Journal of Financial Economics 15 (1986) 373–394. North-Holland

PERFORMANCE MEASUREMENT WITH THE ARBITRAGE PRICING THEORY
A New Framework for Analysis*

Gregory CONNOR and Robert A. KORAJCZYK

Northwestern University, Evanston, IL 60201, USA

Received August 1984, final version received May 1985

This paper develops a theory and econometric method of portfolio performance measurement using a competitive equilibrium version of the Arbitrage Pricing Theory. We show that the Jensen coefficient and the appraisal ratio of Treynor and Black are theoretically compatible with the Arbitrage Pricing Theory. We construct estimators for the two performance measures using a new principal components technique, and describe their asymptotic distributions. The estimators are computationally feasible using a large number of securities. We also suggest a new approach to testing for the correct number of factors.

1. Introduction

The measurement of portfolio performance is an important practical application of asset pricing theory. Two popular measures of performance are the 'Jensen coefficient' and Treynor and Black's 'appraisal ratio'. Using the Capital Asset Pricing Model (CAPM), Jensen (1968) suggests that a positive deviation of a portfolio's average return from that predicted by the security market line (the Jensen coefficient) indicates superior performance. The appraisal ratio is a refinement of Jensen's measure and is equal to the ratio of the Jensen coefficient to the amount of non-market risk undertaken by the manager.

This paper develops analogous performance measures in an Arbitrage Pricing Theory (APT) framework by extending Connor's (1984) equilibrium version of the APT to include a small set of investors with superior information. Estimators of the performance measures are suggested and their asymptotic distributions are derived. The paper shows that:

(1) The Jensen coefficient is an appropriate indicator of superior performance in our equilibrium APT model. That is, an investor's portfolio expected

*We would like to thank Douglas Diamond, Robert Hodrick, Ravi Jagannathan, Larry Jones, Robert McDonald, Carol Mershon, Peter Rossi, Daniel Siegel, Sheridan Titman, and especially the referee, Jay Shanken, for helpful comments.

0304-405X/86/$3.50

return plots above the security market line if and only if he has superior information.

(2) Treynor–Black's appraisal ratio is a valid measure of level of performance, after making additional preference and distributional assumptions. Given that information signals are normally distributed and investors have constant absolute risk aversion, one informed investor has a higher appraisal ratio than another if and only if he has an information set which both investors would strictly prefer.

(3) The two performance measures can be consistently estimated from observable variables. These estimators do not require that the econometrician observe all the assets in the economy. Our new estimators belong to a new class of beta pricing model statistics, based on the recent asymptotic principal components theory of Chamberlain and Rothschild (1983). Unlike Chamberlain and Rothschild, our procedure does not assume knowledge of the true covariance matrix of asset returns.

(4) A significant advantage of our estimators is that large numbers of securities can be used in estimating factor returns. Previously, it was common to use small subsets of securities or to group securities into portfolios. The former alternative reduces the precision of the estimates while the latter can mask the underlying factor structure. The estimators require only an approximate factor structure; previous APT estimation models generally use an exact factor structure. Normality of asset returns is not necessary for our procedure.

(5) We propose a new approach to testing whether the correct number of factors has been extracted. The test is asymptotically valid (as the number of assets approaches infinity) for both exact and approximate factor structures.

Section 2 presents the definitions and assumptions of the model and extends the pricing theory of Connor to an economy in which a small number of investors have superior information. In section 3 we prove the validity of the two portfolio performance measures in the model. Section 4 describes consistent estimators for the two performance measures and their asymptotic distributions. Section 5 provides a summary of our results and a brief description of possible improvements, extensions, and applications of the model. The appendix contains the proofs. In a future paper we will evaluate the performance of a sample of mutual funds with the techniques derived here.

2. The pricing model

This section extends the pricing theory presented in Connor to permit a small number of investors with superior information. Since the number of informed investors is 'very small' (infinitesimal) relative to the total number, these investors have no impact on competitive equilibrium prices. This simplifies the analysis by eliminating the rational inferences that uninformed investors would try to make about the information set of informed investors by observing equilibrium prices.

The pricing theory is a direct application of the competitive equilibrium version of Ross's APT derived in Connor. In the model, there exists a countable infinity of assets whose unconditional per dollar returns follow an approximate factor model:[1]

$$\tilde{r}_i = \mu_i + b_{i1}\tilde{f}_1 + b_{i2}\tilde{f}_2 + \cdots + b_{ik}\tilde{f}_k + \tilde{\varepsilon}_i, \qquad i = 1,2,3,\ldots, \tag{1}$$

or, in vector notation: $\tilde{r} = \mu + B\tilde{f} + \tilde{\varepsilon}$. We impose $\mathrm{E}[\tilde{f}_j] = 0$, $\mathrm{E}[\tilde{\varepsilon}_i] = 0$, and $\mathrm{E}[\tilde{f}_j\tilde{\varepsilon}_i] = 0$, for all i, j. An exact factor model requires $V \equiv \mathrm{E}[\tilde{\varepsilon}\tilde{\varepsilon}']$ to be diagonal with bounded diagonal elements, whereas an approximate factor model allows V to have a non-diagonal form and requires bounded eigenvalues. By judicious choice of B, we normalize the factors $\tilde{f}$ to give $\mathrm{E}[\tilde{f}\tilde{f}'] = I_k$, where I_k is the $k \times k$ identity matrix. Define $B^* = (\mu, B)$. The matrices B and B^* will be restricted to have full column rank.

We assume that k, the number of factors, is known. (We will discuss a test for the appropriate number of factors later in the paper.)

A *portfolio* α is a linear functional on R^∞. Unless noted otherwise, portfolios are assumed to have unit cost (i.e., $\alpha'e = 1$, where e is an R^∞ vector of ones). The product of a portfolio with the vector of asset returns is the *portfolio return*. The set of portfolios is restricted to those linear functionals which have a finite mean-squared return, and this mean-square defines a norm on the space of portfolios.[2] Any portfolio α which entirely eliminates idiosyncratic risk, that is, in which $\mathrm{E}[(\alpha'\tilde{\varepsilon})^2] = 0$, is called *well diversified*.

The *market portfolio*, m, is the per-capita supply of assets held by the uninformed investors. The informed investors can be ignored in constructing the market portfolio since they constitute a set of measure zero.

Expected returns are *beta linear* if there exists a scalar γ_0 and k-vector γ such that

$$\mu = \gamma_0 e + B\gamma. \tag{2}$$

The model places restrictions on the special information of informed investors. Informed investors have superior information about asset-specific events

[1] A tilda ($\tilde{}$) is used to denote random variables. Sample values of random variables are shown without tildas. Asset returns are denoted by $\tilde{r}$, while asset excess returns (i.e., returns in excess of the riskless return) are denoted by $\tilde{R}$.

[2] The norm is given by $\|\alpha\|_* = \mathrm{E}[(\alpha'r)^2]^{1/2}$. Using this norm creates equivalence classes containing all linear functionals whose difference in return is uniformly zero. That is, if α and β are such that $\mathrm{E}[(\alpha'r - \beta'r)^2] = 0$, then $\alpha = \beta$.

only; no investor has special information about market-wide events. For simplicity (this assumption can be weakened with only minor changes to the analysis), each informed investor has special information about one asset. The informed investors are indexed by the asset about which they have special information. Informed investor i receives a signal s_i about the return of asset i. The conditional distribution of ε_i obeys

$$\begin{aligned} &\tilde{\varepsilon}_i = \tilde{\eta}_i + \tilde{s}_i, \\ &\tilde{\eta}_i \quad \text{independent of } \tilde{s}_i, \\ &\mathrm{E}_i[\tilde{\eta}_i] = 0, \\ &\sigma_{\eta i} > 0, \qquad \sigma_{si} > 0. \end{aligned} \tag{3}$$

We use E[ ] to denote the unconditional (uninformed) expectation and E_i[ ] to denote the conditional expectation after an informed investor i receives a signal. The public information is assumed to be consistent with the signals under iterated expectations, so that $\mathrm{E}[\tilde{r}_i] = \mu_i$ implies $\mathrm{E}[\tilde{s}_i] = 0$.

Summarizing the assumptions of the model:

Assumption 1. There exists an uncountable infinity of each of N types of uninformed investors and a countable infinity of informed investors, all of whom have risk-averse von Neumann–Morgenstern utility functions.

Assumption 2. There exists a countable infinity of risky assets with bounded variances whose unconditional per-dollar returns obey the approximate factor model (1).

Assumption 3. Informed investor i receives a signal $\tilde{s}_i$ which is independent of $\tilde{f}$ and $\tilde{\varepsilon}_j$, $j \neq i$, and which obeys (3).

Assumption 4. $\mathrm{E}[\tilde{\varepsilon}|f] = 0$ and $\mathrm{E}[\tilde{\varepsilon}\tilde{\varepsilon}'|f] = V$.

Assumption 5. $\mathrm{E}[(m'\tilde{\varepsilon})^2] = 0$.

Let B^{*n} denote the first n rows of B^*, and let V^n denote the first n rows and n columns of V. Throughout the paper, $\| \; \|$ denotes the L^2-norm.[3]

Assumption 6. There exists a $c_1 < \infty$ such that $\|((1/n)B^{*n\prime}B^{*n})^{-1}\| \leq c_1$ for all n.

[3] Given X is a square matrix, then $\|X\| = \max_{g'g=1}|g'Xg|$. For symmetric, positive semi-definite matrices, the L^2-norm equals the largest eigenvalue of the matrix.

Assumption 7. There exists a $c_2 < \infty$ such that $\|V^n\| \le c_2$ for all n.

Assumptions 1–3 are self-explanatory. Assumption 4 guarantees that the idiosyncratic risks are 'increasing risk' in the Rothschild–Stiglitz sense, and that the idiosyncratic returns are conditionally homoskedastic. Assumption 5 states that the market portfolio is well-diversified. Assumption 6 guarantees that the market factors are pervasive forces in the economy, that is, each market factor affects many assets non-negligibly. Assumption 7 comes from Chamberlain (1983) and is symmetric to Assumption 6; it guarantees that the idiosyncratic risks are non-pervasive, so that they disappear from a portfolio with weights spread evenly across the assets [see Chamberlain (1983, theorem 3)].

Connor uses assumptions similar to those above except that all investors are uninformed. That paper uses a competitive equilibrium argument to show that all investors hold well-diversified portfolios. It is also shown that the expected returns of assets are beta linear. Theorem 1 below is an application of Connor.

A *competitive equilibrium* consists of a set of portfolios for the N types of uninformed investors $\{\beta^1, \beta^2, \dots, \beta^N\}$ and a set of portfolios for the informed investors $\{\alpha^1, \alpha^2, \alpha^3, \dots\}$ such that the portfolios are budget-constrained optimal for every i, j, and supply equals demand:

$$\alpha^i \in \arg\max \mathrm{E}_i\left[u^i\left(w_i \alpha^{i\prime}\tilde{r}\right)\right], \qquad i = 1, 2, \dots,$$

$$\beta^j \in \arg\max \mathrm{E}\left[u^j\left(w_j \beta^{j\prime}\tilde{r}\right)\right], \qquad j = 1, 2, \dots, N,$$

$$m = \sum_{j=1}^{N} \left(w_j/\bar{w}\right)\beta^j,$$

where w_i, w_j are initial wealth levels, and $\bar{w} = \sum_{j=1}^{N} w_j$.

Theorem 1. In competitive equilibrium, unconditional expected returns are beta linear and all uninformed investors hold well-diversified portfolios. (All proofs are presented in the appendix).

A portfolio α has a *long position* (*short position*) in asset i as $\alpha' z^i > 0$ (< 0), where z^i is a vector with a one in the ith place and zeros elsewhere. Note that any well-diversified portfolio δ has a zero position in each individual asset i ($\delta' z^i = 0$). Informed investors do not necessarily choose well-diversified portfolios. The following theorem partially describes their optimal choices.

Theorem 2. Informed investor i holds an equilibrium portfolio with a long (short) position in asset i if and only if s_i is strictly positive (strictly negative).

3. Performance measurement in the model

This section shows that the Jensen coefficient is a valid indicator of superior performance for this beta pricing model: an investor's portfolio return will have a positive Jensen coefficient if and only if he has superior information. Furthermore, Treynor–Black's appraisal ratio is also valid and is more precise than Jensen's measure; it balances the gain from exploiting special information against the risk of holding a non-diversified portfolio. Under additional assumptions, the appraisal ratio gives a complete ordering of the values of information sets.

Define a *portfolio rule* $\alpha(\ )$ as a mapping from signals s_i to portfolios α. Since $\alpha(\ )$ is conditioned on s_i, it is possible that $\mathrm{E}[\alpha(\tilde{s}_i)'\tilde{\varepsilon}] > 0$ even though $\mathrm{E}[\tilde{\varepsilon}] = 0$. These expectations will be given the usual statistical time series interpretations. We assume that the economy is repeated many times with independent realizations of the random variates, and investor i uses the portfolio rule $\alpha(\ )$ each period. The sample arithmetic mean return for investor i will approach $\mathrm{E}[\alpha(\tilde{s}_i)'\tilde{r}]$ as the number of time series observations grows large. We seek to measure the difference between this expected return and the expected return to passive (i.e., uninformed) portfolios with the same amount of market risk.

We define the *Jensen coefficient* for the ith investor as $a_\alpha = \mathrm{E}[\alpha(\tilde{s}_i)'\tilde{\varepsilon}]$, which is the APT analog of the CAPM-based performance measure suggested by Jensen. We can describe the Jensen coefficient in the form of a 'security market line' (in our multi-beta model, a security market hyperplane). If there exists a riskless asset or portfolio with return γ_0, then combining (1) and (2) we can write

$$\tilde{R} = \tilde{r} - \gamma_0 e = B(\gamma + \tilde{f}) + \tilde{\varepsilon}, \tag{4}$$

which expresses (2) in terms of 'excess returns' (returns above the riskless rate). If an investor is uninformed, then the conditional expected excess return on every asset given the market factor returns will be the beta vector of the asset times the excess return on the market factors. Using (4), the excess return of a portfolio rule $\alpha(\tilde{s})$, $\tilde{R}_\alpha = \alpha(\tilde{s})'\tilde{R}$ is

$$\tilde{R}_\alpha = a_\alpha + b_\alpha(\gamma + \tilde{f}) + \tilde{\varepsilon}_\alpha, \tag{5}$$

where

$$b_\alpha = \mathrm{E}[\alpha(\tilde{s})'B], \qquad a_\alpha = \mathrm{E}[\alpha(\tilde{s})'\varepsilon],$$

and

$$\tilde{\varepsilon}_\alpha = \alpha(\tilde{s})'\tilde{\varepsilon} - a_\alpha + (\alpha(\tilde{s})'B - b_\alpha)(\gamma + \tilde{f}).$$

For an uninformed investor, a_α is zero. We also prove the converse for an informed investor.

Theorem 3. An investor's equilibrium portfolio rule has a positive Jensen coefficient ($a_\alpha > 0$) *if and only if he is informed.*

An analogous result in a mean-variance framework is derived in Dybvig and Ross (1984, theorem 2). They assume that informed investors receive information about two or more assets such that the signals cancel out and, therefore, give no information about the return on the market portfolio.[4] This is equivalent to our assumption of asset-specific information in an economy with a large number of assets. They show that an investor who chooses conditional mean-variance efficient portfolios will exhibit a positive Jensen coefficient.

Theorem 3 shows that the Jensen coefficient is appropriate for comparing a portfolio's performance against an absolute benchmark of 'no superior performance'. However, it is not appropriate for relative rankings of portfolios since differences in a_α may be due merely to differences in risk exposure (for example, differences caused by levering the same underlying portfolio rule). A performance measure that allows relative ranking is the Treynor–Black appraisal ratio. It is valid in this model under the assumptions that investors have constant absolute risk aversion and that the joint distribution of idiosyncratic risk and information signals is multivariate normal. Recall that $\tilde{\varepsilon}_i = \tilde{\eta}_i + \tilde{s}_i$ for all i.

Assumption 8. $\tilde{\eta}_i, \tilde{s}_i$, $i = 1, 2, \ldots,$ are normal and $\tilde{\eta}_i, \tilde{s}_i$ are independent for each i.

Assumption 9. All investors have constant absolute risk aversion.

Define the *t-ratio* as the Jensen coefficient divided by the standard deviation of the portfolio's idiosyncratic variation:

$$t_\alpha = a_\alpha / \sigma_{\varepsilon\alpha}.$$

The t-ratio is the APT analog of the appraisal ratio that Treynor and Black suggest for a CAPM model. Given constant absolute risk-averse preferences, all investors have t-ratios equal to $\sigma_{si}/\sigma_{\eta i}$. Since these normally distributed signals are completely characterized by this ratio, the t-ratio gives a complete ordering of the superior information of informed investors. The following theorem codifies these ideas.

[4]An example of non-market information, suggested in Dybvig and Ross (1984), is knowledge about which of two firms will win a lawsuit. Since the unexpected gain to one firm equals the unexpected loss to the other firm the net effect on the market portfolio return is zero.

Theorem 4. Assume Assumptions 8 and 9 in addition to Assumptions 1–7. Consider two informed investors i and j. The investors may have differing wealth levels and coefficients of absolute risk aversion. Ex-ante, investor j would strictly prefer to trade his signal for the signal of investor i if and only if $t_{\alpha i} > t_{\alpha j}$.

To apply this theory to the measurement of mutual fund performance, we assume that fund managers act to maximize the expected utility of their clients. Although the model is stated in terms of one actively managed asset position, it is simple to generalize the results to a more realistic environment where a mutual fund manager follows many securities. We have assumed that the informed trader only receives asset-specific information. This precludes information about the realizations of the k factors. Hence market timing (or 'factor timing') activities are not included in the model. The assumption of no market timing ability may be justified on the basis of empirical evidence. Chang and Lewellen (1984), Henriksson (1984) and Kon (1983) do not find evidence consistent with positive timing performance of mutual funds.

4. A new class of estimators

In section 3, we developed a general equilibrium model which produces a multi-beta extension of Jensen's and Treynor and Black's measures of performance. Now we describe an econometric procedure to estimate these measures with asset return data.

Recalling eq. (5), we can write

$$\tilde{R}_{\alpha} = a_{\alpha} + b_{\alpha}(\gamma + \tilde{f}) + \tilde{\varepsilon}_{\alpha}. \tag{6}$$

Given the independence of $\tilde{s}_i$ and $\tilde{f}$, and the assumption that $\mathrm{E}[\tilde{\varepsilon}|f] = 0$, it follows[5] that $\mathrm{E}[\tilde{\varepsilon}_{\alpha}|f] = 0$. Therefore, we can consistently estimate the coefficients a_{α} and b_{α} in (6) by ordinary least squares if we observe $\gamma + f$. If we also assume that $\sigma_{\varepsilon\alpha}$ does not depend upon f (conditional homoskedasticity[6]), then the usual techniques may be used to compute standard errors for the estimated coefficients.

[5] Note that $\mathrm{E}[\alpha(\tilde{s})'Bf|f] = \mathrm{E}[\alpha(\tilde{s})'B]f = b_{\alpha}f$ and the result follows easily.

[6] Despite the assumption of conditional homoskedasticity of the asset's idiosyncratic variates (Assumption 4), the idiosyncratic variate of a managed portfolio can be heteroskedastic, due to the term $(\alpha(\tilde{s})'B - b_{\alpha})(f + \gamma)$ included in $\tilde{\varepsilon}_{\alpha}$. The usual standard error estimates require conditional homoskedasticity.

One solution is to assume constant absolute risk aversion, in which case $\alpha(\tilde{s})'B$ is constant for all s as shown in the proof of Theorem 4. In this case, the offensive term disappears from ε_{α} and it is homoskedastic. Alternatively, one can estimate the standard errors allowing for the conditional heteroskedasticity of ε_{α}. Using Hsieh's (1983) procedure, the covariance matrix of the parameter estimates can be easily adjusted.

Two main difficulties confront us in estimating (6). First, in order to test any hypotheses about a_α and t_α we must rely on asymptotic results (as $T \to \infty$) because the error terms in the above regression need not be normal even if we assume that the factors and idiosyncratic terms have a multivariate normal distribution. The returns on a managed portfolio are the product of the portfolio rule $\alpha(\tilde{s}_i)$ and the returns on assets. These portfolio returns need not be normal even if individual asset returns are normal.[7]

The second, and more consequential, difficulty is that we do not observe the k market factors. To handle this problem, we develop asymptotic principal components estimators based on the recent work of Chamberlain and Rothschild. The asymptotic principal components theory is similar to standard principal components theory except that it uses statistical approximations which are valid as the number of cross-sectional observations grows large.

Our technique assumes that we observe a large collection of asset returns from the economy. We show that the sample principal components of the excess returns on the n observed assets converge to the realized k market factors as n grows large. Thus, our performance measurement procedures require that both n and T be large. As n grows large, we can approximate the realized k market factors with negligible error; as T grows large, we obtain consistent and asymptotically normal least squares estimates of the parameters in (6).[8]

For simplicity of notation, we suppress the T superscript when it is not needed. Let R be the $\infty \times T$ matrix of realizations of the excess returns. We observe R^n, which is the $n \times T$ matrix of excess returns on the first n assets. We show that our statistics based on R^n converge to statistics based on R as n approaches infinity.

Recalling (4), we can write $R^n = B^n F + \varepsilon^n$, where F contains the $k \times T$ realized values of $\gamma + \tilde{f}$ in the sample and ε^n contains the $n \times T$ realized values

[7]As an example, take the case discussed in section 3 where it was assumed that signals and idiosyncratic risks are multivariate normal (Assumption 8) and investors have constant absolute risk aversion (Assumption 9). In this example, the errors in the regression of portfolio returns on market factor excess returns have a distribution that is a mixture of normal distribution with a gamma distribution. That is, each period the errors are drawn from a normal distribution whose variance is drawn from a gamma distribution:

$$p(\sigma^2) = \exp(-\sigma^2/\psi)/(\pi\sigma^2\psi)^{1/2} \quad \text{where} \quad \psi = 2\sigma_s^2/\sigma_\eta^2 A_i^2.$$

This is a rather complicated distribution. Its characteristic function is given by

$$\phi(t) = (1 + \psi t^2/2)^{-1/2}.$$

Different assumptions about the utility functions of investors will lead to different distributions of the error terms. However, as long as the portfolio rule is such that the errors are independent with finite variance we can rely on the asymptotic distributions of $\hat{a}_\alpha$ and $\hat{t}_\alpha$.

[8]The use of the asymptotic distribution, for T large, does not require that we assume constancy of the factor structure over time. It is possible to use the principal components technique to extract excess returns on market factors over subperiods. A switching regression can then be employed to estimate a_α while allowing b_α to change across subperiods.

of $\tilde{\varepsilon}^n$. Define

$$\Omega^n = (1/n)R^{n\prime}R^n,$$

$$A^n = (1/n)F'B^{n\prime}B^nF,$$

$$Y^n = (1/n)(F'B^{n\prime}\varepsilon^n + \varepsilon^{n\prime}B^nF),$$

$$Z^n = (1/n)\varepsilon^{n\prime}\varepsilon^n.$$

The observable matrix Ω^n is the $T \times T$ cross-product matrix of asset excess returns. The unobservable matrix A^n is a $T \times T$ cross-product matrix of factor related returns. Note that $\Omega^n = A^n + Z^n + Y^n$.

Our estimation strategy basically works as follows. We would like to run the following regression:

$$R_\alpha = \hat{a}_\alpha e_T' + \hat{b}_\alpha F + \hat{\varepsilon}_\alpha, \tag{7}$$

where R_α is the T-vector of returns on portfolio α, and e_T is a T-element unit vector. We cannot run (7) because we cannot observe F. We use a proxy for F consisting of the first k eigenvectors of Ω^n. We show that this proxy gives asymptotically identical estimates (as $n \to \infty$) to those obtained if we were able to use the unobservable F.

For any symmetric positive semi-definite $T \times T$ matrix X, let the *principal components matrix* be the orthogonal $k \times T$ matrix of the k eigenvectors of X corresponding to the k largest eigenvalues. We normalize the rows of the principal components matrix to have mean-square of 1. If J is the principal components matrix of X and Λ_J is the diagonal matrix of the corresponding eigenvalues, then by definition:

$$JX = \Lambda_J J \quad \text{and} \quad (1/T)JJ' = I_k.$$

Let G^n denote the (observed) principal components matrix of Ω^n and H^n the (unobserved) principal components matrix of A^n. Let Λ^n_G and Λ^n_H be the matrices of corresponding eigenvalues. One can show that H^n is a non-singular linear transformation of F. That is, there exists a non-singular $k \times k$ transformation $\bar{L}^n$ such that $\bar{L}^nF = H^n$.

Lemma 1. For every n, there exists a non-singular $k \times k$ matrix $\bar{L}^n$ such that $H^n = \bar{L}^nF$.

Temporarily assume that we can observe H^n, and consider the regression (7) using H^n in place of F. Let $\hat{a}^T_\alpha$ and $\hat{\varepsilon}^T_\alpha$ be the estimates from this hypothetical regression, and define $\hat{t}^T_\alpha = \hat{a}^T_\alpha/\hat{\sigma}^T_{\varepsilon\alpha}$. Although our original regression equation used F, replacing F with H^n has no effect on the estimates. The reader can check, using equations (7.18) and (7.23) in Theil (1971, pp. 39–41), that the non-singular $\bar{L}^n$ disappears from the least squares estimates of $\hat{a}_\alpha^T$ and $\hat{\varepsilon}_\alpha^T$. This is intuitively clear: a non-singular $k \times k$ transformation is merely a 'rescaling' of the independent variables, and so will not affect the estimated intercept term or residuals. After adjusting for this non-singular linear transformation, the following theorem is the well-known result about the asymptotic normality of ordinary least squares estimates.

Theorem 5. Given Assumptions 1–7,

$$\operatorname*{dlim}_{T \to \infty} T^{1/2}\left(\hat{a}_\alpha^T - a_\alpha\right) \sim \mathrm{N}\left(0, \sigma_{\varepsilon\alpha}^2(1 + \gamma'\gamma)\right), \tag{8}$$

and

$$\operatorname*{dlim}_{T \to \infty} T^{1/2}\left(\hat{t}_\alpha^T - t_\alpha\right) \sim \mathrm{N}(0, 1 + \gamma'\gamma), \tag{9}$$

where dlim *denotes convergence in distribution.*

Next, we prove that G^n is approximately equal to a non-singular $k \times k$ transformation of F. We need an additional assumption.

Assumption 10. There exists an average cross-sectional idiosyncratic variance,

$$\sigma^2 = \operatorname*{plim}_{n \to \infty} (1/n) \tilde{\varepsilon}^{n\prime} \tilde{\varepsilon}^n.$$

Intuitively, the proof of the next theorem works as follows. We wish to show that G^n (the principal components matrix of Ω^n) is approximately equal to a non-singular transformation of F. We have shown that H^n (the principal components matrix of A^n) is a non-singular transformation of F. We have $\Omega^n = A^n + Z^n + Y^n$. It should be clear that the matrix Y^n goes to a zero matrix with n. The matrix Z^n consists of the time autocovariance of idiosyncratic variates. The cross-terms and the diagonal terms of this matrix are, respectively, of the form

$$(1/n) \sum_{i=1}^{n} \varepsilon_{it} \varepsilon_{i\tau} \quad \text{and} \quad (1/n) \sum_{i=1}^{n} \varepsilon_{it}^2.$$

In Lemma 3 we show that the off-diagonal terms go to zero with n, and the diagonal terms approach σ^2. This implies that Z^n goes to $\sigma^2 I_T$ for large n. Therefore, for large n, Ω^n is approximately equal to $A^n + \sigma^2 I_T$. The eigenvectors of a matrix are unaffected by the addition of a scalar matrix. Therefore, for large n, G^n is approximately equal to H^n (up to a non-singular transformation). Since H^n is a non-singular transformation of F, G^n is approximately a non-singular transformation of F.

Theorem 6. Given Assumptions 1–7 and 10, there exists Φ^n *and non-singular* L^n *such that* $G^n = L^n F + \Phi^n$ *and* $\operatorname{plim}_{n\to\infty}\Phi^n = 0$, *the null matrix.*

Consider running regression (7) using the observable G^n in place of the unobservable H^n. Let $\hat{a}_\alpha^{nT}$ and $\hat{t}_\alpha^{nT}$ be the values calculated from this regression. We combine Theorems 5 and 6 to derive the asymptotic distributions of these observable estimates.

Corollary 1. Given Assumptions 1–7 and 10,

$$\operatorname*{dlim}_{T\to\infty}\ \operatorname*{plim}_{n\to\infty} T^{1/2}\left(\hat{a}_\alpha^{nT} - a_\alpha\right) \sim \mathrm{N}\left(0, \sigma_{\varepsilon\alpha}^2(1+\gamma'\gamma)\right), \tag{10}$$

and

$$\operatorname*{dlim}_{T\to\infty}\ \operatorname*{plim}_{n\to\infty} T^{1/2}\left(\hat{t}_\alpha^{nT} - t_\alpha\right) \sim \mathrm{N}(0, 1+\gamma'\gamma). \tag{11}$$

Note that, for finite n, the use of G^n rather than F in (7) corresponds to a regression with errors in variables. Thus the coefficient estimates are biased. However, Corollary 1 shows that this bias disappears as n becomes large. The mean square error of the parameter estimates will be equal to the sum of the squared bias and the variance of the parameter estimates. The usual computation of the standard errors of the parameter estimates will approach the second component. Thus, for finite n the usual reported standard errors will understate the mean square error by the bias squared. Monte Carlo simulation may provide an indication of the size of n required for our approximation to be reasonable.

To summarize the results of this section, we present an algorithm for testing the performance of a mutual fund. Suppose one observes a T-vector of excess returns on a mutual fund, R_α, and a $n \times T$ matrix of excess returns on a large collection of securities R^n. To test the performance of the fund:

(1) Compute a $k \times T$ principal components matrix of $(1/n)R^{n\prime}R^n$. There are many efficient routines for performing this calculation.
(2) Run the time series regression $R_\alpha = \hat{a}_\alpha + \hat{b}_\alpha G^n + \hat{\varepsilon}_\alpha$. Also calculate $\hat{t}_\alpha = \hat{a}_\alpha / \hat{\sigma}_{\varepsilon\alpha}$.
(3) For large n and T, the statistics $\hat{a}_\alpha^{nT}$ and $\hat{t}_\alpha^{nT}$ approximately have the distributions given by (10) and (11).

The algorithm works with the time-series cross-product matrix of excess returns. For most time period/cross-sectional samples, the time-series matrix has a much smaller dimension than the corresponding cross-sectional cross-product matrix used by standard factor analysis routines. Furthermore, the algorithm calculates the principal components (eigenvectors) of this matrix, which is computationally much easier than calculating factor variates for a matrix of equal dimension.

This algorithm tests the performance of a single fund. Our pricing model allows multiple informed portfolio holders. One can easily adjust the statistics of Corollary 1 to perform tests against a collection of estimates; for instance, to test whether all of the true Jensen coefficients equal zero.

The results of this section also provide a new approach to testing the assumption that the correct number of factors has been extracted. We have shown that $\operatorname{plim}_{n\to\infty}(\Omega^n - A^n) = \sigma^2 I_T$. To test for an approximate factor structure with k factors, one can check whether the last $T-k$ eigenvalues of Ω^n are *equal*.[9] This promising approach to finding the value of k is investigated further in Connor and Korajczyk (1985).

5. Summary

The Capital Asset Pricing Model (CAPM) has played a central role in performance evaluation by providing a benchmark against which actively managed portfolios can be compared. Two of the most influential applications of the CAPM to performance evaluation are the Jensen coefficient and Treynor and Black's appraisal ratio. However, the use of these CAPM-based measures has been called into question on the basis of a number of theoretical considerations and empirical findings. A major criticism is based on the infeasibility of observing the true market portfolio.

We propose performance measures based on an equilibrium version of the Arbitrage Pricing Theory. We derive APT analogs for both Jensen's measure and the Treynor–Black measure, prove their theoretical compatibility with the model, and develop consistent estimators for the performance measures.

[9]Given that X is a positive semi-definite matrix of rank h, let $\operatorname{eig}(X)$ denote the h-vector of the eigenvalues of X ordered from largest to smallest. Note that A^n is positive semi-definite with rank k and its smallest eigenvalue is bounded above zero (let c_3 denote this bound). The scalar matrix $\sigma^2 I_T$ has T eigenvalues all equal to σ^2. It follows that

$$\begin{aligned} \operatorname{eig}_j\left(A^n + \sigma^2 I_T\right) &= \operatorname{eig}_j(A^n) + \sigma^2, && j \le k, \\ &= \sigma^2, && T \ge j > k. \end{aligned}$$

Since $\operatorname{eig}_k(A^n) > c_3 > 0$ for all n and the eigenvalues of the plim equal the plim of the eigenvalues, we have

$$\operatorname*{plim}_{n\to\infty}\left[\operatorname{eig}_k(\Omega^n) - \operatorname{eig}_{k+1}(\Omega^n)\right] > 0, \quad \operatorname*{plim}_{n\to\infty}\left[\operatorname{eig}_{k+1}(\Omega^n) - \operatorname{eig}_T(\Omega^n)\right] = 0.$$

In this model an investor's portfolio return has a positive Jensen coefficient on average if and only if he has superior information. Using added assumptions, Treynor and Black's appraisal ratio has an even stronger justification. We show that one investor has a higher appraisal ratio than another if and only if he has a strictly preferable information set.

We derive consistent estimators for the two performance measures, and show their asymptotic distributions. These estimators are an extension of the asymptotic principal components technique of Chamberlain and Rothschild. The asymptotic principal components theory is similar to standard principal components analysis except that it uses approximations which hold as the number of cross-sections becomes large.

The technique requires only an approximate factor structure, whereas previous APT estimation algorithms have assumed an exact factor structure. Our statistical technique does not assume normally distributed returns, as do previous APT estimation techniques. The method is computationally efficient and therefore does not require that the analyst form portfolios or otherwise restrict the number of securities under consideration. We feel that this is a significant advantage since much of the debate about the applicability of the APT has revolved around the questions of whether current techniques provide 'good' estimates of the model's parameters [e.g., Dhrymes et al. (1984) and Roll and Ross (1984)]. Some ambiguity remains with our APT-based approach since uncertainty about the true number of pervasive factors replaces the uncertainty about the return on the true market portfolio in CAPM-based measures [see Shanken (1982, 1985)]. However, we can test the assumption that the true number of pervasive factors is equal to k by investigating the last $T-k$ eigenvalues of Ω^n, even when the economy only follows an approximate factor structure. That is, the last $T-k$ eigenvalues of Ω^n approach σ^2 as n becomes large. Since our technique allows the use of large numbers of securities, reliance on asymptotic results is not unreasonable.

Additional work could take several directions. Our asymptotic results involve n approaching infinity before T is allowed to become large. This does correspond to a common case in finance in which many more cross-sections are available than time series observations. Ideally we would like to allow n and T to grow simultaneously (possibly with their ratio approaching some limit). We know of no straightforward technique for solving this problem and leave it for future endeavors.

We have ignored the delegated management problem by assuming that mutual fund managers act completely in the interest of their clients. One might wish to modify this assumption, or give a competitive entry model which justifies it. We have eliminated the problem of rational price inference by treating the informed investors as a set of measure zero. It may be possible to extend the analysis to permit truly heterogeneous information. [See Admati and Ross (1985) for discussion of this problem in a CAPM context.]

Further research will include estimation of the model. In a subsequent paper, we intend to apply the measures derived here to the evaluation of open-ended mutual funds. Also, we plan to use this analytical framework to perform tests of the APT. Individual assets or constant composition portfolios should have Jensen coefficients equal to zero if the APT is valid. Thus a direct test of the APT, in our framework, is a test for zero abnormal performance across assets or constant composition portfolios. These future papers will treat a number of remaining empirical and econometric issues.

Appendix

This appendix gives the proofs of the theorems.

Proof of Theorem 1. Since the informed investors have no effect on prices or supply clearing, the proof of this theorem is identical to Connor's Theorems 2 and 3. We will show that Assumptions 1, 2, and 4–7 of our paper imply Assumptions 1*–6* used there. Assumptions 1*–4* are identical to Assumptions 1, 2, 4 and 5. We must show that Assumptions 6 and 7 imply Assumptions 5* and 6*.

Assumption 5.* For any portfolio α there exists a factor-equivalent (i.e., $\alpha' B^* = \delta' B^*$) well-diversified portfolio δ, where δ is not subject to the unit cost constraint.

Assumption 6.* There exists a riskless portfolio.

Consider an arbitrary portfolio α. Construct the n-vector $\delta^n = B^{*n}(B^{*n\prime}B^{*n})^{-1}(\alpha' B^*)'$. Next, construct the portfolio represented by the portfolio weights $\delta^{*n\prime} = (\delta^{n\prime}, 0, 0, 0, \ldots)$. Consider the portfolio $\delta = \lim_{n\to\infty}\delta^{*n}$. It is easy to show (following Connor's Lemma 3) that this limit exists and that the resulting portfolio has the properties $\mathrm{E}[(\delta'\tilde{\varepsilon})^2] = 0$ and $\delta' B^* = \alpha' B^*$. This proves Assumption 5*. Next, consider any 'zero beta' portfolio α_0, i.e., $\alpha_0' B = (0, 0, \ldots, 0)$; such a portfolio exists since B has full column rank by assumption. Construct a well-diversified portfolio δ which is factor-equivalent to α_0. Note that the portfolio is riskless. This proves Assumption 6*. Theorem 1 of this paper then follows from Theorems 2 and 3 of Connor.

It should be noted that, given Theorem 1, factor-equivalent portfolios have equal cost [as is easy to show using (2)]. A portfolio which is factor-equivalent to a unit cost portfolio will have unit cost; a portfolio which is factor-equivalent to a zero cost portfolio will have zero cost.

Proof of Theorem 2. Let α be the equilibrium portfolio of investor i. Construct a well-diversified portfolio δ which is factor-equivalent to α. Thus:

$\alpha'\tilde{r} \equiv \delta'\tilde{r} + \alpha'\tilde{\varepsilon}$. Since $\mathrm{E}[\tilde{\varepsilon}|f] = 0$, the portfolio δ is strictly preferred to α for any risk-averse preferences unless $\mathrm{E}_i[\alpha'\tilde{\varepsilon}] > 0$ or $\mathrm{E}_i[(\alpha'\tilde{\varepsilon})^2] = 0$. The budget-optimality of α ensures that one of these two conditions holds. Since $\mathrm{E}_i[\tilde{\varepsilon}_j] = 0$, $j \neq i$, we know that $\mathrm{E}_i[\alpha'\tilde{\varepsilon}] > 0$ if and only if $(\alpha' z^i)s_i > 0$. We have shown that $(\alpha' z^i)s_i \geq 0$ and that $\mathrm{E}[(\alpha'\tilde{\varepsilon})^2] > 0$ implies $(\alpha' z^i)s_i > 0$.

The last step is to show that the inequality $(\alpha' z^i)s_i \geq 0$ is strict when $s_i \neq 0$. Suppose that $\mathrm{E}[(\alpha'\tilde{\varepsilon})^2] = 0$ when $s_i \neq 0$. Construct a well-diversified portfolio ϕ which has the same factor risk and unconditional expected payoff as asset i:

$$\phi' B^* = z^{i\prime} B^*, \qquad \mathrm{E}\left[(\phi'\tilde{\varepsilon})^2\right] = 0.$$

Construct the costless portfolio $\eta = z^i - \phi$. The portfolio η has a random return of $\tilde{r}_i - \phi'\tilde{r} = \tilde{\varepsilon}_i$. Consider adding a small increment of η to the investor's equilibrium portfolio α. By the budget-optimality of α, the marginal effect on expected utility must be zero:

$$(\mathrm{d}/\mathrm{d}x)\mathrm{E}_i\left[u(\alpha'\tilde{r} + x\tilde{\varepsilon}_i)\right] = 0 \quad \text{at} \quad x = 0. \tag{12}$$

Solving for the derivative in (12),

$$s_i \mathrm{E}_i\left[u'(\alpha'\tilde{r})\right] + \mathrm{E}_i\left[(\tilde{\varepsilon}_i - s_i)u'(\alpha'\mu + \alpha' B\tilde{f})\right]$$

$$= s_i \mathrm{E}_i\left[u'(\alpha'\tilde{r})\right] \neq 0, \quad \text{a contradiction.}$$

Proof of Theorem 3. If an investor j is uninformed, then he cannot condition on any information more precise than the public information set, $\mathrm{E}[\tilde{\varepsilon}] = 0$, and therefore $\mathrm{E}[\alpha_j'\tilde{\varepsilon}] = 0$. If an investor i is informed, then by Assumptions 3 and 4, $\mathrm{E}[\alpha_i(\tilde{s}_i)'\tilde{\varepsilon}] = \mathrm{E}[\alpha_i(\tilde{s}_i)' z^i \tilde{s}_i]$. By Theorem 2, $\alpha_i(s_i)' z^i s_i > 0$ whenever $s_i \neq 0$; therefore $\mathrm{E}[\tilde{s}_i^2] \neq 0$ implies $\mathrm{E}[(\alpha_i(\tilde{s}_i)' z^i)\tilde{s}_i] > 0$.

Proof of Theorem 4. First we show that $t_\alpha = \sigma_{si}/\sigma_{\eta i}$. Consider the portfolio choice problem of the investor given an observed signal s_i,

$$\max_{\alpha' e = 1} \mathrm{E}_i\left[-\exp(-A_i w_i(\alpha' B^* \tilde{f}^* + \alpha'\tilde{\varepsilon}))\right], \tag{13}$$

where $A_i > 0$ is the coefficient of absolute risk aversion of investor i and $\tilde{f}^{*\prime} = (1, \tilde{f}')$. Using the mean independence of $\tilde{\varepsilon}$ given $\tilde{f}$, (13) is equivalent to

$$\max_{\alpha' e = 1} \mathrm{E}\left[-\exp(-A_i w_i \alpha' B^* \tilde{f}^*)\right]\mathrm{E}_i\left[-\exp(-A_i w_i \alpha'\tilde{\varepsilon})\right]. \tag{14}$$

For any portfolio α there exists a portfolio δ and a zero-cost portfolio β such that $\alpha'\tilde{r} = \delta' B^* \tilde{f}^* + \beta'\tilde{\varepsilon}$. Hence the second multiplicative term in (14) is not subject to the cost constraint and an unconstrained subproblem of (14) is

$$\max_{\beta} \mathrm{E}_i[-\exp(-A_i w_i \beta'\tilde{\varepsilon})]. \tag{15}$$

Note that (14)–(15) and the independence of s_i and f imply that $\alpha(s)'B = b_\alpha$ for all s. That is, the factor risk of the portfolio is constant. By Assumptions 4 and 8, $\tilde{\varepsilon}_h$, $h = 1, 2, \ldots$, are independent variates and all except $\tilde{\varepsilon}_i$ have zero mean. Applying Jensen's inequality, the optimal β for (15) will have $\beta'\varepsilon \equiv (\beta' z^i)\varepsilon_i$. Define $\theta = \beta' z^i$ and rewrite (15) as

$$\max_{\theta} \mathrm{E}_i[-\exp(-A_i w_i \theta(s_i + \tilde{\eta}_i))]. \tag{16}$$

Applying the formula for the mean of a lognormal random variate to (16) and setting the derivative equal to zero, we get a necessary and sufficient first-order condition:

$$\theta(s_i) = s_i/\sigma_{\eta_i}^2 w_i A_i.$$

We have shown that the return on the optimal portfolio $\alpha(s_i)'r$ can be written in the form $\alpha(s_i)'r = \delta' B^* f^* + (s_i/\sigma_{\eta_i}^2 A_i w_i)(s_i + \eta_i)$. Therefore, $\alpha(s_i)'\varepsilon = (s_i/\sigma_{\eta_i}^2 w_i A_i)(s_i + \eta_i)$ and computing t_α (using the mean-zero normal property $\mathrm{E}[\tilde{s}_i^4] = 3\sigma_{si}^4$) gives $t_\alpha = \sigma_{s_i}/\sigma_{\eta_i}$.

Next we show that $t_{\alpha i} > t_{\alpha j}$ if and only if investor j would ex ante prefer to observe s_i instead of s_j. Since t_α does not depend upon A_i or w_i, this holds if unconditional expected utility is increasing in t_α. Since, as shown above, the optimal portfolio $\alpha(s_i)$ is separable into δ and $\theta(s_i)$ where δ does not depend on s_i, expected utility is proportional to

$$\mathrm{E}\left[\max_{\theta(s_i)} \mathrm{E}_i[-\exp(-A_i w_i \theta(\tilde{s}_i)(\tilde{s}_i + \tilde{\eta}_i))]\right].$$

Using the optimal rule for $\theta(s_i)$,

$$= \mathrm{E}\left(\mathrm{E}_i\left[-\exp\left(-\left(\tilde{s}_i^2 + \tilde{s}_i\tilde{\eta}_i\right)/\sigma_\eta^2\right)\right]\right).$$

Bringing a term which does not depend upon η_i outside of $\mathrm{E}_i[\ ]$,

$$= \mathrm{E}\left[-\exp\left(-\tilde{s}_i^2/\sigma_{\eta_i}^2 \mathrm{E}_i\left[\exp\left(-\tilde{s}_i\tilde{\eta}_i/\sigma_{\eta_i}^2\right)\right]\right)\right].$$

Calculating the expectation over η_i,

$$= \mathrm{E}\left[-\exp\left(-\tilde{s}_i^2/\sigma_{\eta_i}^2 - \tfrac{1}{2}\tilde{s}_i^2/\sigma_{\eta_i}^2\right)\right].$$

Combining terms and writing each side in terms of a standard normal random variate $\tilde{z}$,

$$= \mathrm{E}\left[-\exp\left(-\tfrac{3}{2}\tilde{z}^2\sigma_{si}^2/\sigma_{\eta i}^2\right)\right] = -\left(3t_\alpha^2 + 1\right)^{-1/2},$$

which is an increasing function of t_α.

Proof of Lemma 1. By the definition of A^n, we have $(1/n)F'B^{n\prime}B^nF = A^n$. Together with the definition of the principal components matrix this gives

$$(1/n)F'B^{n\prime}B^nFH^{n\prime} = H^{n\prime}\Lambda_H. \tag{17}$$

Postmultiplying both sides of (17) by $n(FH^{n\prime})^{-1}(B^{n\prime}B^n)^{-1}$ gives

$$F' = H^{n\prime}n\Lambda_H(FH^{n\prime})^{-1}(B^{n\prime}B^n)^{-1}.$$

This completes the proof, except that we must verify that $(FH^{n\prime})$ is non-singular, so that its inverse exists. This follows from (17). Both F and $H^{n\prime}\Lambda_H$ have rank k. The $k \times k$ matrix $(FH^{n\prime})$ must have rank k; otherwise the left-hand side of (17) will have rank strictly less than k.

Proof of Theorem 5. By Lemma 1 we have that

$$R_\alpha = a_\alpha + b_\alpha' H^n + \varepsilon_\alpha,$$

where H^n is a rotation of the original factors ($H^n = \bar{L}^n F$) and b_α is the inverse rotation of the original factor loadings of the portfolio. By assumption, $\tilde{\varepsilon}_{\alpha t}$ is independently distributed through time with finite second moment. Let Q_H denote the limit (as $T \to \infty$) of the cross-product matrix of the independent variables (including the unit vector) divided by T,

$$Q_H = \lim_{T\to\infty} (1/T)\begin{bmatrix} e_T'e_T & e_T'H' \\ He_T & HH' \end{bmatrix}.$$

Under these conditions we have [see Schmidt (1976, theorem 5, p. 60)]

$$\operatorname*{dlim}_{T\to\infty} T^{1/2}\left(\hat{a}_\alpha^T - a_\alpha\right) = \mathrm{N}\left(0, \sigma_{\varepsilon\alpha}^2 Q_H^{11}\right),$$

where $Q_H^{11} \equiv$ the (1,1) element of Q_H^{-1}. A consistent estimate of Q_H^{11} can be obtained from the (1,1) element of $\hat{Q}_H^{-1}$, where $\hat{Q}_H$ is the sample cross-product matrix. Applying the formula for the inverse of a partitioned matrix [Theil (1971, p. 18)] we can show that $Q_H^{11} = 1 + \gamma'\gamma$. The details are not presented here.

Thus,

$$\underset{T\to\infty}{\text{dlim}}\, T^{1/2}\left(\hat{a}_\alpha^T - a_\alpha\right) = \mathrm{N}\left(0, \sigma_{\varepsilon\alpha}^2(1+\gamma'\gamma)\right).$$

To show the limiting distribution of the t-ratio note that $\hat{\sigma}_{\varepsilon\alpha}^T$ converges in probability to $\sigma_{\varepsilon\alpha}$ [by Proposition 1 of Schmidt (1976, p. 55)]. Since $T^{1/2}(\hat{a}_\alpha^T - a_\alpha)$ converges in distribution to $\mathrm{N}(0, \sigma_{\varepsilon\alpha}^2(1+\gamma'\gamma))$ and $\hat{\sigma}_{\varepsilon\alpha}^T$ in probability to $\sigma_{\varepsilon\alpha}$, then $T^{1/2}(\hat{t}_\alpha^T - t_\alpha)$ converges in distribution to $\mathrm{N}(0, 1+\gamma'\gamma)$ [see Rao (1973, p. 122)].

Before proving Theorem 6 we need two lemmas. Since the results are technical, we relegate the proofs to an unpublished appendix which is available from the authors.

Lemma 2. Let A^n and Ω^n be (non-random) sequences of positive semi-definite $T\times T$ matrices with rank equal to k and greater than or equal to k, respectively. Define G^n, Λ_G^n, H^n, Λ_H^n to be the principal components matrices and associated eigenvalue matrices of Ω^n, A^n, respectively. Define $W^n = \Omega^n - A^n$ and assume that $\lim_{n\to\infty}\|W^n - \sigma^2 I_T\| = 0$ and there exists $c_3 < \infty$ such that $\|(\Lambda_G^n - \sigma^2 I_k)^{-1}\| < c_3$ for all n. Then $G^n = L^n H^n + \Phi^n$, where $L^n = (\Lambda_G^n - \sigma^2 I_k)^{-1} \times (1/T) G^n H^{n\prime} \Lambda_H^n H^n$ and $\lim_{n\to\infty}\Phi^n = 0$.

Lemma 3. Define $W^n = \Omega^n - A^n$. Given Assumptions 1–7 and 10,

$$\underset{n\to\infty}{\text{plim}}\, \|W^n - \sigma^2 I_T\| = 0,$$

and

$$\lim_{n\to\infty} \text{prob}\left\{\left\|\left(\Lambda_G^n - \sigma^2 I_k\right)^{-1}\right\| > c_3\right\} = 0 \quad \textit{for some} \quad c_3 < \infty.$$

Proof of Theorem 6. By Lemma 2 we know that for any two sequences of non-random matrices Ω^n, A^n, with $\lim_{n\to\infty}\|\Omega^n - A^n - \sigma^2 I_T\| = 0$ and $\|(\Lambda_G^n - \sigma^2 I_k)^{-1}\| < c_3$, for all n, we have $G^n = L^n H^n + \Phi^n$, where $\lim_{n\to\infty}\Phi^n = 0$. We must convert this limit on fixed matrices into a probability limit on the random realizations of G^n and H^n.

It follows from Lemma 3 that there exists a sequence δ_n such that $\lim_{n\to\infty}\delta_n = 0$ and, for every n, $\text{prob}\{\|\Omega^n - A^n - \sigma^2 I_T\| < \delta_n\} > 1 - \delta_n$ and $\text{prob}\{\|(\Lambda^n_G - \sigma^2 I_k)^{-1}\| \leq c_3\} > 1 - \delta_n$. For each n, consider the set of possible realizations of A^n and Ω^n for which $\|\Omega^n - A^n - \sigma^2 I_T\| < \delta_n$ and $\|(\Lambda^n_G - \sigma^2 I_k)^{-1}\| < c_3$. The chosen set of events at point n has probability greater than or equal to $1 - 2\delta_n$. For any sequence of realizations H^n, G^n from the chosen set, we have $\lim_{n\to\infty}\Phi^n = 0$. Since the probability of this set of events approaches 1 as n approaches infinity, this means that $\text{plim}_{n\to\infty}\Phi^n = 0$.

Proof of Corollary 1. We show that

$$\operatorname*{plim}_{n\to\infty} \hat{a}^{nT}_\alpha = \hat{a}^T_\alpha \quad \text{and} \quad \operatorname*{plim}_{n\to\infty} \hat{\sigma}^{nT}_{\varepsilon\alpha} = \hat{\sigma}^T_{\varepsilon\alpha}.$$

This result combined with Theorem 6 proves the corollary.

Theorem 5 shows that $G^n = L^n F + \Phi^n$, where $\text{plim}_{n\to\infty}\Phi^n = 0$. By the definition of OLS intercept estimates [see Theil (1971, p. 39)] the difference between $\hat{a}^{nT}_\alpha$ and $\hat{a}^T_\alpha$ is given by

$$\bar{e}'_T\Big[F'(FDF')^{-1}F - (L^nF+\Phi^n)'((L^nF+\Phi^n)D(L^nF+\Phi^n)')^{-1} \times (L^nF+\Phi^n)\Big]DR'_\alpha, \tag{18}$$

where

$$D = (I_T - (1/T)e_T e'_T) \quad \text{and} \quad \bar{e}_T = (1/T)e_T.$$

Now the term $(L^nF+\Phi^n)D(L^nF+\Phi^n)'$ is the covariance matrix of the sample principal components, which is non-singular with probability one when T is greater than k [see Arnold (1981, p. 438)]. Therefore (18) is a continuous function of Φ^n in a neighborhood of the sample principal components. This allows us to use the property that $\text{plim}_{n\to\infty} g(\Phi^n) = g(\text{plim}_{n\to\infty}\Phi^n)$ if $g(\cdot)$ is continuous. Therefore,

$$\begin{aligned}
&\operatorname*{plim}_{n\to\infty} \hat{a}^{nT}_\alpha - \hat{a}^T_\alpha \\
&= \bar{e}'_T\Big[F'(FDF')^{-1}F - F'L^{n\prime}(L^nFDF'L^{n\prime})^{-1}L^nF\Big]DR'_\alpha \\
&= \bar{e}_T\Big[F'(FDF')^{-1}F - F'L^{n\prime}L^{n\prime -1}(FDF')^{-1}L^{n-1}L^nF\Big]DR'_\alpha = 0.
\end{aligned} \tag{19}$$

Since

$$\hat{a}_{\alpha}^{nT} - a_{\alpha} = \left(\hat{a}_{\alpha}^{nT} - \hat{a}_{\alpha}^{T}\right) + \left(\hat{a}_{\alpha}^{T} - a_{\alpha}\right),$$

we have that

$$\operatorname*{dlim}_{T\to\infty} \operatorname*{plim}_{n\to\infty} T^{1/2}\left(\hat{a}_{\alpha}^{nT} - a_{\alpha}\right) = \operatorname*{dlim}_{T\to\infty} \left[0 + T^{1/2}\left(\hat{a}_{\alpha}^{T} - a_{\alpha}\right)\right]$$

$$= \mathrm{N}\left(0, \sigma_{\varepsilon\alpha}^{2}(1 + \gamma'\gamma)\right),$$

where the last equality is by Theorem 6 and the penultimate equality is by (19). The proof that $\operatorname{plim}_{n\to\infty}\hat{\sigma}_{\varepsilon\alpha}^{nT} = \hat{\sigma}_{\varepsilon\alpha}^{T}$ and hence that $\operatorname{dlim}_{T\to\infty}\operatorname{plim}_{n\to\infty}(\hat{t}_{\alpha}^{nT} - t_{\alpha}) = \mathrm{N}(0, 1 + \gamma'\gamma)$ is essentially identical to the proof for $\hat{a}_{\alpha}^{nT}$. We omit it here.

References

Admati, Anat R. and Stephen A. Ross, 1985, Measuring investment performance in a rational expectations equilibrium model, Journal of Business 58, 1–26.

Arnold, Steven F., 1981, The theory of linear models and multivariate analysis (Wiley, New York).

Chamberlain, Gary, 1983, Funds, factors, and diversification in arbitrage pricing models, Econometrica 51, 1303–1324.

Chamberlain, Gary and Michael C. Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, Econometrica 51, 1281–1304.

Chang, Eric C. and Wilbur G. Lewellen, 1984, Market timing and mutual fund investment performance, Journal of Business 56, 57–72.

Chen, Nai-fu, 1983, Some empirical tests of the theory of arbitrage pricing, Journal of Finance 38, 1393–1414.

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

Connor, Gregory and Robert A. Korajczyk, 1985, A test for the number of factors in an approximate factor model, Unpublished working paper (Kellogg Graduate School of Management, Northwestern University, Evanston, IL).

Dhrymes, Phoebus J., Irwin Friend and N. Bulent Gultekin, 1984, A critical reexamination of the empirical evidence on the arbitrage pricing theory, Journal of Finance 39, 323–346.

Dybvig, Philip H. and Stephen A. Ross, 1984, Performance measurement using excess returns relative to a security market line: Differential information, Unpublished working paper (School of Organization and Management, Yale University, New Haven, CT).

Henriksson, Roy D., 1984, Market timing and mutual fund performance: An empirical investigation, Journal of Business 57, 73–96.

Hsieh, David A., 1983, A heteroskedasticity-consistent covariance matrix estimator for time series regressions, Journal of Econometrics 22, 281–290.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, Journal of Finance 23, 389–416.

Kon, Stanley J., 1983, The market-timing performance of mutual fund managers, Journal of Business 56, 323–347.

Luedecke, Bernd P., 1984, An empirical investigation into arbitrage and approximate k-factor structure on large asset markets, Unpublished doctoral dissertation (Department of Economics, University of Wisconsin, Madison, WI).

Rao, C.R., 1973, Linear statistical inference and its applications, 2nd ed. (Wiley, New York).

Roll, Richard, 1978, Ambiguity when performance is measured by the securities market line, Journal of Finance 33, 1031–1069.
Roll, Richard and Stephen A. Ross, 1984, A critical reexamination of the empirical evidence on the arbitrage pricing theory: A reply, Journal of Finance 39, 347–350.
Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, Journal of Economic Theory 13, 341–360.
Schmidt, Peter, 1976, Econometrics (Marcel Decker, New York).
Searle, Shayle R., 1982, Matrix algebra useful for statistics (Wiley, New York).
Shanken, Jay, 1982, The arbitrage pricing theory: Is it testable?, Journal of Finance 37, 1129–1140.
Shanken, Jay, 1985, Multi-beta CAPM or equilibrium-APT? A reply to Dybvig and Ross, Unpublished working paper (Graduate School of Business, University of California, Berkeley, CA).
Stewart, G.W., 1973, Introduction to matrix computations (Academic Press, New York).
Theil, Henri, 1971, Principles of econometrics (Wiley, New York).
Treynor, Jack L. and Fischer Black, 1973, How to use security analysis to improve portfolio selection, Journal of Business 46, 66–86.