this is the most convenient norm with which to work. However, if there is a riskless asset (defined rigorously in Section 2 below), it is sometimes more convenient to work with the more conventional mean-square norm.[6] Again for $\underset{\sim}{p},\underset{\sim}{q} \in F$, let

(1.10) $$(\underset{\sim}{p},\underset{\sim}{q})_* = \mathrm{Cov}(\underset{\sim}{p},\underset{\sim}{q}) + E(\underset{\sim}{p})\, E(\underset{\sim}{q})\ ;$$

$(\underset{\sim}{p},\underset{\sim}{q})_*$ is an inner product so that

(1.11) $$||p||_* = (\underset{\sim}{p},\underset{\sim}{p})_*^{1/2}$$

is a norm.

In either of the norms $||\ ||$ or $||\ ||_*$, F is a pre-Hilbert space. The closure of F in the topology defined by $||\ ||$ will be denoted P, and the closure of F in the topology defined by $||\ ||_*$ will be denoted P_*. Both P and P_* are Hilbert spaces. We shall call elements of P and P^* limit portfolios and mean-square limit portfolios respectively. As the name implies, any limit portfolio or any mean-square limit portfolio can be approximated arbitrarily closely by a finite linear combination of assets.

Note that the cost of a portfolio, $C(\underset{\sim}{p})$, is a linear functional on F as is $E(\underset{\sim}{p})$, the mean of a portfolio. Much of the analysis that follows is a straightforward application of the following elementary facts about linear functionals on Hilbert spaces. For an entertaining and economical presentation of these facts, see Reed and Simon [12]. A linear functional, L( ), on a linear space F (with norm d( ) and inner product <,>) is continuous if and only if

$$\sup_{\underset{\sim}{y} \neq \underset{\sim}{0}} \frac{|L(\underset{\sim}{y})|}{d(\underset{\sim}{y})} < \infty .$$

An equivalent condition is that

$$d(\underset{\sim}{y}_N) \to 0 \quad \text{implies} \quad |L(\underset{\sim}{y}_N)| \to 0 .$$

We shall often have occasion to use:

<u>The B.L.T. Theorem</u>: If L is a continuous linear functional on a normed linear space, F, L can be extended uniquely to a continuous linear functional on $\overline{F}$, the closure of F;

and

<u>The Reisz Representation Theorem</u>: If L is a continuous linear functional on a Hilbert space H with inner product <,>, then there is a unique element $\underset{\sim}{h} \in H$ such that for all $\underset{\sim}{y} \in H$

$$L(\underset{\sim}{y}) = \langle \underset{\sim}{y}, \underset{\sim}{h} \rangle .$$

We also remind the reader that any closed linear subspace of a Hilbert space is itself a Hilbert space. Together these facts imply the following result, which we use repeatedly below.

If L is a linear functional on a linear space F contained in a Hilbert space H, the following statements are equivalent.

(1.12) If $\{y_N\}$ is a sequence in F such that $d(y_N) \to 0$, then $L(y_N) \to 0$;

(1.13) There is a unique $f \in \overline{F}$ such that for $y \in F$, $L(y) = \langle y, f \rangle$.

As we shall be considering two norms, it will be necessary to distinguish between convergence with respect to the two norms. By $p_N \to p$ we denote convergence in the $\| \ \|$ norm. By $p_N \to_* p$ we denote convergence in mean square or convergence in the $\| \ \|_*$ norm. Similarly, when we write that a linear functional on F is continuous, we mean continuous with respect to the $\| \ \|$ norm. If L is a linear functional on F, if $\|p_N\|_* \to 0$ implies $L(p_N) \to 0$, then we shall write that L is continuous$_*$.

2. ARBITRAGE

We consider now what it means for there to be no arbitrage opportunities on the asset market. Consider first finite portfolios. A minimal definition would be the absence of opportunities to make positive returns with no risk and no investment. That is to say, if $\{p_N\}$ is a sequence of finite portfolios,

$$\text{(A)} \qquad V(\underset{\sim}{p}_N) \to 0 \text{ and } C(\underset{\sim}{p}_N) \to 0 \text{ imply } E(\underset{\sim}{p}_N) \to 0 .$$

We shall show that this definition is equivalent to some other possibly broader definitions of the absence of arbitrage profits. We show first that (A) is equivalent to an apparently somewhat weaker condition. Let $H_F = \{\underset{\sim}{p} \in F \mid C(\underset{\sim}{p}) = 0\}$ and let $\bar{H}_F$ be the closure of H_F in P. We shall call H_F the hedge space as it consists of portfolios with no net cost.

PROPOSITION 1: Condition (A) is equivalent to $E(\ )$ being a continuous linear functional on H_F.

PROOF: Obviously (A) implies $E(\ )$ is continuous on H_F. To prove the converse, let $\underset{\sim}{p}_N$ be a sequence of portfolios such that $\|p_N\| \to 0$ and $C(p_N) \to 0$. Let $\underset{\sim}{z}$ be the portfolio formed by investing \$1 in the first asset and consider $\underset{\sim}{q}_N = \underset{\sim}{p}_N - C(\underset{\sim}{p}_N)\underset{\sim}{z}$. Clearly $C(\underset{\sim}{q}_N) = 0$ and $\underset{\sim}{q}_N \in H_F$. However, $\|\underset{\sim}{q}_N\| \le \|\underset{\sim}{p}_N\| + |C(\underset{\sim}{p}_N)| \, \|z\| \to 0$. Thus, since $E(\ )$ is continuous on H_F, $E(\underset{\sim}{q}_N) = E(\underset{\sim}{p}_N) - C(\underset{\sim}{p}_N)\, E(\underset{\sim}{z}) \to 0$. This implies $E(\underset{\sim}{p}_N) = E(\underset{\sim}{q}_N) + C(\underset{\sim}{p}_N)\, E(\underset{\sim}{z}) \to 0$.

Q.E.D.

The implications of condition (A) depend on whether or not there is a riskless asset, which we define as follows:

(R) There is a sequence of finite portfolios $\{\underset{\sim}{q}_N\}$ such that $V(\underset{\sim}{q}_N) \to 0$ while $C(\underset{\sim}{q}_N) \to 1$.

If (R) does not hold, we shall say there is no riskless asset. Necessary and sufficient conditions for the existence of a riskless asset are given in Proposition 17. If $E(\ )$ is continuous on H_F, it can be extended uniquely to a continuous functional $\hat{E}(\ )$ on $\overline{H}_F$, the $\|\ \|$ closure of H_F. However, this functional may not correspond to the expected value of the corresponding portfolio. As we demonstrate in Proposition 6 below, if there is a limiting riskless asset, the $\|\ \|$ closure of H_F is P. Suppose the x_i's are i.i.d. random variables with variance σ^2 and mean m. Then as Proposition 17 shows, there is a limiting riskless asset. Consider the limit portfolio $\underset{\sim}{p} = \Sigma_1^\infty x_i i^{-1}$. Then $\underset{\sim}{p} \in P$ since $\|\underset{\sim}{p}\| = (\sigma^2 \Sigma_1^\infty i^{-2})^{1/2} = \sigma\pi/\sqrt{6} < \infty$; but $E(\underset{\sim}{p}) = m\, \Sigma_1^\infty i^{-1} = \infty$, so $E(\ )$ is not a continuous linear functional on P. Thus $E(\ )$ and $\hat{E}(\ )$ must be different. Note also that in this case, $E(\ )$ is not continuous on F. If there is no riskless asset, the situation is very different.

PROPOSITION 2: Suppose (A) holds. There is no riskless asset if and only if $E(\ )$ and $C(\ )$ are both continuous on F.

PROOF: The negation of (R) is equivalent to the continuity of $C(\ )$ on F. Condition (A) implies that whenever $C(\ )$ is continuous on F, $E(\ )$ is also continuous on F.

Q.E.D.

Note that in contrast to this result, $E(\ )$ is always continuous$_*$ as $(p_N, p_N)_* = (p_N, p_N) + [E(p_N)]^2 \to 0$ implies $E(p_N) \to 0$.

Condition (A) implies there is no way to make a riskless positive return unless there is a riskless asset. If there is a riskless asset, the situation is different.

PROPOSITION 3: There is a riskless asset if $\|q_N\| \to 0$ and $C(q_N) \to 1$; then $E(q_N)$ converges to a unique limit r if and only if (A) holds.

PROOF: Suppose (A) holds. Note first that $|E(q_N)|$ is uniformly bounded. Suppose $|E(q_N)| \uparrow \infty$ (perhaps on a subsequence). Consider $p_N = |E(q_N)|^{-1} q_N$. Clearly $|E(p_N)| = 1$ for all N. For N sufficiently large, $\|p_N\| < \|q_N\| \to 0$. But $C(p_N) = |E(q_N)|^{-1} C(q_N) \to 0$. Thus (A) implies $E(p_N) \to 0$, a contradiction.

Since $E(q_N)$ is bounded, it either converges or it has convergent subsequences converging to different limits. Suppose $\{q_N\}$ and $\{\hat{q}_N\}$ are sequences satisfying $\|q_N\| \to 0$, $\|\hat{q}_N\| \to 0$, $C(q_N) \to 1$, $C(\hat{q}_N) \to 1$, and $E(q_N) \to r$ while $E(\hat{q}_N) \to \hat{r}$. Consider $p_N = q_N - \hat{q}_N$. $\|p_N\| \leq \|q_N\| + \|\hat{q}_N\| \to 0$. But $C(p_N) = C(q_N) - C(\hat{q}_N) \to 0$. Thus (A) implies $E(p_N) \to 0$. But $E(p_N) \to r - \hat{r}$.

To prove the converse, suppose that there is a sequence $\underset{\sim}{p}_N$ such that $V(\underset{\sim}{p}_N) \to 0$ and $C(\underset{\sim}{p}_N) \to 0$; we must show that $E(\underset{\sim}{p}_N) \to 0$. Let $\underset{\sim}{q}_N$ be any sequence such that $V(\underset{\sim}{q}_N) \to 0$ and $C(\underset{\sim}{q}_N) \to 1$. Then, by assumption $E(\underset{\sim}{q}_N) \to r$. Consider $\underset{\sim}{z}_N = \underset{\sim}{p}_N + \underset{\sim}{q}_N$. $||\underset{\sim}{z}_N|| \leq ||\underset{\sim}{p}_N|| + ||\underset{\sim}{q}_N|| \to 0$ and $C(\underset{\sim}{z}_N) = C(\underset{\sim}{p}_N) + C(\underset{\sim}{q}_N) \to 1$. Thus $E(\underset{\sim}{z}_N) \to r$. Since $E(\underset{\sim}{z}_N) = E(\underset{\sim}{p}_N) + E(\underset{\sim}{q}_N)$ and $E(\underset{\sim}{q}_N) \to r$, it follows that $E(\underset{\sim}{p}_N) \to 0$.

Q.E.D.

This result allows us to define r as the rate of return on the riskless asset.

It is desirable now to extend our definition of no arbitrage opportunities. If the rate of return r is zero, then investors face no budget constraints. It is possible by selling the riskless asset short to generate arbitrarily large amounts of cash, which can be used to purchase investments or, in a complete model, for current consumption, while incurring no future obligations. Thus we define the lack of arbitrage opportunities as

(A_*) If there is a riskless asset, its rate of return is unique and positive.

All sequences $\underset{\sim}{q}_N$ such that $||\underset{\sim}{q}_N|| \to 0$ have as their limit the vector zero. In P, the riskless asset is indistinguishable from $\underset{\sim}{0}$. Thus it is very hard to speak of a riskless asset in P. However, the sequence whose existence is asserted in (R) converges in mean square to a well-defined limit in P_*, which is distinct from $\underset{\sim}{0}$ if $r \neq 0$.

COROLLARY 1: If $||\underset{\sim}{g}_N|| \to 0$, $C(\underset{\sim}{g}_N) \to 1$, then (A_*) implies there is a mean-square limit portfolio $\underset{\sim}{s} \in P_*$ such that $\underset{\sim}{g}_N \to_* \underset{\sim}{s}$. Furthermore $\underset{\sim}{s} \neq \underset{\sim}{0}$.

PROOF: Since $E(\underset{\sim}{g}_N) \to r$ and $||\underset{\sim}{g}_N||_* = (||\underset{\sim}{g}_N||^2 + [E(\underset{\sim}{g}_N)]^2)^{1/2}$, it follows that $\{\underset{\sim}{g}_N\}$ is a Cauchy sequence in P_*. Thus there is an $\underset{\sim}{s} \in P_*$ such that $\underset{\sim}{g}_N \to_* \underset{\sim}{s}$. That $r \neq 0$ implies $\underset{\sim}{s} \neq \underset{\sim}{0}$ follows from the fact that $E(\underset{\sim}{g}_N) \to E(\underset{\sim}{s})$ ($E(\ )$ is continuous$_*$).

Q.E.D.

It is worth noting that $\underset{\sim}{s}/r$ represents $E(\ )$ in the $\|\ \|_*$ norm; hence $\underset{\sim}{s}$ is unique.

COROLLARY 2: $E(\underset{\sim}{p}) = \frac{1}{r}(\underset{\sim}{p}, \underset{\sim}{s})_*$.

PROOF: Since $E(\underset{\sim}{p})$ is continuous$_*$, there is a unique $\underset{\sim}{a}_*$ belonging to P_* such that

$$(2.1) \qquad E(\underset{\sim}{p}) = (\underset{\sim}{p}, \underset{\sim}{a}_*)_* = \mathrm{Cov}(\underset{\sim}{p}, \underset{\sim}{a}_*) + E(\underset{\sim}{p})\, E(\underset{\sim}{a}_*) .$$

But $\underset{\sim}{a}_* = \underset{\sim}{s}/r$ satisfies (2.1) for all $\underset{\sim}{p}$.

Q.E.D.

PROPOSITION 4: Suppose (A) holds. If there is a riskless asset, then the linear functional $L_\rho(\ )$ given by

$$L_\rho(\underset{\sim}{p}) = E(\underset{\sim}{p}) - \rho\, C(\underset{\sim}{p})$$

is continuous on F if and only if ρ is the rate of return on the riskless asset.

Since $L_\rho(\ )$ is continuous if and only if $||\underset{\sim}{p}_N|| \to 0$ implies $L_\rho(\underset{\sim}{p}_N) \to 0$, this proposition implies

COROLLARY 3: Suppose (A) holds. If there is a riskless asset with rate of return r, $\|\underset{\sim}{p}_N\| \to 0$ implies $E(\underset{\sim}{p}_N) - r\,C(\underset{\sim}{p}_N) \to 0$.

PROOF OF PROPOSITION 4: Since $E(\ )$ is continuous on the hedge space (H_F), there is a unique limit portfolio $\underset{\sim}{h} \in \overline{H}_F$ such that $E(\underset{\sim}{z}) = (\underset{\sim}{z}, \underset{\sim}{h})$ for $\underset{\sim}{z} \in H_F$. Let $\{\underset{\sim}{q}_N\}$ be a sequence of portfolios in F such that $\|\underset{\sim}{q}_N\| \to 0$ and $C(\underset{\sim}{q}_N) = 1$. Proposition 3 implies $E(\underset{\sim}{q}_N) \to r$. Let $\underset{\sim}{p}$ be an arbitrary element of F and define $\underset{\sim}{z}_N = \underset{\sim}{p} - C(\underset{\sim}{p})\,\underset{\sim}{q}_N$. Then
$E(\underset{\sim}{z}_N) = E(\underset{\sim}{p}) - C(\underset{\sim}{p})\,E(\underset{\sim}{q}_N) \to E(p) - r\,C(\underset{\sim}{p})$. Since $\underset{\sim}{z}_N \in H_F$, we have
$E(\underset{\sim}{z}_N) = (\underset{\sim}{z}_N, \underset{\sim}{h}) = (p, \underset{\sim}{h}) - C(\underset{\sim}{p})\,(\underset{\sim}{q}_N, \underset{\sim}{h}) \to (\underset{\sim}{p}, \underset{\sim}{h})$; hence $E(\underset{\sim}{p}) - r\,C(\underset{\sim}{p}) = (\underset{\sim}{p}, \underset{\sim}{h})$, and so $L_r(\ )$ is continuous on F. If $\rho \neq r$, then
$L_\rho(\underset{\sim}{p}) = L_r(\underset{\sim}{p}) + (r - \rho)C(\underset{\sim}{p})$ is continuous on F only if $C(\ )$ is continuous on F; but $C(\ )$ is not continuous if there is a riskless asset.

Q.E.D.

PROPOSITION 5: Condition (A_*) implies $C(\ )$ is continuous$_*$.

PROOF: Consider a sequence $\underset{\sim}{p}_N \in F$ such that $\|\underset{\sim}{p}_N\|_* \to 0$. If $C(\underset{\sim}{p}_N)$ does not converge to 0, there is an $\varepsilon > 0$ and a subsequence with $|C(\underset{\sim}{p}_N)| \geq \varepsilon$. Let $\underset{\sim}{q}_N = \underset{\sim}{p}_N / C(\underset{\sim}{p}_N)$. Then along the subsequence we have $C(\underset{\sim}{q}_N) = 1$, $\|\underset{\sim}{q}_N\|_* = \|\underset{\sim}{p}_N\|_* / |C(\underset{\sim}{p}_N)| \leq \|\underset{\sim}{p}_N\|_* / \varepsilon \to 0$; hence there is a riskless asset with a zero rate of return, which violates (A_*).

Q.E.D.

If there is a riskless asset, any portfolio may be approximated (in the $\| \ \|$ norm) by a hedge portfolio.

PROPOSITION 6: If there is a riskless asset, the $\| \ \|$ closure of H_F is equal to P.

PROOF: Let $\underset{\sim}{p} \in F$ and consider a sequence of finite portfolios $\underset{\sim}{q}_N$ such that $\|\underset{\sim}{q}_N\| \to 0$ while $C(\underset{\sim}{q}_N) = 1$. Define

$$\underset{\sim}{y}_N = \underset{\sim}{p} - C(\underset{\sim}{p})\, \underset{\sim}{q}_N\, ;$$

then $C(\underset{\sim}{y}_N) = C(\underset{\sim}{p}) - C(\underset{\sim}{p}) = 0$, so $\underset{\sim}{y}_N \in H_F$. However $\|\underset{\sim}{y}_N - \underset{\sim}{p}\| = |C(\underset{\sim}{p})|\ \|\underset{\sim}{q}_N\|$; so $\|\underset{\sim}{y}_N - \underset{\sim}{p}\| \to 0$, which implies $\underset{\sim}{y}_N \to \underset{\sim}{p}$.

Q.E.D.

We close this section by noting that the lack of arbitrage opportunities is equivalent to the distance between the vector of means $\underset{\sim}{m}_N$ and a vector of ones $\underset{\sim}{e}_N$ being uniformly bounded in the $\underset{\sim}{\Sigma}_N^{-1}$ norm. In the next section below, we shall show how this bound, which we denote δ, is related to (and can be computed from) the mean-variance efficiency frontier. The bound δ is just the risk premium on an efficient risky portfolio.

If $\underset{\sim}{y}$ and $\underset{\sim}{z}$ are vectors in R^N and $\underset{\sim}{Q}$ is a positive-definite N×N matrix, then let

$$\|\underset{\sim}{y}\|_Q = (\underset{\sim}{y}'\, \underset{\sim}{Q}\, \underset{\sim}{y})^{1/2}$$

and

$$\|\underset{\sim}{y} - [\underset{\sim}{z}]\|_{\underset{\sim}{Q}} = \underset{\alpha}{\text{Min}}\ [(\underset{\sim}{y} - \alpha \underset{\sim}{z})'\, \underset{\sim}{Q}(\underset{\sim}{y} - \alpha \underset{\sim}{z})]^{1/2} .$$

PROPOSITION 7: Condition (A) is equivalent to the existence of a least upper bound δ such that

$$\|\underset{\sim}{m}_N - [\underset{\sim}{e}_N]\|_{\underset{\sim}{\Sigma}_N^{-1}} \leq \delta$$

for all N.

PROOF: It is easy to check that the values of the solutions to the following two problems are the same:

$$\text{(2.2)} \qquad \underset{\alpha}{\text{Min}} \left[(\underset{\sim}{m}_N - \alpha\, \underset{\sim}{e}_N)' \, \underset{\sim}{\Sigma}_N^{-1} \, (\underset{\sim}{m}_N - \alpha\, \underset{\sim}{e}_N)\right]^{1/2}$$

and

$$\text{(2.3)} \qquad \text{Max}\,|E(\underset{\sim}{p}_N)| \text{ subject to } \underset{\sim}{p}_N \in R^N, \quad C(\underset{\sim}{p}_N) = 0, \quad \|\underset{\sim}{p}_N\| = 1 .$$

However, since (A) is equivalent to E( ) being continuous on H_F, the values of solutions to (2.3) will remain bounded if and only if (A) holds.

Q.E.D.

In the proof of Proposition 4, we defined $\underset{\sim}{h}$ as the unique limit portfolio in $\bar{H}_F$ that represents E( ) on H_F. It is important to realize that $\underset{\sim}{h}$ is well defined if there is a riskless asset (as in Proposition 4) or if there is not. If there is no riskless asset, we can explicitly calculate that

$$\text{(2.4)} \qquad \underset{\sim}{h} = \underset{\sim}{a} - [(\underset{\sim}{a},\underset{\sim}{c})/(\underset{\sim}{c},\underset{\sim}{c})]\ \underset{\sim}{c},$$

where $\underset{\sim}{a}$ and $\underset{\sim}{c}$ are the limit portfolios in P that represent E( ) and C( ).

PROPOSITION 8: $||\underset{\sim}{h}|| = \delta$, where δ is the bound given in Proposition 7.

PROOF: (1) Let $\delta_N = \text{Max } |E(\underset{\sim}{p}_N)|$ subject to $\underset{\sim}{p}_N \in R^N$, $||\underset{\sim}{p}_N|| = 1$, $C(\underset{\sim}{p}_N) = 0$. We shall show that $\delta_N \to ||\underset{\sim}{h}||$. Let $\underset{\sim}{z}_n \in H_F$ be a sequence of hedge portfolios converging to $\underset{\sim}{h}$ as $n \to \infty$. Consider $\underset{\sim}{g}_n = \underset{\sim}{z}_n/||\underset{\sim}{z}_n||$. Thus $C(\underset{\sim}{g}_n) = 0$ and $||\underset{\sim}{g}_n|| = 1$; hence $\delta_N \geq E(\underset{\sim}{g}_n)$ for some N. But $E(\underset{\sim}{g}_n) = (\underset{\sim}{g}_n,\underset{\sim}{h}) = (\underset{\sim}{z}_n,\underset{\sim}{h})/||\underset{\sim}{z}_n|| \to (\underset{\sim}{h},\underset{\sim}{h})/||\underset{\sim}{h}|| = ||\underset{\sim}{h}||$.

(2) For any hedge portfolio $\underset{\sim}{p}$, $E(\underset{\sim}{p}) = (\underset{\sim}{p},\underset{\sim}{h})$ so that $|E(\underset{\sim}{p})| \leq ||\underset{\sim}{p}|| \; ||\underset{\sim}{h}||$; hence $||\underset{\sim}{h}|| \geq \delta_N$.

Q.E.D.

In the next section we shall see that $||\underset{\sim}{h}||$ or δ is the slope of the tradeoff between risk and return on the mean-variance efficient set.

3. MEAN-VARIANCE EFFICIENCY

Roll [14] and Ross [17] have shown that the empirical content of the CAPM is contained in the observation that the market portfolio is on the mean-variance efficiency frontier. If arbitrage opportunities exist on an infinite market -- that is, if (A) does not hold -- then there is no trade-off between mean and variance. If the mean functional is not continuous on the hedge space, then there exist finite hedge portfolios with arbitrarily large means and arbitrarily small variances. If (A) does hold, there is a well-defined tradeoff between mean and variance; the set of portfolios which are mean-variance efficient is a subset of P. In this section we show that the relationship between the absence of arbitrage opportunities and the mean-variance efficient set is even closer. We show in Proposition 9 that if there is not a riskless asset, the mean-variance efficient set is a cone generated by the limit portfolios that represent the linear functionals E( ), -E( ), and -C( ) in the $\| \ \|$ norm. As was proved in Proposition 2 above, the existence of limit portfolios that represent these functionals is, if there is no riskless asset, equivalent to condition (A).

If there is a riskless asset, the same result holds. We showed in Proposition 5 above that the extended arbitrage condition (A_*) implies that C( ) is continuous$_*$, so that there is an element $\tilde{c}_* \in P_*$ such that $C(\tilde{p}) = (\tilde{c}_*, \tilde{p})_*$ for all $\tilde{p} \in P_*$. In Corollary 1 we showed that (A_*) implies that there is a riskless asset $\tilde{s} \in P_*$ and that $\tilde{s} \neq \tilde{0}$. In Corollary 2 we showed that $\tilde{a}_* = \tilde{s}/r$ represents E( ) in the $\| \ \|_*$ norm. In Proposition 10 we show that if there is a riskless asset, the mean-variance efficient set is the cone generated by $\tilde{s}$, $-\tilde{s}$, and $-\tilde{c}_*$; of course this is the same as the cone generated by $\tilde{a}_*$, $-\tilde{a}_*$, and $-\tilde{c}_*$. Proposition 11 gives another representation of the mean-variance efficient set. It is the cone generated by $\tilde{s}$, $-\tilde{s}$, and the limit portfolio $\tilde{h}$ which represents the continuous linear functional $L_r( ) = E( ) - rC( )$.

In either case, the efficient set is a very small subset of P; the CAPM is, in principle, as empirically restrictive on an infinite market as on a finite market. The absence of arbitrage opportunities does not cause asset prices on a large market to exhibit relations like those predicted by the CAPM; condition (A) makes it possible to describe the set of limit portfolios which are mean-variance efficient. When (A) does not hold, the efficiency frontier is trivial; there is no meaningful tradeoff between mean and variance.

The representations of Propositions 9 and 11 make it easy to compute the constant δ of Propositions 7 and 8.

PROPOSITION 9: If there is no riskless asset and (A) holds, then for any portfolio $\underset{\sim}{p}$, there is a limit portfolio $\underset{\sim}{p}^o = \alpha\underset{\sim}{c} + \beta\underset{\sim}{a}$ with $\alpha \leq 0$ such that $E(\underset{\sim}{p}^o) = E(\underset{\sim}{p})$, $V(\underset{\sim}{p}^o) \leq V(\underset{\sim}{p})$, and $C(\underset{\sim}{p}^o) \leq C(\underset{\sim}{p})$.

PROOF: If $\underset{\sim}{c}$ is proportional to $\underset{\sim}{a}$ this is obvious; so assume $\underset{\sim}{c}$ is not proportional to $\underset{\sim}{a}$. Let $\underset{\sim}{d} = \underset{\sim}{c} - \psi\underset{\sim}{a}$, where ψ is chosen so that $(\underset{\sim}{d},\underset{\sim}{a}) = 0$. Consider $\underset{\sim}{p}^o = \alpha\underset{\sim}{c} + \beta\underset{\sim}{a} = \alpha\underset{\sim}{d} + \gamma\underset{\sim}{a}$ where $\gamma = \beta + \alpha\psi$. Choose γ so that

$$E(\underset{\sim}{p}^o) = \gamma(\underset{\sim}{a},\underset{\sim}{a}) = (\underset{\sim}{p},\underset{\sim}{a}) = E(\underset{\sim}{p}) .$$

Let $\underset{\sim}{q}_2 = \underset{\sim}{p} - \gamma\underset{\sim}{a}$ and note that $(\underset{\sim}{q}_2,\underset{\sim}{a}) = 0$. Choose α so that

$$V(\underset{\sim}{p}^o) = \alpha^2(\underset{\sim}{d},\underset{\sim}{d}) + \gamma^2(\underset{\sim}{a},\underset{\sim}{a}) = (\underset{\sim}{q}_2,\underset{\sim}{q}_2) + \gamma^2(\underset{\sim}{a},\underset{\sim}{a}) = V(\underset{\sim}{p})$$

with $\alpha \leq 0$. Then

$$C(\underset{\sim}{p}) - \psi E(\underset{\sim}{p}) = (\underset{\sim}{q}_2,\underset{\sim}{d}) ,$$

$$C(\underset{\sim}{p}^o) - \psi E(\underset{\sim}{p}^o) = \alpha(\underset{\sim}{d},\underset{\sim}{d}) = -(\underset{\sim}{q}_2,\underset{\sim}{q}_2)^{1/2}(\underset{\sim}{d},\underset{\sim}{d})^{1/2} ,$$

so that $E(\underset{\sim}{p}) = E(\underset{\sim}{p}^o)$ implies

$$C(\underset{\sim}{p}) - C(\underset{\sim}{p}^o) = (\underset{\sim}{q}_2,\underset{\sim}{d}) + (\underset{\sim}{q}_2,\underset{\sim}{q}_2)^{1/2}(\underset{\sim}{d},\underset{\sim}{d})^{1/2} \geq 0$$

by the Cauchy-Schwarz inequality. The inequality is strict unless $\underset{\sim}{p} = \underset{\sim}{p}^o$.

Q.E.D.

We have shown that (mean-variance) efficient limit portfolios have the form $\tilde{p} = \alpha\tilde{c} + \beta\tilde{a}$ when there is no riskless asset. As in (2.4), we can set $\tilde{h} = \tilde{a} - [(\tilde{a},\tilde{c})/(\tilde{c},\tilde{c})]\, \tilde{c}$; so if $\tilde{p}$ and $\tilde{q}$ are efficient limit portfolios with $C(\tilde{p}) = C(\tilde{q})$, it follows that $\tilde{p} - \tilde{q} = \psi\, \tilde{h}$ for some ψ, which implies that $V(\tilde{p}-\tilde{q}) = \psi^2||\tilde{h}||^2$. Since $\tilde{p}-\tilde{q} \in \bar{H}_F$, we have $E(\tilde{p}-\tilde{q}) = (\tilde{p}-\tilde{q}, \tilde{h})$; hence

$$(3.1) \qquad |E(\tilde{p}-\tilde{q})| = \delta[V(\tilde{p}-\tilde{q})]^{1/2} .$$

PROPOSITION 10: If there is a riskless asset and (A_*) holds, then for any portfolio $\tilde{p}$ there is a mean-square limit portfolio $\tilde{p}^o = \alpha\tilde{c}_* + \beta\tilde{s}$ $(= \alpha\tilde{c}_* + \beta\, r\tilde{a}_*)$ with $\alpha \leq 0$ such that $E(\tilde{p}^o) = E(\tilde{p})$, $V(\tilde{p}^o) \leq V(\tilde{p})$, and $C(\tilde{p}^o) \leq C(\tilde{p})$.

PROOF: The result follows from Proposition 11.

Since (A) implies that $E(\ )$ is continuous on H_F, there is a unique $\tilde{h} \in \bar{H}_F$ such that $E(\tilde{p}) = (\tilde{p},\tilde{h})$ for $\tilde{p} \in H_F$. If $\tilde{q}_N \in H_F$ and $\tilde{q}_N \to \tilde{h}$, then $E(\tilde{q}_N)$ is a Cauchy sequence; hence $\tilde{q}_N \to_* \tilde{h}$ and $\tilde{h} \in P_*$. We know from the proof of Proposition 4 that $E(\tilde{p}) - r\,C(\tilde{p}) = (\tilde{p},\tilde{h})$ for $\tilde{p} \in F$. An explicit representation for $\tilde{h} \in P_*$ is $\tilde{h} = -r\tilde{c}_* + rC(\tilde{c}_*)\tilde{s}$. For then $C(\tilde{h}) = 0$ and

$$\begin{aligned} E(\tilde{p}) - rC(\tilde{p}) &= (\tilde{p}, \frac{\tilde{s}}{r} - r\tilde{c}_*)_* \\ &= (\tilde{p}, -r\tilde{c}_*) + E(\tilde{p})\left[1 - r\,(\tilde{c}_*, \tilde{s}/r)_*\right] \\ &= (\tilde{p}, \tilde{h}) + E(\tilde{p})\left[1 - C(\tilde{s})\right] = (\tilde{p}, \tilde{h}). \end{aligned}$$

Another representation of the efficiency frontier is given by

PROPOSITION 11: If there is a riskless asset and (A_*) holds, then for any portfolio $\underset{\sim}{p}$ there is a mean-square limit portfolio

$$\underset{\sim}{p}^{o} = \alpha \underset{\sim}{h} + \beta \underset{\sim}{s} \tag{3.2}$$

with $\alpha \geq 0$ such that $E(\underset{\sim}{p}^{o}) = E(\underset{\sim}{p})$, $V(\underset{\sim}{p}^{o}) \leq V(\underset{\sim}{p})$, and $C(\underset{\sim}{p}^{o}) \leq C(\underset{\sim}{p})$.

PROOF: Observe first that if $E(\underset{\sim}{p}) \leq r\,C(\underset{\sim}{p})$, the constraints can be satisfied by a portfolio consisting solely of the safe asset $\underset{\sim}{s}$. Assume now that $E(\underset{\sim}{p}) > r\,C(\underset{\sim}{p})$, so that $\alpha = (E(\underset{\sim}{p}) - r\,C(\underset{\sim}{p}))/(\underset{\sim}{h},\underset{\sim}{h}) = (\underset{\sim}{p},\underset{\sim}{h})/(\underset{\sim}{h},\underset{\sim}{h}) \geq 0$. Let $\beta = C(\underset{\sim}{p})$. Then $E(p^{o}) = E(p)$ and $C(\underset{\sim}{p}^{o}) = C(\underset{\sim}{p})$, but $V(\underset{\sim}{p}^{o}) = (\underset{\sim}{p},\underset{\sim}{h})^2/(\underset{\sim}{h},\underset{\sim}{h}) \leq (\underset{\sim}{p},\underset{\sim}{p}) = V(\underset{\sim}{p})$. The inequality is strict unless $\underset{\sim}{p} = \underset{\sim}{p}^{o}$.

Q.E.D.

Consider the tradeoff between risk and return available to an investor if he invests a dollar in a portfolio of the form (3.2). Since a dollar is to be invested, $\beta = 1$, but $\alpha \geq 0$ is arbitrary. The return on such a portfolio is $\mu = r + \alpha\, E(\underset{\sim}{h}) = r + \alpha \|\underset{\sim}{h}\|^2$; risk, as measured by standard deviation, is $\sigma = \alpha \|\underset{\sim}{h}\|$; thus, for an efficient portfolio,

$$\mu = r + \sigma \|\underset{\sim}{h}\| ; \tag{3.3}$$

$\|\underset{\sim}{h}\|$ or δ is the slope of the tradeoff between risk and return.

The mean-variance efficient set, constructed from the risky assets alone, is always drawn (in mean-standard deviation space) as a curve. If there is a limiting riskless asset, then representation (3.3) implies that the curve is a straight line.

4. FACTOR STRUCTURE AND ROSS' THEOREM

Ross [16,18] proved that if there were no arbitrage opportunities, then expected returns would be approximately linear in factor loadings. In this section we show that our definition of there being no arbitrage opportunities is, with some minor additional assumptions, equivalent to Ross' result. For an alternative proof of Ross' theorem, see Huberman [7].

We shall say that there is a factor structure if, for every N, the variance-covariance matrix can be decomposed as follows:

$$(4.1) \qquad \underset{\sim}{\Sigma}_N = \underset{\sim}{D}_N + \underset{\sim}{B}_N \underset{\sim}{B}'_N ,$$

where $\underset{\sim}{D}_N$ is a diagonal matrix and $\underset{\sim}{B}_N$ is an N by K matrix. This decomposition will hold when the return on the i^{th} asset is generated by

$$(4.2) \qquad x_i = m_i + b_{i1} v_1 + \ldots + b_{iK} v_K + \varepsilon_i ,$$

where $v_1, \ldots, v_K$ are random factors with zero mean uncorrelated with the ε_i's. The ε_i's are called idiosyncratic disturbances and are uncorrelated with each other.

Suppose that $\underset{\sim}{y}$ is a N×1 vector, $\underset{\sim}{Z}$ is a N×J matrix, $\underset{\sim}{Q}$ is a N×N positive-definite matrix, and $\underset{\sim}{\tau}$ is a J×1 vector; then define

$$||\underset{\sim}{y} - [\underset{\sim}{Z}]||_{\underset{\sim}{Q}} = \min_{\underset{\sim}{\tau}} [(\underset{\sim}{y}-\underset{\sim}{Z}\underset{\sim}{\tau})' \underset{\sim}{Q}(\underset{\sim}{y}-\underset{\sim}{Z}\underset{\sim}{\tau})]^{1/2}.$$

Let $\underset{\sim}{m}'_N = (m_1, \ldots, m_N)$.

PROPOSITION 12: If there is a factor structure, that is, if (4.2) holds, then condition (A) implies there is a number M_1 such that for all N

(4.3) $$\|\underset{\sim}{m}_N - [\underset{\sim}{e}_N, \underset{\sim}{B}_N]\|_{\underset{\sim}{I}_N} \leq M_1 .$$

($\underset{\sim}{I}_N$ is the identity matrix of order N.)

Proposition 12 is a special case of Proposition 14, whose proof is given below. It follows from Proposition 14 that $\|\underset{\sim}{m}_N - [\underset{\sim}{e}_N, \underset{\sim}{B}_N]\|_{\underset{\sim}{I}_N} \leq M_N^{1/2}\delta$, where $M_N = \max\{\sigma_{ii}, i = 1,\ldots,N\}$ and δ is given in Proposition 7. So $M_1 = M^{1/2}\delta$, where $M = \lim_{N\to\infty} M_N$ is finite by assumption. We shall give a numerical example to illustrate this bound; similar calculations are presented in Ross [16].

Table 1 in Roll [14] gives a riskless return (r-1) of .0192 per year, and a mean and standard deviation for a market proxy of .1896 and 1.069. These estimates are taken from Black, Jensen, and Scholes [1] and cover the period 1931 - 1965. If the market proxy is close to a limit portfolio of the form (3.2), then (3.3) and Proposition 8 imply that $\delta \cong (.1896 - .0192)/1.069 = .159$. Roll then uses the mean and standard deviation of the orthogonal (zero-beta) portfolio to calculate that the mean and standard deviation of the Sharpe-Lintner tangent portfolio are .1234 and .5849; the calculation assumes that the market proxy and its orthogonal portfolio are on the mean-variance frontier generated by the risky assets. The Sharpe-Lintner tangent portfolio should converge as $N \to \infty$ to a limit portfolio of the form (3.2); so we obtain

$\delta \cong (.1234 - .0192)/.5849 = .178$. Roll repeats these calculations using data from Morgan [9] for the period July 1962 through December 1972 and a riskless rate of .03. Now the δ estimates are $\delta \cong (.1554 - .03)/1.014 = .124$ and $\delta \cong (.1275 - .03)/.7774 = .125$. Using data from Friend [6] and Myers [10] for the period January 1962 through December 1971, Ross [16] sets the mean and standard deviation on the market proxy at .074 and .123, and he uses .051 for the risk-free rate. He argues that a hedge portfolio earning the market risk premium is unlikely to have less than one-half the market variance; hence a conservative estimate is $\delta \cong \sqrt{2}(.074 - .051)/.123 = \sqrt{2}(.187) = .264$. So the calculations suggest a δ in the .1 to .3 range.

It is instructive to present the bound in terms of the squared multiple correlation coefficient:

$$R_N^2 = 1 - \frac{\| \underset{\sim}{m}_N - [\underset{\sim}{e}_N, \underset{\sim}{B}_N] \|^2_{\underset{\sim}{I}_N}}{\| \underset{\sim}{m}_N - [\underset{\sim}{e}_N] \|^2_{\underset{\sim}{I}_N}} .$$

Table 4.3 in Fama [5] gives the sample means and standard deviations for the thirty common stocks that accounted for the largest fraction of the total value outstanding on the New York Stock Exchange at the end of 1971. The sample period is July 1963 through June 1968. For these thirty stocks, the bound in (4.3) implies a lower bound on R_{30}^2 of .84 if $\delta = .2$ and a lower bound $R_{30}^2 \geq .64$ if $\delta = .3$. Fama's Table 4.4 gives the sample means and standard deviations for thirty randomly selected stocks. The bound for these stocks is $R_{30}^2 \geq .69$ if $\delta = .2$ and $R_{30}^2 \geq .31$ if $\delta = .3$. Combining the two sets of stocks gives $R_{60}^2 \geq .83$ if $\delta = .2$ and $R_{60}^2 \geq .61$ if $\delta = .3$.

The bound in (4.3) implies that the lower bound on R_N^2 tends to one as $N \to \infty$; our calculations indicate that the bound has considerable force even for moderate values of N.

It is natural to ask whether condition (4.3) is equivalent to there being no arbitrage opportunities. That is, we ask whether if (4.2) holds, (4.3) is equivalent to (A). It is easy to see that without further assumptions, (4.3) does not imply (A). It seems to be necessary for the diagonal elements of $\underset{\sim}{D}$ to be bounded away from 0. Recall that by Proposition 7, (A) is equivalent to $||\underset{\sim}{m}_N - [\underset{\sim}{e}_N]||_{\underset{\sim}{\Sigma}_N^{-1}}$ being bounded. Suppose, for example, there are no factors and

$$x_i = (1 - \frac{1}{i}) + \varepsilon_i ,$$

where $V(\varepsilon_i) = 1/i$. Then $||\underset{\sim}{m}_N - [\underset{\sim}{e}_N]||^2_{\underset{\sim}{I}_N} \leq \Sigma_1^N 1/i^2$, which is bounded. However, $||\underset{\sim}{m}_N - [\underset{\sim}{e}_N]||^2_{\underset{\sim}{\Sigma}_N^{-1}} = (\Sigma_1^N 1/i) - 2N/(N+1)$, which is unbounded.

PROPOSITION 13: If (4.2) holds and $\inf_i d_{ii} \geq d > 0$, then (4.3) implies (A).

PROOF: We know that

$$\min_{\theta,\lambda} \|m_N - \theta e_N - B_N \lambda\|_{I_N} \leq M_1 . \tag{4.4}$$

Let θ_N, λ_N solve the minimization problem in (4.4). Clearly it will suffice to show that there is an $M_2 < \infty$ such that

$$\|m_N - \theta_N e_N\|_{\Sigma_N^{-1}} \leq M_2 \tag{4.5}$$

for all N.

Let $r_N = m_N - \theta_N e_N - B_N \lambda_N$. Since $m_N - \theta_N e_N = r_N + B_N \lambda_N$, $\|m_N - \theta_N e_N\|_{\Sigma_N^{-1}} \leq \|r_N\|_{\Sigma_N^{-1}} + \|B_N \lambda_N\|_{\Sigma_N^{-1}}$. We shall use "$A \geq B$" to denote "$A-B$ is positive semi-definite."

$$\Sigma_N = D_N + B_N B_N' \geq c(I + B_N B_N') \geq c\, I,$$

where c = min [1,d]; hence $\Sigma_N^{-1} \leq c^{-1} I$. Since $(I + B B')^{-1} = (I - B(B'B + I)^{-1}B')$, we have

$$\|r_N\|^2_{\Sigma_N^{-1}} \leq c^{-1} \|r_N\|^2_{I_N} \leq c^{-1} M_1^2 ;$$

$$\|B_N \lambda_N\|^2_{\Sigma_N^{-1}} \leq c^{-1} \lambda_N' Q_N \lambda_N ,$$

where $Q_N = B_N' B_N - B_N' B_N (B_N' B_N + I)^{-1} B_N' B_N$. The maximal eigenvalue of Q_N is less than one. Thus

$$\|\underset{\sim}{B}_N \underset{\sim}{\lambda}_N\|^2_{\underset{\sim}{\Sigma}_N^{-1}} \leq c^{-1} \underset{\sim}{\lambda}'_N \underset{\sim}{\lambda}_N .$$

To complete the proof, we need only show that $\underset{\sim}{\lambda}'_N \underset{\sim}{\lambda}_N$ is uniformly bounded.

Let $\underset{\sim}{C}_N = (\underset{\sim}{e}_N, \underset{\sim}{B}_N)$. We shall assume that $\mathrm{rank}(\underset{\sim}{C}_J) = K+1$ for some J. (If $\mathrm{rank}(\underset{\sim}{B}_N) < K$ for all N, then we can respecify the model with a smaller number of factors; if $\underset{\sim}{e}_N$ is in the column space of $\underset{\sim}{B}_N$ for all N, then our argument will apply with $\theta_N = 0$.) Let $\underset{\sim}{\gamma}'_N = (\theta_N, \underset{\sim}{\lambda}'_N)$. For $N \geqq J$ we have

$$M_1 \geq \|\underset{\sim}{m}_N - \underset{\sim}{C}_N \underset{\sim}{\gamma}_N\|_{\underset{\sim}{I}_N} \geq \|\underset{\sim}{m}_J - \underset{\sim}{C}_J \underset{\sim}{\gamma}_N\|_{\underset{\sim}{I}_J}$$

$$\geq \|\underset{\sim}{C}_J \underset{\sim}{\gamma}_N\|_{\underset{\sim}{I}_J} - \|\underset{\sim}{m}_J\|_{\underset{\sim}{I}_J}$$

$$\geq (\rho\, \underset{\sim}{\gamma}'_N \underset{\sim}{\gamma}_N)^{1/2} - \|\underset{\sim}{m}_J\|_{\underset{\sim}{I}_J} ,$$

where ρ is the minimal eigenvalue of $\underset{\sim}{C}'_J \underset{\sim}{C}_J$; $\rho > 0$ since $\underset{\sim}{C}'_J \underset{\sim}{C}_J$ is positive definite. So we have

$$\underset{\sim}{\lambda}'_N \underset{\sim}{\lambda}_N \leq \underset{\sim}{\gamma}'_N \underset{\sim}{\gamma}_N \leq (M_1 + \|\underset{\sim}{m}_J\|_{\underset{\sim}{I}_J})^2/\rho . \qquad \text{Q.E.D.}$$

Up to this point we have assumed, as did Ross [16], that the factor structure is exact. This is both an overly strong assumption and one which is at odds with the way in which factor models are used and tested. It seems more reasonable to suppose that, for any N, $\underset{\sim}{\Sigma}_N$ can be approximated by a factor structure with K_N factors. That is, we suppose that

$$(4.6) \qquad \underset{\sim}{\Sigma}_N = \underset{\sim}{B}_N \underset{\sim}{B}_N' + \underset{\sim}{D}_N + \underset{\sim}{R}_N ,$$

where $\underset{\sim}{B}_N$ is an N by K_N matrix, and $\underset{\sim}{R}_N$ is a residual matrix. In the representation (4.6), the number of factors depends on N, as do the factor loadings. That is, the first row of $\underset{\sim}{B}_N$ may change as N changes, reflecting the fact that as N changes, then so may the best approximating factor structure. This is to be contrasted with the exact factor structure (4.2), where the first N-1 rows of $\underset{\sim}{B}_N$ are $\underset{\sim}{B}_{N-1}$.

Of course the representation (4.6) is only interesting if K_N is much less than N and the approximation error is small. (In applications this requirement generally determines K_N.) We can measure the approximation error by

$$\|\underline{\text{vec}}\ \underset{\sim}{R}_N\|_{\underset{\sim}{\Omega}_N} ,$$

where $\underline{\text{vec}}\ \underset{\sim}{R}_N$ is the column vector formed from the lower triangle of the symmetric matrix $\underset{\sim}{R}_N$ and $\underset{\sim}{\Omega}_N$ is a positive-definite matrix of dimension $N(N+1)/2$. This corresponds to the fitting criteria used in the estimation of factor models. A minimum distance (generalized least squares) approach will, for given K, choose $\underset{\sim}{D}_N$ and $\underset{\sim}{B}_N \underset{\sim}{B}_N'$ to minimize $||\underline{\text{vec}}(\hat{\underset{\sim}{\Sigma}}_N - \underset{\sim}{B}_N \underset{\sim}{B}_N' - \underset{\sim}{D}_N)||_{\underset{\sim}{\Omega}_N}$, where $\hat{\underset{\sim}{\Sigma}}_N$ is the sample covariance matrix and $\underset{\sim}{\Omega}_N$ is an estimate of the inverse of the asymptotic covariance matrix of $\underline{\text{vec}}(\hat{\underset{\sim}{\Sigma}}_N)$. If we have a random sample of T observations from a multivariate normal distribution, then $\underset{\sim}{\Omega}_N$ is based on $\hat{\underset{\sim}{\Sigma}}_N^{-1} \otimes \hat{\underset{\sim}{\Sigma}}_N^{-1}/2$; or we can drop the normality assumption and use the sample fourth moments. A maximum likelihood estimator based on normality assumptions is asymptotically (as $T \to \infty$) equivalent to a minimum distance estimator that uses $\hat{\underset{\sim}{\Sigma}}_N^{-1} \otimes \hat{\underset{\sim}{\Sigma}}_N^{-1}/2$. This equivalence holds even if the assumption of multivariate normality is false (see Chamberlain [2]).

It is natural to ask whether a version of Ross' theorem (Proposition 12) holds for this more general setting. The answer is yes and is given by

PROPOSITION 14: If $||\underline{\text{vec}}\ \underset{\sim}{R}_N||_{\underset{\sim}{\Omega}_N} \leq \alpha_N$, then condition (A) implies that

$$||\underset{\sim}{m}_N - [\underset{\sim}{e}_N, \underset{\sim}{B}_N]||_{\underset{\sim}{I}_N} \leq \sqrt{\ell_N}\ \delta, \tag{4.7}$$

where δ is as in Proposition 7,

$$\ell_N = M_N + (1 + \sqrt{2})\ \alpha_N \rho_N^{-1/2},$$

$$M_N = \max[\sigma_{ii},\ i = 1, \ldots, N],$$

and ρ_N is the smallest eigenvalue of $\underset{\sim}{\Omega}_N$.

PROOF: Since $\underset{\sim}{R}_N = \underset{\sim}{\Sigma}_N - (\underset{\sim}{D}_N + \underset{\sim}{B}_N \underset{\sim}{B}_N')$,

$$\alpha_N^2 \geq ||\underline{\text{vec}}\ (\underset{\sim}{R}_N)||^2_{\underset{\sim}{\Omega}_N} \geq \rho_N\ ||\underline{\text{vec}}\ (\underset{\sim}{R}_N)||^2_{\underset{\sim}{I}_N}$$

$$\geq \frac{1}{2}\ \rho_N\ \text{tr}(\underset{\sim}{R}_N \underset{\sim}{R}_N) = \frac{1}{2}\ \rho_N \sum_1^N \lambda_{iN}^2 \geq \frac{1}{2}\ \rho_N \lambda_{1N}^2 ,$$

where $\lambda_{1N}^2 \geq \cdots \geq \lambda_{NN}^2$ are the squared eigenvalues of $\underset{\sim}{R}_N$. Hence $\lambda_{1N}^2 \leq 2\ \alpha_N^2/\rho_N$ and so

$$\underset{\sim}{R}_N \leq |\lambda_{1N}|\ \underset{\sim}{I}_N \leq \sqrt{2}\ \alpha_N\ \rho_N^{-1/2}\ \underset{\sim}{I}_N .$$

Let g_{iN} be the i^{th} diagonal element of $\underset{\sim}{D}_N + \underset{\sim}{B}_N\underset{\sim}{B}_N'$. Then

$$\alpha_N^2 \geq \rho_N\ ||\underline{\text{vec}}\ (\underset{\sim}{R}_N)||^2_{\underset{\sim}{I}_N} \geq \rho_N \sum_1^N (\sigma_{ii} - g_{iN})^2$$

$$\geq \rho_N \max_{1 \leq i \leq N} (\sigma_{ii} - g_{iN})^2 ,$$

$$g_{iN} \leq \sigma_{ii} + \alpha_N\ \rho_N^{-1/2} \leq M_N + \alpha_N\ \rho_N^{-1/2} ,$$

and so

$$\underset{\sim}{D}_N \leq (M_N + \alpha_N\ \rho_N^{-1/2})\underset{\sim}{I}_N .$$

Hence

$$\underset{\sim}{\Sigma}_N = \underset{\sim}{D}_N + \underset{\sim}{B}_N\ \underset{\sim}{B}_N' + \underset{\sim}{R}_N \leq \ell_N(\underset{\sim}{I}_N + \underset{\sim}{C}_N\ \underset{\sim}{C}_N') ,$$

where $\ell_N = M_N + (1+\sqrt{2})\alpha_N\rho_N^{-1/2}$ and $\underset{\sim}{C}_N = \ell_N^{-1/2}\ \underset{\sim}{B}_N$.

Let $\underset{\sim}{C}_N'\underset{\sim}{C}_N = \underset{\sim}{S}_N\underset{\sim}{\Psi}_N\underset{\sim}{S}_N'$, where $\underset{\sim}{\Psi}_N$ is a diagonal matrix containing the positive eigenvalues and $\underset{\sim}{S}_N'\underset{\sim}{S}_N = \underset{\sim}{I}$. Let $(\underset{\sim}{C}_N'\underset{\sim}{C}_N)^- = \underset{\sim}{S}_N\underset{\sim}{\Psi}_N^{-1}\underset{\sim}{S}_N'$ and note that $\underset{\sim}{C}_N(\underset{\sim}{C}_N'\underset{\sim}{C}_N)^-\underset{\sim}{C}_N'\underset{\sim}{z}$ is the orthogonal projection (relative to $\underset{\sim}{I}_N$) of $\underset{\sim}{z}$ onto the column space of $\underset{\sim}{C}_N$; this is clearly identical to the orthogonal projection onto the column space of $\underset{\sim}{B}_N$.

$$\underset{\sim}{\Sigma}_N^{-1} \geq \ell_N^{-1}(\underset{\sim}{I}_N + \underset{\sim}{C}_N\underset{\sim}{C}_N')^{-1}$$

$$= \ell_N^{-1}(\underset{\sim}{I}_N - \underset{\sim}{C}_N(\underset{\sim}{C}_N'\underset{\sim}{C}_N + \underset{\sim}{I})^{-1}\underset{\sim}{C}_N')$$

$$\geq \ell_N^{-1}(\underset{\sim}{I}_N - \underset{\sim}{C}_N(\underset{\sim}{C}_N'\underset{\sim}{C}_N)^-\underset{\sim}{C}_N') = \ell_N^{-1}(\underset{\sim}{I}_N - \underset{\sim}{B}_N(\underset{\sim}{B}_N'\underset{\sim}{B}_N)^-\underset{\sim}{B}_N').$$

$$\delta^2 \geq \|\underset{\sim}{m}_N - [\underset{\sim}{e}_N]\|^2_{\underset{\sim}{\Sigma}_N^{-1}}$$

$$\geq \ell_N^{-1} \min_{\alpha}(\underset{\sim}{m}_N - \alpha\,\underset{\sim}{e}_N)'(\underset{\sim}{I} - \underset{\sim}{B}_N(\underset{\sim}{B}_N'\underset{\sim}{B}_N)^-\underset{\sim}{B}_N')(\underset{\sim}{m}_N - \alpha\,\underset{\sim}{e}_N)$$

$$= \ell_N^{-1} \min_{\alpha} \min_{\underset{\sim}{\beta}}(\underset{\sim}{m}_N - \alpha\,\underset{\sim}{e}_N - \underset{\sim}{B}_N\underset{\sim}{\beta})'(\underset{\sim}{m}_N - \alpha\,\underset{\sim}{e}_N - \underset{\sim}{B}_N\underset{\sim}{\beta})$$

$$= \ell_N^{-1}\|\underset{\sim}{m}_N - [\underset{\sim}{e}_N, \underset{\sim}{B}_N]\|^2_{\underset{\sim}{I}_N}.$$

Q.E.D.

Note that we have not shown that $\underset{\sim}{B}_N\underset{\sim}{B}_N'$ converges in some appropriate sense as $N \to \infty$. In the one-factor case, Proposition 14 implies that if $\rho_N \geq \rho > 0$ and $\alpha_N \leq \alpha < \infty$ with $K_N = 1$ for all N, then $\Sigma_{i=1}^N (m_i - \tau_{0N} - \tau_{1N}\, b_{iN})^2$ is uniformly bounded as $N \to \infty$. This result would be more useful if there were a sequence $\{b_i, i = 1,2,\ldots\}$ such that $\lim_{N\to\infty} b_{iN} = b_i$. We shall return to this important convergence question in the next section.

A problem with using $\alpha_N = \|\underline{\text{vec}}\ \underset{\sim}{R}_N\|_{\underset{\sim}{\Omega}_N}$ as a measure of approximation error is that apparently minor departures from an exact factor structure can result in α_N being unbounded as $N \to \infty$. For example, suppose that $x_i = m_i + b_i v + \varepsilon_i$, where the ε_i are "almost" uncorrelated: $\text{Cov}(\varepsilon_i,\varepsilon_j) = 0$ if $|i - j| > 1$. Then α_N will be unbounded as $N \to \infty$ unless $K_N \to \infty$.

In fact the condition that $\alpha_N \leq \alpha < \infty$ is overly strong. Let $\lambda_j(\underset{\sim}{A})$ denote the j^{th} largest eigenvalue of the symmetric matrix $\underset{\sim}{A}$. It follows from the proof of Proposition 14 that if $\rho_N \geq \rho > 0$ and $\alpha_N \leq \alpha < \infty$, then $\Sigma_{j=1}^{N}[\lambda_j(\underset{\sim}{R}_N)]^2$ and the diagonal elements of $\underset{\sim}{D}_N$ are uniformly bounded as $N \to \infty$. We shall show in the next section that a sufficient condition for Ross' result is that only the K largest eigenvalues of $\underset{\sim}{\Sigma}_N$ are unbounded as $N \to \infty$. This condition is satisfied if $\lambda_1(\underset{\sim}{R}_N)$ and the diagonal elements of $\underset{\sim}{D}_N$ are uniformly bounded as $N \to \infty$ with $K_N = K$. For we have

$$\lambda_{K+1}(\underset{\sim}{\Sigma}_N) \leq \lambda_{K+1}(\underset{\sim}{B}_N \underset{\sim}{B}_N') + \lambda_1(\underset{\sim}{D}_N + \underset{\sim}{R}_N)$$

$$\leq \lambda_1(\underset{\sim}{D}_N) + \lambda_1(\underset{\sim}{R}_N) .$$

(See Rao [11], Exercise 1.f.1.9.)

5. WHEN IS THERE A FACTOR STRUCTURE ASYMPTOTICALLY?

Suppose that we do not have an exact factor structure. Under what conditions can we find a sequence $\{(b_{i1},\ldots,b_{iK}),\ i = 1,2,\ldots\}$ and scalars $\tau_0,\tau_1,\ldots,\tau_K$ such that

$$\sum_{i=1}^{\infty} (m_i - \tau_0 - \tau_1 b_{i1} - \ldots - \tau_K b_{iK})^2 < \infty\,?$$

We shall show in Proposition 15 that a sufficient condition is that only K of the eigenvalues of $\underset{\sim}{\Sigma}_N$ are unbounded as $N \to \infty$. In addition, we shall show that in a relevant sense, the corresponding K eigenvectors converge to the b's.

Let $\underset{\sim}{\Sigma}_N = \Sigma_{j=1}^{N} \lambda_{jN}\, \underset{\sim}{t}_{jN}\, \underset{\sim}{t}'_{jN}$, where $\underset{\sim}{t}'_{jN}\, \underset{\sim}{t}_{jN} = 1$, $\underset{\sim}{t}'_{jN}\, \underset{\sim}{t}_{kN} = 0$ $(j \neq k;\ j,k = 1,\ldots,N)$, and $\lambda_{1N} \geq \lambda_{2N} \geq \cdots \geq \lambda_{NN}$; let $\underset{\sim}{t}'_{jN} = (t_{1jN}, \ldots, t_{NjN})$.

PROPOSITION 15: Assume that λ_{KN} is unbounded as $N \to \infty$ and that $\lambda_{K+1,N} \leq \lambda_{K+1} < \infty$, $\lambda_{NN} \geq \lambda_\infty > 0$ for all N. Then condition (A) implies:

I. There exists a sequence $\{(b_{i1},\ldots,b_{iK}),\ i = 1,2,\ldots\}$ and scalars $\tau_0,\tau_1,\ldots,\tau_K$ such that

$$\sum_{i=1}^{\infty} (m_i - \tau_0 - \tau_1 b_{i1} - \ldots - \tau_K b_{iK})^2 \leq \lambda_{K+1}\, \delta^2,$$

where δ is the bound given in Proposition 7;

II. For any m, let $\underset{\sim}{B}$ be the m×K matrix with i,j element b_{ij} and let $\underset{\sim}{B}_N$ be the m×K matrix with i,j element $\sqrt{\lambda_{jN}}\; t_{ijN}$ $(i = 1,\ldots,m;\ j = 1,\ldots,K)$; then each element of $\underset{\sim}{B}_N\underset{\sim}{B}'_N$ converges as $N \to \infty$ to the corresponding element of $\underset{\sim}{B}\,\underset{\sim}{B}'$.

PROOF: I.1 Let P_d be the space dual to P and let $\underset{\sim}{g} \in P$ represent a bounded linear functional in P_d. Let $(x_i, \underset{\sim}{g}) = \gamma_i$, $\underset{\sim}{\alpha}_N' = (\alpha_1, \ldots, \alpha_N)$, $\underset{\sim}{\gamma}_N' = (\gamma_1, \ldots, \gamma_N)$, and $\underset{\sim}{x}_N' = (x_1, \ldots, x_N)$. Then if $\underset{\sim}{q} = \underset{\sim}{\alpha}_N' \underset{\sim}{x}_N$, we have $(\underset{\sim}{q}, \underset{\sim}{g}) = \underset{\sim}{\alpha}_N' \underset{\sim}{\gamma}_N$ and the maximum value of $|(\underset{\sim}{q}, \underset{\sim}{g})|$ subject to $(\underset{\sim}{q}, \underset{\sim}{q}) = \underset{\sim}{\alpha}_N' \underset{\sim}{\Sigma}_N \underset{\sim}{\alpha}_N = 1$ is $(\underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\gamma}_N)^{1/2}$. Hence $(\underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\gamma}_N)$ is an increasing sequence that converges as $N \to \infty$ to $\|\underset{\sim}{g}\|^2$. Let P_{2d} consist of all the linear functionals defined on F with $\Sigma_{i=1}^{\infty} \gamma_i^2 < \infty$; i.e., $G(\ ) \in P_{2d}$ if $G(\underset{\sim}{q}) = \Sigma_i \alpha_i \gamma_i$ for $\underset{\sim}{q} = \Sigma_i \alpha_i x_i \in F$, where $\Sigma_{i=1}^{\infty} \gamma_i^2 < \infty$. We show in Lemma 1 that P_{2d} is a closed linear subspace of P_d. Hence $P_d = P_{2d}^{\perp} \oplus P_{2d}$. Let $P_{1d} = P_{2d}^{\perp}$.

I.2 Let P_j consist of the $\underset{\sim}{g} \in P$ that represent linear functionals in P_{jd} $(j = 1,2)$; then $P = P_1 \oplus P_2$. If $\underset{\sim}{p} \in P_1$, then there is a sequence $\underset{\sim}{p}_N = \Sigma_{i=1}^{N} \alpha_{iN} x_i \in F$ such that $\underset{\sim}{p}_N \to \underset{\sim}{p}$ as $N \to \infty$. We shall show that $\Sigma_{i=1}^{N} \alpha_{iN}^2 \to 0$.

$$\|\underset{\sim}{p}_n - \underset{\sim}{p}_m\|^2 = \sum_{i,j=1}^{n} (\alpha_{in} - \alpha_{im})(\alpha_{jn} - \alpha_{jm})\sigma_{ij}$$
$$\geq \rho_\infty \sum_{i=1}^{n} (\alpha_{in} - \alpha_{im})^2 ,$$

where $n \geq m$ and we set $\alpha_{im} = 0$ for $i > m$. Hence $\|\underset{\sim}{p}_n - \underset{\sim}{p}_m\| \to 0$ as $n,m \to \infty$ implies that there exists a sequence $\{\gamma_i,\ i = 1,2,\ldots\}$ with $\Sigma_{i=1}^{\infty} \gamma_i^2 < \infty$ such that $\Sigma_{i=1}^{N} (\alpha_{iN} - \gamma_i)^2 \to 0$ as $N \to \infty$. Consider the linear functional $G(\underset{\sim}{q}) = \Sigma_i \alpha_i \gamma_i$ for $\underset{\sim}{q} = \Sigma_i \alpha_i x_i \in F$. Since $G(\ ) \in P_{2d}$, there is a $\underset{\sim}{g} \in P_2$ that represents $G(\ )$. $\underset{\sim}{p}_N \to \underset{\sim}{p} \in P_1$ implies that

$$(\underset{\sim}{p}_N, \underset{\sim}{g}) = \sum_{i=1}^{N} \alpha_{iN} \gamma_i \to (\underset{\sim}{p}, \underset{\sim}{g}) = 0 .$$

Hence

$$\sum_{i=1}^{N} (\alpha_{iN} - \gamma_i)^2 = \sum_{i=1}^{N} \alpha_{iN}^2 - 2 \sum_{i=1}^{N} \alpha_{iN}\gamma_i + \sum_{i=1}^{N} \gamma_i^2 \to 0$$

implies that $\Sigma_{i=1}^{N} \alpha_{iN}^2 \to 0$ as $N \to \infty$.

I.3 We shall show that $\dim P_1 \leq K$. Suppose that $\dim P_1 > K$. Then we can choose $\underset{\sim}{p}_j \in P_1$ such that $\|\underset{\sim}{p}_j\| = 1$, $(\underset{\sim}{p}_j, \underset{\sim}{p}_k) = 0$ $(j \neq k;\ j,k = 1,\ldots,K+1)$. There is a sequence $\underset{\sim}{p}_{jN} = \Sigma_{i=1}^{N} \alpha_{ijN} x_i$ with $\underset{\sim}{p}_{jN} \to \underset{\sim}{p}_j$ as $N \to \infty$. Let $\underset{\sim}{\alpha}'_{jN} = (\alpha_{1jN},\ldots,\alpha_{NjN})$ and let $\underset{\sim}{\Sigma}_{1N} = \Sigma_{j=1}^{K} \lambda_{jN} \underset{\sim}{t}_{jN} \underset{\sim}{t}'_{jN}$, $\underset{\sim}{\Sigma}_{2N} = \underset{\sim}{\Sigma}_N - \underset{\sim}{\Sigma}_{1N}$; then we have

$$(\underset{\sim}{p}_{jN}, \underset{\sim}{p}_{kN}) = \underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{1N} \underset{\sim}{\alpha}_{kN} + \underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{2N} \underset{\sim}{\alpha}_{kN},$$

$$|\underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{2N} \underset{\sim}{\alpha}_{kN}| \leq (\underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{2N} \underset{\sim}{\alpha}_{jN})^{1/2} (\underset{\sim}{\alpha}'_{kN} \underset{\sim}{\Sigma}_{2N} \underset{\sim}{\alpha}_{kN})^{1/2},$$

$$\underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{2N} \underset{\sim}{\alpha}_{jN} \leq \lambda_{K+1}\, \underset{\sim}{\alpha}'_{jN} \underset{\sim}{\alpha}_{jN}.$$

We know from I.2 that $\underset{\sim}{p}_{jN} \to \underset{\sim}{p}_j \in P_1$ implies that $\underset{\sim}{\alpha}'_{jN} \underset{\sim}{\alpha}_{jN} \to 0$. Since $(\underset{\sim}{p}_{jN}, \underset{\sim}{p}_{kN}) \to (\underset{\sim}{p}_j, \underset{\sim}{p}_k)$, we have $\underset{\sim}{\alpha}'_{jN} \underset{\sim}{\Sigma}_{1N} \underset{\sim}{\alpha}_{kN} \to (\underset{\sim}{p}_j, \underset{\sim}{p}_k)$. Let $\underset{\sim}{A}_N = (\underset{\sim}{\alpha}_{1N},\ldots,\underset{\sim}{\alpha}_{K+1,N})$ and let $\underset{\sim}{F}_N = \underset{\sim}{A}'_N \underset{\sim}{\Sigma}_{1N} \underset{\sim}{A}_N$. Then $\operatorname{rank}(\underset{\sim}{F}_N) \leq K$ but $\underset{\sim}{F}_N$ converges to a $(K+1)\times(K+1)$ identity matrix. This is a contradiction and so $\dim P_1 \leq K$.

I.4 If there is a riskless asset, then condition (A) implies that $E(\ ) - r\,C(\ )$ is a bounded linear functional. In that case we set $\tau_0 = r$. If there is no riskless asset, then condition (A) implies that $E(\ )$ and $C(\ )$ are bounded linear functionals represented by $\underset{\sim}{a}, \underset{\sim}{c} \in P$; in that case we choose τ_0 so that $(\underset{\sim}{a} - \tau_0 \underset{\sim}{c}, \underset{\sim}{c}) = 0$. Set $\eta_i = m_i - \tau_0$; then $H(\underset{\sim}{q}) = \Sigma_i\, \alpha_i \eta_i$ for $\underset{\sim}{q} = \Sigma_i\, \alpha_i x_i \in F$ defines a bounded linear

functional with a norm of δ; it is represented by $\underset{\sim}{h} \in P$. Let $\underset{\sim}{h} = \underset{\sim}{f} + \underset{\sim}{g}$, where $\underset{\sim}{f} \in P_1$ and $\underset{\sim}{g} \in P_2$. Since $\dim P_1 \leq K$, we have $\underset{\sim}{f} = \Sigma_{j=1}^{K} \tau_j \underset{\sim}{p}_j$ for $\underset{\sim}{p}_j \in P_1$. Set $b_{ij} = (x_i, \underset{\sim}{p}_j)$.
Then we have

$$(\underset{\sim}{q}, \underset{\sim}{h} - \underset{\sim}{f}) = \sum_i \alpha_i(m_i - \tau_0 - \tau_1 b_{i1} - \ldots - \tau_K b_{iK})$$

for $\underset{\sim}{q} = \Sigma_i \alpha_i x_i \in F$. Since $\underset{\sim}{h} - \underset{\sim}{f} \in P_2$ represents a linear functional in P_{2d}, we have

$$\sum_{i=1}^{\infty} (m_i - \tau_0 - \tau_1 b_{i1} - \ldots - \tau_K b_{iK})^2 < \infty .$$

Using a result from the proof of Lemma 1, we can give an explicit bound:

$$\sum_{i=1}^{\infty} (m_i - \tau_0 - \tau_1 b_{i1} - \ldots - \tau_K b_{iK})^2 \leq \lambda_{K+1} \|\underset{\sim}{h} - \underset{\sim}{f}\|^2$$

$$= \lambda_{K+1}(\|\underset{\sim}{h}\|^2 - \|\underset{\sim}{f}\|^2) \leq \lambda_{K+1} \|\underset{\sim}{h}\|^2 = \lambda_{K+1} \delta^2 .$$

II.1 It remains to show that $\underset{\sim}{B}_N \underset{\sim}{B}_N'$ converges to $\underset{\sim}{B}\underset{\sim}{B}'$ as $N \to \infty$. Let $\underset{\sim}{p} = \Sigma_{i=1}^{N} \alpha_i x_i$ and $\underset{\sim}{q} = \Sigma_{i=1}^{N} \beta_i x_i$ be portfolios in F. Define $(\underset{\sim}{p},\underset{\sim}{q})_{\ell_2} = \Sigma_{i=1}^{N} \alpha_i \beta_i$.

$$(\underset{\sim}{p},\underset{\sim}{q})_{\ell_2}^2 \leq (\sum_{i=1}^{N} \alpha_i^2)(\sum_{i=1}^{N} \beta_i^2) ,$$

$$\|\underset{\sim}{p}\|^2 \geq \lambda_\infty \sum_{i=1}^{N} \alpha_i^2 ,$$

$$(\underset{\sim}{p},\underset{\sim}{q})_{\ell_2}^2 \leq \lambda_\infty^{-2} \|\underset{\sim}{p}\|^2 \|\underset{\sim}{q}\|^2 .$$

Then it is a consequence of the Riesz Representation Theorem that there is a unique bounded linear operator $\underset{\sim}{S}: P \to P$ such that $(\underset{\sim}{p},\underset{\sim}{q})_{\ell_2} = (\underset{\sim}{p}, \underset{\sim}{S}\underset{\sim}{q})$ for $\underset{\sim}{p},\underset{\sim}{q} \in F$. So we can extend the definition of $(\ ,\ )_{\ell_2}$ to P by continuity. $\underset{\sim}{S}$ is positive $((\underset{\sim}{p}, \underset{\sim}{S}\underset{\sim}{p}) \geq 0)$ and symmetric:

$$(\underset{\sim}{p}, \underset{\sim}{S}\underset{\sim}{q}) = (\underset{\sim}{p},\underset{\sim}{q})_{\ell_2} = (\underset{\sim}{q},\underset{\sim}{p})_{\ell_2} = (\underset{\sim}{q}, \underset{\sim}{S}\underset{\sim}{p})$$

$$= (\underset{\sim}{S}\underset{\sim}{p}, \underset{\sim}{q}) .$$

II.2 We shall show that there is a $\psi > 0$ such that $(\underset{\sim}{g}, \underset{\sim}{S}\underset{\sim}{g}) \geq \psi(\underset{\sim}{g},\underset{\sim}{g})$ for $\underset{\sim}{g} \in P_2$. If not, then there is a sequence $\underset{\sim}{g}_n \in P_2$ with $\|\underset{\sim}{g}_n\| = 1$ and $(\underset{\sim}{g}_n, \underset{\sim}{S}\underset{\sim}{g}_n) \to 0$ as $n \to \infty$. Since $\underset{\sim}{g}_n$ represents a linear functional in P_{2d}, we have $(\underset{\sim}{q}, \underset{\sim}{g}_n) = \Sigma_i \alpha_i \gamma_{in}$ for $\underset{\sim}{q} = \Sigma_i \alpha_i x_i \in F$, where $\Sigma_{i=1}^{\infty} \gamma_{in}^2 < \infty$. It follows from the proof of Lemma 1 that

$$\sum_{i=1}^{\infty} \gamma_{in}^2 \leq \lambda_{K+1} \|\underset{\sim}{g}_n\|^2 = \lambda_{K+1} .$$

We can construct a sequence $\underset{\sim}{q}_n$ such that

$$\underset{\sim}{q}_n = \sum_{i=1}^{N_n} \alpha_{in} x_i \in F, \quad \|\underset{\sim}{q}_n - \underset{\sim}{g}_n\| \to 0$$

as $n \to \infty$. Then $\|\underset{\sim}{q}_n\| \to 1$, $(\underset{\sim}{q}_n, \underset{\sim}{S}\underset{\sim}{q}_n) \to 0$, and $(\underset{\sim}{q}_n,\underset{\sim}{g}_n) \to 1$. But we also have

$$(\underset{\sim}{q}_n,\underset{\sim}{g}_n) = \sum_{i=1}^{N_n} \alpha_{in}\gamma_{in} \leq \Big(\sum_{i=1}^{N_n} \alpha_{in}^2\Big)^{1/2} \Big(\sum_{i=1}^{\infty} \gamma_{in}^2\Big)^{1/2}$$

$$\leq \lambda_{K+1}^{1/2} (\underset{\sim}{q}_n, \underset{\sim}{S}\underset{\sim}{q}_n)^{1/2} \to 0 ,$$

a contradiction. Hence $(\underset{\sim}{g}, \underset{\sim}{S}\underset{\sim}{g}) \geq \psi(\underset{\sim}{g},\underset{\sim}{g})$ for $\underset{\sim}{g} \in P_2$.

II.3 Let $\underset{\sim}{t}'_{jN} = (t_{1jN},\ldots,t_{NjN})$ and let $\underset{\sim}{p}_{jN} = \Sigma_{i=1}^{N} t_{ijN} x_i / \sqrt{\lambda_{jN}}$, $j = 1,\ldots,K$. Then $\|\underset{\sim}{p}_{jN}\| = 1$, $(\underset{\sim}{p}_{jN}, \underset{\sim}{p}_{kN}) = 0$ $(j \neq k)$, and $(\underset{\sim}{p}_{jN}, \underset{\sim}{S}\,\underset{\sim}{p}_{jN}) = \lambda_{jN}^{-1} \to 0$ as $N \to \infty$ $(j = 1,\ldots,K)$. Consider the orthogonal projection $\underset{\sim}{Q}_N$: $\mathcal{P} \to \mathcal{P}$ defined by $\underset{\sim}{Q}_N \underset{\sim}{p} = \Sigma_{j=1}^{K} (\underset{\sim}{p}, \underset{\sim}{p}_{jN})\underset{\sim}{p}_{jN}$. $\underset{\sim}{Q}_N$ projects onto the linear subspace spanned by $\underset{\sim}{p}_{1N},\ldots,\underset{\sim}{p}_{KN}$. Any $\underset{\sim}{p} \in \mathcal{P}$ can be uniquely decomposed into $\underset{\sim}{p} = \underset{\sim}{p}_1 + \underset{\sim}{p}_2$, where $\underset{\sim}{p}_1 \in \mathcal{P}_1$ and $\underset{\sim}{p}_2 \in \mathcal{P}_2$; so $\underset{\sim}{Q}\,\underset{\sim}{p} = \underset{\sim}{p}_1$ defines the orthogonal projection onto $\mathcal{P}_1$. We shall show that $\|\underset{\sim}{Q}_N \underset{\sim}{p} - \underset{\sim}{Q}\,\underset{\sim}{p}\| \to 0$ as $N \to \infty$ for any $\underset{\sim}{p} \in \mathcal{P}$, so that $\underset{\sim}{Q}_N$ converges strongly to $\underset{\sim}{Q}$.

Let $\underset{\sim}{p}_{jN} = \underset{\sim}{p}_{1jN} + \underset{\sim}{p}_{2jN}$, where $\underset{\sim}{p}_{1jN} \in \mathcal{P}_1$, $\underset{\sim}{p}_{2jN} \in \mathcal{P}_2$. We shall show first that $\underset{\sim}{p}_{2jN} \to \underset{\sim}{0}$ as $N \to \infty$. We know from I.2 that if $\underset{\sim}{q}_N \to \underset{\sim}{q} \in \mathcal{P}_1$ for $\underset{\sim}{q}_N \in F$, then $(\underset{\sim}{q}_N, \underset{\sim}{S}\,\underset{\sim}{q}_N) \to 0$. Since $\underset{\sim}{S}\,\underset{\sim}{q}_N \to \underset{\sim}{S}\,\underset{\sim}{q}$, we also have $(\underset{\sim}{q}_N, \underset{\sim}{S}\,\underset{\sim}{q}_N) \to (\underset{\sim}{q}, \underset{\sim}{S}\,\underset{\sim}{q})$, which implies that $(\underset{\sim}{q}, \underset{\sim}{S}\,\underset{\sim}{q}) = 0$ for $\underset{\sim}{q} \in \mathcal{P}_1$. For any $\underset{\sim}{p} \in \mathcal{P}$,

$$(\underset{\sim}{p}, \underset{\sim}{S}\,\underset{\sim}{q})^2 \leq (\underset{\sim}{p}, \underset{\sim}{S}\,\underset{\sim}{p})(\underset{\sim}{q}, \underset{\sim}{S}\,\underset{\sim}{q}) = 0$$

(Riesz and Sz.-Nagy [13] p. 262); hence $\underset{\sim}{S}\,\underset{\sim}{q} = \underset{\sim}{0}$ for $\underset{\sim}{q} \in \mathcal{P}_1$.

$$\begin{aligned}(\underset{\sim}{p}_{jN}, \underset{\sim}{S}\,\underset{\sim}{p}_{jN}) &= (\underset{\sim}{p}_{1jN} + \underset{\sim}{p}_{2jN}, \underset{\sim}{S}\,\underset{\sim}{p}_{2jN}) \\ &= (\underset{\sim}{p}_{2jN}, \underset{\sim}{S}\,\underset{\sim}{p}_{2jN}) \geq \psi \|\underset{\sim}{p}_{2jN}\|^2 .\end{aligned}$$

So $(\underset{\sim}{p}_{jN}, \underset{\sim}{S}\,\underset{\sim}{p}_{jN}) = \lambda_{jN}^{-1} \to 0$ implies that $\underset{\sim}{p}_{2jN} \to \underset{\sim}{0}$ $(j = 1,\ldots,K)$.

Since $\underset{\sim}{p}_{jN} = \underset{\sim}{p}_{1jN} + \underset{\sim}{p}_{2jN}$, $\underset{\sim}{p}_{1jN} \in \mathcal{P}_1$, $\underset{\sim}{p}_{2jN} \in \mathcal{P}_2$, we have $\underset{\sim}{Q}_N = \underset{\sim}{Q}_{1n} + \underset{\sim}{Q}_{2N}$ where

$$Q_{1N}\,p = \sum_{j=1}^{K} (p, p_{jN})\,p_{1jN} ,$$

$$Q_{2N}\,p = \sum_{j=1}^{K} (p, p_{jN})\,p_{2jN} .$$

$Q_{2N}\,p \to 0$ since $p_{2jN} \to 0$ $(j = 1,\ldots,K)$. So we need to show only that $Q_{1N}\,p \to Q\,p$. Consider the K×K matrix F_N with j,k element (p_{1jN}, p_{1kN}). F_N converges to a K×K identity matrix, and so there is an N^* such that p_{1jN} $(j = 1,\ldots,K)$ are linearly independent for $N \geq N^*$. We showed in I.3 that $\dim P_1 \leq K$, and so p_{1jN} $(j = 1,\ldots,K)$ form a basis for P_1 if $N \geq N^*$. Hence $Qp = \sum_{j,k=1}^{K} (p, p_{1kN})\, f_N^{jk}\, p_{1jN}$ for $N \geq N^*$, where f_N^{jk} is the j,k element of F_N^{-1}. Then $\lim_{N\to\infty} F_N = I_K$ implies that $\|Q\,p - Q_{1N}p\| \to 0$ for any $p \in P$.

II.4 Let p_j $(j = 1,\ldots,K)$ be an orthonormal basis for P_1. As in I.4 we set $b_{ij} = (x_i, p_j)$. Then the orthogonal projection onto P_1 is given by $Q\,p = \Sigma_{j=1}^{K} (p, p_j)p_j$. We know from II.3 that $Q_N\,p \to Q\,p$.

$$(x_i, Q_N\,x_k) = \sum_{j=1}^{K} (x_k, p_{jN})(x_i, p_{jN})$$

$$= \sum_{j=1}^{K} \lambda_{jN}\, t_{kjN}\, t_{ijN} \qquad (i,k \leq N),$$

$$(x_i, Q\,x_k) = \sum_{j=1}^{K} (x_k, p_j)(x_i, p_j)$$

$$= \sum_{j=1}^{K} b_{kj}\, b_{ij} .$$

Hence $(x_i, Q_N x_k) \to (x_i, Q x_k)$ implies that the i,k element of $B_N B_N'$ converges as $N \to \infty$ to the i,k element of $B\,B'$. Q.E.D.

LEMMA 1: Assume that λ_{KN} is unbounded as $N \to \infty$ and that $\lambda_{K+1,N} \leq \lambda_{K+1} < \infty$, $\lambda_{NN} \geq \lambda_\infty > 0$ for all N. Then P_{2d} is a closed linear subspace of P_d.

PROOF: 1. Let $G(\underset{\sim}{q}) = \Sigma_{i=1}^{N} \alpha_i \gamma_i$ for $\underset{\sim}{q} = \Sigma_{i=1}^{N} \alpha_i x_i \in F$, where $\Sigma_{i=1}^{\infty} \gamma_i^2 < \infty$. Then $G \in P_{2d}$. Let $\underset{\sim}{\alpha}_N' = (\alpha_1,\ldots,\alpha_N)$ and $\underset{\sim}{\gamma}_N' = (\gamma_1,\ldots,\gamma_N)$. The maximum value of $|G(\underset{\sim}{q})|$ subject to $||\underset{\sim}{q}|| = (\underset{\sim}{\alpha}_N' \underset{\sim}{\Sigma}_N \underset{\sim}{\alpha}_N)^{1/2} = 1$ is $(\underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\gamma}_N)^{1/2}$.

$$\underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\gamma}_N \leq \lambda_\infty^{-1} \underset{\sim}{\gamma}_N' \underset{\sim}{\gamma}_N \leq \lambda_\infty^{-1} \sum_{i=1}^{\infty} \gamma_i^2 .$$

So $G \in P_d$ and $P_{2d} \subset P_d$. P_{2d} is a linear subspace since the square-summable sequences form a linear space.

2. Let G_1 and G_2 be linear functionals in P_d. The inner product on P induces an inner product on P_d: $(G_1,G_2)_d = (\underset{\sim}{g}_1,\underset{\sim}{g}_2)$, where $G_j(\underset{\sim}{q}) = (\underset{\sim}{q}, \underset{\sim}{g}_j)$ and $\underset{\sim}{q},\underset{\sim}{g}_j \in P$ $(j = 1,2)$. Let $\gamma_i = G(x_i)$. We shall complete the proof by showing that

$$(G,G)_d \geq \lambda_{K+1}^{-1} \sum_{i=1}^{\infty} \gamma_i^2$$

if $G \in P_{2d}$. It then follows that P_{2d} is closed under $\| \ \|_d$. For suppose that $G_n \in P_{2d}$ and $||G_n - G_m\|_d \to 0$ as $n,m \to \infty$. Let $\gamma_{in} = G_n(x_i)$. Then

$$||G_n - G_m\|_d^2 \geq \lambda_{K+1}^{-1} \sum_{i=1}^{\infty} (\gamma_{in} - \gamma_{im})^2$$

implies that there exists a sequence $\{\gamma_i^*\}$ such that $\Sigma_{i=1}^{\infty} \gamma_i^{*2} < \infty$ and $\Sigma_{i=1}^{\infty} (\gamma_{in} - \gamma_i^*)^2 \to 0$ as $n \to \infty$. Let $G^*(x_i) = \gamma_i^*$; then $G^* \in P_{2d}$ and

$$\|G_n - G^*\|_d^2 \leq \lambda_\infty^{-1} \sum_{i=1}^{\infty} (\gamma_{in} - \gamma_i^*)^2 \to 0$$

as $n \to \infty$. Hence $(G,G)_d \geq \lambda_{K+1}^{-1} \Sigma_{i=1}^{\infty} [G(x_i)]^2$ implies that P_{2d} is closed.

3. Consider the Hilbert space ℓ_2 of square-summable sequences: $\underset{\sim}{\gamma} = \{\gamma_1, \gamma_2, \ldots\} \in \ell_2$ iff $\Sigma_{i=1}^{\infty} \gamma_i^2 < \infty$; the inner product is $\langle \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle = \Sigma_{i=1}^{\infty} \beta_i \gamma_i$ for $\underset{\sim}{\beta}, \underset{\sim}{\gamma} \in \ell_2$. Every $\underset{\sim}{\gamma} \in \ell_2$ corresponds to a $G \in P_{2d}$, where $G(x_i) = \gamma_i$. We use this correspondence to define a quadratic form on ℓ_2: $\langle \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle_d = (B,G)_d$, where $B(x_i) = \beta_i$ and $G(x_i) = \gamma_i$, $i = 1,2,\ldots$.

$$\langle \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle_d^2 \leq \langle \underset{\sim}{\beta}, \underset{\sim}{\beta} \rangle_d \langle \underset{\sim}{\gamma}, \underset{\sim}{\gamma} \rangle_d \leq \lambda_\infty^{-2} \langle \underset{\sim}{\beta}, \underset{\sim}{\beta} \rangle \langle \underset{\sim}{\gamma}, \underset{\sim}{\gamma} \rangle .$$

So $\langle \ , \ \rangle_d$ is a bounded quadratic form, and it is a consequence of the Riesz Representation Theorem that there exists a unique bounded linear operator $\underset{\sim}{T}: \ell_2 \to \ell_2$ such that $\langle \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle_d = \langle \underset{\sim}{\beta}, \underset{\sim}{T} \underset{\sim}{\gamma} \rangle$. T is positive ($\langle \underset{\sim}{\gamma}, \underset{\sim}{T} \underset{\sim}{\gamma} \rangle \geq 0$) and symmetric:

$$\begin{aligned} \langle \underset{\sim}{\beta}, \underset{\sim}{T} \underset{\sim}{\gamma} \rangle &= \langle \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle_d = \langle \underset{\sim}{\gamma}, \underset{\sim}{\beta} \rangle_d \\ &= \langle \underset{\sim}{\gamma}, \underset{\sim}{T} \underset{\sim}{\beta} \rangle = \langle \underset{\sim}{T} \underset{\sim}{\beta}, \underset{\sim}{\gamma} \rangle . \end{aligned}$$

4. Let $\underset{\sim}{\beta}, \underset{\sim}{\gamma} \in \ell_2$ and $\underset{\sim}{\beta}_N' = (\beta_1, \ldots, \beta_N)$, $\underset{\sim}{\gamma}_N' = (\gamma_1, \ldots, \gamma_N)$. Define the linear operator $\underset{\sim}{T}_N: \ell_2 \to \ell_2$ by

$$\underset{\sim}{T}_N \underset{\sim}{\gamma} = \{\underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1}, 0, 0, \ldots\}$$

for $N = 1,2,\ldots$. $\underset{\sim}{T}_N$ is positive and symmetric.

$$\begin{aligned} \langle \underset{\sim}{T}_N \underset{\sim}{\gamma}, \underset{\sim}{T}_N \underset{\sim}{\gamma} \rangle &= \underset{\sim}{\gamma}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{\gamma}_N \\ &\leq \lambda_\infty^{-2} \langle \underset{\sim}{\gamma}, \underset{\sim}{\gamma} \rangle . \end{aligned}$$

Hence this sequence of operators is uniformly bounded as $N \to \infty$. Since $\langle \gamma, T_N \gamma \rangle = \gamma_N' \Sigma_N^{-1} \gamma_N$ is increasing in N, the sequence of operators is monotonic. It follows that T_N converges strongly to a bounded, symmetric operator T_∞; i.e., $\langle T_N \gamma - T_\infty \gamma, T_N \gamma - T_\infty \gamma \rangle \to 0$ as $N \to \infty$ for every $\gamma \in \ell_2$ (Riesz and Sz.-Nagy [13], p. 263). As $N \to \infty$,

$$\langle \gamma, T_N \gamma \rangle = \gamma_N' \Sigma_N^{-1} \gamma_N \to \langle \gamma, \gamma \rangle_d = \langle \gamma, T \gamma \rangle.$$

Since $T_N \gamma \to T_\infty \gamma$, we also have

$$\langle \gamma, T_N \gamma \rangle \to \langle \gamma, T_\infty \gamma \rangle ,$$

and so $\langle \gamma, (T - T_\infty)\gamma \rangle = 0$ for all $\gamma \in \ell_2$. Since $T - T_\infty$ is a symmetric operator, its norm is equal to $\sup|\langle \gamma, (T - T_\infty)\gamma \rangle|$ subject to $\langle \gamma, \gamma \rangle = 1$ (Riesz and Sz.-Nagy [13], p. 230). Hence $T_\infty = T$.

5. Let $\sigma(T_N)$ be the spectrum of T_N; it consists of the points $0, \lambda_{1N}^{-1}, \ldots, \lambda_{NN}^{-1}$. For any $\varepsilon > 0$, there is an N^* such that $\sigma(T_N)$ has no points in the open interval $(\varepsilon, \lambda_{K+1}^{-1})$ for $N \geq N^*$. Since $\{T_N\}$ is a sequence of symmetric operators that converges strongly to the symmetric operator T, we have $\sigma(T) \cap (\varepsilon, \lambda_{K+1}^{-1}) = \emptyset$ (Riesz and Sz.-Nagy [13], p. 369 or Reed and Simon [12], Theorem VIII.24).

6. Note that o cannot be an eigenvalue of T: $T\gamma = 0$ implies that $\langle \gamma, \gamma \rangle_d = \langle \gamma, T\gamma \rangle = 0$; $0 = \langle \gamma, \gamma \rangle_d \geq \gamma_N' \Sigma_N^{-1} \gamma_N$ for all N implies that $\langle \gamma, \gamma \rangle = \Sigma_{i=1}^{\infty} \gamma_i^2 = 0$. Hence the spectrum of T is contained in the interval $[\lambda_{K+1}^{-1}, \lambda_\infty^{-1}]$, and so $\langle \gamma, \gamma \rangle_d = \langle \gamma, T\gamma \rangle \geq \lambda_{K+1}^{-1} \langle \gamma, \gamma \rangle$. Hence $(G,G)_d \geq \lambda_{K+1}^{-1} \Sigma_{i=1}^{\infty} [G(x_i)]^2$ if $G \in P_{2d}$. Then the argument in step 2 shows that P_{2d} is closed.

Q.E.D.

We have focused on the eigenvalues (λ_{jN}) of $\underset{\sim}{\Sigma}_N$ relative to an identity matrix $\underset{\sim}{I}_N$. Suppose that we used instead the eigenvalues of $\underset{\sim}{\Sigma}_N$ relative to some positive-definite matrix $\underset{\sim}{\Psi}_N$: $\underset{\sim}{\Sigma}_N \underset{\sim}{s}_{jN} = \rho_{jN} \underset{\sim}{\Psi}_N \underset{\sim}{s}_{jN}$, $\rho_{1N} \geq \rho_{2N} \geq \cdots \geq \rho_{NN}$, $\underset{\sim}{s}'_{jN} \underset{\sim}{\Psi}_N \underset{\sim}{s}_{jN} = 1$ $(j = 1,\ldots,N)$. Let the eigenvalues of $\underset{\sim}{\Psi}_N$ relative to $\underset{\sim}{I}_N$ be $\kappa_{1N} \geq \kappa_{2N} \geq \cdots \geq \kappa_{NN}$ and assume that these eigenvalues are uniformly bounded away from zero and infinity:

$$0 < \kappa_\infty \leq \kappa_{NN} \leq \kappa_{1N} \leq \kappa_1 < \infty$$

for all N. Then $\{\underset{\sim}{\Psi}_N\}$ is equivalent to $\{\underset{\sim}{I}_N\}$ for the purposes of Proposition 15; if ρ_{KN} is unbounded and $\rho_{K+1,N} \leq \rho_{K+1} < \infty$, $\rho_{NN} \geq \rho_\infty > 0$ for all N, it follows that the λ_{jN} have the same properties, and so there exists a sequence of b's such that

$$\sum_{i=1}^{\infty} (m_i - \tau_0 - \tau_1 b_{i1} - \cdots - \tau_K b_{iK})^2 < \infty .$$

The convergence part of our proof goes through as before, with $(\underset{\sim}{p},\underset{\sim}{q})_{\ell_2}$ in II.1 replaced by $(\underset{\sim}{p},\underset{\sim}{q})_{\underset{\sim}{\Psi}} = \underset{\sim}{\alpha}'_N \underset{\sim}{\Psi}_N \underset{\sim}{\beta}_N$ for $\underset{\sim}{p} = \underset{\sim}{\alpha}'_N \underset{\sim}{x}_N$, $\underset{\sim}{q} = \underset{\sim}{\beta}'_N \underset{\sim}{x}_N$, where $\underset{\sim}{x}'_N = (x_1,\ldots,x_N)$. In II.3 we set $\underset{\sim}{p}_{jN} = \Sigma_{i=1}^N s_{ijN} x_i/\sqrt{\rho_{jN}}$ $(j = 1,\ldots,K)$, where $\underset{\sim}{s}'_{jN} = (s_{1jN},\ldots,s_{NjN})$. The only difference is that in II.4 we have

$$(x_i, \underset{\sim}{Q}_N x_k) = \sum_{j=1}^{K} (x_k, \underset{\sim}{p}_{jN})(x_i, \underset{\sim}{p}_{jN})$$

$$= \sum_{j=1}^{K} \rho_{jN} w_{kjN} w_{ijN} \qquad (i,k \leq N),$$

where $\underset{\sim}{w}'_{jN} = (w_{1jN},\ldots,w_{NjN}) = \underset{\sim}{s}'_{jN}\underset{\sim}{\Psi}_N$. So now we let $\underset{\sim}{B}_N$ be the m×K matrix with i,j element $\sqrt{\rho_{jN}}\, w_{ijN}$. As before we let $\underset{\sim}{B}$ be the m×K matrix with i,j element b_{ij}. Then $\lim_{N\to\infty} \underset{\sim}{B}_N \underset{\sim}{B}'_N = \underset{\sim}{B}\underset{\sim}{B}'$.

In factor analysis, one considers the eigenvalues of $\underset{\sim}{\Sigma}_N$ relative to a diagonal matrix $\underset{\sim}{D}_N$ $(\underset{\sim}{\Sigma}_N \underset{\sim}{s}_{jN} = \rho_{jN}\underset{\sim}{D}_N \underset{\sim}{s}_{jN})$. So if the diagonal elements of $\underset{\sim}{D}_N$ are uniformly bounded away from zero and infinity, a sufficient condition for Ross' result is that ρ_{KN} be unbounded, $\rho_{K+1,N} \le \rho_{K+1} < \infty$, and $\rho_{NN} \ge \rho_\infty > 0$. Furthermore, the factor loadings are obtained from $\underset{\sim}{D}_N \underset{\sim}{s}_{jN}$ $(j = 1,\ldots,K)$; hence the columns of $\underset{\sim}{B}_N$ are proportional to the factor loadings.

However, we depart from standard factor analysis in our criterion for the appropriate number of factors. Standard goodness-of-fit tests correspond to the distance measure used in Proposition 14; they measure how much the eigenvalues $\rho_{K+1,N},\ldots,\rho_{NN}$ differ from unity (Lawley and Maxwell [8], p. 36). Our discussion of Proposition 14 argued that "small" departures from an exact K factor model will result in these measures being unbounded as $N \to \infty$. Proposition 15 shows that these measures do not correspond to the relevant population criterion; we require only that $\rho_{K+1,N}$ be uniformly bounded.

Furthermore, the precise choice of $\underset{\sim}{D}_N$ is of no consequence for our results. So computational costs in empirical work can be reduced by dispensing with iterative methods, such as maximum likelihood, that are used to estimate $\underset{\sim}{D}_N$ from a sample covariance matrix $\hat{\underset{\sim}{\Sigma}}_N$. The implementation of Proposition 15 requires only a principal components analysis of $\hat{\underset{\sim}{\Sigma}}_N$. Our results provide one of the few rigorous justifications for this technique and pose some challenging problems in statistical inference.

6. RISKLESS PORTFOLIOS

We shall say a portfolio $\underset{\sim}{p} = \Sigma_1^N \alpha_i x_i$ is <u>nontrivial</u> if $\Sigma_1^N |\alpha_i| \geq 1$. Nontrivial portfolios are those in which the total amount invested and disinvested is nontrivial. Nontrivial limit portfolios are limits of sequences of nontrivial portfolios.

PROPOSITION 16: There are nontrivial riskless limit portfolios.

To prove Proposition 16, it is necessary to construct a sequence of portfolios $\{\underset{\sim}{p}_N\}$ such that $\Sigma_1^N |\alpha_{iN}| = 1$ and $V(\underset{\sim}{p}_N) = ||\underset{\sim}{p}_N||^2 \to 0$. Since the solution to

$$\underset{\underset{\sim}{\alpha}}{\text{Min}}\ \underset{\sim}{\alpha}' \underset{\sim}{\Sigma}_N \underset{\sim}{\alpha} \quad \text{subject to} \quad \sum_1^N |\alpha_i| = 1$$

is bounded away from zero when $\underset{\sim}{\Sigma}_N$ is nonsingular, there are no riskless portfolios on markets with finite numbers of assets. The existence of nontrivial riskless portfolios distinguishes infinite from finite markets.

PROOF OF PROPOSITION 16: Let d_i, $i = 1,\ldots,N$ be a sequence of i.i.d. random variables independent of the x_i's. The d_i's are binomial and $\Pr\{d_i = +1\} = \Pr\{d_i = -1\} = 1/2$. Let $z_i = d_i x_i$; then $E(z_i) = E(d_i)\,E(x_i) = 0$ and $E(z_i z_j) = E(x_i x_j)\,E(d_i d_j) = 0$ $(i \neq j)$ while $V(z_i) = E(x_i^2) \leq B$. Let $\underset{\sim}{d}$ be any realization of the i.i.d. binary sequence. For each $\underset{\sim}{d}$ and each N there is a portfolio in which a net investment of $\$d_i N^{-1}$ is made in asset i, $i = 1,2,\ldots,N$. Let $p_N = N^{-1} \Sigma_1^N z_i$. Clearly $E(p_N) = 0$, and $V(p_N) \to 0$ since $V(p_N) = N^{-2} \Sigma_1^N E(z_i^2) \leq N^{-2} \Sigma_1^N B = B/N \to 0$.

We want to show that there is a riskless portfolio, that is, some realization of the d process, say $\hat{d}$, such that $V(p_N|\hat{d}) \to 0$. Now

$$(5.1) \qquad V(p_N) = E(V(p_N|d)) + V(E(p_N|d)) .$$

Since $V(p_N) \to 0$ while both terms on the right-hand side of (5.1) are non-negative, these terms must both converge to 0. Let $y_N(d) = V(p_N|d)$ and let P be the measure on sequences of +1's and -1's generated by the d_i's. Then we have shown that $\int y_N(d)dP \to 0$. But since $y_N(d) \geq 0$, Fatou's Lemma states that

$$\int \liminf y_N(d)dP \leq \liminf \int y_N(d)dP = 0 .$$

Therefore $\liminf y_N(d) = 0$ almost surely P, so that for almost any d, there is a subsequence, say N', such that

$$(5.2) \qquad y_{N'}(d) = V(p_{N'}|d) \to 0 .$$

By construction, the sequence of portfolios $p_{N'}$ is nontrivial, so (5.2) states that there is a riskless limit portfolio. Q.E.D.

The net investment in such a portfolio is $\lim_{N\to\infty} N^{-1} \Sigma_1^N d_i$, which is almost surely 0, so that $\underset{\sim}{d}$ generates a sequence of portfolios which almost surely converges to a hedge portfolio.

In the last section we showed that the implications of arbitrage depended on whether or not there was a riskless asset, which we defined as a limit portfolio which cost $1 and had no risk. Whether or not such a portfolio can be approximated arbitrarily closely by a finite portfolio is a relatively simple question to pose, and to answer. The problem of finding the portfolio of the first N assets that has minimum variance among all portfolios costing $1 is simply:

$$\underset{\underset{\sim}{\alpha}}{\text{Min}} \quad \underset{\sim}{\alpha}'\underset{\sim}{\Sigma}_N\underset{\sim}{\alpha} \qquad \text{subject to} \qquad \underset{\sim}{\alpha}'\,\underset{\sim}{e}_N = 1\,,$$

where $\underset{\sim}{e}_N' = (1,\ldots,1)$, the vector in R^N all of whose components are equal to one. It is straightforward to calculate that the value of the solution to this problem is simply $[\underset{\sim}{e}_N'\,\underset{\sim}{\Sigma}_N^{-1}\,\underset{\sim}{e}_N]^{-1}$, so that we have proved

PROPOSITION 17: There is a riskless asset if and only if the sequence $\left\{\underset{\sim}{e}_N'\underset{\sim}{\Sigma}_N^{-1}\,\underset{\sim}{e}_N\right\}$ is unbounded.

Note that it is immediate from the structure of the problem that $\left\{\underset{\sim}{e}_N'\,\underset{\sim}{\Sigma}_N^{-1}\,\underset{\sim}{e}_N\right\}$ is an increasing sequence. It is easy to give examples where $\left\{\underset{\sim}{e}_N'\,\underset{\sim}{\Sigma}_N^{-1}\,\underset{\sim}{e}_N\right\}$ is bounded and examples where it diverges. If the assets are uncorrelated, then $\underset{\sim}{e}_N'\,\underset{\sim}{\Sigma}_N^{-1}\,\underset{\sim}{e}_N = \Sigma_1^N\,\sigma_{ii}^{-1} \geq N/M \to \infty$. If on the other hand assets are equally correlated, so $\sigma_{ii} = \sigma^2$ and $\sigma_{ij} = \rho\sigma^2$ $(i \neq j,\ \ 0 < \rho < 1)$, then $e_N' \Sigma_N^{-1} e_N \to (\rho\sigma^2)^{-1} < \infty$. This example is illustrative of the general case. In a sense the nonexistence of a riskless asset is equivalent to the variance-covariance matrix of the x_i's having an equally correlated component. When assets are equally correlated,

$$\Sigma_N = \alpha I_N + \beta e_N e_N'$$

for some scalars α and β. We shall say the sequence of random variables $\{x_1, x_2, \ldots\}$ has an equally correlated component if $\tau_N = \sup\{\tau \mid \Sigma_N - \tau e_N e_N'$ is positive definite$\}$ is bounded away from zero.

PROPOSITION 18: There is a riskless asset if and only if $\{x_1, x_2, \ldots\}$ does not have an equally correlated component.

PROOF: Recall that a matrix is positive definite if and only if its inverse is positive definite. The proof is based on the following identity:

$$(5.3) \qquad (\Sigma_N - \tau e_N e_N')^{-1} = \Sigma_N^{-1} + \frac{\tau \Sigma_N^{-1} e_N e_N' \Sigma_N^{-1}}{1 - \tau e_N' \Sigma_N^{-1} e_N} .$$

Suppose $\{x_1, x_2, \ldots\}$ has an equally correlated component. Then the left-hand side of (5.3) is a positive-definite matrix for $\tau \in [0, \underline{\tau})$ for some $\underline{\tau} > 0$. Thus

$$0 < e_N'(\Sigma_N - \tau e_N e_N')^{-1} e_N = \frac{e_N' \Sigma_N^{-1} e_N}{1 - \tau e_N' \Sigma_N^{-1} e_N} .$$

But this can only happen if $1 > \tau e_N' \Sigma_N^{-1} e_N$ or if $e_N' \Sigma_N^{-1} e_N < \tau^{-1}$ for all N.

To prove the converse, note that if $\underset{\sim}{e}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{e}_N$ is bounded, there is a $\underline{\tau}$ such that $1 - \tau \underset{\sim}{e}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{e}_N > 0$ for $0 < \tau < \underline{\tau}$. For $0 < \tau < \underline{\tau}$, the right-hand side of (5.3) is the sum of a positive-definite matrix and a positive semi-definite matrix; hence it is positive definite. Q.E.D.

It is commonly remarked that an asset market does not permit perfect diversification if all the assets share a common risk, such as the risk of war. Proposition 18 is a precise statement of when a large market permits perfect diversification. We now give two examples which show that the condition that $\{x_1, x_2, \ldots\}$ have an equally correlated component is a very stringent one. It is not enough for the assets $\{x_1, x_2, \ldots\}$ all to be positively correlated with the same random factor. Suppose, for example, that

$$x_i = r + \alpha v + \varepsilon_i \qquad \text{(i odd)}$$

$$x_i = v + \varepsilon_i \qquad \text{(i even)},$$

where $0 < \alpha < 1$, $r > 0$, and v, and all the ε_i's are zero-mean uncorrelated random variables. Form an asymptotically riskless portfolio by investing $1/(2N)$ in each of the first N odd assets and $-\alpha/(2N)$ in each of the first N even assets. Net investment is $(1-\alpha)/2$ and the random return is

$$r/2 + \frac{1}{2}\left(\sum_{i=1}^{N} \varepsilon_{2i-1}/N - \alpha \sum_{i=1}^{N} \varepsilon_{2i}/N\right),$$

which converges in quadratic mean to $r/2$. Thus, there is a riskless asset. If $\alpha = 1$, there is an equally correlated component, but there is also a nontrivial riskless hedge portfolio which yields a positive return, so condition (A) is violated.

For a second example, consider the exact K-factor model:

$$\underset{\sim}{\Sigma}_N = \underset{\sim}{B}_N \underset{\sim}{B}_N' + \underset{\sim}{D}_N , \tag{5.4}$$

where $\underset{\sim}{D}_N$ is diagonal and $\underset{\sim}{B}_N$ is N×K. Let

$$\lambda_{1N} \geq \lambda_{2N} \geq \cdots \geq \lambda_{NN}$$

be the eigenvalues of $\underset{\sim}{\Sigma}_N$; then it is an immediate consequence of Exercise 1.f.1.9 in Rao [11] that $\lambda_{K+1,N} \leq \underset{i \leq N}{\text{Max}}\ d_{ii} \leq M$, so that

$$\lim_{N \to \infty} \frac{\lambda_{K+1,N}}{N} = 0 . \tag{5.5}$$

PROPOSITION 19: If (5.5) holds, there is a riskless asset if the multiple correlation between $\underset{\sim}{e}_N$ and the eigenvectors corresponding to the K largest eigenvalues of $\underset{\sim}{\Sigma}_N$ does not converge to ± 1.

PROOF: If $\underset{\sim}{t}_{jN}$ $(j = 1,\ldots,N)$ are eigenvectors corresponding to λ_{jN} $(j = 1,\ldots,N)$,

$$\underset{\sim}{e}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{e}_N = \sum_{j=1}^{N} \lambda_{jN}^{-1} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2$$

$$\geq \sum_{j=K+1}^{N} \lambda_{jN}^{-1} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2$$

$$\geq \frac{N}{\lambda_{K+1,N}} \sum_{j=K+1}^{N} \frac{(\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2}{N} .$$

Since $N \lambda_{K+1,N}^{-1} \to \infty$, $\underset{\sim}{e}_N' \underset{\sim}{\Sigma}_N^{-1} \underset{\sim}{e}_N$ will diverge unless $\Sigma_{j=K+1}^{N} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2/N \to 0$. However, $\Sigma_{j=1}^{N} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2 = (\underset{\sim}{e}_N' \underset{\sim}{e}_N) = N$, so $\Sigma_{j=K+1}^{N} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2/N$ will not converge to zero if $\Sigma_{j=1}^{K} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2/N$ does not converge to one. Since $\underset{\sim}{t}_{jN}' \underset{\sim}{t}_{jN} = 1$, $\underset{\sim}{t}_{jN}' \underset{\sim}{t}_{kN} = 0$ $(j \neq k)$, we see that $\Sigma_{j=1}^{K} (\underset{\sim}{e}_N' \underset{\sim}{t}_{jN})^2/N$ is the squared multiple correlation coefficient between $\underset{\sim}{e}_N$ and the eigenvectors corresponding to the largest K eigenvalues of $\underset{\sim}{\Sigma}_N$.

Q.E.D.

Propositions 17, 18, and 19 state the conditions for the existence of a riskless asset in three different ways. Whether these conditions are met is an empirical question. While Propositions 18 and 19 may make it appear that the sufficient conditions are quite weak, it should be remembered that it is an immediate consequence of Proposition 11 that a necessary condition for the existence of a riskless asset is that the mean-variance efficient frontier be a straight line. This seems to be a strong condition. It is also one which can be checked empirically.

FOOTNOTES

[1]We are indebted to Bob Anderson, Steve Ross, and Jose Scheinkman for helpful suggestions and to the National Science Foundation for research support.

[2]For precise definitions see Section 4 below.

[3]A simple example demonstrates that his conclusions do not hold without further assumptions. Consider a market on which N+1 assets are traded. Let x_i, $i = 0,1,\ldots,N$, be the random return on the i^{th} asset and suppose that

$$x_0 = 2 + \varepsilon_0$$
$$x_i = 1 + \varepsilon_i \qquad (i = 1,2,\ldots),$$

where the ε_i are i.i.d. random variables with $E(\varepsilon_i) = 0$, $i = 0,1,\ldots,N$. Then all risk-averse investors will choose portfolios consisting of a combination of x_0 and the diversified portfolio $\overline{x}_N = N^{-1}\Sigma_1^N x_i$. As $N \to \infty$, $\overline{x}_N$ becomes a riskless asset with return 1. The relative price of x_0 and $\overline{x}_N$ will be determined by investors' attitudes towards risk. Arbitrage alone cannot determine the relative prices. The same conclusions would hold if there were a simple factor structure so that

$$x_0 = 2 + \beta_0 \delta + \varepsilon_0$$
$$x_i = 1 + \beta_i \delta + \varepsilon_i ,$$

where δ is a mean-zero common factor. Consider for simplicity the case where $\beta_0 = \beta$, $\beta_i = +1$ if i is even, and $\beta_i = -1$ if i is odd.

[4]Connor requires that the supply of assets be "well diversified." In the example of the previous footnote, they are not. As the number of assets increases, the fraction of asset $i = 1,2,\ldots$ in everyone's portfolio shrinks to insignificance. Asset 0 on the other hand will continue to be a significant fraction of everyone's portfolio. If this is to be true in

equilibrium, there must be much more of asset 0 than assets 1,2,... . Supply is not well diversified.

[5]In this paper, linear combinations are finite linear combinations. $\Sigma_1^\infty \alpha_i x_i$ is not a linear combination of $x_1, x_2, \ldots$.

[6]Bob Anderson and Jose Scheinkman suggested that we work with $\| \|_*$.

REFERENCES

[1] Black, F., M. C. Jensen, and I. Friend: "The Capital Asset Pricing Model: Some Empirical Tests," in M. C. Jensen (ed.), Studies in the Theory of Capital Markets (New York: Praeger, 1972).

[2] Chamberlain, G.: "A Characterization of the Distributions that Imply Mean-Variance Utility Functions," University of Wisconsin-Madison, Social Systems Research Institute Workshop Series No. 8027, 1980; Journal of Economic Theory, forthcoming.

[3] __________: "Multivariate Regression Models for Panel Data," University of Wisconsin-Madison, Social Systems Research Institute Workshop Series No. 8103, 1981.

[4] Connor, G.: "Asset Prices in a Well-Diversified Economy," Yale University Technical Paper No. 47, 1980.

[5] Fama, E. F.: Foundations of Finance (New York: Basic Books, 1976).

[6] Friend, I.: "Rates of Return on Bonds and Stocks, the Market Price of Risk, and the Cost of Capital," Working Paper No. 23-73, Rodney L. White Center for Financial Research, 1973.

[7] Huberman, G.: "A Simple Approach to Arbitrage Pricing Theory," Journal of Economic Theory, 1981 forthcoming.

[8] Lawley, D. N. and A. E. Maxwell: Factor Analysis as a Statistical Method (New York: Elsevier, 1971).

[9] Morgan, I. G.: "Prediction of Return with the Minimum Variance Zero-Beta Portfolio," Journal of Financial Economics, 2 (1975), 361-376.

[10] Myers, S.: "A Reexamination of Market and Industry Factors in Stock Price Behavior," Journal of Finance (1973), 695-705.

[11] Rao, C. R.: Linear Statistical Inference and Its Applications, 2nd edition (New York: John Wiley & Sons, 1973).

[12] Reed, M. and B. Simon: Methods of Modern Mathematical Physics. I: Functional Analysis (New York: Academic Press, 1972).

[13] Riesz, F. and B. Sz.-Nagy: Functional Analysis (New York: Ungar, 1955).

[14] Roll, R.: "A Critique of the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory," Journal of Financial Economics, 4 (1977), 129-176.

[15] Roll, R. and S. A. Ross: "An Empirical Investigation of the Arbitrage Pricing Theory," Journal of Finance, 5 (1980), 1073-1103.

[16] Ross, S. A., "The Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory, 13 (1976), 341-360.

[17] ________: "The Capital Asset Pricing Model (CAPM), Short-Sale Restrictions, and Related Issues," Journal of Finance, 32 (1977), 177-183.

[18] ________: "Return, Risk, and Arbitrage," in I. Friend and J. L. Bicksler (eds.), Risk and Return in Finance, I (Cambridge, Mass.: Ballinger, 1977).