Reference Variables, Factor Structure, and the Approximate Multibeta Representation
Author(s): Haim Reisman
Source: *The Journal of Finance*, Vol. 47, No. 4 (Sep., 1992), pp. 1303-1314
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328941
Accessed: 25/11/2009 11:57

THE JOURNAL OF FINANCE • VOL. XLVII, NO. 4 • SEPTEMBER 1992

Reference Variables, Factor Structure, and the Approximate Multibeta Representation

HAIM REISMAN*

ABSTRACT

The Arbitrage Pricing Theory (APT) implies that if asset returns have a factor structure, then an approximate multibeta representation holds with respect to the factors as reference variables. This paper assumes that asset returns satisfy a factor structure and derives a condition under which the approximate multibeta representation holds with respect to a set of reference variables which may not be the factors. This condition is that the regression matrix of the reference variables on the factors is nonsingular. Implications for the testability of the APT are also discussed.

MANY ASSET PRICING MODELS imply that the expected return of each asset is a linear function of the covariances between the asset return and some reference variables. This relation is called an exact multibeta representation. The Arbitrage Pricing Theory (APT)[1] implies that if an infinite set of asset returns satisfy a factor structure, then the multibeta representation holds approximately with respect to the factors as reference variables. The approximation is in the sense that the sum of the squares of the pricing mistakes is finite.

This paper derives conditions under which the multibeta representation holds approximately with respect to a given set of reference variables which may be viewed as instruments or proxies for the factors. The approximation is in the same sense as the approximation in the usual APT. These conditions are that asset returns have a factor structure and that the regression matrix of the reference variables with respect to the factors is nonsingular. The result obtained is surprisingly strong. For example, in the special case in which there is only one factor the result implies that any random variable with non-zero correlation with the factor may be used as a reference variable.

Although the proof relies on some mathematics with which many readers will not be familiar, the basic idea is easy to describe (see also Shanken

* Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, Haifa 32000, Israel. I received helpful comments from seminar participants at the University of Maryland, the University of Rochester, Washington University, and Yale University. I am especially grateful to J. Shanken for suggestions and comments that improved the exposition of the paper.

[1] By the APT we mean the version of the APT studied by Ross (1976), Huberman (1982), Chamberlain and Rothschild (1983), Reisman (1988), and others. This version of the theory is called the Arbitrage-APT. We do not discuss in this paper the Equilibrium-APT studied by Connor (1984), Dybvig (1983), and Grinblatt and Titman (1983).

(1992) for an alternative proof). For simplicity assume that there is a riskless asset and that excess returns, x_i, conform to a one-factor model, i.e.,

$$x_i = E[x_i] + b_i f + e_i \text{ for all } i, \tag{1}$$

where $E[e_i] = E[f] = \text{cov}(e_i, f) = 0$ and the e_i's are weakly correlated in a sense made precise in the paper. The intuition underlying the APT is that, if two assets have the same factor betas, then they are (approximate) substitutes and therefore should have (approximately) the same expected returns, i.e.,

$$E[x_i] \approx \gamma b_i, \tag{2}$$

where γ is the price of risk and "$\approx$" indicates that the approximation is in the sense that the infinite sum of squared differences is finite.

Reisman (1988) extends this intuition, showing that a similar result holds, not only for expected returns, but for any return attribute that satisfies a certain continuity requirement. In this paper we focus on a particular attribute, an asset's beta, call it k_i, with respect to the given reference variable g. Thus, it can be shown that

$$k_i \approx \text{cov}(f, g) b_i, \tag{3}$$

with "$\approx$" defined as in (2). If $\text{cov}(f, g) \neq 0$, we also have

$$\gamma b_i \approx \lambda k_i, \tag{4}$$

where $\lambda = \gamma/\text{cov}(f, g)$. Noting that

$$E[x_i] - \lambda k_i = (E[x_i] - \gamma b_i) + (\gamma b_i - \lambda k_i), \tag{5}$$

and combining (2), (3), and (4), it follows that

$$E[x_i] \approx \lambda k_i. \tag{6}$$

Thus expected returns are also approximately linear in the betas on the reference variable g, which can be any variable correlated with the factor f.

Shanken (1982, 1985) has questioned whether it is appropriate to interpret the usual APT approximation as roughly comparable to the exact expected return relation of the capital asset pricing model (CAPM). He emphasizes that, while the APT implies that pricing mistakes must be small for "most" assets in an infinite asset economy, the theory is consistent with deviations which may not be small for a finite subset of assets. It is interesting to reconsider this point in light of our new results.

If the APT approximation is to be interpreted, for practical purposes, as an equality, then we are led to the uncomfortable conclusion that expected returns should be a linear function of the betas with respect to virtually any set of reference variables. For example, if there are five true factors, then almost any set of five individual securities will satisfy the weak invertibility condition and thus could serve as reference variables. Insofar as this implication appears unreasonable, we must question the initial premise, i.e., the traditional view of the theoretical APT approximation. This is the same conclusion reached by Shanken, although by a different logical route.

Further insight into the puzzle posed above is obtained by noting that an upper bound (derived in the paper) on the APT approximation increases as the correlation between the true factors and the reference variables decreases. Thus an arbitrary set of reference variables may not do as good a job at pricing as the true factors. Shanken (1982), however, noted that the bound used in the original APT derivation is a mathematical tautology for any finite set of assets, i.e., if we estimate the bound for a finite set of assets, then there is no point in estimating the approximation mistakes and testing if the bound holds. This is because the mathematics implies that it must hold. The economic content of the theory lies in the fact that the series of bounds does not diverge as the number of assets approaches infinity. Similar conclusions apply to the bounds derived here.

The plan of the paper is as follows. Section I introduces the economic setting of the paper. The definition of the approximate multibeta representation is given in Section II. Section III describes well known results about approximations of linear functionals. The main results of the paper are introduced in Section IV. Further discussion on the testability of the model is given in the concluding Section V. The proofs are given in the Appendix.

I. Factor Structure

This section describes an economy with a sequence of assets with payoffs that satisfy a factor structure and introduces some notations needed later in the paper. Uncertainty in our economy is modeled by a probability space (Ω, F, P), where Ω is the set of all states of the economy and F is the set of all distinguishable events at some terminal time T. $L = L^2(\Omega, F, P)$ is the space of all square integrable random variables. Under consideration is a sequence of time T payoffs $\{x_i \in L\colon i \geq 1\}$ that will remain fixed throughout the paper. Let $g_1, \cdots, g_N \in L$ be given. Denote $g_0 = 1$ and assume without loss of generality, that for every $i \neq j$, $E[g_i g_j] = 0$, and for each i, $E[g_i^2] = 1$. For each i, define

$$x_i - \text{proj}(x_i | g_0, \cdots, g_N) = e_i, \tag{7}$$

where $\text{proj}(x_i | g_0, \ldots, g_N)$ is the orthogonal projection of x_i on span $\{g_0, \ldots, g_N\}$. Denote the largest eigenvalue of the covariance matrix of $e_1, \ldots, e_M$ by $\Lambda_M^2 = \Lambda_M^2(\{g_1, \ldots, g_N\})$ and let

$$\sup_{M \geq 1} \Lambda_M(\{g_1, \ldots, g_N\}) = \Lambda(\{g_1, \ldots, g_N\}). \tag{8}$$

$\Lambda = \Lambda(\{g_1, \ldots, g_N\})$ is called the *correlation module* of the sequence $\{x_i; i \geq 1\}$, with respect to $\{g_1, \ldots, g_N\}$. If $\Lambda(\{g_1, \ldots, g_N\})$ is finite, then the residuals e_i are weakly correlated in the sense of Chamberlain and Rothschild (1983). In this case we say that the x_i's have a *factor structure*, and we call $g_1, \ldots, g_N$ factors.

Examples 1, 2, and 3 provide some intuition on the interpretation of the definition of the correlation module. In Example 1, the e_i's are uncorrelated and we have a finite correlation module.

Example 1: Assume that $\text{cov}(e_i, e_j) = 0$ for every $i \neq j$ and $\text{var}(e_i) = 1$ for each i. Then, for each M, the covariance matrix of $e_1, \ldots, e_M$ is diagonal and $\Lambda(\{g_1, \ldots, g_N\}) = 1 < \infty$.

In Example 2 the e_i's are perfectly correlated and we have an infinite correlation module.

Example 2: Assume that $e_i = e$ for every i and $\text{var}(e) = 1$. Then, $\text{cov}(e_i, e_j) = 1$ for every i and j. This implies that for each M, the largest eigenvalue of the covariance matrix of $e_1, \ldots, e_M$ is M. Therefore, $\Lambda_M^2(\{g_1, \ldots, g_N\}) = M$ and we get that $\Lambda(\{g_1, \ldots, g_N\}) = \infty$.

Let $f_1, \ldots, f_N \in L$ be given. Denote $f_0 = 1$, and assume without loss of generality that, for every $i \neq j$, $E[f_i f_j] = 0$, and for each i, $E[f_i^2] = 1$.

In Example 3 below we have a factor structure, i.e., the correlation module relative to some factors is finite. However, the correlation module relative to a proxy is not finite. This is possible even in cases where the proxy is a small perturbation to the factors.

Example 3: Assume that $N = 1$ and that for each i, $x_i = f_1$. All the residuals relative to f_1 are zero. Therefore, $\Lambda(\{f_1\}) = 0$. Assume that g_1 and f_1 are not perfectly correlated. There exist constants a and b such that $f_1 = a + bg_1 + e$, where $\text{cov}(g_1, e) = 0$, $E[e] = 0$, and $e \neq 0$. This implies that for each i, $x_i = a + bg_1 + e$. Therefore, all the residuals relative to g_1 are equal to e. Example 2 implies that in this case $\Lambda(\{g_1\}) = \infty$.

The *projection matrix*, G, between $\{g_1, \ldots, g_N\}$ and $\{f_1, \ldots, f_N\}$ is defined as the $N \times N$ matrix

$$G = (G_{n,m})_{n,m=1,\ldots,N} \tag{9}$$

where for each n and m

$$G_{n,m} = E[f_n g_m]. \tag{10}$$

Assuming G is nonsingular, the *projection module*, K^2, between $\{g_1, \ldots, g_N\}$ and $\{f_1, \ldots, f_N\}$ is defined as the largest eigenvalue of $(G^{-1})^t G^{-1}$, where G^t is the transpose of G. The definition of G implies that

$$(\hat{g}_1, \ldots, \hat{g}_n) = (f_1, \ldots, f_N)G, \tag{11}$$

where each $\hat{g}_n$ is the orthogonal projection of g_n on $\text{span}\{f_1, \ldots, f_N\}$, the linear space spanned by $\{f_1, \ldots, f_N\}$. (11) implies[2] that $1 \leq K < \infty$. In addition, $K = 1$, if $\{g_1, \ldots, g_N\}$ is a basis for $\text{span}\{f_1, \ldots, f_N\}$.

[2] (11) implies that $N = (\Sigma_{n=1}^N \|f_n\|^2) \leq K(\Sigma_{n=1}^N \|\hat{g}\|^2) \leq KN$, where $\| \, \|$ denotes the norm in L. Therefore, $K \geq 1$ and $K = 1$ if $\|\hat{g}_n\| = 1$ for each n. This is equivalent to $\{f_1, \ldots, f_N\}$ being a basis for $\text{span}\{g_1, \ldots, g_N\}$.

Example 4: In the case where $N = 1$, the projection matrix is given by

$$G = (\operatorname{cov}(f_1, g_1)). \tag{12}$$

G is nonsingular if $\operatorname{cov}(f_1, g_1) \neq 0$. In addition, the projection module between $\{f_1\}$ and $\{g_1\}$ is given by

$$K^2 = 1/\operatorname{cov}^2(f_1, g_1). \tag{13}$$

Note that K^2 approaches infinity as the correlation between f_1 and g_1 approaches zero.

II. The Approximate Multibeta Representation

Let Φ be a functional on L. We say that the *approximate multibeta representation* (AMBR) holds for the sequence $\{\Phi(x_i);\ i \geq 1\}$ with respect to $\{g_1, \ldots, g_N\}$ as reference variables, if there exist $\lambda_0, \ldots, \lambda_N \in R$ such that

$$\Phi(x_i) = \Sigma_{n=0}^{N} k_{i,n} \lambda_n + \xi_i; \text{ for each } i, \tag{14}$$

and

$$\Sigma_{i=1}^{\infty} \xi_i^2 < \infty, \tag{15}$$

where for each i and each n, $k_{i,n}$, the beta of x_i with respect to g_n, is defined as

$$k_{i,n} = E[g_n x_i], \tag{16}$$

In particular $k_{i,0} = E[x_i]$ since $g_0 = 1$.

The AMBR implies that for each i, $\Phi(x_i)$ can be approximated by a linear combination, with coefficients $\lambda_0, \ldots, \lambda_N$, of the betas of x_i with the reference variables, i.e.,

$$\Phi(x_i) \approx \Sigma_{n=0}^{N} k_{i,n} \lambda_n, \tag{17}$$

where ξ_i is the approximation mistake. The approximation mistakes are small in the sense that the infinite sum of the squares of the approximation mistakes is bounded. This implies that most of the approximation mistakes are small, although nothing is implied about any specific approximation mistake.

III. The AMBR with Factors as Reference Variables

This section reviews some results about the AMBR in the case where the reference variables are factors. These results follow from Reisman (1988).

Theorem 1: *Let Φ be a continuous linear functional on L. Then for each i*

$$\Phi(x_i) = \Sigma_{n=0}^{N} k_{i,n} v_n + \zeta_i, \tag{18}$$

where for each n,

$$\nu_n = \Phi(g_n), \tag{19}$$

$$k_{i,n} = E[x_i g_n], \tag{20}$$

and for each M

$$\Sigma_{i=1}^{M} \zeta_i^2 \le C_*^2 \Lambda_M^2(\{g_1, \dots, g_N\}). \tag{21}$$

where C_* is the norm[3] *of the restriction of the functional Φ to the orthogonal complement of span$\{g_0, \dots, g_N\}$. In addition, if $\Lambda(\{g_1, \dots, g_N\})$ is finite, then*

$$\Sigma_{i=1}^{\infty} \zeta_i^2 \le C_*^2 \Lambda^2(\{g_1, \dots, g_N\}) < \infty. \tag{22}$$

The Theorem implies that if the x_i's have a factor structure, then the AMBR holds with the factors as reference variables.

The APT follows from Theorem 1. Assume there is a riskless asset and that each x_i is an excess return. Thus, x_i can be viewed as a payoff with zero cost. Denote by $\hat{L}$ the span of the x_i and 1. We assume that there exists a continuous linear functional $\hat{\Phi}$ on $\hat{L}$ such that for each i, $\hat{\Phi}(x_i) = 0$, $\hat{\Phi}(1)$ is equal to the price of the payoff 1, and $\hat{\Phi}(1) > 0$. $\hat{\Phi}$ is the restriction of the price functional to $\hat{L}$. This assumption is equivalent to the assumption about no asymptotic arbitrage made in Huberman (1982).[4] The assumption that $\hat{\Phi}$ as above exists implies also that the slope of the mean variance efficient frontier is finite. This is proved in the proof of Corollary 1 below. However, the assumption that $\hat{\Phi}$ as above exists does not imply that $\hat{\Phi}$ is positive.

Corollary 1: *Let Φ be as above. Then there exist $v_1', \dots, v_N' \in R$ such that*

$$E[x_i] = \Sigma_{n=1}^{N} k_{i,n} v_n' + \zeta_i', \tag{23}$$

where for each M

$$\Sigma_{i=1}^{M} \zeta_i'^2 \le S^2 \Lambda_M^2(\{g_1, \dots, g_N\}). \tag{24}$$

and S is the slope of the mean variance efficient frontier of portfolios in the x_i's. In addition, if $\Lambda(\{g_1, \dots, g_N\}) < \infty$, then

$$\Sigma_{i=1}^{\infty} \zeta_i'^2 \le S^2 \Lambda^2(\{g_1, \dots, g_N\}). \tag{25}$$

The proof of the corollary is given in Appendix A. Related bounds were derived by Chamberlain and Rothschild (1983) and by Shanken (1987).

[3] $C_*^2 = \|\Phi\|^2 - \Sigma_{n=0}^{N} \Phi(g_n)^2$, where $\|\Phi\|$ denotes the norm of Φ.

[4] The Hahn Banach Theorem implies that $\hat{\Phi}$ as above exists iff 1 is not in the closure of the linear span of the x_i. This is true iff there exists no sequence of payoffs of zero cost portfolios that converges in L to 1. This holds iff there exists no sequence of payoffs of zero cost portfolios with variances converging to one and with expectations converging to zero. This holds iff the expectations of every sequence of zero cost payoffs converge to zero if the variances of the payoffs converge to zero. This implies that Huberman's condition holds because it is equivalent to the condition that the expectations of every sequence of portfolio returns converge to the risk-free rate if the variances of the returns converge to zero.

To get an idea about the size of the bound, assume that there is a strict factor structure and assume that the standard deviation of all x_i's is the same and is equal to σ. Then $\Lambda \leq \sigma$. We denote the expected excess return of a mean variance efficient portfolio with a standard deviation of σ by v. This implies that $S = v/\sigma$. In this case we get that the bound on the sum of the squares of the pricing mistakes is smaller than v^2.

IV. The AMBR with General Reference Variables

Theorem 2 below is the main result of the paper.

Theorem 2: *Let Φ be a continuous linear functional on L. Assume that the projection matrix, G, between $g_1, \ldots, g_N$ and $f_1, \ldots, f_N$ is nonsingular. Define*

$$(\mu_1, \ldots, \mu_N)G^{-1} = (\lambda_1, \ldots, \lambda_N);\ \mu_0 = \lambda_0, \tag{26}$$

where for each n,

$$\Phi(f_n) = \mu_n, \tag{27}$$

and

$$E[x_i g_n] = k_{i,n}. \tag{28}$$

Then

$$\Phi(x_i) = \Sigma_{n=0}^{N} k_{i,n}\lambda_n + \xi_i, \tag{29}$$

where for each M,

$$\Sigma_{i=1}^{M} \xi_i^2 \leq C(K)^2 C_{**}^2 \Lambda_M^2(\{f_1, \ldots, f_N\}), \tag{30}$$

where

$$C_{**}^2 = \|\Phi\|^2 - \Phi(1)^2, \tag{31}$$

$$C(K) = 1 + (N(K^2 - 1))^{1/2}, \tag{32}$$

K^2 is the projection module defined in Section ii, and N is the number of reference variables. In addition, if $\Lambda(\{f_1, \ldots, f_N\})$ is finite, then

$$\Sigma_{i=1}^{\infty} \xi_i^2 \leq C(K)^2 C_{**}^2 \Lambda^2(\{f_1, \ldots, f_N\}). \tag{33}$$

The proof is given in Appendix B.

Theorem 2 implies that if there exists a factor structure, then the AMBR holds for every set of reference variables such that, G, the matrix of coefficients from regressing the reference variables on the factors is nonsingular. The Theorem also derives a bound on the sum of the squared deviations from exact pricing relative to a proxy for the factors.

The difference between Theorem 2 and Theorem 1 is that, if we apply Theorem 1 to a proxy for the factors, the bound obtained there depends on the correlation module of the residuals obtained from regressing the returns on the proxy. On the other hand, the bound implied by our theory is in terms of the correlation module for the true factor model and the projection module between the proxy and the factors. Example 3 demonstrates that the bound implied by Theorem 1 may not be finite even if the proxy is a small perturbation of the factors. This example also demonstrates that the bound provided by Theorem 1 may be substantially larger than the bound obtained in Theorem 2, when the number of assets is large but finite.

To discuss the implications of Theorem 2 to the APT, assume that the x_i's are excess returns.

Corollary 2: *Assume that the conditions of Corollary* 1 *and Theorem* 2 *hold. Then there exist* $\lambda'_1, \ldots, \lambda'_N \in R$ *such that*

$$E[x_i] = \Sigma_{n=1}^{N} k_{i,n}\lambda'_n + \xi'_i, \tag{34}$$

where for each M

$$\Sigma_{i=1}^{M} \xi'^2_i \leq C(K)^2 S^2 \Lambda_M^2(\{f_1, \ldots, f_N\}), \tag{35}$$

where S is as in Corollary 1 *and* $C(K)$ *and* $k_{i,n}$ *are as in Theorem* 2. *In addition, if* $\Lambda(\{f_1, \ldots, f_N\})$ *is finite, then*

$$\Sigma_{i=1}^{\infty} \xi'^2_i \leq C(K)^2 S^2 \Lambda(\{f_1, \ldots, f_N\}). \tag{36}$$

The proof follows from Theorem 2 and is similar to the proof of Corollary 1. The bound on the sum of the squared deviations from exact pricing derived here improves the bound derived in Corollary 1. That bound is usually not finite if the reference variables are a proxy for the factors, while the bound derived here is finite as long as the projection matrix of the proxy on the factors is nonsingular.

To illustrate the size of the bound, assume as in the discussion to Corollary 1 that there is a strict factor structure, that the standard deviation of all x_i's is the same and is equal to σ, and that the expected excess return of a mean variance efficient portfolio with a standard deviation of σ is given by v. Then Theorem 2 implies that the sum of the squared deviations from exact pricing relative to the proxy is bounded by $C(K)^2 v^2$.

V. Conclusion

Chamberlain and Rothschild (1983) derived a bound on the sum of the squared deviations from exact pricing relative to any set of reference variables. This bound is finite if the reference variables are the factors. This bound is usually not finite if the factors are replaced by a proxy, even in cases where the proxy is close to the factors. The contribution of this paper is in improving the existing bound and deriving a finite bound for any multivariate proxy provided the regression matrix of the proxy on the factors is nonsingular.

The bound derived here is given in terms of the slope of the mean variance efficient frontier, the correlation module relative to the true factors, and the projection module between the proxy and the factors. We can test jointly our prior belief regarding the size of the above parameters by testing the bound.

If the slope of the mean variance efficient frontier is large, then there exists a portfolio with a large expected return and a small variance. Therefore, the size of slope of the mean variance efficient frontier may be viewed as a measure of the extent to which there exist trading opportunities which may be interpreted as being close to arbitrage. If there exists a strict factor structure, then the correlation module with respect to the true factors is equal to the supremum of the variances of the asset returns. Therefore, the correlation module with respect to the factors may be viewed as a measure of the extent to which there exists a factor structure. The projection module is a measure of the extent to which the proxy and the factors are correlated.

Conditional on a prior belief on the correlations between the proxy and the factors, one might use the bound derived here to test one's prior belief about the extent to which there exists a factor structure and the extent to which there exist arbitrarge opportunities.

Shanken (1987) and Kandel and Stambaugh (1987), considered the general question of how to test an *exact* equilibrium pricing theory when one has a proxy for the equilibrium benchmark (market portfolio return in the basic CAPM), assuming a prior belief about the correlation between the benchmark and the proxy. Shanken (1987) derived a bound on the sum of the squared deviations from exact pricing relative to a proxy for a mean variance efficient return. This bound is in terms of the correlation module of the residuals obtained from projecting the returns on the proxy, the slope of the mean variance efficient frontier, and a function of the correlation of the proxy and the mean variance efficient return. Conditional on a prior belief concerning the correlation between the proxy and the true market, one can test the equilibrium restriction that the market portfolio is efficient by testing implications of the bound.

The bound derived here might be used as a way of testing jointly one's prior beliefs about the extent to which the proxy and the factors are correlated, the extent to which there exists a factor structure, and the extent to which there exist arbitrage opportunities. However, such test should not be interpreted as a test of the APT conditional on a prior belief concerning the bound. As was pointed out by Shanken (1982), the approximate APT relationship is a mathematical tautology for any finite set of assets. If we know the bound, then there is nothing to test.

Appendix

A. *Proof of Corollary 1*

By the Riesz representation theorem it follows that there exists a ϕ in the closure of $\hat{L}$ such that $\hat{\Phi}(x) = E[\phi x]$ for every x in $\hat{L}$. Define $\Phi(x) = E[\phi x]$ for every $x \in L$. From Theorem 1

$$0 = \Sigma_{n=0}^{N} k_{i,n} \Phi(g_n) + \zeta_i, \tag{A1}$$

where

$$\Sigma_{i=1}^{M} \zeta_i^{2} \leq C_*^2 \Lambda_M^2. \tag{A2}$$

Since $E[x_i] = k_{i,0}$, then

$$E[x_i] = \Sigma_{n=1}^{N} - k_{i,n}\Phi(g_n)/\Phi(1) + \zeta_i', \tag{A3}$$

where $\zeta_i' = \zeta_i/\Phi(1)$. Therefore

$$\Sigma_{i=1}^{M} \zeta'^2_i \leq \left(C^2{}_*/\Phi(1)^2\right)\Lambda_M^2. \tag{A4}$$

From the definition of C_*, we have that

$$C_*^2 = \|\Phi\|^2 - \Sigma_{n=0}^{N}\Phi(g_n)^2 \leq \left(\|\Phi\|^2 - \Phi(1)^2\right). \tag{A5}$$

The Corollary follows from Proposition 1 below. □

Proposition 1: *Let* Φ *be as above. Then*

$$\|\Phi\|^2/\Phi(1)^2 - 1 = S^2. \tag{A6}$$

Proof of Proposition 1: Chamberlain and Rothschild (1983) showed that among all x in the closure of $\hat{L}$ such that $\Phi(x) = 0$, $\phi - \Phi(\phi)/\Phi(1)$ is the one with the smallest variance for a given mean. This implies that

$$(E[\phi - \Phi(\phi)/\Phi(1)])^2/\mathrm{var}(\phi - \Phi(\phi)/\Phi(1)) = S^2. \tag{A7}$$

First we observe that

$$\|\Phi\|^2 = E[\phi^2] = \Phi(\phi);\ E[\phi] = \Phi(1). \tag{A8}$$

By a straightforward computation

$$E[\phi - \Phi(\phi)/\Phi(1)] = \Phi(1) - \Phi(\phi)/\Phi(1) = -\left(\Phi(\phi) - \Phi(1)^2\right)/\Phi(1). \tag{A9}$$

and

$$\mathrm{var}(\phi - \Phi/\Phi(1)) = E\left[(\phi - \Phi(1))^2\right] = \Phi(\phi) - \Phi(1).^2 \tag{A10}$$

This implies that

$$S^2 = \Phi(\phi)/\Phi(1)^2 - 1, \tag{A11}$$

and the result follows. □

The identification of $\|\Phi\|^2/\Phi(1)^2 - 1$ in terms of the slope of the mean variance efficient frontier was discussed in Chamberlain and Rothschild (1983), Shanken (1982), and Shanken (1987).

B. Proof of Theorem 2

Let M be fixed. Denote $b_{i,n} = E[x_i f_n]$, and $\Lambda_M = \Lambda_M(\{f_1, \dots, f_N\})$. From Theorem 1.

$$\Phi(x_i) = \Sigma_{n=0}^N b_{i,n}\mu_n + \rho_i, \tag{B1}$$

where for each n, $\mu_n = \Phi(f_n)$. Since $C_* \leq C_{**}$, Theorem 1 implies that

$$\Sigma_{i=1}^M \rho_i^2 \leq C_{**}^2 \Lambda_M^2. \tag{B2}$$

For each m, define

$$\psi_m(x) = E[g_m x]; \text{ for every } x \in L. \tag{B3}$$

ψ_m is a continuous linear functional on L. From Theorem 1 (with ψ_m instead of Φ) we have that for each m and each i,

$$k_{i,m} = \Sigma_{n=0}^N E[g_m f_n] b_{i,n} + \rho_{i,m}, \tag{B4}$$

where

$$\Sigma_{i=1}^M \rho_{i,m}^2 \leq \|g_m - \hat{g}_m\|^2 \Lambda_M^2, \tag{B5}$$

Using the definition of G we get that for each i,

$$(k_{i,1}, \dots, k_{i,N})G^{-1} = (b_{i,1}, \dots, b_{i,N}) + (\rho_{i,1}, \dots, \rho_{i,N})G^{-1}. \tag{B6}$$

Substituting (B6) in (B1) we get that

$$\Sigma_{n=0}^N b_{i,n}\mu_n + \rho_i = \Sigma_{n=0}^N k_{i,n}\lambda_n + \xi_i, \tag{B7}$$

where

$$\rho_i - \Sigma_{n=0}^N \rho_{i,n}\lambda_n = \xi_i. \tag{B8}$$

(B8) together with the fact that $\rho_{i,0} = 0$ (follows since $g_0 = 1$) implies

$$\xi_i^2 \leq \rho_i^2 + 2\rho_i\left(\Sigma_{n=1}^N \lambda_n^2\right)^{1/2}\left(\Sigma_{n=1}^N \rho_{i,n}^2\right)^{1/2} + \left(\Sigma_{n=1}^N \lambda_n^2\right)\left(\Sigma_{n=1}^N \rho_{i,n}^2\right). \tag{B9}$$

Since $(\mu_1, \dots, \mu_N)G^{-1} = (\lambda_1, \dots, \lambda_N)$ we have that

$$\Sigma_{i=1}^N \lambda_i^2 \leq K^2\left(\Sigma_{n=1}^N \mu_n^2\right). \tag{B10}$$

Since $\{f_0, \dots, f_N\}$ is an orthonormal set and for each n, $\mu_n = \Phi(f_n)$, then

$$\Sigma_{n=1}^N \mu_n^2 \leq \|\Phi\|^2 - \Phi(f_0)^2 = C_{**}^2. \tag{B11}$$

(B9), (B10), and (B11) imply that

$$\Sigma_{i=1}^M \xi_i^2 \leq \Sigma_{i=1}^M \rho_i^2 + 2KC_{**}\Sigma_{i=1}^M\left(\rho_i\left(\Sigma_{n=1}^N \rho_{i,n}^2\right)^{1/2}\right) + K^2C_{**}^2\Sigma_{i=1}^M\left(\Sigma_{n=1}^N \rho_{i,n}^2\right). \tag{B12}$$

Since

$$\Sigma_{i=1}^M\left(\rho_i\left(\Sigma_{n=1}^N \rho_{i,n}^2\right)^{1/2}\right) \leq \left(\Sigma_{i=1}^M \rho_i^2\right)^{1/2}\left(\Sigma_{n=1}^N\Sigma_{i=1}^M \rho_{i,n}^2\right)^{1/2}, \tag{B13}$$

then the above together with (B2) and (B5) implies that

$$\Sigma_{i=1}^{M} \xi_i^2 \leq \left(1 + 2K\left(\Sigma_{n=1}^{N} \|g_n - \hat{g}_n\|^2\right)^{1/2} + K^2\left(\Sigma_{n=1}^{N} \|g_n - \hat{g}_n\|^2\right)\right) C_{**}^2 \Lambda_M^2. \tag{B14}$$

The Theorem follows since

$$\Sigma_{n=1}^{N} \|g_n - \hat{g}_n\|^2 \leq N(1 - K^{-2}). \tag{B15}$$

The last inequality follows since

$$(g_1 - \hat{g}_1, \ldots, g_N - \hat{g}_N) = (f_1, \ldots, f_N)(I - G), \tag{B16}$$

the largest eigenvalue of $(I - G)^t(I - G)$ is $1 - K^{-2}$; and, in addition, $\|g_n - \hat{g}_n\| \leq 1$ and $\|f_n\| = 1$. □

REFERENCES

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

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

Dybvig, P., 1983, An explicit bound on the deviations from the APT, *Journal of Financial Economics* 12, 483–496.

Grinblatt, M. and S. Titman, 1983, Factor pricing in a finite economy, *Journal of Financial Economics* 12, 497–507.

Huberman, G., 1982, A simple approach to arbitrage pricing theory, *Journal of Economic Theory* 28, 183–191.

Kandel, S. and R. Stambaugh, 1987, On correlation and inferences about mean variance efficiency, *Journal of Financial Economics* 18, 61–90.

Reisman, H., 1988, A general approach to the APT, *Econometrica* 56, 473–476.

Ross, S., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Shanken, J., 1982, The APT: is it testable? *Journal of Finance* 37, 1129–1140.

———, 1985, A multibeta CAPM or an equilibrium APT? *Journal of Finance* 40, 1173–1188.

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

———, 1992, The current state of the arbitrage pricing theory, *Journal of Finance* 47, 1569–1574.