Sequential Tests of the Arbitrage Pricing Theory: A Comparison of Principal Components and Maximum Likelihood Factors
Author(s): Ravi Shukla and Charles Trzcinka
Source: *The Journal of Finance*, Vol. 45, No. 5 (Dec., 1990), pp. 1541-1564
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2328749
Accessed: 26/11/2009 07:22

THE JOURNAL OF FINANCE • VOL. XLV, NO. 5 • DECEMBER 1990

Sequential Tests of the Arbitrage Pricing Theory: A Comparison of Principal Components and Maximum Likelihood Factors

RAVI SHUKLA and CHARLES TRZCINKA*

ABSTRACT

We examine the cross-sectional pricing equation of the APT using the elements of eigenvectors and the maximum likelihood factor loadings of the covariance matrix of returns as measures of risk. The results indicate that, for data assumed stationary over twenty years, the first vector is a surprisingly good measure of risk when compared with either a one- or a five-factor model or a five-vector model. We conclude that in some circumstances principal components analysis may be preferred to factor analysis.

THE ARBITRAGE PRICING THEORY (APT) of Ross (1976) hypothesizes that the cross-sectional distribution of expected returns of financial assets can be approximately measured by their sensitivities, usually factor loadings, to k unknown economic factors. In an important development, Chamberlain and Rothschild (1983) extend Ross's APT by proving that, if k eigenvalues of the population covariance matrix increase without bound as the number of securities in the population increases, then the elements of the corresponding k eigenvectors (the weights in principal components) of the covariance matrix can be used as the factor sensitivities. Connor and Korajczyk (1988) show that this conclusion holds for the sample covariance matrix as well.

However, the proof that the eigenvectors can be used instead of statistical factor loadings in the returns-generating model for a large (with infinitely many assets) economy does not necessarily imply that the eigenvectors can be used in a cross-sectional model of security pricing in finite economies. Principal components analysis is equivalent to factor analysis only when the idiosyncratic risks of all firms are equal. If idiosyncratic risks vary, factor analysis is less constrained than the components analysis because factor analysis estimates idiosyncratic risks simultaneously with factor loadings while idiosyncratic risks are ignored when components are estimated. In practice this means that any nonstationarity and measurement errors in the data will affect the estimation of components more than the factors because k factor loadings are always estimated with more degrees of freedom than k components. Thus, there appears to be a reasonable a priori belief by some researchers that k eigenvectors simply will not perform as well as k factor loadings in measuring systematic risk.

* Department of Business Administration and Economics, College at Brockport, SUNY and School of Management, University at Buffalo, SUNY, respectively. We would like to thank the participants in the Finance Workshops of Indiana University and the University at Buffalo for their comments. This paper was substantially improved by the thoughtful attention of the anonymous referee. Nevertheless, we do not hold anyone but ourselves responsible for the errors that undoubtedly remain.

In spite of this belief, the question of whether eigenvectors are adequate is still an empirical issue. It is of interest to address this issue because principal components analysis has several distinct advantages over factor analysis. First, if the eigenvectors are an acceptable measure of systematic risk, then principal components have the advantage that they are unique up to scalar transformation and are not affected by rotation while factors are only unique up to orthogonal rotation. This implies, for example, that t-statistics are meaningful for components but are useless for factors. Second, for a given significance level, a test of the APT which uses eigenvectors is more likely to reject the null hypothesis because of measurement errors than a test which uses factor analysis. This is because factor analysis imposes fewer restrictions on the data. Thus, a finding of no evidence against an eigenvectors-based model is stronger evidence in favor of the APT than the same finding with a factor-loadings-based model. Third, principal components are much easier to compute than factors. Thus, if eigenvectors are satisfactory measures of risk, they may be preferred in many applications.

To compare the principal components and factor analyses, this paper estimates the eigenvectors and factor loadings of a sequence of covariance matrices and uses them as measures of systematic risk in explaining the cross-sectional variation in mean returns. The sequence approach has been used empirically in Linn and Chang (1985) and Trzcinka (1986). Trzcinka (1986) provides strong evidence that the first eigenvalue of the covariance matrix grows without bound with the number of securities in the sample and somewhat weaker evidence that the second through fifth eigenvalues are unbounded. This suggests that one or at most five eigenvectors should be sufficient to explain the cross-sectional variation in the mean returns of assets. Therefore, using one and five vectors, we examine a principal components model and compare it to similar factor models under different assumptions about stationarity. We also compare the performance of the eigenvector and factor based pricing equations against the benchmark model that uses the market model betas as measures of risk.

The paper is organized as follows. Section I gives a brief review of the APT, focusing on the importance of comparing eigenvectors with factors. Section II discusses the data and the specific methodology for the cross-sectional tests. It also examines the issue of measurement error in some detail. Section III gives the results, and Section IV concludes the paper with a summary.

I. The Arbitrage Pricing Theory

Theoretical derivations of the APT begin with the assumption that the $N \times 1$ vector of asset returns $\tilde{\mathbf{r}}$ is linearly related to the $k \times 1$ vector of realizations of common factors $\tilde{\delta}$:

$$\tilde{\mathbf{r}} = \mathbf{e} + \mathbf{b}\tilde{\delta} + \tilde{\epsilon}, \tag{1}$$

where $\mathbf{e}$ is an $N \times 1$ vector of means, $\mathbf{b}$ is an $N \times k$ matrix with $b_{i\kappa}$ the sensitivity of the ith asset to the κth factor, and $\tilde{\epsilon}$ is the $N \times 1$ vector of idiosyncratic returns. If $\operatorname{cov}(\tilde{\epsilon}, \tilde{\delta}) = 0$, then the covariance matrix of returns, $\mathbf{\Sigma}$, can be written as

$$\mathbf{\Sigma} = \mathbf{bb}' + \mathbf{\Phi}, \tag{2}$$

where $E(\tilde{\epsilon}\tilde{\epsilon}') = \mathbf{\Phi}$. The original work of Ross assumes that equation (1) follows a *strict factor structure*, i.e., $\mathbf{\Phi}$ is diagonal. Dybvig and Ross (1985) discuss the theoretical development of the APT prediction that $\mathbf{e}$ is spanned (approximately) by the columns of $\mathbf{b}$ and $\mathbf{l}$, an $N \times 1$ vector of ones, or

$$\mathbf{e} \simeq \mathbf{l}\gamma_0 + \mathbf{b}\boldsymbol{\gamma}, \tag{3}$$

where γ_0 is the return on the zero-beta (or risk-free) portfolio and $\boldsymbol{\gamma}$ is a $k \times 1$ vector of factor risk premia. All of the ways to derive this prediction from assumption (1) are based on the intuition that the firm specific return ($\tilde{\epsilon}$) represents diversifiable risk which should have a zero price in an economy with no arbitrage opportunities. Ross shows that the sum of squared approximation errors (the *pricing errors*) is finite as the number of securities in the economy approaches infinity.

In one of the most general approaches, Chamberlain and Rothschild (CR) weaken the assumption on (1) to an *approximate factor structure* where $\mathbf{\Phi}$ need not be diagonal. They show that the pricing equation (3) still holds in the sense that a *weighted* sum of the squared pricing errors is bounded as the number of securities in the economy grows large and the columns of $\mathbf{b}$ are the k eigenvectors corresponding to the k largest eigenvalues of $\mathbf{\Sigma}$. This is the definition of principal components analysis which assumes that the smallest $N - k$ eigenvalues are approximately equal and uses the eigenvectors corresponding to the k largest eigenvalues to represent the covariance matrix.[1] As CR observe,

> A common objection to principal components analysis is that it is arbitrary to examine the eigenvectors of $\mathbf{\Sigma}$ relative to an identity matrix rather than some other positive definite matrix—one which is in some sense natural for the problem at hand. We show that for the problem of investigating the approximate factor structure of an asset market this objection is groundless.

If this theoretical approach accurately describes security pricing, the computational complexity of obtaining the $N \times k$ matrix $\mathbf{b}$ is drastically reduced. The eigenvectors corresponding to the k largest eigenvalues constitute the first k principal components and are computed by simply finding the eigenvectors of the sample estimate of $\mathbf{\Sigma}$. If components are used to represent security returns in equations (1) to (3), then the residual covariance matrix $\mathbf{\Phi}$ is not estimated and the investigator acts as if it could be ignored. This contrasts sharply with

[1] Principal components models decompose the covariance matrix as $\mathbf{\Sigma} = \mathbf{GD}_\lambda\mathbf{G}'$, where $\mathbf{G}$ is the $N \times N$ matrix of eigenvectors and $\mathbf{D}_\lambda$ is a diagonal matrix of eigenvalues. To fully describe the $N \times N$ covariance matrix, one needs N eigenvectors. Partitioning the matrices, we can write this equation as $\mathbf{\Sigma} = \mathbf{G}_k\mathbf{D}_{\lambda_k}\mathbf{G}'_k + \mathbf{G}_{N-k}\mathbf{D}_{\lambda_{N-k}}\mathbf{G}'_{N-k}$. Now, if the k principal components are sufficient, then the differences between the remaining $N - k$ eigenvalues are insignificant, so that they can be treated as approximately equal to each other and we get $\mathbf{D}_{\lambda_{N-k}} = \lambda^*\mathbf{I}_{N-k}$, and, since the eigenvectors are orthonormal, we can write $\mathbf{\Sigma} = \mathbf{G}_k\mathbf{D}_{\lambda_k}\mathbf{G}'_k + \lambda^*\mathbf{I}_N$.

the factor analysis where the joint estimation of $\mathbf{\Phi}$ and $\mathbf{b}$ causes the technique to be far more computationally difficult.

Using components has the additional advantage that they are uniquely defined with respect to the criteria that the size of the eigenvalue indicates the relative importance of the component.[2] It is well known that factor analysis produces a $\mathbf{b}$ matrix which is unique only up to *orthogonal rotation.* This means that it is impossible to interpret any factor since practically any set of numbers will work as long as the dimension k is preserved. Principal components, on the other hand, are uniquely defined up to scalar transformation. For a stationary population distribution, one can compare, for example, the second component of one sample of return data to the second component of another. This comparison cannot be made with factors since even the order of the factors is not unique. However, the sampling error in the covariance matrix may be large enough so that the confidence interval for component i of one sample does not contain component i from another sample.

Principal components have one serious disadvantage which has discouraged their use. The result of not estimating $\mathbf{\Phi}$ is that the principal components estimates of $\mathbf{b}$ must explain much more of firm-specific variation in returns than factor analytic estimates of $\mathbf{b}$ for the same k. This will be a problem when firm-specific variation either is a large percentage of total variation or is very different between firms. Trzcinka (1986) provides evidence that both are the case for CRSP data. Moreover, firm-specific variation can be large not only because idiosyncratic risks are high but because measurement error and nonstationarity differ between firms. Thus, previous researchers, quite reasonably, have avoided principal component estimation of $\mathbf{b}$ in spite of its advantages.

However, the question of whether the principal components perform as well as the maximum likelihood factors has not been addressed in much detail. This paper will provide a comparison of the two techniques by regressing average returns on the factor loadings and the eigenvectors of sequentially larger covariance matrices. The average returns are computed using weekly returns over a twenty-year period and five-year subperiods. The stationarity assumption of these average returns should give factors an advantage over components since there is no question that the mean returns are nonstationary. We will examine both the pricing error and the measurement error of the two methods. The analysis presented below shows that the impact of measurement error will converge as the number of securities in the sample increases, and thus we can observe the approximate empirical pricing error of each method. A superior model, among other things, will have lower *empirical pricing errors*. If principal components analysis is a valid technique in the context of APT, then the eigenvectors and the factor loadings should have equivalent pricing errors when the measurement errors converge. Even if factor analysis is superior, our study

[2] CR observe that both factor analysis (FA) and principal components analysis (PCA) compute the eigenvectors of $\mathbf{\Sigma} - \mathbf{A}$, where $\mathbf{A}$ is defined as an identity matrix for PCA and is estimated for FA. CR prove that in the limit all choices of $\mathbf{A}$ are equivalent to the identity matrix. Since $\mathbf{b}$ is unique when $\mathbf{A}$ is the identity matrix, PCA produces unique estimates. Thus, the criterion of equal residual risks is justified for large economies, and the empirical question is whether our economy is large enough.

will provide some guidance as to whether this superiority is worth the computational cost and the indefiniteness of factors.

Finally, our tests *do not* assume that a small size economy is represented by a small sample of securities. Taking samples from a population is not what theorists have in mind when they let the size of an economy grow without bound. For an economy with N assets, the theoretical pricing error of asset j is the same regardless of the sample in which the security is included. Empirical researchers must analyze subsets, and the patterns that emerge will clearly be of interest for any practical applications of the APT. Furthermore, as we discuss in detail below, an increasing sequence of matrices allows us to isolate the pricing error of a group of securities.

II. Methodology

A. Data

Weekly data are used from the 1983 Daily CRSP tapes to estimate the covariance matrices. There are 1,069 weeks of data on the tape and 865 firms listed for all twenty years having greater than 865 observations. The minimum number of observations is 906, and only nine firms have between 906 and 1,000 observations. We could have expanded the number of observations and firms by using daily data, but recent evidence suggests that there are a variety of problems with using daily data for our purpose. First, Roll and Ross (1980) found that the results of their cross-sectional pricing tests were improved when every other observation was deleted, and they argued that it was due to skewness in daily data. Second, Perry (1982) suggests that daily data are best modeled by an unstable variance process. Also, it is well known that daily returns are not well described by the normal distribution (Fama (1976)). Furthermore, covariance matrices estimated using daily data produce poor estimates of the factor structure (Shanken (1987)). Thus, we felt that the gain in observations and firms was not worth the cost of nonstationarity of the data. Nevertheless, we did not feel that monthly data would provide us with enough degrees of freedom as there are only 696 months on the 1983 Monthly CRSP tape, and consequently we settled on weekly data.[3]

We employ a *random* sequence of covariance matrices where a smaller sized sample is not necessarily a subset of a larger sized sample. This sequence begins with a sample covariance matrix of size 30, and we increase each sample by 10 percent so that $n = 30, 33, 36, \cdots, 596$. There are 32 portfolios in the sequence.[4] This sequence will determine the behavior of eigenvectors in a situation similar to previous empirical studies which use factor analysis on randomly selected covariance matrices and will shed light on the result of factor analyzing larger and larger matrices. Also, it will allow us to determine whether the impact

[3] Performing our tests for monthly, weekly, and daily data is prohibitively expensive.

[4] We also analyzed a *nested* sequence where a smaller sized sample is a subset of the larger sized sample. The nested sequence grows by simply adding rows and columns to the smallest size matrix. Our results were very similar for both the sequences, and we will only report the results for the random sequence.

of measurement error is growing, decreasing, or constant as the sample size increases.

B. Pricing Tests

The general approach of our tests is conceptually straightforward. The eigenvectors, factors, and market model betas are used as independent variables to explain the cross-sectional variation in the mean returns of all n securities which comprise the sample (portfolio). The first five eigenvectors, a one-factor model, and a five-factor model will be computed for each covariance matrix in the sequence. The mean returns of all securities in all matrices will be estimated for the entire twenty years and four five-year subperiods. These mean returns will be used as a proxy for the expected returns. The mean returns for securities composing a sample will be regressed against the eigenvectors or the factor loadings. The resulting equations will be compared to determine which measure of systematic risk results in a lower pricing error and a greater explanatory power. We will also compare the APT models with the cross-sectional pricing equations using the market model betas to determine whether the APT measures of risk perform at the same level as the market model betas.[5]

To test the APT we will continue to make the assumption of most factor analytic tests (see, for example, Brown and Weinstein (1983) and Roll and Ross (1980)) that the return distribution is stationary over time. Estimating measures of systematic risk from a covariance matrix calculated using twenty years of data obviously makes the heroic assumption that systematic risk is constant for twenty years. There is no shortage of studies which amply demonstrate that the distribution characteristics of security returns are not stationary, and most asset pricing tests take great care to control for the problems of nonstationarity.[6] While it is easy for us to alter the number of weeks used to compute the mean return, we cannot alter the estimates of systematic risk because we need all twenty years of data to ensure that the covariance matrix has positive degrees of freedom. Therefore, our regression coefficients will likely reflect measurement error. Moreover, as discussed above, since factor analysis imposes fewer constraints on the measure of systematic risk than principal components analysis, it is likely that measurement error will affect the eigenvector regression more than the factor regressions. Thus, our procedures are biased against finding evidence in favor of principal components analysis.

The stationarity assumption on the returns generating process implies that $\mathbf{e}$ is an $n \times 1$ vector of returns averaged over time. If we take $\mathbf{e}^{(n)}$ as the vector of

[5] Strictly speaking, the usual derivation of the CAPM pricing equation does not allow for theoretical pricing errors as the APT does. However, pricing errors may be observed in CAPM because of the use of an improper proxy for the market portfolio. Alternatively, one may view CAPM as a one-factor APT model, with the market model being the return generation process. In this view the pricing equation based on the market model will be an approximate one and will contain theoretical pricing errors.

[6] See Ferson, Kandel, and Stambaugh (1987) for recent evidence that expected risk premia and asset betas vary over time. Dhrymes, Friend, and Gultekin (1984) argue that the typical methodology to control for nonstationarity is inadequate.

n average returns, the hypothesis of the APT can be tested by the cross-sectional regression equation:

$$\mathbf{e}^{(n)} = \mathbf{l}\gamma_0 + \mathbf{b}^{(n)}\boldsymbol{\gamma}^{(n)} + \mathbf{u}^{(\mathbf{n})} = \mathbf{l}\alpha^{(n)} + \mathbf{b}^{(n)}\boldsymbol{\gamma}^{(n)} + \mathbf{v}^{(n)}, \tag{4}$$

where $\mathbf{b}^{(n)}$ is the matrix of $n \times k$ measures of systematic risk, $\boldsymbol{\gamma}^{(n)}$ is the $k \times 1$ vector of factor risk premiums, $\mathbf{u}^{(n)}$ is an $n \times 1$ vector of pricing error terms, and γ_0 is the zero-beta return. The intercept from the regression of $\mathbf{e}^{(n)}$ on $\mathbf{b}^{(n)}$ is $\alpha^{(n)} = \gamma_0 + \bar{u}^{(n)}$, where $\bar{u}^{(n)} = (1/n) \sum_{i=1}^{n} u_i^{(n)}$ is the *average* pricing error for the portfolio securities.

A better measure of systematic risk should result in a better fit of the pricing equation (4). The better measure should also have lower pricing errors. (APT assumes that the pricing errors are negligible.) This can be tested by examining the magnitudes of intercepts across the models. Since the average pricing errors may not necessarily be positive, we cannot compare the intercepts directly. The magnitude of deviation of the intercept from the zero-beta rate will be higher for a poorer measure of systematic risk. For such a comparison, we need to have an exogenous estimate of γ_0, the zero-beta rate.

If $\mathbf{b}$ is measured correctly, we do not expect $\alpha^{(n)}$ to have any systematic relation with n because $\alpha^{(n)}$ is made up of the zero-beta rate, γ_0, which is a constant across various samples, and the average pricing error $\bar{u}$, and there is no reason to believe that the average pricing error depends in a systematic way on the number of securities in random samples. To test whether the intercepts are different in different size portfolios, we assume that the adjacent portfolios of size n and m are independent random samples. Let

$$S = \frac{1}{n + m - 2} (a^{(n)} + a^{(m)}), \tag{5}$$

where $a^{(j)} = \sum_{t=1}^{j} (v_t^{(j)} - \bar{v}^{(j)})^2$ (note that $\bar{v}^{(j)} = 0$) for $j = m, n$. Then, if the samples are drawn from a normal population, an F-statistic which tests the hypothesis that the intercepts are different is

$$F = \frac{nm}{n + m - 2} \frac{(\alpha^{(n)} - \alpha^{(m)})^2}{S} = nm \frac{(\alpha^{(n)} - \alpha^{(m)})^2}{(a^{(n)} + a^{(m)})}, \tag{6}$$

which has the F-distribution with 1 and $n + m - 2$ degrees of freedom.[7] We will refer to this statistic as *pricing error F*.

The assumption of this test is that $\alpha^{(n)}$ is independent of $\alpha^{(m)}$. For large n this will not be the case if the CRSP tapes are regarded as the population. Drawing 596 securities randomly from 865 is not constructing a simple random sample if the 865 securities are considered the population. However, if the time series of returns on 865 securities is considered as one of many sample paths from an

[7] This F-ratio is simply a multivariate Hotelling T^2 test for $p = 1$. See Morrison (1976), pp. 128–138, for development of the statistic. The denominator of our statistic would be larger if we accounted for the sampling error of α. We assume that all variation in α is due to variance in pricing errors across the securities.

infinite population, then the sample is random. The independence assumption may be violated in another manner because the covariance between $\alpha^{(n)}$ and $\alpha^{(m)}$ is dependent on whether the off-diagonal elements of $\mathbf{\Phi}$ influence the pricing errors. If $\mathbf{\Phi}$ is diagonal, or if $\mathbf{\Phi}$ is not diagonal but the covariances between idiosyncratic components have no influence on pricing errors, then $\alpha^{(n)}$ will be independent of $\alpha^{(m)}$. Otherwise the pricing error F-statistic will be either biased upward or downward depending on the sign of the covariance between the intercepts. To see this bias, note that the statistic is the ratio of the variance of the intercepts *between* portfolios to the variance of the intercepts *within* the portfolio. The covariance between $\alpha^{(m)}$ and $\alpha^{(n)}$ will change the numerator but not the denominator since the denominator considers only the variation within the portfolio. Fortunately, the residual covariance matrix is closer to being diagonal for large k rather than small k, and for factor analysis rather than principal components analysis. Thus, the biases are observable, and we can determine the direction by comparing the F-statistic for estimates of $\mathbf{b}$ using the various models and specifications of k. Nevertheless, if the biases in the F-statistic are small, an F-value significantly different from zero implies that the average pricing error for a group of securities is changing with the number of securities in the group or that the zero-beta rate γ_0 is changing across groups. Either conclusion is evidence against the APT.

C. Isolating the Measurement Error

The discussion above assumed that the systematic risk is measured without any errors. This is certainly not true given our stationarity assumptions and our use of sample averages. While our estimates will be subject to measurement error problems, we should be able to observe the impact of errors in the variables by analyzing the sequence of covariance matrices. To see that the sequence can isolate measurement error, suppose there is only one factor. If the APT is correct, the true model to be tested is

$$e_i = \gamma_0 + b_i\gamma_1 + u_i, \tag{7}$$

where γ_1 is the factor risk premium, γ_0 is the risk-free rate, $u_i = C_i/N$ is the pricing error, and N is the number of securities in the economy. We assume that C_i is a bounded, positive constant. This specification is in the spirit of Dybvig (1983) and Grinblatt and Titman (1983), where C_i is a function of relative asset supplies and idiosyncratic variance. When we estimate $(\hat{\gamma}_0, \hat{\gamma}_1)$, we will not have b_i but $b_i^* = b_i + \zeta_i$, where ζ_i is the measurement error of our estimate b_i^*. The actual cross-sectional regression is

$$e_i = \gamma_0 + \bar{u} + b_i^*\gamma_1 + v_i = \alpha + b_i^*\gamma_1 + v_i, \tag{8}$$

where $u_i = \bar{u} + v_i$ and v_i is mean zero. The OLS estimator of α is

$$\hat{\alpha} = \bar{e} - \overline{b}^*\hat{\gamma}_1, \tag{9}$$

where

$$\bar{e} = \frac{1}{n} \sum_{i=1}^{n} e_i,$$

$$\bar{b}^* = \frac{1}{n} \sum_{i=1}^{n} b_i^* = \frac{1}{n} \sum_{i=1}^{n} b_i + \frac{1}{n} \sum_{i=1}^{n} \zeta_i = \bar{b} + \bar{\zeta},$$

$$\hat{\gamma}_1 = \sum_{i=1}^{n} w_i^* e_i = \gamma_1 \sum_{i=1}^{n} w_i^* b_i + \sum_{i=1}^{n} w_i^* v_i,$$

and

$$w_i^* = \frac{b_i^* - \bar{b}^*}{\sum_{i=1}^{n} (b_i^* - \bar{b}^*)^2}.$$

Since $e_i = \alpha + b_i^* \gamma_1 + v_i$, we can write $\hat{\alpha}$ as

$$\hat{\alpha} = \alpha + \bar{b}\gamma_1 + \frac{1}{n} \sum_{i=1}^{n} v_i - \left(\gamma_1 \sum_{i=1}^{n} w_i^* b_i + \sum_{i=1}^{n} w_i^* v_i\right)(\bar{b} + \bar{\zeta}). \tag{10}$$

Taking expectations,

$$\mathrm{E}(\hat{\alpha}) = \alpha + \bar{b}\gamma_1\left(1 - \sum_{i=1}^{n} w_i^* b_i\right) - \bar{\zeta}\gamma_1 \sum_{i=1}^{n} w_i^* b_i. \tag{11}$$

Obviously, the measurement error biases the OLS estimate of the intercept. Write $\sum_{i=1}^{n} w_i^* b_i$ as

$$\begin{aligned} \sum_{i=1}^{n} w_i^* b_i &= \sum_{i=1}^{n} w_i^* (b_i - \bar{b}) \\ &= \frac{\sum_{i=1}^{n} (b_i - \bar{b})^2 + \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})(b_i - \bar{b})}{\sum_{i=1}^{n} (b_i - \bar{b})^2 + 2 \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})(b_i - \bar{b}) + \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2}, \end{aligned} \tag{12}$$

and consider the limit of this expression as n grows large.

If the covariance between the measurement error, ζ, and the coefficient, b_i, is negligible, then

$$\sum_{i=1}^{n} w_i^* b_i \simeq \frac{\sum_{i=1}^{n} (b_i - \bar{b})^2}{\sum_{i=1}^{n} (b_i - \bar{b})^2 + \sum_{i=1}^{n} (\zeta_i - \bar{\zeta})^2} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_\zeta^2}, \tag{13}$$

where σ_b^2 and σ_ζ^2 are the variances of b and ζ, respectively. Now, if the variance of ζ is large relative to that of b, then $\lim_{n \to N} \sum_{i=1}^{n} w_i^* b_i = 0$ and we get

$$\lim_{n \to N} \mathrm{E}(\hat{\alpha}) = \alpha + \bar{b}\gamma_1 = \gamma_0 + \bar{u} + \bar{b}\gamma_1 = \gamma_0 + \frac{\bar{C}}{N} + \bar{b}\gamma_1. \tag{14}$$

On the other hand, if the variance of ζ is small relative to that of b, then $\lim_{n\to N}\sum_{i=1}^{n} w_i^* b_i = 1$ and we get

$$\lim_{n\to N} \mathrm{E}(\hat{\alpha}) = \alpha - \bar{\zeta}\gamma_1 = \gamma_0 + \bar{u} - \bar{\zeta}\gamma_1 = \gamma_0 + \frac{\bar{C}}{N} - \bar{\zeta}\gamma_1. \quad (15)$$

If $\bar{b}$ and $\bar{\zeta}$ are bounded, then the intercept will converge to one of the two values after n grows sufficiently large. The assumption of no correlation between the coefficient, b, and its measurement error, ζ, is reasonable and is typically made in the *error in the variables* econometric literature. (See Johnston (1984), p. 429.)

Thus, measurement error has two implications for our tests. First, our comparison of intercepts across the models will be affected by the measurement error, with a smaller deviation of the intercept from the zero-beta rate implying lower pricing error plus measurement error. In practice the pricing errors and the measurement errors will be inseparable, and a model with a lower total error (*empirical pricing error*) is the one that is more useful. Therefore, we can still arrive at meaningful conclusions by comparing the intercepts.

Second, the pricing error F-statistic discussed in Section II.*B* will be affected by the measurement error. For small sized portfolios the measurement errors in the two portfolios to be compared may be different, causing the F-value to be significant even if the *true* intercepts are not different. However, if measurement error converges, its impact on the intercept will be constant beyond some n, in which case F will be identically zero under the null hypothesis of equal average pricing errors.

Even if measurement error does not converge but continuously falls, say as $1/n$, the F-statistics will tend toward zero. To see that the ratio should vanish, consider $m < n$ and m and n are large enough so that $\alpha^{(n)}$ is on the order of $1/n$. Then the F-ratio can be written as

$$F = nm\,\frac{(\alpha^{(n)} - \alpha^{(m)})^2}{(a^{(n)} + a^{(m)})} = nm\,\frac{(1/n - 1/m)^2}{(a^{(n)} + a^{(m)})} = \frac{1}{nm}\cdot\frac{(m - n)^2}{(a^{(n)} + a^{(m)})}. \quad (16)$$

By Chebychev's inequality the probability that $v_i^{(n)} - \bar{v}^{(n)}$ is greater than $n\sigma$ is at most $1/n^2$, so that for large enough n and m we may write

$$a^{(n)} \le \sum_{i=1}^{n} (n\sigma)^2 = n^3\sigma^2, \qquad \text{and} \qquad a^{(m)} \le \sum_{i=1}^{m} (m\sigma)^2 = m^3\sigma^2. \quad (17)$$

Thus, $a^{(n)} = \mathrm{O}(n^3\sigma^2)$ and $a^{(m)} = \mathrm{O}(m^3\sigma^2)$, where $\mathrm{O}(\cdot)$ indicates order of magnitude. The denominator of F will grow large quickly with n, implying that F should fall to zero for large n.

D. Summary of Hypotheses

To compare the factors with the components, we will observe the behavior of the intercept, the R^2, the regression f, and the *pricing error F* of the cross-sectional pricing regression, equation (4), as n grows large. A better measure of the $\mathbf{b}^{(n)}$ should result in lower measurement and pricing error than a poor measure of $\mathbf{b}^{(n)}$. Hence, deviation of $\hat{\alpha}^{(n)}$ from the zero-beta rate should be smaller for better measures of $\mathbf{b}^{(n)}$. Note that, if the intercepts are higher than the zero-beta rate, then the empirical pricing errors are positive and a higher intercept would imply a higher empirical pricing error. A better $\mathbf{b}^{(n)}$ should produce higher R^2, higher regression f, and lower pricing error F than poor measures of $\mathbf{b}^{(n)}$. We hypothesize that factor loadings will be better measures of $\mathbf{b}^{(n)}$ than eigenvectors for reasons discussed above and that a larger k will do better than a smaller k because there are more dimensions to capture risk. Note that a larger k also increases the potential sources of measurement error, so increasing k beyond some point may not be worth the cost.

III. Cross-Sectional Test Results

Using twenty years of weekly data for each portfolio in the sequence, we estimate the eigenvectors and the one- and five-factor models for the sample covariance matrix, as well as the market model *betas* for the securities in the portfolios using the equal weighted and value weighted CRSP indices.[8] We use the average, realized Treasury bill rate from the Ibbotson Associates (1987) data as an estimate of the zero-beta rate. For the twenty-year period, 1962–1983, the estimate of the zero-beta rate is 6.4% per year, which translates to 0.0012 on a weekly basis. Tables reported here present the results for the first two portfolios in the sequence (30, 33), the midpoint of the sequence (129), and the largest portfolio (596). Average intercept and slope obtained using the individual coefficients for the portfolios in the sequence are also reported. The tables of complete results are available from the authors on request.

Table I reports the magnitudes of correlations among the first factor of the one-factor model, the first eigenvector, and the market model betas. Only magnitudes are reported because the signs of the correlations are irrelevant owing to scaling and rotation considerations.[9] The correlation between the first vector and first factor is high, growing from 0.923 in the 30-firm sample to 0.990 in the 596-firm sample. The first factor is highly correlated with both the equal and value weighted market betas, but the first vector has a much higher correlation with the equal weighted betas than with the value weighted. This supports Brown's (1989) theoretical argument that the first principal component *is* the equal weighted market index if the idiosyncratic risks are equal across firms.

The mean returns computed over the entire twenty-year period were regressed on the first eigenvector and the first five eigenvectors.[10] This was repeated for

[8] The eigenvectors reported were computed using CDC 815 and a standard IMSL software. The vectors of matrices up to size 600 were checked against the output of an IBM 3081 with no discrepancies.

[9] The *security market line* for the eigenvector case lies in the fourth quadrant. The magnitude of the slope coefficient is getting larger with n, indicating that the risk premium for the first eigenvector is growing larger.

[10] By *first* eigenvector we mean the eigenvector corresponding to the largest eigenvalue. The *first five* refers to the vectors associated with the five largest eigenvalues.

Table I

Magnitudes of Correlations Between the Risk Measures

Correlations among the one-dimensional measures of risk—the first factor loading, the first eigenvector, the value weighted index beta, and the equal weighted index beta—were calculated for the securities in every portfolio in the sequence. The four panels of this table report the results for the portfolios consisting of 30, 33, 129, and 596 securities.

	First Factor	Value Wtd. Mkt.	Equal Wtd. Mkt.
	A. Size 30		
First Vector	0.923	0.846	0.953
First Factor		0.969	0.990
Value Wtd. Mkt.			0.949
	B. Size 33		
First Vector	0.931	0.849	0.956
First Factor		0.963	0.989
Value Wtd. Mkt.			0.949
	C. Size 129		
First Vector	0.973	0.931	0.996
First Factor		0.980	0.975
Value Wtd. Mkt.			0.933
	D. Size 596		
First Vector	0.990	0.950	0.995
First Factor		0.979	0.957
Value Wtd. Mkt.			0.927

one factor, five factors, and the market model betas. Five-year averages for four subperiods were also regressed on these measures of systematic risk.

A. Twenty-Year Mean Returns

Table II reports the regression results for the twenty-year mean returns. The intercept and the risk premium have both been scaled by multiplying by 1,000. Note that the intercept is always larger than the zero-beta rate of 1.2×10^{-3}. The column marked R^2 reports the *adjusted* R^2 for the regression, f reports the usual f-statistic for the regression, and the column headed F reports the *pricing error* F. Panels A and B report the regressions using the first eigenvector and one factor. Panels C and D show the results using the two market model betas. Comparing the numbers in Panels A and B, we find that the first eigenvector better describes the cross-sectional distribution than the first factor loading. The R^2's are higher, intercepts lower, regression f's higher, and pricing error F's lower for the vector regressions than for the factor regressions. In fact, this is true for every portfolio. Interestingly, the first vector also dominates both market model betas for smaller portfolios but is somewhat worse than the value weighted beta for the larger portfolios. It is not worse than the equal weighted beta, however.

Table II

Cross-Sectional Regressions for One-Dimensional Measures of Risk

This table reports the regression of twenty-year average security returns on the first eigenvector (Panel A), the first factor loading (Panel B), and the value (Panel C) and equal weighted index betas (Panel D). The column marked n denotes the number of securities in the cross-section. $\hat{\alpha}$ and $\hat{\gamma}_1$ denote the intercept and the systematic risk premium, respectively. The numbers in parentheses are the t-statistics. R^2 is the coefficient of determination, f is the usual regression F-statistic, and F tests for the convergence of the pricing errors. $\hat{\alpha}$ and $\hat{\gamma}_1$ have been multiplied by 10^3. Therefore, the numbers in the first row are actually 0.001903 and −0.006058.

n	$\hat{\alpha}(\times 10^3)$	$\hat{\gamma}_1(\times 10^3)$	R^2	f	F
A. First Eigenvector as Systematic Risk					
30	1.903	−6.058	0.3417	16.05	
	(6.90)	(−4.01)			
33	1.949	−6.471	0.2501	11.67	0.0428
	(5.91)	(−3.42)			
129	1.518	−16.814	0.2372	40.80	0.0614
	(6.55)	(−6.39)			
596	1.328	−43.976	0.2533	203.14	1.3484
	(10.51)	(−14.25)			
Average	1.534	−20.900			
	(41.12)	(−9.87)			
B. First Factor as Systematic Risk					
30	1.754	58.052	0.1851	7.59	
	(4.10)	(2.75)			
33	1.865	56.838	0.1176	5.27	0.2061
	(3.77)	(2.30)			
129	1.600	66.353	0.1537	24.25	0.2090
	(6.03)	(4.93)			
596	1.375	80.167	0.2074	156.73	1.2646
	(9.90)	(12.52)			
Average	1.602	68.188			
	(42.66)	(33.93)			
C. Market Model Beta (Using Value Weighted Market) as Systematic Risk					
30	1.800	1.300	0.2100	8.90	
	(4.60)	(2.98)			
33	1.900	1.300	0.1600	6.90	
	(4.19)	(1.62)			
129	1.500	1.600	0.2200	36.20	
	(6.28)	(6.02)			
596	1.300	2.000	0.2700	218.20	
	(10.11)	(14.77)			
Average	1.521	1.671			
	(39.57)	(36.53)			
D. Market Model Beta (Using Equal Weighted Market) as Systematic Risk					
30	1.900	1.100	0.1300	5.20	
	(4.46)	(2.28)			
33	2.100	1.000	0.0800	3.70	
	(4.22)	(1.92)			
129	1.800	1.200	0.1400	21.30	
	(6.72)	(4.62)			
596	1.500	1.600	0.1900	136.30	
	(11.14)	(11.67)			
Average	1.756	1.284			
	(47.86)	(30.62)			

The power of the models is confirmed by Zellner's (1971) posterior odds test of the models specifications. Assuming that the mean returns are normally distributed and the prior information about the models is diffuse, Table III reports the odds in favor of model 1 versus model 2. The number 24.560 means that the odds are 24.560 to 1 that a one-vector model better represents the data than a one-factor model. As Zellner observes, there is an unavoidable degree of arbitrariness in such numbers since the definition of *significant* odds is not obvious. However, it is clear from the table that any reasonable definition of significant odds will result in the conclusion that the value weighted market betas are favored or are equal to the one-vector model, the one-vector model is favored over the one-factor model, and the equal weighted market betas do the worst of all.

This result about the powers of the models is surprising given the high correlations reported in Table I and the results of recent research. Chen, Copeland, and Mayers (1987) and Brown and Weinstein (1985) provide evidence that four factors contribute little over the use of an equal weighted market index, and Chen, Roll, and Ross (1986) argue that, if a *market portfolio* is priced, it is the equal weighted and not the value weighted index. Yet, it is clear from Table II that, when n is large, the first eigenvector has more explanatory power, a lower intercept, a higher regression f, and better odds than the equal weighted market beta and is roughly equivalent to the CRSP value weighted beta. Thus, while the first vector may be highly correlated with the equal weighted beta, it is more like, but slightly worse than, the value weighted beta when used as a measure of systematic risk in a pricing regression. It is worth noting, however, that these betas are assumed stationary over the twenty-year period. Consequently, the results with beta are not necessarily comparable to typical asset pricing betas, which are estimated using shorter periods. However, the measurement errors caused by the heroic assumption of twenty-year mean stationarity should favor the one-factor model, but the one-vector model is clearly empirically superior.

Table II and Figure 1 show that the intercepts clearly fall with n for all three measures of systematic risks. This convergence behavior of the intercept does

Table III

Posterior Odds Ratio Test

This table gives the results of Zellner's posterior odds ratio test to compare the power of the cross-sectional regression equations. For example, for size 30 portfolio, the odds are 24.560 to 1 in favor of the one-vector model over the one-factor model.

Odds in favor of		Portfolio Size			
Model 1 vs.	Model 2	30	33	129	596
One-Vector	One-Factor	24.560	14.638	807.540	5.122×10^{7}
One-Vector	Value Wtd. Mkt.	14.460	7.109	6.039	0.003
One-Vector	Equal Wtd. Mkt.	69.375	30.668	2817.387	1.911×10^{11}
One-Factor	Value Wtd. Mkt.	0.589	0.486	0.007	0.000
One-Factor	Equal Wtd. Mkt.	2.825	2.096	3.482	3813.560
Value Wtd. Mkt.	Equal Wtd. Mkt.	1.359	1.333	4.181	5212.017
Five-Vector	Five-Factor	0.378	0.046	0.001	4.64×10^{-7}

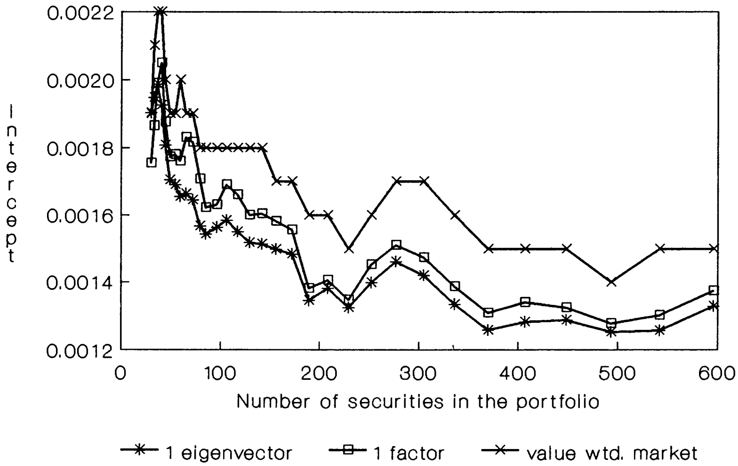

Figure 1. The behavior of cross-sectional pricing regression intercept with respect to the number of securities. One eigenvector, one factor, and the value weighted index beta are used as measures of risk to estimate the cross-sectional regressions for the portfolios in the sequence. The regression intercept is plotted against the number of securities to illustrate its convergence.

not necessarily mean that the measurement error in individual securities' systematic risks is falling as the sample size is increasing.[11] We have shown above that this behavior would be expected if the covariance between the systematic risk and its measurement error were small and the cross-sectional variance of measurement error were small compared to that of the measurement error. The intercepts in the factor and the eigenvector regression appear to be converging to about 0.00125, or about 7% per year, which is reasonably close to our estimate for the return on the zero-beta portfolio. This implies that the average pricing errors in APT may be quite small regardless of the measure of systematic risk utilized. (Remember that the intercept is equal to the zero-beta rate plus the average pricing error in the absence of any measurement error.)

The pricing error F-statistic shows no pattern of convergence but is never significant, indicating that the adjacent portfolios never have significantly different mean pricing errors. The hypothesis is that as measurement errors converge we expect the differences between the intercepts to be insignificant. Table II appears to indicate that large n is 30, implying a very fast convergence.[12] However, there is a significant difference between the empirical pricing error for

[11] This may be true for the maximum likelihood factor analysis and the principal components analysis models, where the estimate of the systematic risk changes—and maybe becomes more precise—as the size of the covariance matrix increases. However, this would not be true for the market model because the market model betas for individual securities do not change as the sample size is changed.

[12] Large measurement or pricing errors do not affect the pricing error F-statistic since it is the ratio of the difference between the average errors in adjacent portfolios to the sum of the variance within each portfolio. Large errors will simply net out. However, highly volatile measurement error with a relatively constant mean will cause this statistic to be insignificant. Nevertheless, it is likely that volatile errors will result in volatile means between portfolios.

Table IV

Cross-Sectional Regressions for Five-Dimensional Measures of Risk

This table reports the regression of twenty-year average security returns on the first five eigenvectors (Panel A) and the five factor loadings (Panel B). The column marked n denotes the number of securities in the cross-section. $\hat{\alpha}$ is the intercept. The numbers in parentheses are the t-statistics. R^2 is the coefficient of determination, f is the usual regression F-statistic, and F tests for the convergence of the pricing errors. $\hat{\alpha}$ has been multiplied by 10^3. Therefore, the number in the first row is actually 0.001868.

n	$\hat{\alpha}(\times 10^3)$	R^2	f	F
A. First Five Eigenvectors as Systematic Risk				
30	1.868	0.3284	3.84	
	(6.60)			
33	1.920	0.2220	2.83	0.0528
	(5.70)			
129	1.773	0.3676	15.88	0.1538
	(8.00)			
96	1.899	0.3688	70.50	6.3007
	(12.80)			
Average	1.764			
	(77.11)			
B. First Five Factors as Systematic Risk				
30	2.790	0.3706	4.42	
	(5.25)			
33	3.319	0.3543	4.51	0.3050
	(5.30)			
129	2.326	0.3587	15.32	0.0025
	(7.90)			
596	2.333	0.3990	80.00	4.4477
	(14.70)			
Average	2.431			
	(40.191)			

a portfolio of size 30 and the largest portfolio of size 596. This indicates that the fall in intercepts seen in Figure 1 is not a statistical artifact and that average measurement errors are indeed falling. Nevertheless, the intercepts of the largest portfolios are insignificantly different from each other, implying that there will be little gain from adding more data.[13] As the average pricing error is growing smaller, it appears that the risk premium for the first vector or the market risk premium, if Brown (1989) is correct, is capturing more of the average return for each security.

Table IV and Figure 2 show the results of regressing the mean return on the first five eigenvectors (Panel A of Table IV) and the first five factors (Panel B). The adjusted R^2 shows an improvement over Table II, but the remaining statistics are not improved. Figure 2 is particularly striking in showing that the intercepts of the five-vector regressions do not converge as fast as the intercept of the

[13] The intercepts begin to level off after portfolio 142. The pricing error F-statistic for portfolio sizes 142 and 596 is 3.638, which is significant, but the difference between the next portfolio, 156, and 596 is 3.280, which is not significant at the 5% level.

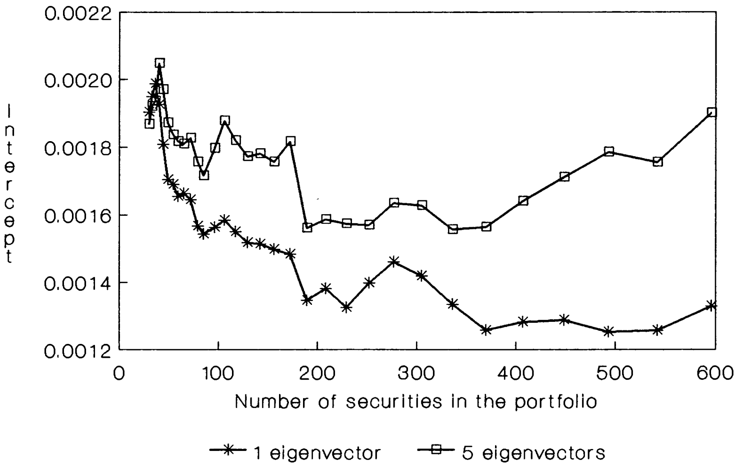

Figure 2. The behavior of cross-sectional pricing regression intercept with respect to the number of securities. One eigenvector and five eigenvectors are used as measures of risk to estimate the cross-sectional regressions for the portfolios in the sequence. The regression intercept is plotted against the number of securities to illustrate its convergence.

one-vector regression, if they converge at all. The pricing error F-statistic tends to be higher for the five-vector regression than the one-vector case, but it is still never significant.

The surprise in Tables II and IV is the larger intercepts of the factor loading regressions over the vector regressions. This is unexpectedly good news for principal components analysis. The stationarity assumption of these tables is that means, variances, and covariances are constant for twenty years. Numerous studies have found otherwise, and the resultant large empirical pricing errors will appear as large intercepts and large estimates of idiosyncratic risk. Maximum likelihood factor analysis, in principle, is much better suited to accommodating such large errors because it jointly estimates the residual covariance matrix and the factor loadings that maximize the likelihood function. Principal components analysis simply chooses the eigenvectors according to the arbitrary criterion that their eigenvalues be the largest. Chi-squared statistics testing the hypothesis that k components capture the variation in these covariance matrices reject the hypothesis for values of k smaller than $\frac{1}{2}n$. For example, it takes 95 components to represent the variation in a covariance matrix of 189 securities. The large chi-squared statistics for components imply that the idiosyncratic risks of firms are a high percentage of the variation of returns and that there are larger residual risk differences between firms. This is exactly the circumstance that favors factor analysis. Yet the empirical pricing error as measured by the intercept is at least as large for factor loadings analysis as for the vectors.

This interpretation is based on a theoretical pricing error that is either negligible[14] or varies in sign. If these assumptions are not true and the theoretical

[14] Ross (1976) and Dybvig (1983) show that the pricing error in our economy is very small. Dybvig's simple calculation shows that the magnitude of pricing error for an average security is on the order of 0.04% per year.

pricing error is negative for all samples, then the lower intercept implies a higher empirical pricing error. In this case, we believe that the magnitude of the intercepts indicates that the principal components analysis is only slightly worse than the maximum likelihood factor analysis.

One plausible explanation for larger intercepts with factor analysis is that the algorithms used to compute eigenvectors are less sensitive to measurement errors than the iterative algorithms used to obtain maximum likelihood estimates. In fact, the factor analysis algorithms—such as the Joreskog algorithm used here—employ eigenvalue and eigenvector routines in an iterative scheme. Since principal components analysis uses the eigenvector and eigenvalue routine only once, they could hardly be computationally less stable than factor analysis. Furthermore, maximum likelihood estimates have desirable properties in large samples where T, the number of observations, grows large. We, like the authors of many other studies, have a growing number of firms but not necessarily a large T. This drawback may cause maximum likelihood factors to fare poorly.[15]

In spite of the intercepts in Table IV, there is evidence that factor loadings are equivalent or better than eigenvectors as measures of systematic risk. First, the posterior odds favor factor loadings in the large portfolios. The odds that factor loadings are a better measure of systematic risk are about equal for the smaller samples but are overwhelming for the larger. This is clearly due to higher R^2's in the factor regressions. Second, the regression f's are higher for the factor case for larger sized samples. Finally, average empirical pricing errors may not adequately represent the empirical distribution of pricing errors for many applications. If an asset pricing model is used to measure the effect of an anomaly or the importance of time variation in risk premia, the average pricing error may be of little interest. Thus, our comparisons of intercepts would not be of much use for such applications.

B. Five-Year Means

To examine the impact of nonstationary means but still assuming stationary covariances, we regressed average returns over four five-year subperiods on the measure of systematic risk. The results are roughly consistent with the finding using twenty-year means, but there are differences. Regression statistics in the first and fourth subperiods (7/19/62–10/9/67 and 11/6/77–12/27/82) are equivalent to that in the twenty-year period, but the middle subperiods produce very poor results. Adjusted R^2's are zero or nearly zero, and the regression f's are often not significant in the middle subperiods. Table V compares the regression statistics of the four models. Panel A compares $\hat{\alpha}$ between the four models. The 29 reported at the intersection of the five-vector column and one-vector row means that $\hat{\alpha}$ is higher in 29 out of 32 regressions for the five-vector model than for the one-vector model. The comparison is not statistical since we are not reporting which $\hat{\alpha}$ is statistically higher, just which coefficient is larger. Our prior

[15] We do not believe that the differences in intercepts are caused by software or hardware problems. We computed all eigenvectors and factor loadings using standard IMSL software and checked them against SAS output. The results were identical. We also computed all eigenvectors on a CYBER 815 and on an IBM 3081 GX and found no differences.

Table V

Comparison of Various Models for the Four Subperiods

These tables show for how many portfolios out of 32 the quantity of interest ($\hat{\alpha}$, R^2, or f) was *larger* for the case denoted by the column compared to the case denoted by the row. For example, 19 times out of 32 $\hat{\alpha}$ was higher for one factor case compared to one vector case in subperiod 4.

	Five Vectors				One Factor				Five Factors			
Subperiod →	1	2	3	4	1	2	3	4	1	2	3	4
					A. $\hat{\alpha}$							
One Vector	29	0	29	17	5	5	32	19	32	27	19	15
Five Vectors					0	26	21	15	32	29	13	17
One Factor									32	28	6	22
					B. R^2							
One Vector	32	21	32	24	2	24	12	0	32	25	29	26
Five Vectors					0	10	0	0	14	23	13	20
One Factor									32	25	29	31
					C. f							
One Vector	3	32	32	0	2	22	12	0	5	31	30	22
Five Vectors					22	0	0	32	17	23	12	22
One Factor									10	31	32	0

probability is that ½ should be higher and ½ lower if there are no differences between the models. For a sample size of 30 observations, a binomial test of the hypothesis that the probability is ½ will be rejected if the observed frequency is outside the range 0.31–0.64 or about 10–20. This simple test assumes independence across samples, which may not be completely satisfied. However, using these criteria, it is reasonable to conclude that the one-vector model is equivalent to the one-factor model in terms of having a lower $\hat{\alpha}$ and higher adjusted R^2, and the one-vector and one-factor models appear to have at least as low an $\hat{\alpha}$ as the five-vector and five-factor models. This conclusion is also confirmed by the posterior odds tests, which favor the one-vector model in subperiods 1 and 4 and neither model in the other two periods. Thus, the unexpected conclusion that a single eigenvector does at least as well as a multiple factor or vector model in terms of measurement and pricing error is as well founded in subperiods as in the overall period.

The multifactor or multivector models do better in terms of explanatory power (adjusted R^2) than the single factor or vector models. The single vector model has a higher adjusted R^2 in three out of four subperiods when compared with a single factor model. The regression f-statistics tend to follow the R^2 results, but the one-vector model has a higher regression f than the five-vector and five-factor models in the first subperiod. Finally, the posterior odds favor the factor model in the first period, the five-vector model in the third period, and neither in the other periods.

The one conclusion that is clearly supported by this table is that eigenvector models are at least as good as the factor models. The vector models tend to have lower $\hat{\alpha}$, equivalent R^2, and equivalent regression f-statistics, and the posterior odds favor the vector models as much as they favor the factor models.

C. The Size Effect

To determine whether the patterns observed in Tables II–IV are the result of an omitted size variable, we computed the total market value of equity for all 865 firms at the beginning of the twenty-year period and at the end of each of the twenty years. Our size variable is the average total market value of these 21 observations for each firm. First, we regressed the average twenty-year return on the size variable for every portfolio in the sequence. The size variable has a significant negative coefficient, and the regression f is always significant at the 0.01 level.

Next, we included firm size as an additional variable in the cross-sectional regressions discussed in Section III.*A*. Table VI reports representative results from these regressions. Size is significant in 11 of 32 one-vector regressions and 20 of 32 one-factor regressions. All the significant size coefficients occur for larger portfolios. The one-factor measure of systematic risk always has lower R^2's and larger intercepts for $n > 33$. Thus, adding size does not alter our basic conclusion that one-vector models appear superior to one-factor models. Similarly, the five-vector models appear equivalent to five-factor models. The size variable is significant in only four of 32 five-vector regressions and in none of the five-factor regressions. The five-vector intercepts are always smaller, and 22 of the five-factor R^2's are larger. Thus, by the criterion of R^2, the five-factor models are better, but, if intercepts are compared, the five-vector models are superior. The one-vector model still has lower intercepts than the one-factor, five-vector, or five-factor models. The five-factor models capture the size variable slightly better than the five-vector models. Finally, the size variable does not alter the negative relationship between the intercept and the portfolio size.

D. Residual Variance

One of the important advantages of factor analysis over principal components is that idiosyncratic risk is explicitly modeled and estimated. As a result, k components will have no greater explanatory power than k factors and can often have much less. In those circumstances where idiosyncratic risks are larger and varied, a small number of components will capture much less variation in the data, and the estimated residual variance implied by the components will be large. Thus, in a cross-sectional pricing equation, we expect that the residual variance implied by components is more likely to be significant than the residual variance of a factor analysis because factor analysis should provide a much better partitioning of total risk into systematic and residual.

To test the differences in the two estimation techniques, we used the residual variances implied by each technique in the cross-sectional pricing equations.[16] The representative results are reported in Table VII. Residual variance is positive and significant for all portfolios for the one-vector, one-factor, and five-factor cases and 22 portfolios for the five-vector case. Thus, while in some cases vectors explain residual variance, most do not. The use of twenty-year means is likely to

[16] The residual variance for security i is the (i, i) element of the residual covariance matrix $\mathbf{\Phi}$, which is calculated as $\mathbf{\Sigma} - \mathbf{bb}'$ for factor analysis and as $\mathbf{\Sigma} - \mathbf{bD}_\lambda\mathbf{b}'$ for vector analysis, where $\mathbf{D}_\lambda$ is the k-dimensional diagonal matrix with eigenvalues along the diagonal.

Table VI

Twenty Years Average Returns Regressed on Measures of Systematic Risk and Size

This table reports the regression of twenty-year average security returns on firm size and the first eigenvector (Panel A), the first factor loading (Panel B), the five eigenvectors (Panel C), and the five factor loadings (Panel D). The column marked n denotes the number of securities in the cross-section. $\hat{\alpha}$, $\hat{\gamma}_1$, and $\hat{\theta}$ denote the intercept, the systematic risk premium, and the coefficient of size, respectively. The numbers in parentheses are the t-statistics. R^2 is the coefficient of determination, and f is the usual regression F-statistic. $\hat{\alpha}$ and $\hat{\gamma}_1$ have been multiplied by 10^3, and $\hat{\theta}$ by 10^7. Therefore, the numbers in the first row are actually 0.002128, −0.005710, and −0.0000003879.

n	$\hat{\alpha}(\times 10^3)$	$\hat{\gamma}_1(\times 10^3)$	$\hat{\theta}(\times 10^7)$	R^2	f
		A. First Eigenvector as Systematic Risk			
30	2.128	−5.710	−3.879	0.3285	7.36
	(5.84)	(−3.36)	(−0.99)		
33	2.242	−5.981	−5.544	0.2515	5.87
	(5.37)	(−2.87)	(−1.21)		
129	1.666	−16.120	−1.336	0.2326	19.19
	(6.48)	(−5.76)	(−1.59)		
596	1.406	−43.438	−0.476	0.2691	104.44
	(10.67)	(−13.80)	(−2.62)		
		B. First Factor as Systematic Risk			
30	2.064	55.734	−6.703	0.2098	4.45
	(4.20)	(2.45)	(−1.63)		
33	2.215	54.307	−8.197	0.1559	3.68
	(3.99)	(2.06)	(−1.72)		
129	1.765	63.792	−1.738	0.1611	12.52
	(5.89)	(4.50)	(−1.99)		
596	1.466	78.893	−0.573	0.2237	82.00
	(10.18)	(12.10)	(−3.07)		
		C. Five Eigenvectors as Systematic Risk			
30	2.013		−2.473	0.2938	2.80
	(5.22)		(−0.60)		
33	2.150		−4.310	0.2002	2.21
	(4.83)		(−0.88)		
129	1.851		−0.618	0.3634	12.42
	(7.60)		(−0.78)		
596	1.959		−0.285	0.3875	60.25
	(12.88)		(−1.67)		
		D. Five factors as Systematic Risk			
30	3.022		−4.651	0.3829	4.00
	(5.40)		(−1.22)		
33	3.631		−6.299	0.3801	4.27
	(5.59)		(−1.46)		
129	2.375		−0.664	0.3568	12.83
	(7.88)		(−0.79)		
596	2.391		−0.335	0.4018	67.61
	(14.84)		(−1.95)		

Table VII

Twenty Years Average Returns Regressed on Measures of Systematic Risk and Residual Variance

This table reports the regressions of twenty-year average security returns on residual variance and the first eigenvector (Panel A), the first factor loading (Panel B), the five eigenvectors (Panel C), and the five factor loadings (Panel D). The column marked n denotes the number of securities in the cross-section. $\hat{\alpha}$, $\hat{\gamma}_1$, and $\hat{\phi}$ denote the intercept, the systematic risk premium, and the coefficient of residual variance, respectively. The numbers in parentheses are the t-statistics. R^2 is the coefficient of determination, and f is the usual regression F-statistic. $\hat{\alpha}$ and $\hat{\gamma}_1$ have been multiplied by 10^3. Therefore, the numbers in the first row are actually 0.001904, −0.002338, and 0.266.

n	$\hat{\alpha}(\times 10^3)$	$\hat{\gamma}_1(\times 10^3)$	$\hat{\phi}$	R^2	f
		A. First Eigenvectors as Systematic Risk			
30	1.904	−2.338	0.266	0.4181	11.42
	(7.34)	(−1.05)	(2.16)		
33	1.956	−0.918	0.376	0.3763	10.65
	(6.50)	(−0.34)	(2.70)		
129	1.649	−8.088	0.262	0.3388	33.79
	(7.58)	(−2.60)	(4.53)		
596	1.486	−23.906	0.270	0.3578	166.76
	(12.56)	(−6.81)	(9.88)		
		B. First Factor as Systematic Risk			
30	1.924	15.412	0.271	0.4063	10.92
	(5.22)	(0.70)	(3.38)		
33	2.077	3.097	0.336	0.3603	10.01
	(4.89)	(0.12)	(3.57)		
129	1.647	30.590	0.287	0.3327	32.90
	(6.68)	(2.28)	(5.92)		
596	1.471	43.171	0.289	0.3547	164.51
	(11.71)	(6.55)	(1.69)		
		C. Five Eigenvectors as Systematic Risk			
30	1.423		0.400	0.4422	4.83
	(4.51)		(3.67)		
33	1.247		0.601	0.4681	5.69
	(3.73)		(3.67)		
129	1.805		0.126	0.3717	13.62
	(8.16)		(1.35)		
596	1.867		0.320	0.4056	68.66
	(12.97)		(6.12)		
		D. Five Factors as Systematic Risk			
30	2.914		0.254	0.5151	6.13
	(6.23)		(2.86)		
33	3.474		0.307	0.4947	6.22
	(6.25)		(2.92)		
129	2.382		0.299	0.4320	17.22
	(8.58)		(4.11)		
596	2.223		0.261	0.4387	78.50
	(14.50)		(6.54)		

introduce more measurement error than in the typical asset pricing study, where the researchers usually make considerable efforts to reduce measurement error.

The vector models do at least as well as the factor models in terms of R^2 and the size of the intercepts. Vector models tend to have lower intercepts and smaller R^2's than factor models. The intercepts of all models converge, and, remarkably, the intercept of the one-vector model is smaller than the five-vector and five-factor models. Thus, we conclude that the relative strength of the models is not affected by including the residual variance.

IV. Summary

This paper examines the behavior of the empirical pricing error (the true pricing error plus the measurement error) from cross-sectional regressions of mean returns on the first and the first five eigenvectors, the factors, and the market model betas. We compute the eigenvectors and factors of sample covariance matrices in a sequence of increasingly larger matrices using data on all the firms with a twenty-year trading history on the 1983 CRSP tape. The twenty-year mean returns are then regressed cross-sectionally against these measures. For empirical tests we assume stationarity for twenty years. This introduces measurement errors which are likely to favor factor analysis. We show that, if measurement error is uncorrelated with the systematic risk, then it must converge as the size of the covariance matrix grows larger, thus allowing us to isolate the theoretical pricing error. The results show that the empirical intercepts of the one-vector models converge to about 7% for the twenty-year means, implying low pricing errors. Remarkably, the one-vector model has an empirical pricing error smaller than or equivalent to the five-vector model or the five-factor model using maximum likelihood factor analysis for twenty-year mean returns. Results for the five-year means show that the eigenvector models do at least as well as the factor models. However, the estimates of the risk premia and the explanatory power of the regression are highly volatile between subperiods, implying that the nonstationarity problem must be dealt with in a more sophisticated manner if the estimates of the risk-return relationship are important. We examine whether these conclusions are due to size and residual variance. Both factors and vectors capture the size effect equally well, while neither is able to explain the pricing of residual variance. The significance of residual variance is likely due to the large measurement errors not normally found in asset pricing tests.

These results should not be taken as a straightforward advocacy of the eigenvector models. In some applications the tendency of factors to have a higher R^2 may be an important advantage. The point of this study is that eigenvectors should not be ruled out a priori. The power of either methodology is an empirical issue which should be examined in specific circumstances. It may be that the factor analysis is superior to principal components analysis, but the results of this study are a clear indication that the superiority of any technique must be decided on a case-by-case basis.

Finally, it is clear from the cross-sectional regression results that the APT has considerable empirical power. A five-dimensional measure of systematic risk explains as much as 40% of the variation in mean returns of our largest sample.

This occurred in a regression which made the heroic assumption that twenty-year means, variances, and covariances are stationary. Moreover, the regression had an intercept of about 7% per year in spite of the measurement and pricing errors which are certainly impounded in the intercept. The explanatory power of the APT over standard estimates of beta occurs because it allows more dimensions. When the model specification is confined to the set of one-dimensional measures, the value weighted beta is the better measure of risk.

REFERENCES

Brown, Stephen J., 1989, The number of factors in security returns, *Journal of Finance* 44, 1247–1262.

——— and Mark I. Weinstein, 1983, A new approach to testing asset pricing models: The bilinear paradigm, *Journal of Finance* 38, 711–743.

——— and Mark I. Weinstein, 1985, Derived factors in event studies, *Journal of Financial Economics* 14, 491–495.

Chamberlain, Gary and Michael Rothschild, 1983, Arbitrage, factor structures, and mean-variance analysis on large asset markets, *Econometrica* 51, 1281–1301.

Chen, Nai-fu, Thomas Copeland, and David Mayers, 1987, A comparison of single and multifactor portfolio performance methodologies, *Journal of Financial and Quantitative Analysis* 22, 401–417.

———, Richard Roll, and Stephen Ross, 1986, Economic forces and the stock market: Testing the APT and alternative pricing theories, *Journal of Business* 59, 383–403.

Connor, Gregory and Robert Korajczyk, 1988, Risk and return in an equilibrium APT: Application of a new methodology, *Journal of Financial Economics* 21, 255–289.

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

Dybvig, Philip, 1983, An explicit bound on individual asset's deviations from APT pricing in a finite economy, *Journal of Financial Economics* 12, 483–496.

——— and Stephen Ross, 1985, Yes, the APT is testable, *Journal of Finance* 40, 1173–1188.

Fama, Eugene F., 1976, *Foundations of Finance* (Basic Books, New York).

Ferson, W., S. Kandel, and R. Stambaugh, 1987, Tests of asset pricing with time-varying expected risk premiums and market betas, *Journal of Finance* 42, 201–220.

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

Ibbotson Associates, 1987, *Stocks, Bonds, Bills and Inflation: 1987 Yearbook.*

Johnston, J., 1984, *Econometric Methods* (McGraw-Hill, New York).

Linn, Scott and E. Chang, 1985, The empirical implications of approximate factor structure in large asset markets, Unpublished working paper, University of Iowa.

Morrison, D., 1976, *Multivariate Statistical Methods* (McGraw-Hill, New York).

Perry, P., 1982, The time variance relationship of security returns: Implications for the return generating stochastic process, *Journal of Finance* 37, 857–869.

Roll, Richard and Stephen Ross, 1980, An empirical examination of the Arbitrage Pricing Theory, *Journal of Finance* 35, 1073–1103.

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

Shanken, Jay, 1987, Nonsynchronous data and the covariance-factor structure of returns, *Journal of Finance* 42, 221–231.

Trzcinka, Charles, 1986, On the number of factors in the arbitrage pricing model, *Journal of Finance* 41, 347–368.

Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley & Sons, New York).