NBER WORKING PAPER SERIES

THE EMPIRICAL FOUNDATIONS OF THE ARBITRAGE PRICING THEORY II: THE OPTIMAL CONSTRUCTION OF BASIS PORTFOLIOS

Bruce N. Lehmann

David M. Modest

Working Paper No. 1726

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1985

We would like to express our gratitude to the Columbia Business School for providing us with the computational resources to carry out this analysis, and to Arthur Warga for helpful comments. The research reported here is part of the NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #1726
October 1985

The Empirical Foundations of the Arbitrage Pricing Theory II: The Optimal Construction of Basis Portfolios

ABSTRACT

The Arbitrage Pricing Theory (APT) of Ross (1976) presumes that a factor model describes security returns. In this paper, we provide a comprehensive examination of the merits of various strategies for constructing basis portfolios that are, in principle, highly correlated with the common factors affecting security returns. Three main conclusions emerge from our study. First, increasing the number of securities included in the analysis dramatically improves basis portfolio performance. Our results indicate that factor models involving 750 securities provide markedly superior performance to those involving 30 or 250 securities. Second, comparatively efficient estimation procedures such as maximum likelihood and restricted maximum likelihood factor analysis (which imposes the APT mean restriction) significantly outperform the less efficient instrumental variables and principal components procedures that have been proposed in the literature. Third, a variant of the usual Fama-MacBeth portfolio formation procedure, which we call the minimum idiosyncratic risk portfolio formation procedure, outperformed the Fama-MacBeth procedure and proved equal to or better than more expensive quadratic programming procedures.

Professor Bruce Lehmann
Graduate School of Business
Columbia University
New York, NY 10027

Professor David M. Modest
Graduate School of Business
Columbia University
New York, NY 10027

I. Introduction

Despite the considerable empirical success of capital asset pricing theory, there remains a disturbing lack of scientific consensus concerning the validity of the various *state of the art* asset pricing models. The Capital Asset Pricing Model (**CAPM**) has run into several roadblocks such as Roll's (1977) suggestion that it is not a testable scientific theory and a plethora of empirical anomalies which provide empirical evidence that the usual market proxies are not mean-variance efficient.[1] Empirical and theoretical research has proceeded in several directions, including consideration of the effects of personal taxes and other market imperfections, the possibility of dynamic effects arising through shifts in the investment opportunity set (intertemporal asset pricing models), and stronger distributional assumptions about the underlying structure of security returns. Each of these approaches has prominent adherents and is, at present, the subject of considerable theoretical and empirical attention.

One of the main lines of current empirical research in asset pricing is the Arbitrage Pricing Theory (**APT**) of Ross(1976,1977). The basic assumptions of this model are that security returns are generated by a small number of common factors plus an additional random component that can be diversified away in large portfolios and that capital markets are well-functioning in the sense that riskless zero net investment portfolios should earn zero profits. Since the theory does not require *a priori* specification of these sources of systematic risk, empirical implementation of the **APT** usually involves the implicit measurement of the common factors underlying security returns.

As a consequence, empirical studies of the **APT** have typically constructed basis or reference portfolios to mimic the factors. These constructed portfolios have been used for a variety of purposes such as testing the **APT** mean restriction, testing the significance of factor risk premia, evaluating the performance of managed portfolios, comparing the explanatory power of covariance measures of risk with other risk measures (such as own standard deviation), and providing the basis for exploratory efforts to determine the macroeconomic variables underlying asset pricing relations. In theory, these portfolios are

[1] See, for instance, Cannistraro(1973), Basu(1977), Litzenberger and Ramaswamy(1979), Banz(1981), Reinganum(1981). The *small firm effect* in particular has received much attention, including an entire issue in the *Journal of Financial Economics*.

supposed to be highly correlated with the common factors affecting security returns and to be relatively free of unsystematic risk. In practice, there is no guarantee that a particular basis portfolio construction procedure will mimic the factors sufficiently well. Obviously basis portfolios which are poorly correlated with the common factors can lead to incorrect inferences about the validity of the **APT** or the interpretation of its application to capital budgeting, performance evaluation, and macroeconomic activity. It is clearly important to know which basis portfolio construction procedures do a good job of mimicking the factors and which do not.

There is an embarrassingly large number of ways to construct such basis portfolios. There are not only several viable methods for forming these portfolios but also there are different procedures for estimating the factor models of security returns which underly these computations. In addition, the performance of these strategies might be expected to vary with the number of securities included and/or the number of factors being considered. Most previous empirical studies, for instance, have used Fama-MacBeth type portfolios coupled with maximum likelihood factor analysis of thirty to sixty securities. However, for reasons discussed below, the use of such a limited number of securities might fail to produce reliable reference portfolios to mimic the factors. Such concerns led Chen(1983) to use an instrumental variables procedure to provide inexpensive estimates of factor models using larger numbers of securities and sophisticated mathematical programming procedures to ensure that the basis portfolios were well-diversified. The less expensive instrumental variables procedures, which have also been advocated by Madansky and Marsh(1985), are less efficient in a statistical sense than maximum likelihood factor analysis and, hence, might yield basis portfolios which are not highly correlated with the underlying factors. Similarly, Chamberlain and Rothschild(1983) and Connor and Korajczyk(1984) have advocated the use of principal components as an inexpensive alternative to maximum likelihood factor analysis. For that matter, randomly selected well-diversified portfolios of large numbers of securities could, in principle, provide the least expensive potentially acceptable alternative.

Which of these strategies is the best one? This is an empirical question which has not been addressed in previous work. In this paper we intend to remedy this omission by providing a comprehensive examination of different basis portfolio formation strategies. In particular, we provide a detailed analysis of the performance of variants of all of the portfolio formation procedures and estimation methods that have been proposed in the literature, as well as an examination of the efficiency to be gained from considering portfolios of up to 750 securities. While it is doubtless possible to find some combination of estimation method and portfolio formation strategy which we have not considered, this study does provide a thorough examination of the major contenders.

The paper is organized as follows. The following section provides a brief review of the **APT**. The third and fourth sections contain a discussion of the two different steps involved in basis portfolio construction. The third section delineates the different portfolio formation procedures considered while the fourth describes the estimation methods used to generate the inputs for forming the reference portfolios. The comparison of the performance of alternative basis portfolio construction strategies is not merely a simple technical problem. Section V provides our solution to the problem of basis portfolio performance evaluation. In Section VI. we report on the results we obtained concerning the merits of different estimation methods, different portfolio formation procedures, and different numbers of securities. The final section provides some concluding remarks.

II. The Arbitrage Pricing Theory

The Arbitrage Pricing Theory (**APT**) of Ross(1976,1977) begins with the assumption that **K** common factors are the dominant sources of covariation among security returns and that other sources of risk impinging on security returns can be removed in large well-diversified portfolios. Formally, Ross assumed that these common factors affect security returns in a linear fashion and that securities returns are generated by the model:

$$\tilde{R}_{it} = E_i + \sum_{k=1}^{K} b_{ik}\tilde{\delta}_{kt} + \tilde{\epsilon}_{it} \tag{1}$$

$$\mathbf{E}[\tilde{\delta}_{kt}] = \mathbf{E}[\tilde{\epsilon}_{it}|\delta_{kt}] = 0$$

where:

$\tilde{R}_{it} \equiv$ Return on security i between time t-1 and time t for i=1,...,N

$E_i \equiv$ Expected return on security i

$\tilde{\delta}_{kt} \equiv$ Realization of the k^{th} common factor { i.e source of systematic risk } between time $t-1$ and t

$b_{ik} \equiv$ sensitivity of the return of security i to the k^{th} common factor { called the factor loading } and

$\tilde{\epsilon}_{it} \equiv$ the idiosyncratic or residual risk of the return on the i^{th} security between time $t-1$ and time t. These residual risks are assumed to have zero mean, finite variance and to be sufficiently independent across securities for a law of large numbers to apply.

How should expected returns be determined in a well-functioning capital market if security returns satisfy these assumptions and there are no taxes, transaction costs, or constraints on short sales? Ross argued that investors should be compensated only for bearing the systematic risk inherent in the **K** common factors since idiosyncratic risk can be virtually eliminated in large and well-diversified portfolios. Suppose we examine zero net investment portfolios and, in particular, the set of such portfolios which are constructed to be well-diversified and to contain no systematic risk. As the number of securities grows large, these portfolios will contain no risk at all and so should earn zero profits to prevent the occurrence of riskless arbitrage opportunities. Since, in these circumstances the number of such arbitrage portfolios tends toward infinity as well, Ross and many others proved that, in order to insure that these arbitrage portfolios do not earn positive profits, expected returns must satisfy (*approximately*):

$$E_i \approx \lambda_0 + b_{i1}\lambda_1 + ... + b_{ik}\lambda_k \tag{2}$$

where:

$\lambda_0 \equiv$ the intercept in the pricing relation and

$\lambda_k \equiv$ the risk premium on the k^{th} common factor, $k = 1, \dots, K$.

Obviously, for empirical purposes, it is desirable to treat equation (2) as an equality. In what follows, we assume sufficient regularity in the economy so that expected returns on the subset of risky securities we study (listed stocks on the New York and American Stock Exchanges) exactly satisfy the expected return condition (2).[2]

[2] Numerous investigators have examined the circumstances in which equation (2) holds as an equality in large economies and have provided explicit bounds on the deviations from (2) in finite economies. For the conditions needed in an infinite economy setting see, for instance, Chamberlain and Rothschild(1983), Connor(1984) and Shanken(1983). As for

Finally, it is important to distinguish between *riskless rate* and *zero beta* versions of the **APT**. As Ingersoll(1984) has emphasized, the distinction between the two does not involve the availability of riskless borrowing and lending, but rather depends on whether it is possible to form riskless (positive investment) portfolios from the subset of risky assets under consideration (as the number of assets in the subset tends toward infinity). If it is possible to construct such portfolios, then λ_0 in (2) is the riskless rate and a well-diversified portfolio which costs a dollar and contains no systematic risk should earn the riskless rate of return and should have a zero variance in the limit. If it is not possible to form such a portfolio, then λ_0 should be zero and one of the factors underlying security returns should correspond to a zero beta portfolio with identical factor loadings for all securities under appropriate transformation of the factor space.[3] This distinction will be important for evaluating the performance of alternative portfolio formation procedures and estimation methods.

III. Basis Portfolio Formation

What is the best way to construct portfolios which reflect the behavior of the common factors underlying the APT? At first blush, this would seem to be a statistical question that is best addressed by studying the assumed return generating process (1). Rewritten in more compact matrix notation, the model is:

$$\underline{\tilde{R}}_t = \underline{E} + B\underline{\tilde{\delta}}_t + \underline{\tilde{\epsilon}}_t \tag{3}$$

which is obtained by stacking equation (1) for $i = 1, \ldots, N$. We assume that the random factors $\underline{\tilde{\delta}}_t$ (a $K \times 1$ vector) and the corresponding elements of the factor loading matrix, $B(N \times K)$, have been normalized so that:[4]

the finite economy results. Grinblatt and Titman(1983), Chen and Ingersoll(1983), and Dybvig(1983) provide different settings in which the equilibrium deviations from equation (2) can be calculated.

[3] This would occur, for instance, if one of the common factors is unanticipated inflation and inflation is neutral such that all securities returns are equally affected by unexpected changes in prices. Formally, it is not possible to form a limiting riskless portfolio of risky assets when one of the eigenvectors of the covariance matrix of the countably infinite subset of security returns under consideration contains identical elements.

[4] The elements of B are not yet uniquely determined since for all orthogonal matrices T, any matrix $B^* = BT$ will yield the same return generating process. We will assume that

$$\mathbf{E}[\tilde{\underline{\delta}}_t] = \underline{0}$$
$$\mathbf{E}[\tilde{\underline{\delta}}_t \tilde{\underline{\delta}}_t'] = I \tag{4}$$

The random variables $\tilde{\underline{\epsilon}}_t$ are are assumed to satisfy:

$$\mathbf{E}[\tilde{\underline{\epsilon}}_t | \tilde{\underline{\delta}}_t] = \underline{0}$$
$$\mathbf{E}[\tilde{\underline{\epsilon}}_t \tilde{\underline{\epsilon}}_t' | \tilde{\underline{\delta}}_t] = \Omega \tag{5}$$

where Ω is a positive definite symmetric matrix.[5]

In these circumstances, a statistician's natural method for estimating $\tilde{\underline{\delta}}_t$, given knowledge of B, $\underline{E}$, Ω, and $\tilde{\underline{R}}_t$, is to employ the generalized least squares estimator :[6]

$$\hat{\underline{\delta}}_t^{GLS} = (B'\Omega^{-1}B)^{-1}B'\Omega^{-1}[\tilde{\underline{R}}_t - \underline{E}] \tag{6}$$

where the covariance matrix for the factor estimates is given by $[B'\Omega^{-1}B]^{-1}$. As is well known, this estimator has the desirable property of being the minimum variance linear unbiased estimator of $\tilde{\underline{\delta}}_t$. In addition as the number of assets grows large :

$$\underset{N\to\infty}{\text{plim}}\ \hat{\underline{\delta}}_t^{GLS} = \tilde{\underline{\delta}}_t \tag{7a}$$

since:

$$\lim_{N\to\infty} [B'\Omega^{-1}B]^{-1} = 0 \tag{7b}$$

It is clear from (7) that as the number of securities grows large this estimator will converge to $\tilde{\underline{\delta}}_t$ at a rate determined by the speed with which the largest eigenvalue of $(B'\Omega^{-1}B)^{-1}$ converges to zero. This, in turn, hinges on two factors: (a) the magnitudes of the variances and covariances of $\tilde{\underline{\epsilon}}_t$ and (b) the degree of dispersion among the responses of individual security returns to the common factors (e.g. the variances and covariances of the rows of B). In particular, if securities typically display similar responses to the common factors, we

---

the necessary $K(K-1)/2$ constraints required to ensure that $T = I$ have been imposed arbitrarily. For example, it is conventional in factor analysis to require $B'\Omega^{-1}B$ to be a diagonal matrix.

[5] When we assume that Ω is positive definite, we are implicitly assuming that no asset in the analysis contains *only* factor risk.

[6] The economic interpretation of (6) has been discussed by Ingersoll(1984) and Grinblatt and Titman(1983b).

would have to include many securities in the cross-section in order to estimate the factors with precision.

The importance of both of these factors deserves careful attention. Much of the recent **APT** literature has emphasized the ease of forming portfolios which mimic the common factors with negligible error as the number of securities in these portfolios tends toward infinity. This observation usually involves the intuition that idiosyncratic risk is likely to be virtually eliminated in portfolios of a moderate number of securities so long as the idiosyncratic disturbances are sufficiently independent.

What is seldom appreciated is the importance of factor (b) above—the dispersion of security responses to common factors. This is a well understood problem in a regression setting. Precise estimation of the covariance between the dependent and independent variables cannot be obtained if there is little variation in the independent variables over time. Similarly, examination of the covariation between individual security returns and their factor loadings cannot lead to accurate measurements of the underlying common factors unless there is sufficient dispersion among the factor loadings. Suppose, for example, that two of the common factors are unexpected changes in expected inflation and unanticipated inflation and that most security returns exhibit equal sensitivity to these two common factors (i.e. inflation has a neutral impact on most security returns). In particular, suppose that only five percent of the securities under consideration exhibit different responses to these two common factors. In this case, a much larger cross-section would be necessary to measure accurately these two common factors than would typically be needed merely to eliminate idiosyncratic risk.

In practice, the choice among basis portfolio construction methods is further complicated since B and, perhaps, Ω must be estimated. This introduces the usual problem of sampling error in the construction of the estimates and the possibility that alternative procedures using different portfolio formation techniques and differing numbers of securities may exhibit differing sensitivity to sampling error. For instance, since the precision of the estimate of $\underline{\tilde{\delta}}_t$ hinges on the estimates of B and Ω that are used, large numbers of securities may be required to mitigate the effects of sampling error. In addition, there is the problem introduced by the need to specify constraints on Ω in order to proceed with estimation. Two popular choices are the statistical factor analysis model, which assumes that Ω is diagonal, and the principal components model, which assumes that Ω is a diagonal matrix with equal variances. Different combinations of portfolio formation procedures and estimation methods might yield different results.

In what follows, we will consider four procedures for constructing basis portfolios. Two of the methods involve biased and unbiased versions of the generalized least squares estimator discussed above. We also consider biased estimators of $\underline{\tilde{\delta}}_t$ since there is no particular virtue associated with the unbiasedness of the generalized least squares estimator in the presence of measurement error. Hence, it may be desirable to seek a biased estimator with potentially lower variance in such circumstances. Such an estimator will be described below. The other two methods utilize a variant of the mathematical programming procedures employed by Chen(1983), which constrain the reference portfolios to be well-diversified. These procedures might mitigate some of the harmful effects created by sampling error and the imposition of constraints on Ω.

Before considering alternative portfolio formation procedures, it is useful to translate the statistical formulation of the generalized least squares estimator given above into the language of optimal portfolio construction after the fashion of Litzenberger and Ramaswamy(1979) and Rosenberg and Marathe(1979). In this language, the generalized least squares estimator provides what we usually refer to as Fama-MacBeth portfolios, after suitable rescaling so that the portfolios have unit net investment.[7] In our formulation, we choose the N portfolio weights $\underline{w}_j$ to mimic the j^{th} factor so that they:

$$\min_{\underline{w}_j} \underline{w}_j{}' D \underline{w}_j \tag{8a}$$

subject to:

$$\begin{aligned} \underline{w}_j' \underline{b}_k &= 0 \qquad \forall j \neq k \\ &= 1 \qquad j = k \end{aligned} \tag{8b}$$

where $\underline{b}_k$ is the k^{th} column of the sample factor loading matrix B and D is the diagonal matrix consisting of the sample variances of the idiosyncratic risk vector $\underline{\tilde{\epsilon}}_t$.[8] This portfolio

[7] Of course, Fama and MacBeth(1973) used the ordinary least squares estimator. The usage in the text is, however, common.

[8] Note that we are now ignoring off-diagonal elements of Ω such as industry effects. As a consequence, our procedures actually are better characterized as weighted least squares or diagonal generalized least squares.

provides the unbiased minimum idiosyncratic risk portfolio which mimics the j^{th} unobservable common factor. We rescale the portfolio weights $\underline{w}_j$ so that they sum to one (i.e. so that the portfolio costs one dollar) in order to maintain comparability with other basis portfolio formation procedures.

An alternative method, which produces what we term minimum idiosyncratic risk portfolios, involves choosing portfolio weights $\underline{w}_j$ so that:

$$\min_{\underline{w}_j} \underline{w}_j{}' D \underline{w}_j \tag{9a}$$

subject to:

$$\begin{aligned} \underline{w}_j' \underline{b}_k &= 0 \qquad \forall j \neq k \\ \underline{w}_j' \underline{\iota} &= 1 \end{aligned} \tag{9b}$$

where $\underline{\iota}$ is a vector of ones. This procedure should, in principle, produce minimum idiosyncratic risk portfolios whose fluctuations are proportional to the j^{th} common factor. In contradistinction to the unbiased GLS estimator, the proportionality factor need not equal one.[9] This is easily seen by comparing equations (8b) and (9b).

It is easy to distinguish these minimum idiosyncratic risk portfolios from the more familiar Fama-MacBeth portfolios in the one factor case. Assume for simplicity that the idiosyncratic variances are identical (i.e. $D = \sigma_\epsilon^2 I$). In this instance, the Fama-MacBeth portfolio solves the programming problem:

$$\min_{\underline{w}} \underline{w}' \underline{w} \tag{10a}$$

subject to:

$$\underline{w}' \underline{b} = 1 \tag{10b}$$

with solution:[10]

$$\underline{w} = (\underline{b}' \underline{b})^{-1} \underline{b} \tag{10c}$$

[9] This estimator can be computed as follows. Let $B = (\underline{b}_1 \underline{b}_2 \ldots \underline{b}_k)$ and suppose we are interested in mimicking the j^{th} factor. The minimum idiosyncratic risk estimator is $D^{-1} B^* [B^{*\prime} D^{-1} B^*]^{-1} \underline{e}_j$ where $B^* = (\underline{b}_1 \underline{b}_2 \ldots \underline{\iota} \ldots \underline{b}_k)$ and $\underline{\iota}$ is a vector of ones in the j^{th} column.

[10] Prior to rescaling so that the portfolio weights sum to one.

where $\underline{b}$ is the vector of sample factor loadings. Similarly, minimum idiosyncratic risk portfolios satisfy:

$$\min_{\underline{w}} \underline{w}'\underline{w} \tag{11a}$$

subject to:

$$\underline{w}'\underline{\iota} = 1 \tag{11b}$$

with the simple equally weighted solution:

$$\underline{w} = \frac{1}{N}\underline{\iota} \tag{11c}$$

Thus Fama-MacBeth portfolio weights are proportional to the sample factor loadings (i.e. the betas) of the individual securities and, as a consequence, take advantage of the differing information content of individual securities regarding the fluctuations in the common factor. Minimum idiosyncratic risk portfolios, however, are merely well diversified and do not take explicit advantage of such information. Note that the factor loading of the minimum idiosyncratic portfolio is the average beta of the securities (i.e. $\bar{b} = \frac{1}{N}\underline{b}'\underline{\iota}$) while that of the Fama-MacBeth portfolio is unity prior to rescaling.

The second thing to note is that the diversification properties of the Fama-MacBeth portfolios depend on the normalization of the common factors. If the factors are normalized so that factor loadings are typically close to one,[11] the two procedures both will yield portfolio weights of order $1/N$.[12] However when average factor loadings are on the order of .001 to .0001, as in the case of typical factor loading estimates from daily data when the factor variance is normalized to unity as in (4), the minimum idiosyncratic risk procedure still will yield small weights while the Fama-MacBeth method will produce very large portfolio weights in finite cross-sections. This does not present a problem when the factor loadings and idiosyncratic variances are measured without error but it is a potentially serious source of difficulty when large factor loadings can reflect measurement error as

[11] It is worth noting that no study we are aware of normalizes the factors to ensure that the typical loading is unity.

[12] The scaling of the loadings so that the natural loading is one can be accomplished with the following transformation. Transform B so that $B'D^{-1}B$ is a diagonal matrix and denote the i^{th} diagonal element as γ_i. Let the vector $\underline{\varsigma} = B'D^{-1}\underline{\iota}$. Then the transformation $B^* = BA$ where A is a diagonal matrix with ς_i/γ_i along the diagonal yields B^*, a normalization of B so that the typical portfolio loading is one.

well as responsiveness to common factors. We turn now to a more complete discussion of the effect of measurement error on the performance of basis portfolios constructed from estimated factor loadings.

The comparative merits of minimum idiosyncratic risk and Fama-MacBeth portfolios in the presence of measurement error can be investigated more fully by again considering the one factor model:

$$\underline{\tilde{R}}_t = \underline{\beta}\tilde{R}_{mt} + \underline{\tilde{\epsilon}}_t \tag{12}$$

where $\underline{\beta}$ is the $N \times 1$ vector of true factor loadings and $\underline{\tilde{\epsilon}}_t$ is the vector of idiosyncratic disturbances which is assumed, for simplicity, to have elements with zero means, common variances σ_ϵ^2, and independently distributed of one another.[13] Suppose that we measure the factor loadings $\underline{\beta}$ with error:

$$\underline{b} = \underline{\beta} + \underline{v} \tag{13}$$

where $\underline{v}$ is an $N \times 1$ vector of the deviations of the true factor loadings from their sample values. $\underline{v}$ is independent of $\tilde{R}_{mt}$ and $\underline{\tilde{\epsilon}}_t \quad \forall t$, and $\underline{v}$ satisfies:[14]

$$\begin{aligned} \mathbf{E}[\underline{v}] &= \underline{0} \\ \mathbf{E}[\underline{v}\underline{v}'] &= \Sigma_v \end{aligned} \tag{14}$$

We also normalize the estimated factor loadings $\underline{b}$ so that $\underline{b}'\underline{b} = \underline{b}'\underline{\iota}$. Under this normalization, the Fama-MacBeth procedure will yield a basis portfolio with a sample loading of unity and whose weights will sum to one.[15] Finally we assume that the cross-section is sufficiently large so that $\bar{b} \approx \bar{\beta}$ [i.e. $\frac{1}{N}\sum_{i=1}^N v_i \approx 0$] to simplify the arithmetic.

[13] This means that we are implicitly ignoring the impact of measurement error in the disturbance variances D.

[14] We assume that our estimates are unbiased for simplicity. The assumption that $\underline{v}$ is independent of $\underline{\tilde{\epsilon}}_t$ and R_{mt} is less innocuous since we typically estimate factor loadings during the same period that we form returns on the basis portfolios. Accounting for such problems would complicate the analysis considerably and would not alter the basic insights gained in the present exercise. For example, if b_i were estimated from an ordinary least squares regression, then $\mathrm{cov}[v_i, \tilde{\epsilon}_t] = \frac{\bar{R}_m \sigma_\epsilon^2}{\sum_{t=1}^T R_{mt}^2} \approx \frac{\bar{R}_m \sigma_\epsilon^2}{T[\sigma_m^2 + \bar{R}_m^2]}$ which would typically be trivial in moderately large samples. In addition, $\mathrm{cov}[v_i, R_{mt}] = 0$ under the assumption that $\underline{\tilde{\epsilon}}_t$ and R_{mt} are independent since $\mathrm{cov}[R_{mt}^2, \underline{\tilde{\epsilon}}_t] = 0$.

[15] Note that this implies that $\bar{b} = \sum_{i=1}^N b_i/N$ will satisfy $0 < \bar{b} < 1$ to ensure that the sample variance of the b_i's will be positive.

In this setting it is easy to evaluate the behavior of the Fama-MacBeth portfolio and its relationship to returns on the common factor. The portfolio returns are:

$$
\begin{aligned}
\tilde{R}_t^{FM} &= (\underline{b}'\underline{b})^{-1}\underline{b}'[\underline{\beta}\tilde{R}_{mt} + \underline{\tilde{\epsilon}}_t] \\
&= (\underline{b}'\underline{b})^{-1}[\underline{\beta}'\underline{\beta}\tilde{R}_{mt} + \underline{v}'\underline{\beta}\tilde{R}_{mt} + \underline{b}'\underline{\tilde{\epsilon}}_t] \\
&\approx \frac{1}{N\overline{\beta}}[\underline{\beta}'\underline{\beta}\tilde{R}_{mt} + \underline{v}'\underline{\beta}\tilde{R}_{mt} + \underline{\beta}'\underline{\tilde{\epsilon}}_t + \underline{v}'\underline{\tilde{\epsilon}}_t]
\end{aligned}
\tag{15}
$$

where the approximation arises from the assumption that $\overline{b} \approx \overline{\beta}$, implying that $N\overline{\beta} \approx \underline{b}'\underline{b}$. The mean and variance of the Fama-MacBeth portfolio as well as its squared correlation with the common factor are given by:

$$
\begin{aligned}
\mathbf{E}[\tilde{R}_t^{FM}] &\approx \frac{1}{N\overline{\beta}}\underline{\beta}'\underline{\beta}\overline{R}_m \\
&\approx \frac{\underline{\beta}'\underline{\beta}}{\underline{b}'\underline{b}}\overline{R}_m \\
\mathbf{Var}[\tilde{R}_t^{FM}] &\approx \frac{1}{N^2\overline{\beta}^2}\mathbf{E}[\underline{\beta}'\underline{\beta}(\tilde{R}_{mt} - \overline{R}_m) + \underline{v}'\underline{\beta}\tilde{R}_{mt} + \underline{\beta}'\underline{\tilde{\epsilon}}_t + \underline{v}'\underline{\tilde{\epsilon}}_t]^2 \\
&\approx \frac{1}{(\underline{b}'\underline{b})^2}[(\underline{\beta}'\underline{\beta})^2\sigma_m^2 + \underline{\beta}'\Sigma_v\underline{\beta}(\sigma_m^2 + \overline{R}_m^2) + \underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \mathbf{tr}(\Sigma_v)\sigma_\epsilon^2] \\
\mathbf{Corr}(\tilde{R}_t^{FM}, \tilde{R}_{mt})^2 &\approx \frac{\frac{1}{(N\overline{\beta})^2}(\underline{\beta}'\underline{\beta})^2\sigma_m^4}{\frac{\sigma_m^2}{N^2\overline{\beta}^2}[(\underline{\beta}'\underline{\beta})^2\sigma_m^2 + \underline{\beta}'\Sigma_v\underline{\beta}(\sigma_m^2 + \overline{R}_m^2) + \underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \mathbf{tr}(\Sigma_v)\sigma_\epsilon^2]} \\
&\approx \frac{(\underline{\beta}'\underline{\beta})^2\sigma_m^2}{(\underline{\beta}'\underline{\beta})^2\sigma_m^2 + \underline{\beta}'\Sigma_v\underline{\beta}(\sigma_m^2 + \overline{R}_m^2) + \underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \mathbf{tr}(\Sigma_v)\sigma_\epsilon^2}
\end{aligned}
\tag{16}
$$

where $\overline{R}_m$ is the mean return on the common factor and σ_m^2 is its variance. In the special case where $\underline{\beta}$ is measured without error [i.e. $\underline{v} = \underline{0}$], the squared correlation reduces to:

$$
\mathbf{Corr}(\tilde{R}_t^{FM}, \tilde{R}_{mt})^2 = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_\epsilon^2/(\underline{\beta}'\underline{\beta})}
\tag{17}
$$

It is easier to evaluate the corresponding quantities for the minimum idiosyncratic risk procedure since the portfolios do not involve the measurement error in the loadings. Hence the portfolio returns are simply:

$$
\begin{aligned}
\tilde{R}_t^{MIRP} &= \frac{1}{N}\sum_{i=1}^{N}(\beta_i\tilde{R}_{mt} + \tilde{\epsilon}_{it}) \\
&= \overline{\beta}\tilde{R}_{mt} + \overline{\epsilon}_t
\end{aligned}
\tag{18}
$$

where $\bar{\epsilon}_t$ is the cross-sectional mean of the idiosyncratic risks at time t. Similarly, the corresponding moments of these portfolio returns are:

$$\begin{aligned}\mathbf{E}[\tilde{R}_t^{MIRP}] &= \overline{\beta}\overline{R}_m \\ \mathbf{Var}[\tilde{R}_t^{MIRP}] &= \overline{\beta}^2\sigma_m^2 + \sigma_\epsilon^2/N \\ \mathbf{Corr}(\tilde{R}_t^{MIRP}, \tilde{R}_{mt})^2 &= \frac{\overline{\beta}^2\sigma_m^2}{\overline{\beta}^2\sigma_m^2 + \sigma_\epsilon^2/N}\end{aligned} \tag{19}$$

What are we to make of this tedious arithmetic? Consider first the case where the $\underline{\beta}$ is measured without error. Then the squared correlation of the Fama-MacBeth portfolio with the common factor is larger than that of the minimum idiosyncratic risk procedure since:

$$\frac{\sigma_\epsilon^2}{N\overline{\beta}^2} > \frac{\sigma_\epsilon^2}{\underline{\beta}'\underline{\beta}} \tag{20}$$

so long as all of the elements of $\underline{\beta}$ are not identical [i.e. $\underline{\beta}'\underline{\beta} = \sum_{i=1}^N(\beta_i - \overline{\beta})^2 + N\overline{\beta}^2 > N\overline{\beta}^2$ unless $\beta_i = \overline{\beta} \quad \forall i$]. The same cannot be said when we consider the impact of measurement error on the Fama-MacBeth portfolio. The squared correlation of returns on the minimum idiosyncratic risk portfolio with the common factor is unaffected by the presence of measurement error in $\underline{b}$ while that of the Fama-MacBeth portfolio falls in this eventuality.

For example, consider the special case where $\underline{b}$ is estimated by ordinary least squares regression on $\tilde{R}_{mt}$. As a consequence, the measurement error covariance matrix will be (approximately):

$$\Sigma_v \approx \frac{\sigma_\epsilon^2}{T(\sigma_m^2 + \overline{R}_m^2)}\mathbf{I}_N \tag{21}$$

where T is the sample size, and $\mathbf{I}_N$ is an $N \times N$ identity matrix, and the approximation arises because we replaced the sample mean and variance of $\tilde{R}_{mt}$ with their population values in (21) for ease of exposition. The squared correlation of the Fama-MacBeth portfolio with the common factor is then:

$$\mathbf{Corr}(\tilde{R}_t^{FM}, \tilde{R}_{mt})^2 \approx \frac{(\underline{\beta}'\underline{\beta})^2\sigma_m^2}{(\underline{\beta}'\underline{\beta})^2\sigma_m^2 + \frac{1}{T}\underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \frac{N}{T}\frac{\sigma_\epsilon^4}{(\sigma_m^2 + \overline{R}_m^2)}} \tag{22}$$

This will be smaller than that of the minimum idiosyncratic risk portfolio when:

$$\frac{\sigma_\epsilon^2}{N\overline{\beta}^2} < \sigma_\epsilon^2[\frac{1}{\underline{\beta}'\underline{\beta}} + \frac{1}{T\underline{\beta}'\underline{\beta}} + \frac{N}{T}\frac{\sigma_\epsilon^2}{(\sigma_m^2 + \overline{R}_m^2)(\underline{\beta}'\underline{\beta})^2}] \tag{23}$$

which simplifies to:

$$\frac{\sigma_\beta^2}{\overline{\beta}^2} < \frac{1}{T} + \frac{1}{T}\frac{\sigma_\epsilon^2}{(\sigma_m^2 + \overline{R}_m^2)(\sigma_\beta^2 + \overline{\beta}^2)} \tag{24}$$

where $\sigma_\beta^2 = \frac{1}{N}\sum_{i=1}^N(\beta_i - \overline{\beta})^2$ is the sample variance of the true loadings. This inequality can be easily obtained when the sampling variation in the betas is small or, equivalently (under the present normalization of the loadings) when $\overline{\beta}$ is close to one in moderate sized samples.[16]

While both of these procedures will produce well-diversified basis portfolios in the limit (or in finite cross-sections as in the preceding example), they may not produce such portfolios with a finite cross-section of securities. This possibility led Chen(1983) to employ mathematical programming methods to produce portfolios which were well-diversified and, in principle, highly correlated with only one factor. Large well-diversified portfolios possess minimal idiosyncratic risk and, perhaps, might suffer only marginally from errors in estimating the factor loadings and idiosyncratic variances. The actual procedure employed by Chen(1983) is not in the public domain, being the proprietary software of Glenn Graves of UCLA, and hence was not available to us for this study. Instead, we chose a simple and similar alternative: quadratic programming subject to fixed upper and lower bounds. This involves:

$$\min_{\underline{w}_j} \underline{w}_j{}' D \underline{w}_j \tag{25}$$

subject to:

$$\begin{aligned} \underline{w}_j'\underline{b}_k &= 0 \qquad \forall j \neq k \\ \underline{w}_j'\underline{\iota} &= 1 \\ l_i \leq w_{ij} &\leq d_i \quad i = 1, \ldots, N \end{aligned} \tag{26}$$

where l_i are the fixed lower bounds we place on the portfolio weights, d_i are the corresponding upper bound constraints, and the remaining variables are as defined above. We examined two choices for these upper and lower bounds. Following Chen(1983), we produced portfolios with non-negative weights (i.e. where l_i was equal to zero) which could

[16] Tedious manipulation of (24) coupled with some minor approximations yields the result that (24) will occur when $\overline{b} > T\sigma_b^2$ where σ_b^2 is the sample variance of the sample loadings although this condition is not necessary. It is suggestive to note that that the sample variance of betas computed with respect to the usual market proxies is quite small.

take on a maximum value of one to two per cent. We also studied the properties of portfolios which were merely constrained to be well-diversified with the portfolio weights taking on maximum and minimum values of plus and minus one or two per cent. While this procedure differs from Chen's, which had flexible rather than fixed upper and lower bounds, we felt that this procedure probably would produce similar results and, hence, would provide reliable evidence on the comparative merits of mathematical programming procedures. These two choices for the upper and lower bound constraints constitute the other two portfolio formation procedures whose performance we examined and report below.[17]

Finally, a note is in order concerning the computation of portfolios whose returns are orthogonal to those associated with the common factors. As noted in the preceding section, such portfolios should, in principle, have the same returns as the riskless portfolio of risky assets postulated by the *riskless rate* version of the APT. For each reference portfolio formation method, we constructed minimum idiosyncratic risk portfolios using that method which had weights orthogonal to B and which cost one dollar. Similarly, the positive net investment quadratic programming portfolios have orthogonal portfolios with nonnegative weights which are orthogonal to B that cost one dollar and have minimum idiosyncratic risk subject to this constraint. Finally, the well-diversified quadratic programming portfolios have orthogonal portfolios which have the same properties except that the portfolio weights are well-diversified instead of nonnegative. These portfolios are used to construct the excess return basis portfolios analyzed below.

The differences in the excess return portfolios is again best illustrated in the one factor case. We will again assume for simplicity that the idiosyncratic variances are identical (i.e. $D = \sigma_\epsilon^2 I$). The required orthogonal portfolio solves the programming problem:

$$\min_{\underline{w}_{rf}} \underline{w}_{rf}{}' \underline{w}_{rf} \tag{27}$$

subject to:

$$\begin{aligned} \underline{w}'_{rf}\underline{\beta} &= 0 \\ \underline{w}'_{rf}\underline{\iota} &= 1 \end{aligned} \tag{28}$$

[17] Additional experimentation was done with weights ranging in absolute value up to five percent. The results, however, do not materially differ from the ones presented below and hence are not reported.

with solution:

$$\underline{w}_{rf} = \frac{1}{N\underline{\beta}'\underline{\beta} - (\underline{\iota}'\underline{\beta})^2}[(\underline{\beta}'\underline{\beta})\underline{\iota} - (\underline{\iota}'\underline{\beta})\underline{\beta}] \tag{29}$$

The minimum idiosyncratic risk excess return portfolio weights can now be obtained by subtracting these weights from $\frac{1}{N}$ (the portfolio weights of the minimum idiosyncratic risk portfolio which mimics the factor) which yields:

$$\begin{aligned} \underline{w}_{MIRP} - \underline{w}_{rf} &= \frac{1}{N}\underline{\iota} - \frac{1}{N\underline{\beta}'\underline{\beta} - (\underline{\iota}'\underline{\beta})^2}[(\underline{\beta}'\underline{\beta})\underline{\iota} - (\underline{\iota}'\underline{\beta})\underline{\beta}] \\ &= \frac{\overline{\beta}}{\underline{\beta}'\underline{\beta} - N\overline{\beta}^2}[\underline{\beta} - \overline{\beta}\underline{\iota}] \end{aligned} \tag{30}$$

The Fama-MacBeth excess return portfolio solves the programming problem:

$$\min_{\underline{w}^{ex}_{FM}} \underline{w}^{ex\,\prime}_{FM}\underline{w}^{ex}_{FM} \tag{31}$$

subject to:

$$\begin{aligned} \underline{w}'_{FM}\underline{\beta} &= 1 \\ \underline{w}'_{FM}\underline{\iota} &= 0 \end{aligned} \tag{32}$$

with solution:

$$\underline{w}^{ex}_{FM} = \underline{w}_{FM} - \underline{w}_{rf} = \frac{1}{\underline{\beta}'\underline{\beta} - N\overline{\beta}^2}[\underline{\beta} - \overline{\beta}\underline{\iota}] \tag{33}$$

These tedious manipulations yield one important insight—the minimum idiosyncratic risk procedure produces weights for excess return portfolios which are proportional to the corresponding output from the Fama-MacBeth procedure. Not surprisingly, the factor of proportionality is the average factor loading. Once again, when the average factor loading is typically much less than one, the portfolio weights produced by the Fama-MacBeth procedure will take on very large positive and negative values. By contrast, the minimum idiosyncratic procedure yields a well-diversified excess return portfolio. The same result arises in the multiple factor case.

IV. Estimation Methods

In this section, we describe four methods for estimating the factor loadings and idiosyncratic variances underlying the APT. The choice among estimation methods involves different tradeoffs than the choice among basis portfolio formation methods. Here, the comparison is between statistically efficient but computationally costly methods such as factor analytic techniques, and less efficient but less costly methods such as instrumental variables or principal components. It is obviously of greater than academic interest whether the comparatively inefficient methods provide performance comparable to that produced by the computationally burdensome efficient estimation methods. This could occur because of the large cross-sections of security returns that we employ or because of good small sample properties of the comparatively inefficient estimation methods.

Recall that the return generating process is:

$$\underline{\tilde{R}}_t = \underline{E} + B\underline{\tilde{\delta}}_t + \underline{\tilde{\epsilon}}_t \tag{34}$$

where the idiosyncratic risks $\underline{\tilde{\epsilon}}_t$ have zero means and covariance matrix Ω. For convenience, we will work with returns expressed as deviations from their respective means:

$$\underline{\tilde{r}}_t = \underline{\tilde{R}}_t - \underline{E} = B\underline{\tilde{\delta}}_t + \underline{\tilde{\epsilon}}_t \tag{35}$$

Under the assumption of joint normality of $\underline{\tilde{r}}_t$ and $\underline{\tilde{\delta}}_t$, the sample covariance matrix :

$$S = \frac{1}{T}\sum_{t=1}^{T} \underline{\tilde{r}}_t \underline{\tilde{r}}_t' \tag{36}$$

follows a Wishart distribution which serves as the basis of the log likelihood function:

$$\begin{aligned} \mathcal{L}(\Sigma|S) &= \frac{-NT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{1}{2}\sum_{t=1}^{T}(\underline{\tilde{R}}_t - \overline{\underline{R}})'\Sigma^{-1}(\underline{\tilde{R}}_t - \overline{\underline{R}}) \\ &= \frac{-NT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{T}{2}\mathbf{trace}(S\Sigma^{-1}) \end{aligned} \tag{37}$$

where:

$$\begin{aligned} \Sigma &= \mathbf{E}[\underline{\tilde{r}}_t \underline{\tilde{r}}_t'] \\ &= BB' + \Omega \end{aligned} \tag{38}$$

Unfortunately, it is not possible to proceed with estimation of B and Ω when security returns possess an approximate factor structure without specifying further constraints on Ω. One popular choice is the statistical factor analysis model where the residual covariance matrix Ω is assumed to be a diagonal matrix D. Under this additional assumption, the model for the return covariance matrix Σ is:

$$\Sigma = BB' + D \tag{39}$$

It is now conceptually simple and computationally costly to maximize the log likelihood (37) subject to (39) by setting the derivatives equal to zero:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial B} &\propto \Sigma^{-1}[\Sigma - S]\Sigma^{-1}B = 0 \\
\frac{\partial \mathcal{L}}{\partial D} &\propto \mathrm{Diag}[\Sigma^{-1}(\Sigma - S)\Sigma^{-1}] = 0
\end{aligned}
\tag{40}
$$

where $\mathrm{Diag}[X]$ is a diagonal matrix formed from the diagonal elements of X. The values of B and D which solve equations (40) are the required maximum likelihood estimates. When the number of securities under consideration is large, it is impractical to obtain these estimates by iteratively solving the likelihood equations (40) and so we employed a significantly cheaper alternative: the **EM** algorithm of Dempster, Laird, and Rubin(1977). This procedure, which is described in considerable detail in Lehmann and Modest(1985a), maximizes (37) subject to the constraints (39) using an iterative multivariate regression procedure. Its principal virtue is that it is inexpensive, both in storage requirements and in computational cost.

Maximum likelihood factor analysis provides, in principle, efficient estimates of the factor loadings and idiosyncratic variances. However, if the **APT** is true, there is information in the vector of sample mean security returns concerning the values of the factor loadings. This is because the **APT** implies that expected security returns are linear combinations of the product of their factor loadings and the factor risk premia. Consequently, the sample mean security returns should on average reflect the magnitudes of the factor loadings. In order to exploit this information, we also performed maximum likelihood factor analysis subject to the constraint that expected security returns are spanned by their factor loadings and the factor risk premia. This involves maximizing the log likelihood function:

$$
\begin{aligned}
\mathcal{L}(\Sigma|S) = \frac{-NT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma| - \frac{T}{2}\mathbf{trace}(S\Sigma^{-1}) \\
- \frac{T}{2}(\overline{\underline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})'\Sigma^{-1}(\overline{\underline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})
\end{aligned}
\tag{41}
$$

subject to the constraints:

$$\Sigma = BB' + D \tag{42}$$

where (41) follows from substituting $\underline{\iota}\lambda_0 + B\underline{\lambda}$ for $\overline{\underline{R}}$ in equation (37). Note that this involves maximizing the original log likelihood function (37) plus an additional term involving the weighted average of the deviation of sample mean security returns from the product of the factor loadings and the corresponding risk premia. It is analogous to the maximum likelihood estimation of the zero beta **CAPM** employed, for example, by Gibbons(1982) and Stambaugh(1982) and analyzed in considerable detail by Shanken(1984). The maximum likelihood estimates of the relevant parameters may be obtained by setting the derivatives equal to zero:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \lambda_0} &\propto \underline{\iota}'\Sigma^{-1}(\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda}) = 0 \\
\frac{\partial \mathcal{L}}{\partial \underline{\lambda}} &\propto B'\Sigma^{-1}(\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda}) = 0 \\
\frac{\partial \mathcal{L}}{\partial B} &\propto \Sigma^{-1}[\Sigma - S - (\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})(\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})']\Sigma^{-1}B + \Sigma^{-1}(\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})\underline{\lambda}' = 0 \\
\frac{\partial \mathcal{L}}{\partial D} &\propto \mathrm{Diag}[\Sigma^{-1}(\Sigma - S - (\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})(\underline{\overline{R}} - \underline{\iota}\lambda_0 - B\underline{\lambda})')\Sigma^{-1}] = 0
\end{aligned}
\tag{43}
$$

where as above Diag$[X]$ is a diagonal matrix formed from the diagonal elements of X. It is even less practical to obtain these estimates by solving equations (43). As a consequence, we employ a variant of the EM algorithm in order to obtain these restricted maximum likelihood estimates.

For all of the putative virtues of these theoretically efficient estimation procedures, they do have one significant disadvantage: their cost. In consequence, it seems reasonable to try less costly procedures and hope for only a small loss in efficiency. Chamberlain and Rothschild(1983) and Connor and Korajczyk(1984) have recently advocated the use of principal components as an inexpensive alternative to maximum likelihood factor analysis. Chamberlain and Rothschild(1983) showed that, as the number of securities being analyzed tends toward infinity, the first K eigenvectors obtained from the spectral decomposition of the true covariance matrix of security returns converge to the factor loadings underlying security returns. Connor and Korajczyk(1984) showed that this holds for the sample covariance matrix as well. The one-time extraction of eigenvalues and eigenvectors is roughly as costly as maximum likelihood factor analysis using the EM algorithm in daily data. Since only one principal components run is required to estimate factor models with different numbers of factors, this is certainly a potentially attractive alternative.

The link between maximum likelihood factor analysis and principal components with a finite sample of data is quite simple: principal components is equivalent to maximum likelihood factor analysis when the idiosyncratic variances are assumed to be identical, i.e. when:

$$D = \sigma^2 I \tag{44}$$

This observation highlights the intuitive distinction between factor analysis and principal components—factor analysis provides weighted least squares estimates of the factors and factor loadings (where the weights are the estimated idiosyncratic variances) while principle components provides the corresponding ordinary least squares estimates. Hence the factor analysis model will perform comparatively better, the greater is the cross-sectional variability of the idiosyncratic variances since principal components ignores any information imbedded in these variances.[18]

The estimates of the factor loadings provided by principal components can be obtained by maximizing the likelihood function (37) subject to the constraints (44) which involves solving the equations (40) iteratively. Needless to say there is a more cost effective way to obtain these estimates. Instead, we employ the singular value decomposition algorithm of the NAG Subroutine Library to obtain the required eigenvalues and eigenvectors. Each column of the matrix of eigenvectors was multiplied by the square root of the corresponding eigenvalue in order to scale the factors to have unit variance. Estimates of the idiosyncratic variances were then obtained by solving equation (39) for the required estimates of D by substituting the transformed eigenvectors for B and the sample covariance matrix S for Σ.

Finally, another inexpensive alternative to maximum likelihood factor analysis is the instrumental variables estimator. Instrumental variables estimators have recently been employed by Chen(1983) and Madansky and Marsh(1985). The basic idea of these estimators is quite simple: substitute consistent estimates of the factors $\underline{\tilde{\delta}}_t$ for the factors themselves in equation (35) and then estimate the factor loadings B by the ordinary least squares regression of individual security returns on the estimates of the factors.[19] Chen(1983)

[18] Under the assumption of an approximate factor structure, principal components provides consistent estimates of the factors as $N \to \infty$ and consistent estimates of the factor loadings as $T \to \infty$ even when the returns are not normally distributed. In these circumstances, maximum likelihood factor analysis will provide consistent estimates of the factors and factor loadings as well.

[19] The application of instrumental variables methods to factor analysis models typically

used portfolios formed by mathematical programming based on maximum likelihood factor analysis of 180 securities as the required consistent estimates of the factors. Following Madansky(1964) and Hagglund(1982), we employ a simpler instrumental variables procedure that does not require a preliminary maximum likelihood factor analysis.

Suppose that we normalize the factors so that the factor loadings of the first K securities are the identity matrix. Note that this change implies that the new factors are correlated. In terms of the original representation (34) and the normalization of the factors (4), we have rotated the factors so that they have covariance matrix $\Phi = B_1 B_1'$ where B_1 is the matrix of factor loadings on the first K securities in the original model. Letting $\tilde{\underline{r}}_{1t}$ denote the returns, expressed as deviations from their respective means, on the first K securities, we have:

$$\tilde{\underline{r}}_{1t} = \tilde{\underline{\delta}}_t + \tilde{\underline{\epsilon}}_{1t} \tag{45}$$

where $\tilde{\underline{\epsilon}}_{1t}$ is the vector of residual error terms associated with $\tilde{\underline{r}}_{1t}$. Similarly, letting $\tilde{\underline{r}}_{3t}$ denote the vector of demeaned returns on the last $N-K-1$ securities, we have:

$$\tilde{\underline{r}}_{3t} = \Gamma_3 \tilde{\underline{\delta}}_t + \tilde{\underline{\epsilon}}_{3t} \tag{46}$$

where Γ_3 is the matrix of factor loadings of these securities and $\tilde{\underline{\epsilon}}_{3t}$ are the corresponding idiosyncratic error terms. Finally, the equation for the $K+1^{st}$ security, $\tilde{r}_{2t}$ is:

$$\tilde{r}_{2t} = \Gamma_2' \tilde{\underline{\delta}}_t + \tilde{\epsilon}_{2t} \tag{47}$$

where Γ_2 is the vector of factor loadings of the $K+1^{st}$ security and $\tilde{\epsilon}_{2t}$ is its residual term.

Consider the regression of $\tilde{r}_{2t}$ on $\tilde{\underline{r}}_{1t}$:

$$\begin{aligned} \tilde{r}_{2t} &= \Gamma_2' \tilde{\underline{r}}_{1t} + u_{2t} \\ &= \Gamma_2' \tilde{\underline{r}}_{1t} + (\tilde{\epsilon}_{2t} - \Gamma_2' \tilde{\underline{\epsilon}}_{1t}) \end{aligned} \tag{48}$$

Clearly application of ordinary least squares to this equation will lead to biased and inconsistent estimates of Γ_2 since $\tilde{\underline{r}}_{1t}$ is correlated with its own idiosyncratic disturbance term

involves the assumption that the idiosyncratic disturbances are independent (as in the statistical factor analysis model). The procedures, however, will provide consistent estimates even when the idiosyncratic disturbances are correlated so long as the disturbances are sufficiently independent for a law of large numbers to apply as the number of securities tends toward infinity.

$\tilde{\epsilon}_{1t}$. If instead we first regress $\tilde{\underline{r}}_{1t}$ on $\tilde{\underline{r}}_{3t}$:

$$\tilde{\underline{r}}_{1t} = \Pi \tilde{\underline{r}}_{3t} + \underline{u}_{3t} \tag{49}$$

and then replace $\tilde{\underline{r}}_{1t}$ with the fitted values from this regression in equation (48), ordinary least squares can then be used to estimate Γ_2 consistently. This estimate is consistent because the errors in estimating the common factors by using the fitted values from the regression (49) involve only the idiosyncratic disturbances $\tilde{\epsilon}_{3t}$ which are uncorrelated with the idiosyncratic disturbance $\tilde{\epsilon}_{2t}$ by assumption. Repeated application of this procedure replacing $\tilde{r}_{2t}$ with an element of $\tilde{\underline{r}}_{3t}$ leads to the corresponding estimates of Γ_3. Finally, solution of the matrix equations:

$$\begin{aligned} \Gamma'(S - \Gamma\Phi\Gamma' - D)\Gamma &= 0 \\ \mathrm{Diag}[S - \Gamma\Phi\Gamma' - D] &= 0 \end{aligned} \tag{50}$$

produces estimates of the factor covariance matrix Φ and the matrix of idiosyncratic variances D. The estimate of the factor covariance matrix Φ can be used to transform the factor loading estimates so that the factors are again rescaled to be uncorrelated and have unit variances.

V. Basis Portfolio Comparison

If it were possible to observe the factors underlying security returns, it would be a simple statistical problem to determine which combination of basis portfolio formation procedure and estimation method produced the best reference portfolios. Of course, if we observed the common factors, we would not need to construct basis portfolios to test the **APT** or for use in performance evaluation. Since we do not assume that we have sufficient prescience to identify and measure the factors underlying security returns, the problem of determining which basis portfolios perform best remains.

One non-rigorous approach to comparing these different basis portfolios is to examine the behavior of the basis portfolio weights and the sample means and variances of their returns in order to check whether the results appear to be reasonable. For example, the constructed reference portfolios ought to be well-diversified if they are to mimic the factors with minimal idiosyncratic risk. Although the quadratic programming portfolios are well-diversified by construction, nothing in the Fama-MacBeth or minimum idiosyncratic risk procedures guarantees that the resulting portfolios will be well-diversified. Similarly, examination of the usual summary statistics describing the sample behavior of the constructed reference portfolios can reveal peculiarities in their performance. For example, we have examined basis portfolios which had mean returns as high as 120 percent and standard deviations as great as 450 percent *per month*. Clearly, such behavior is not likely to reflect the performance of good reference portfolios.

Unfortunately, searching for *reasonable* reference portfolios is not likely to eliminate many candidates and hardly constitutes a scientific testing procedure. Fortunately, it is possible to make reasonable assumptions which lead to more scientific comparisons. In particular, suppose that the **APT** is exactly true, i.e. equation (2) holds exactly. Assume that the underlying universe of securities under consideration is sufficiently large so that there exist portfolios whose returns $\underline{\tilde{R}}_{mt}$ are perfectly correlated with the unobservable common factors. If the riskless rate version of the **APT** is true, then security returns satisfy:

$$\underline{\tilde{R}}_t - \underline{\iota} R_{ft} = B(\underline{\tilde{R}}_{mt} - \underline{\iota} R_{ft}) + \underline{\epsilon}_t \tag{51}$$

where R_{ft} is the return on the limiting riskless portfolio of risky assets. If the zero beta version of the **APT** is true, then security returns satisfy:

$$\underline{\tilde{R}}_t = B\underline{\tilde{R}}_{mt} + \underline{\epsilon}_t \tag{52}$$

since the zero beta portfolio corresponds to one of the common factors underlying security returns.

Consider the behavior of a set of **K** basis portfolios whose returns $\underline{\tilde{R}}_{pt}$ are excess returns over the riskless rate when that version of the **APT** is appropriate and are the relevant raw returns in the zero beta case. Similarly, let $\underline{\tilde{R}}_{mt}$ also denote excess or raw returns on the true reference portfolios where appropriate. Then the returns on the basis portfolios can be compactly expressed as:

$$\underline{\tilde{R}}_{pt} = B_p \underline{\tilde{R}}_{mt} + \underline{\tilde{\epsilon}}_{pt} \tag{53}$$

where B_p is the matrix of factor loadings of the reference portfolios and $\underline{\tilde{\epsilon}}_{pt}$ is the vector of their idiosyncratic errors. The vector of sample mean returns of these basis portfolios $\underline{\overline{R}}_p$ is:

$$\underline{\overline{R}}_p = \frac{1}{T} \sum_{t=1}^{T} \underline{\tilde{R}}_{pt} = B_p \underline{\overline{R}}_m + \underline{\overline{\epsilon}}_p \tag{54}$$

where $\underline{\overline{R}}_m$ is the sample mean return vector of $\underline{\tilde{R}}_{mt}$ and $\underline{\overline{\epsilon}}_p$ is the sample mean vector of $\underline{\tilde{\epsilon}}_{pt}$. Similarly, the sample covariance matrix of the basis portfolio returns $\hat{\Sigma}_p$ is given by:

$$\begin{aligned} \hat{\Sigma}_p &= \frac{1}{T} \sum_{t=1}^{T} (\underline{\tilde{R}}_{pt} - \underline{\overline{R}}_p)(\underline{\tilde{R}}_{pt} - \underline{\overline{R}}_p)' \\ &= B_p \hat{\Sigma}_m B_p' + B_p \hat{\Sigma}_{m\epsilon_p} + \hat{\Sigma}_{m\epsilon_p}' B_p' + \hat{\Sigma}_{\epsilon_p} \end{aligned} \tag{55}$$

where $\hat{\Sigma}_m$ is the sample covariance matrix of the returns on $\underline{\tilde{R}}_{mt}$, $\hat{\Sigma}_{m\epsilon_p}$ is the matrix of sample covariances between $\underline{\tilde{R}}_{mt}$ and $\underline{\tilde{\epsilon}}_{pt}$, and $\hat{\Sigma}_{\epsilon_p}$ is the sample covariance matrix of $\underline{\tilde{\epsilon}}_{pt}$.

In this setting, it is possible to contrast the performance of different basis portfolio construction methods under simple assumptions. Under the assumptions set out above, the sample covariance matrix $\hat{\Sigma}_{m\epsilon_p}$ will be close to its theoretical value of zero in large samples (i.e. $\hat{\Sigma}_p \approx B_p \hat{\Sigma}_m B_p' + \hat{\Sigma}_{\epsilon_p}$). Similarly, the sample mean vector $\underline{\overline{\epsilon}}_p$ will be close to zero if the basis portfolios are large, well-diversified, and constructed such that their weights are not systematically related to the realizations of $\underline{\tilde{\epsilon}}_{pt}$ (i.e. $\underline{\overline{R}}_p \approx B_p \underline{\overline{R}}_m$). Note that we will not assume that $\hat{\Sigma}_{\epsilon_p}$ is close to zero so that we are implicitly recognizing that $\underline{\overline{\epsilon}}_p$ will converge to zero faster than $\hat{\Sigma}_{\epsilon_p}$ for well-diversified portfolios of large numbers of securities.

Now consider the usual χ^2 statistic for testing the hypothesis that the mean returns of the basis portfolios are all zero:

$$\begin{aligned} T \underline{\overline{R}}_p' \hat{\Sigma}_p^{-1} \underline{\overline{R}}_p &\approx T \underline{\overline{R}}_m' B_p' [B_p \hat{\Sigma}_m B_p' + \hat{\Sigma}_{\epsilon_p}]^{-1} B_p \underline{\overline{R}}_m \\ &\approx T \underline{\overline{R}}_m' [\hat{\Sigma}_m + B_p'^{-1} \hat{\Sigma}_{\epsilon_p} B_p^{-1}]^{-1} \underline{\overline{R}}_m \\ &\approx T \underline{\overline{R}}_m' \hat{\Sigma}_m^{-1} \underline{\overline{R}}_m - T \underline{\overline{R}}_m' \hat{\Sigma}_m^{-1} [B_p' \hat{\Sigma}_{\epsilon_p}^{-1} B_p + \hat{\Sigma}_m^{-1}]^{-1} \hat{\Sigma}_m^{-1} \underline{\overline{R}}_m \end{aligned} \tag{56}$$

As the analysis in equation (56) indicates, this χ^2 statistic permits simple comparisons of the quality of different basis portfolios in terms of their ability to mimic $\underline{\tilde{R}}_{mt}$.[20] For example, suppose the constructed basis portfolios had negligible idiosyncratic risk (i.e. $\hat{\Sigma}_{\epsilon_p}$ is close to zero). Then the second term in the last line of equation (56) will be close to zero and the χ^2 statistic will be close to the maximum attainable one, that associated with the returns $\underline{\tilde{R}}_{mt}$ (i.e. the first term in the last line of equation (56)). Any increase in the idiosyncratic variances of the basis portfolios will reduce the magnitude of the χ^2 statistic. Similarly, consider the impact of an increase in the factor loadings B_p of the basis portfolios holding the residual risk of the portfolios, $\hat{\Sigma}_{\epsilon_p}$, fixed. This will lead to increase in the percentage of the variation of the basis portfolio returns that is explained by the true reference portfolios and to a larger χ^2 statistic that will be closer to that of the true reference portfolios, $\underline{\tilde{R}}_{mt}$. Hence, the usual χ^2 statistic for testing the hypothesis that all of the reference portfolio mean returns are zero can rank the performance of different combinations of portfolio formation procedures and factor loading estimation methods.

It is worth considering an alternative derivation of this performance criterion. As in the analysis of equation (53) above, let $\underline{\tilde{R}}_t$ be the vector of excess returns on individual securities when the riskless rate version of the **APT** is true and be the corresponding raw returns when the zero beta version is appropriate. Consider the fitted multivariate regression of $\underline{\tilde{R}}_t$ on $\underline{\tilde{R}}_{mt}$ and a constant term:

$$\underline{\tilde{R}}_t = \hat{\underline{\alpha}} + \hat{B}\underline{\tilde{R}}_{mt} + \hat{\underline{\epsilon}}_t \tag{57}$$

where $\hat{\underline{\alpha}}$ is the estimated constant term vector, $\hat{B}$ is the estimated factor loading matrix, and $\hat{\underline{\epsilon}}_t$ is the fitted residual vector. If the **APT** is true, then $\hat{\underline{\alpha}}$ should be statistically insignificantly different from zero. The usual χ^2 for testing this hypothesis is:

$$\begin{aligned} T\hat{\underline{\alpha}}'\hat{\Omega}^{-1}\hat{\underline{\alpha}} &= T(\overline{\underline{R}} - \hat{B}\overline{\underline{R}}_m)'\hat{\Omega}^{-1}(\overline{\underline{R}} - \hat{B}\overline{\underline{R}}_m) \\ &= T(\overline{\underline{R}}'\hat{\Sigma}^{-1}\hat{\Omega})\hat{\Omega}^{-1}(\hat{\Omega}\hat{\Sigma}^{-1}\overline{\underline{R}}) \end{aligned} \tag{58}$$

[20] We implicitly assume in (56) that B_p is invertible. If it is not, then the basis portfolios $\underline{\tilde{R}}_{pt}$ are mimicking a linear combination of the factors. It is easy to see that this would result in a lower χ^2 statistic for $\underline{\tilde{R}}_{pt}$ than the corresponding statistic for $\underline{\tilde{R}}_{mt}$ when the rank of B_p is less than K.

where $\hat{\Omega}$ is the sample residual covariance matrix of $\hat{\underline{\epsilon}}_t$. Manipulation of equation (58) yields:[21]

$$T\hat{\underline{\alpha}}'\hat{\Omega}^{-1}\hat{\underline{\alpha}} = T\overline{\underline{R}}'\hat{\Sigma}^{-1}\overline{\underline{R}} - T\overline{\underline{R}}_m'\hat{\Sigma}_m^{-1}\overline{\underline{R}}_m \tag{59}$$

Thus the χ^2 statistic for testing the statistical significance of $\hat{\underline{\alpha}}$ is the difference of two χ^2 statistics—the χ^2 statistic for testing the joint significance of the individual security mean returns and the χ^2 statistic for testing the joint significance of the mean returns on the factor portfolios.

By analogy with the analysis leading from (57) to (59), suppose that $\hat{\underline{\alpha}}_p$ is the vector of intercepts from the regression of $\tilde{\underline{R}}_t$ on $\tilde{\underline{R}}_{pt}$ and that the corresponding χ^2 statistic for testing the statistical significance of $\hat{\underline{\alpha}}_p$ is:

$$\begin{aligned} T\hat{\underline{\alpha}}_p'\hat{\Omega}_p^{-1}\hat{\underline{\alpha}}_p &= T\overline{\underline{R}}'\hat{\Sigma}^{-1}\overline{\underline{R}} - T\overline{\underline{R}}_p'\hat{\Sigma}_p^{-1}\overline{\underline{R}}_p \\ &= [T\overline{\underline{R}}'\hat{\Sigma}^{-1}\overline{\underline{R}} - T\overline{\underline{R}}_m'\hat{\Sigma}_m^{-1}\overline{\underline{R}}_m] + [T\overline{\underline{R}}_m'\hat{\Sigma}_m^{-1}\overline{\underline{R}}_m - T\overline{\underline{R}}_p'\hat{\Sigma}_p^{-1}\overline{\underline{R}}_p] \end{aligned} \tag{60}$$

where $\hat{\Omega}_p$ is the residual covariance matrix from this regression. As the second line of (60) indicates, this χ^2 statistic has two components: the *correct* χ^2 statistic (58) for testing the null hypothesis using the *true* basis portfolios and a term reflecting the deviation of the measured reference portfolio returns $\tilde{\underline{R}}_{pt}$ from $\tilde{\underline{R}}_{mt}$. The analysis in equation (56) suggests that choosing the basis portfolios with the largest χ^2 statistic will minimize this problem.

This points up a potential problem associated with measuring basis portfolio performance with its χ^2 statistic. Clearly if we choose basis portfolios which maximize the χ^2 statistic (56) then such portfolios minimize the χ^2 statistic (60) for testing the **APT**. This obviously reduces the power of such tests, although the magnitude of such a bias when the **APT** is false cannot be analyzed without further assumptions. Fortunately, this problem can be mitigated to a considerable extent by using known empirical anomalies such as those associated with firm size and dividend yield to increase the power of tests of the **APT**. This occurs because the basis portfolios will tend to be well-diversified while the anomalies

[21] This follows directly from three observations: (1) $\hat{\underline{\alpha}} = \hat{\Omega}\hat{\Sigma}^{-1}\overline{\underline{R}}$; (2) $\hat{B} = \hat{\Sigma}\omega(\omega'\hat{\Sigma}\omega)^{-1}$; and (3) $\hat{\Omega} = \hat{\Sigma} - \hat{\Sigma}\omega'(\omega\hat{\Sigma}\omega')^{-1}\omega\hat{\Sigma}$. Here ω denotes the $N \times K$ matrix of portfolio weights of the *true* basis portfolios, $\tilde{\underline{R}}_{mt}$.

are not. Characteristics such as *small* market capitalization, *zero* dividend yield, or *high* dividend yield are clearly not distributed uniformly over securities.

Finally, this analysis can be linked to the extensive literature on mean variance efficiency tests. Following Ross(1977), Grinblatt and Titman(1983), and Ingersoll(1984), consider the transformation of the basis portfolios so that only one portfolio has non-zero expected return. Let the true values of the mean vector and covariance matrix of the transformed factor returns, $\underline{\tilde{R}}_{pt}^*$, be denoted by $\underline{\bar{R}}_p^*$ and Σ_p^* respectively. Now form the K new basis portfolios:

$$\underline{\tilde{R}}_{pt}^* = C^* T' \Sigma_p^{-1/2} \underline{\tilde{R}}_{pt} \tag{61}$$

where the orthogonal (partitioned) matrix T is chosen so:[22]

$$T = [(\underline{\bar{R}}_p{}' \underline{\bar{R}}_p)^{-1} \underline{\bar{R}}_p \mid T^*] \tag{62}$$

that is, so the last $K-1$ of the new basis portfolios are still orthogonal to the first and have zero expected excess returns, and where C^* is a diagonal matrix with the inverse elements of:

$$T' \Sigma_p^{-1/2} \underline{\iota} \tag{63}$$

along the main diagonal in order to scale the new basis portfolio to have unit investment. Now only the first basis portfolio has a non-zero risk premium and, in terms of the original reference portfolios, it is given by:

$$\tilde{R}_{1t}^* = (\underline{\bar{R}}_p' \Sigma_p^{-1/2} \underline{\iota})^{-1} \underline{\bar{R}}_p' \Sigma_p^{-1/2} \underline{\tilde{R}}_{pt} \tag{64}$$

Consider the squared Sharpe ratio of this portfolio:

$$\frac{\overline{R}_1^{*2}}{\sigma_1^{*2}} = \frac{[(\underline{\bar{R}}_p' \Sigma_p^{-1/2} \underline{\iota})^{-1} \underline{\bar{R}}_p' \Sigma_p^{-1/2} \underline{\bar{R}}_p]^2}{[(\underline{\bar{R}}_p' \Sigma_p^{-1/2} \underline{\iota})^{-2} \underline{\bar{R}}_p' \Sigma_p^{-1/2} \Sigma_p \Sigma^{-1/2} \underline{\bar{R}}_p]} \tag{65}$$

which is proportional to the squared t statistic for the hypothesis that its mean return is zero. Manipulation of (65) yields:

[22] See Appendix A, Section 1.17 of Lawley and Maxwell(1971) for an explicit description of one method for constructing this matrix. Note also that $\Sigma_p^{1/2}$ is taken to be any symmetric square root of Σ_p.

$$\frac{\overline{R}_1^{*2}}{\sigma_1^{*2}} = \underline{\bar{R}}_p' \Sigma_p^{-1} \underline{\bar{R}}_p \tag{66}$$

which is precisely the χ^2 statistic (56) of the original basis portfolios. In consequence, comparison of this Sharpe ratio with the sample Sharpe ratio of the sample mean variance efficient portfolio based on $\underline{\dot{R}}_t$ which has an orthogonal portfolio with zero expected return yields Jobson and Korkie's(1982) test for the potential performance of this constructed reference portfolio. Their analysis links this test to the other mean variance efficiency tests.

VI. Empirical Results

In this section, we provide evidence on the comparative performance of different reference portfolios. This effort requires numerous decisions and technical choices. In particular, we have to choose between daily and monthly data, the time period to be covered, the number of included securities, and the number of postulated factors.

What is the appropriate frequency of observation for estimating factor models of security returns? There is certainly substantial freedom of choice since the **CRSP** monthly file provides returns on all **NYSE** stocks from 1926 to the present, the **CRSP** daily file contains daily data on all **NYSE** and **AMEX** stocks from July 1962 to the present, and minimal computationally skill stands between us and bidaily, weekly, biweekly, or other intermediate frequencies. The primary advantage of daily data is, of course, the potential increase in precision of the estimated variances and covariances, the inputs to the various estimation methods. There are two main disadvantages of daily data: (1) the persistent incidence of non-trading and thin trading which bias the estimates of second order moments and (2) the biases in mean security returns associated with bid-ask spreads that are well-documented in Blume and Stambaugh(1984) and Roll(1983). Following Roll and Ross(1980) and most other investigations of the **APT**, we opted for the putative virtues of a large sample and used daily data to estimate the factor loadings and idiosyncratic variances. In Lehmann and Modest(1985b), we present evidence on the optimum observation frequency.

We estimated factor models for four subperiods covered by the **CRSP** daily returns file: 1963 through 1967, 1968 through 1972, 1973 through 1977, and 1978 through 1982. In each period, we confined our attention to continuously listed firms in order to have the same number of observations for each security and ignored any potential selection bias associated with this choice. This yields a sample of 1001 securities in the first period, 1359 firms in the second period, 1346 in the third period, and 1281 in the final period. The number of daily observations in each of the four subperiods is 1259, 1234, 1263, and 1264, respectively. The **CRSP** daily file (with few exceptions) lists securities in alphabetical order by their most recent name. To guard against any biases induced by the natural progression of letters (General Dynamcis, General Electric, etc.), we randomly reordered the securities in each subperiod. The usual sample covariance matrix of these security returns provided the basic input to our subsequent analysis.

We also made choices as to the number of securities and the number of factors included in the analysis. In order to study the impact of the number of securities on the sampling variation of reference portfolios, we estimated factor models for the first 30, 250, and 750 securities in our randomly sampled data files for each period. We have completed runs involving as many as 1000 securities but this larger number of firms yielded little improvement over the results for 750 securities and proved to be very expensive as the number of factors grew large. We restricted our attention to models containing five, ten, and fifteen factors, although for obvious reasons we did not estimate a fifteen factor model with only thirty securities. We chose to remain agnostic about the ***true*** number of factors underlying security returns.

We first provide evidence on the diversification properties of alternative combinations of basis portfolio formation procedures and estimation methods. As noted in Section III, there is little point in comparing the performance of the quadratic programming portfolios with the Fama-MacBeth and minimum idiosyncratic risk portfolios since the former are well-diversified by construction. Hence, we limit our attention to Fama-MacBeth and minimum idiosyncratic risk portfolios. Previous investigators have employed estimates of factor loadings obtained under the normalization that the covariance matrix of the factors is the identity matrix and that $B'D^{-1}B$ is diagonal. We employed estimates normalized in this fashion due to their obvious empirical relevance although the diversification properties of the Fama-MacBeth portfolios are **not** invariant with respect to normalization.[23]

[23] See, in particular, the discussion on pages 9-11 above.

There are several ways to quantify the diversification properties of reference portfolio weights. We confine our attention to one simple summary measure: the sum of squared portfolio weights. In the analysis underlying the **APT**, a portfolio with weights $\underline{w}_k$ which is perfectly correlated with the k^{th} common factor is well-diversified if, as the number of securities included in the portfolio grows large, the sum of squared portfolio weights $\underline{w}_k'\underline{w}_k$ is close to zero. The minimum sum that can be attained with portfolio weights that sum to one is the inverse of the number of securities. Hence, $.0333\bar{3}$ is the smallest sum of squared portfolio weights that can be attained with thirty securities, .004 is the corresponding total for 250 securities, and $.001333\bar{3}$ is the minimum attainable with 750 securities.

Table 1 reports the comparative degree of diversification of Fama-MacBeth and minimum idiosyncratic risk portfolios coupled with different estimation methods, numbers of securities, and numbers of factors. For each portfolio formation strategy, estimation method, number of securities, and number of factors, we report two numbers. The first number is the average sum of squared portfolio weights across both factors and sample periods. Under each such mean, we also provide the sample standard deviation of the sum of squared portfolio weights. These quantities are given by:

$$\begin{aligned}\overline{\underline{w}_k'\underline{w}_k} &= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{K}\sum_{k=1}^{K}\underline{w}_{kt}'\underline{w}_{kt} = \frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}\sum_{i=1}^{N}w_{ikt}^2 \\ \mathrm{Var}(\overline{\underline{w}_k'\underline{w}_k}) &= \frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}(\underline{w}_{kt}'\underline{w}_{kt} - \overline{\underline{w}_k'\underline{w}_k})^2\end{aligned} \tag{67}$$

where k indexes factors, t refers to time periods (where T=4), and i indexes the firms. Obviously, these quantities represent descriptive measures and are not appropriate for inference without further assumptions. We have examined the sum of squared portfolio weights for each basis portfolio individually and determined that the averaging implicit in Table 1 altered none of the basic conclusions.

The overwhelming conclusion from Table 1 is that Fama-MacBeth portfolios formed under the usual normalization of the factor loadings are extremely poorly diversified compared with minimum idiosyncratic risk portfolios. Moreover, minimum idiosyncratic portfolios proved to be quite well-diversified with mean sums of squared portfolio weights only ten to twenty times the minimum attainable ones. The contrast in performance is striking—

the mean sums of squares associated with Fama-MacBeth portfolios are between 40 and 50,000 times those associated with minimum idiosyncratic risk portfolios. Examination of these magnitudes for individual basis portfolios reveals that similar differences occur in the disaggregated data as well. In fact, in no case is a Fama-MacBeth basis portfolio better diversified than the corresponding minimum idiosyncratic risk portfolio. The evidence is less clear on the comparative performance of different estimation methods. The minimum idiosyncratic risk portfolios formed from principal components estimates appear to be slightly better diversified than those computed from the other three estimation methods but the difference is slight and is reversed often when individual basis portfolios are compared.

What accounts for the sharp contrast in the diversification behavior of the Fama-MacBeth and minimum idiosyncratic risk portfolios? The answer lies in the scaling of the two portfolios as discussed in Section III. Fama-MacBeth portfolios are the minimum idiosyncratic risk portfolios which have a loading of *one* on one factor and loadings of zero on the other factors prior to rescaling to unit net investment. Since, under the usual normalization, factor loading estimates are typically much smaller than one (on the order of .001 to .0001 in daily data), some of the weights have to be very large and positive to insure that the portfolio has a loading of one on the factor being mimicked. By the same token, some of the weights will have to be large and negative in order to have loadings of zero on the other factors. As pointed out in Section III, this problem does not arise in the **CAPM** context since it is natural to require the loading on the proxy for the market portfolio to equal one. By contrast, minimum idiosyncratic risk portfolios need not have any particular loading on the factor being mimicked. Hence, the resulting portfolios need only have small positive and negative weights to insure orthogonality to the other factors.

The remainder of this section is devoted to the evaluation of the comparative merits of alternative methods using the χ^2 statistic (56). Tables 2 through 10 contain aggregate χ^2 statistics summed over the four sample periods. This is possible because the data contained in the tables are independent by assumption, and sums of independent χ^2 statistics are distributed as χ^2 as well. Tables 11 through 43, which are contained in the Appendix,[24] provide the corresponding information from the individual sample periods to illuminate any divergences in performance that are obscured by averaging.

[24] The appendix is available from the authors on request.

Table 2 provides the fitted χ^2 statistics for selected combinations of portfolio formation procedures, estimation methods, and numbers of securities assuming a five factor model of security returns is appropriate. Tables 3 and 4 present the corresponding χ^2 statistics assuming there are ten and fifteen common factors respectively. These statistics were computed using daily returns on individual securities from the same period used to estimate the factor models. One possible source of bias in these tables involves the use of daily returns to calculate $\overline{\underline{R}}_p$ due to the bid-ask spread bias problem alluded to above. As a consequence, Tables 5 through 7 provide the same statistics presented in Tables 2 through 4 computed using weekly returns to calculate $\overline{\underline{R}}_p$ in order to mitigate this problem.[25] Finally, another potential source of bias in these tables is overfitting that might arise since the security returns used to compute the χ^2 statistics are from the same period used to estimate the factor models. We guarded against this possibility in Tables 8 through 10 which provide the same statistics as Tables 2 through 4 computed using daily returns from the subsequent five year period. The aggregate χ^2 statistics reported in these tables are smaller than those reported in the other tables in part because they are based on three sample periods due to the unavailability (as yet) of data covering 1983 through 1987. In addition, each χ^2 statistic was based on a smaller number of securities since not all securities which were continuously listed during the estimation period were also continuously listed during the subsequent five year period.

Each table contains a plenitude of information. The four columns listed at the top of each table correspond to the four portfolio formation methods under consideration: minimum idiosyncratic risk, Fama-MacBeth, positive net investment quadratic programming, and well-diversified quadratic programming. Eight combinations of estimation methods and numbers of securities comprise the rows of each table. These combinations are maximum likelihood and restricted maximum likelihood factor analysis with 30, 250, and 750 securities and principal components and instrumental variables with 750 securities. Four numbers are reported for each portfolio formation method, estimation method, and number of securities. The first quantity is the χ^2 statistic for the raw returns on the K basis portfolios. This χ^2 statistic is appropriate when the zero beta version of the **APT** is true.

[25] In this paper, we do not address the temporal aggregation bias issues investigated in Lehmann and Modest(1985c) involving the optimal periodicity for estimation of the factor model.

The number in parentheses underneath is the marginal significance level associated with the hypothesis that the means of these portfolios are jointly insignificantly different from zero over all of the sample periods. The number next to the χ^2 statistic for the raw return portfolio is the χ^2 statistic for the excess returns on the K basis portfolios. This statistic is relevant when the riskless rate version of the **APT** is appropriate. As discussed in Section III, these excess returns are the difference between the raw returns on the reference portfolios and the returns on the minimum idiosyncratic risk portfolio which is orthogonal to all of the factors. Note that the χ^2 statistics for the excess returns of the minimum idiosyncratic risk and Fama-MacBeth portfolios are identical for the reasons given in Section III on pp. 15-16.[26] The number in parentheses underneath is the corresponding marginal significance level for the χ^2 statistic of the excess return portfolios.

There are three dimensions along which contrasts in performance might be expected: (1) the number of securities; (2) the estimation method; and (3) the portfolio formation procedure. The first two categories are of particular interest due to the large differences in computational cost between factor analysis on 250 securities and on 750 securities and between factor analysis, instrumental variables, and principal components. Differential performance across portfolio formation methods might occur if minimum idiosyncratic risk and Fama-MacBeth portfolios prove to be poorly diversified compared with quadratic programming portfolios. We shall also be concerned with the consistency of the results across time periods, in and out of sample, and the number of factors.

The results strongly suggest the importance of examining large cross-sections of securities when constructing basis portfolios.[27] The 750 security basis portfolios always

[26] In terms of the one factor example in Section III, the portfolio weights of the Fama-MacBeth excess return portfolio are proportional to those of the minimum idiosyncratic risk excess return portfolio where the factor of proportionality is the inverse of the average factor loading. Since the average loading is typically much smaller than one, we would expect the Fama-MacBeth portfolio weights to be much larger. However, this distinction would not affect the χ^2 statistic since the proportionality factor cancels out in its formation.

[27] It is not possible to test whether the differences in chi-squared statistics are significant since the statistics are dependent. We therefore present these results as being suggestive rather than as a formal definitive test. However, in conjunction with the evidence presented in Lehmann and Modest(1985b) it seems safe to conclude that the differences in chi-squared statistics are significant since alternative basis portfolio construction methods lead to sharply dissimilar conclusions about the absolute and relative performance of mutual funds.

outperformed the corresponding 250 security portfolios which, in turn, always dominated 30 security portfolios. These rankings hold across estimation methods, portfolio formation procedures, time periods, and observation intervals with both raw return and excess return portfolios. The differences in performance increase with the number of factors extracted in sample in the daily data, although this finding does not persist in the weekly data or out of sample. Note that this need not have occurred—we have examined randomly selected well-diversified portfolios and found that, quite often, the χ^2 statistics for 250 security portfolios exceeded those for 750 securities. The typical χ^2 statistics for 250 securities were twice as large as those for 30 securities and were roughly three times as large for 750 securities. For example assuming the ten factor riskless rate version of the **APT** was appropriate, the chi-squared statistics for the significance of the mean returns on the basis portfolios using the unrestricted maximum likelihood estimation method in conjunction with the minimum idiosyncratic risk portfolio formation procedure were 35.20 using 30 securities, 81.14 using 250 securities and 116.88 using 750 securities. These differences are also apparent in the individual period results reported in the Appendix. Thirty security raw return basis portfolios were insignificant at the ten per cent level for two out of four time periods in daily data while the corresponding excess returns portfolios were insignificant in three out of four periods. By contrast, the 250 and 750 security raw return portfolios were highly significant in all time periods while the excess return portfolios were highly significant in daily data for all but the five factor model in the third time period.

The four estimation methods exhibited similarly striking contrasts. Restricted maximum likelihood factor analysis systematically outperformed conventional maximum likelihood factor analysis in the daily and weekly data but this dominance did not persist out of sample, where they achieved almost identical performance. Both maximum likelihood and restricted maximum likelihood factor analysis consistently outperformed the less efficient instrumental variables and principal components techniques for both raw and excess return portfolios in both daily data (in and out of sample) and weekly data. For instance, the unrestricted maximum likelihood estimation method in conjunction with the minimum idiosyncratic risk portfolio formation procedure yielded an aggregate in-sample chi-squared statistic of 116.88 using daily data under the assumption that the ten factor riskless rate version of the **APT** was appropriate. This is substantially larger than the aggregate chi-squared statistics of 97.88 and 82.94 that were obtained using the principal components and instrumental variables estimation methods.[28] The superiority of the relatively efficient maximum likelihood procedures also persisted out of sample. Out of sample the three estimation methods generated chi-squared statistics of 53.50, 37.13 and 39.07 respectively. Moreover, these discrepancies do not appear to merely reflect peculiarities of daily data as the differences did not disappear when weekly returns were used to perform the comparisons – using weekly data the corresponding chi-squared statistics were 73.83, 53.08, and 62.48.

An examination of the subperiod results supports the basic conclusion that the efficient estimation methods typically outperform the inefficient methods and shows some interesting intertemporal variation in performance. In the in-sample results, principal components performed well only in the first five year period, achieving performance comparable to the 250 security basis portfolios in the subsequent three periods. Instrumental variables provided more consistent performance but was consistently inferior to the more efficient estimation procedures in sample. The out-of-sample results were broadly consistent with these findings with some idiosyncrasies in the individual subperiods. In five factor models, the efficient methods proved superior in the second subperiod while all methods yielded similar performance in the first and third subperiods. In ten factor models, all methods provided similar performance in the first and second subperiods except for the superior performance of instrumental variables in the first and its inferior performance in the second. The efficient methods outperformed the inefficient methods by a wide margin in the final out-of-sample period. Similar inconsistencies emerged in the fifteen factor runs. The methods provided similar performance in the first out-of-sample period except for the inferior results provided by principal components. In the second subperiod, only instrumental variables provided inferior results to the similar performance of the other three methods. Again, the efficient methods outperformed the inefficient methods in the final subperiod.

The final set of comparisons is between the four portfolio formation methods. The minimum idiosyncratic risk portfolios provided almost identical performance to well-diversified quadratic programming portfolios using 750 securities and provided consistently superior

[28] A perusal of the tables indicates that the differences in chi-squared statistics between the estimation methods is typically much larger for the ten and fifteen factor models than for the five factor model.

performance with 250 securities. Other observations depend on the estimation method and choice of raw or excess returns. For the efficient estimation methods and raw returns, minimum idiosyncratic risk portfolios consistently dominated Fama-MacBeth portfolios[29] in-sample both with daily and weekly data which, in turn, typically outperformed the positive net investment quadratic programming portfolios. For instance under the assumption that a five factor model is appropriate, the chi-squared statistics for the minimum idiosyncratic risk and Fama-MacBeth raw return portfolios are 131.96 and 92.28 respectively for the daily in-sample results using the unrestricted maximum likelihood estimation procedure. The corresponding numbers under the assumption that there are ten common factors are 201.21 and 144.94, and under the assumption of fifteen common factors 227.77 and 168.24. The Fama-MacBeth portfolios, however, performed slightly better than the minimum idiosyncratic risk portfolios out-of-sample while both continued to dominate the positive net investment quadratic programming portfolios. Out of sample the corresponding chi-squared statistics for the minimum idiosyncratic risk and Fama-MacBeth raw return portfolios are 60.29 and 63.32 for the five factor model, 80.67 and 85.09 for the ten factor model, and 101.63 and 99.86 for the fifteen factor model. For the efficient estimation methods and excess return portfolios, the picture about the relative merits of the minimum idiosyncratic risk and Fama-MacBeth basis portfolios versus the quadratic programming portfolios is virtually identical. The only difference is that the positive net investment quadratic programming portfolios provided superior performance for five factor models in-sample, an improvement which did not persist out-of-sample.

The final question considered here is whether less efficient estimation methods or smaller numbers of securities coupled with quadratic programming procedures provides a good substitute for the more computationally expensive alternatives. With two exceptions, no combination of inefficient estimation methods or smaller numbers of securities with quadratic programming procedures improves on the results obtained by efficient es-

[29] As noted above, the minimum idiosyncratic risk and Fama-MacBeth portfolio formation procedures produce identical chi-squared statistics for the excess return reference portfolios when excess returns (over the riskless rate) are computed using constructed riskless rates. However, if excess returns were computed using an exogenously specified riskless rate such as the Treasury bill rate then the differing performance of the two portfolio formation procedures, as indicated by the differences in the χ^2 statistics of the raw return basis portfolios, would be relevant.

timation and minimum idiosyncratic risk portfolio formation with 750 securities. Both exceptions involve the five factor models in-sample using the daily data. The first exception is the combination of positive net investment quadratic programming portfolios and instrumental variables estimation for excess returns which yielded a chi-squared statistic of 77.64 compared to the chi-squared statistic of 64.49 that arose from the unrestricted maximum likelihood estimation method coupled with the minimum idiosyncratic risk portfolio formation procedure. The other aberration is the superior performance of basis portfolios constructed using 250 securities, quadratic programming and the efficient estimation procedures over those basis portfolios consisting of 750 securities that were constructed using the minimum idiosyncratic risk procedure. Again, this combination provided inferior performance out-of-sample.

Why do the minimum idiosyncratic risk portfolios perform better in-sample than those produced by the Fama-MacBeth procedure? We suspect that the answer lies in the deleterious impact of measurement error on the Fama-MacBeth portfolios. For example, consider the one factor case discussed extensively in Section III. From the analysis in equations (16) it is apparent that the χ^2 statistic for the Fama-MacBeth portfolio in the single factor case should be (approximately):

$$\chi^2_{FM} \approx T[\frac{(\underline{\beta}'\underline{\beta})^2\overline{R}_m^2}{(\underline{\beta}'\underline{\beta})^2\sigma_m^2 + \underline{\beta}'\Sigma_v\underline{\beta}(\sigma_m^2 + \overline{R}_m^2) + \underline{\beta}'\underline{\beta}\sigma_\epsilon^2 + \mathbf{tr}(\Sigma_v)\sigma_\epsilon^2}] \tag{68}$$

while that of the minimum idiosyncratic risk portfolio should be (approximately):

$$\chi^2_{MIRP} \approx T[\frac{\overline{\beta}^2\overline{R}_m^2}{\overline{\beta}^2\sigma_m^2 + \sigma_\epsilon^2/N}] \tag{69}$$

where $\underline{\beta}$ is the vector of (true) loadings of the N securities, $\overline{\beta}$ is their sample mean, $\overline{R}_m$ is the mean of the true unobserved basis portfolio, σ_m^2 is its variance, σ_ϵ^2 is the common variance of the idiosyncratic disturbances, and Σ_v is the covariance matrix of the errors in the estimated factor loadings. The approximation arises from replacing sample moments with population moments in the χ^2 expressions. When measurement error in the sample loadings is negligible, the chi-squared statistic of the Fama-MacBeth portfolio will typically exceed that associated with the minimum idiosyncratic risk procedure. In contradistinction, this ordering can easily be reversed in the presence of measurement error. As the analysis in (23) and (24) suggests, this is a likely occurrence when the number of securities is large relative to the number of time series observations.

VII. Conclusion

This paper has provided a comprehensive examination of the merits of different basis portfolio formation strategies. In so doing, the analysis involved the four main estimation methods that have been proposed in the literature and the principal portfolio formation procedures that have been employed as well as some that have not previously been considered. In addition, this study provided a detailed evaluation of the impact of increases in the number of securities underlying the analysis, including the application of maximum likelihood methods to far larger cross-sections than in prior work. The result is a detailed set of data measuring the performance of excess return and raw return basis portfolios both in and out of sample over a variety of time periods and observation frequencies.

Three conclusions emerge from the examination of the more than 2300 statistics reported in this document. First, increasing the number of securities included in the analysis dramatically improves basis portfolio performance. Our results indicate that factor models involving 750 securities provide markedly superior performance to those involving 30 or 250 securities. Second. comparatively efficient estimation procedures such as maximum likelihood and restricted maximum likelihood factor analysis significantly outperform the less efficient instrumental variables and principal components procedures that have been proposed in the literature. In particular, the less efficient estimation procedures typically provided performance comparable to maximum likelihood factor analysis with 250 securities. Third, the minimum idiosyncratic risk portfolio formation procedure proposed in the third section outperformed both the Fama-MacBeth and positive net investment quadratic programming portfolios and proved equal to or better than the more expensive well-diversified quadratic programming procedure. The Fama-MacBeth procedure, which has dominated empirical research on the **APT**, yielded poorly diversified portfolios which provided inferior performance in this context. In sum, if an investigator had to choose one basis portfolio construction strategy from the formidable list considered here, the clear winner is minimum idiosyncratic risk portfolios coupled with maximum likelihood factor analysis of 750 securities.

We think that these conclusions should have a profound influence on empirical research involving the **APT**. These results suggest that the inconclusiveness of the bulk of existing research, which involves maximum likelihood factor analysis of groups of thirty to sixty securities, may reflect the inability of small groups of securities to capture the empirical content of the **APT**—the ability to measure the sources of systematic risk underlying security returns. Moreover, the analysis suggests that refuge cannot be found in the employment of less efficient but less expensive estimation procedures applied to large cross-sections. Instead, subsequent empirical investigations of issues pertaining to the validity of the **APT** should probably incur the costs associated with expensive efficient estimation procedures in the interest of providing more scientific inferences.

Of course, we have left open numerous issues associated with basis portfolio comparisons. In particular, our results may have little bearing on the comparative merits of different procedures in producing portfolios that mimic factors at weekly or monthly observation frequencies. In addition, we have provided little evidence on the quantitative impact of the use of comparatively ineffective portfolio formation procedures on inferences in particular applications such as the evaluation of managed portfolios. We are currently engaged in research along both of these lines (Lehmann and Modest [1985b,1985c]). The evidence presented in Lehmann and Modest(1985b) suggests that statistical differences in basis portfolio performance documented in this paper have an economically and quantitatively significant impact on the evaluation of the performance of mutual funds.

TABLE 1: AVERAGE SUM OF SQUARED PORTFOLIO WEIGHTS
[standard deviation of sums in parentheses]

NUMBER OF SECURITIES	ESTIMATION METHOD	FIVE FACTORS		TEN FACTORS		FIFTEEN FACTORS	
		MINIMUM IDIOSYNCRATIC RISK	FAMA-MACBETH	MINIMUM IDIOSYNCRATIC RISK	FAMA-MACBETH	MINIMUM IDIOSYNCRATIC RISK	FAMA-MACBETH
30	Maximum Likelihood	.397 (.494)	75.011 (459.206)	1.276 (6.063)	98.150 (846.203)	-	-
	Restricted Maximum Likelihood	.399 (.496)	2178.232 (16057.81)	1.257 (5.941)	273.641 (2781.711)	-	-
250	Maximum Likelihood	.048 (.049)	4.491 (28.928)	.063 (.045)	51.643 (551.863)	.071 (.041)	502.614 (4253.059)
	Restricted Maximum Likelihood	.048 (.049)	5.209 (34.645)	.063 (.045)	195.268 (2277.061)	.071 (.041)	81.613 (492.380)
750	Maximum Likelihood	.016 (.016)	.659 (4.297)	.022 (.015)	72.689 (798.788)	.025 (.014)	255.234 (3267.704)
	Restricted Maximum Likelihood	.016 (.016)	.659 (4.310)	.022 (.015)	22.570 (189.855)	.025 (.014)	46.992 (469.129)
	Principal Components	.013 (.019)	1.391 (5.784)	.017 (.035)	26.992 (251.929)	.019 (.043)	244.917 (2912.448)
	Instrumental Variables	.015 (.018)	77.021 (605.504)	.019 (.013)	1087.395 (9194.892)	.022 (.012)	93.953 (1310.661)

TABLE 2: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results-In Sample NUMBER OF FACTORS: 5

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
30	Maximum Likelihood	47.37 (.52E-03)	26.30 (.16)	41.96 (.28E-02)	26.30 (.16)	42.83 (.22E-02)	33.11 (.33E-01)	-	-
	Restricted Maximum Likelihood	50.22 (.21E-03)	29.68 (.75E-01)	44.44 (.13E-02)	29.68 (.75E-01)	42.06 (.27E-02)	32.81 (.35E-01)	-	-
250	Maximum Likelihood	102.10 (.53E-12)	61.53 (.41E-05)	83.76 (.90E-09)	61.53 (.41E-05)	81.38 (.23E-08)	75.61 (.22E-07)	86.40 (.31E-09)	50.89 (.16E-03)
	Restricted Maximum Likelihood	108.26 (.41E-03)	70.09 (.18E-06)	91.08 (.48E-10)	70.09 (.18E-06)	83.32 (.11E-08)	80.03 (.39E-08)	91.18 (.46E-10)	58.29 (.13E-04)
750	Maximum Likelihood	131.96 (.17E-17)	64.49 (.14E-05)	92.28 (.30E-10)	64.49 (.14E-05)	106.85 (.74E-13)	76.32 (.16E-07)	131.93 (.17E-17)	64.32 (.15E-05)
	Restricted Maximum Likelihood	138.26 (.11E-18)	73.89 (.42E-07)	99.66 (.14E-11)	73.89 (.42E-07)	118.77 (.48E-15)	78.42 (.73E-08)	138.13 (.11E-18)	73.77 (.44E-07)
	Principal Components	109.00 (.30E-13)	57.34 (.18E-04)	84.18 (.76E-09)	57.34 (.18E-04)	92.93 (.23E-10)	37.57 (.10E-01)	108.17 (.42E-13)	57.08 (.20E-04)
	Instrumental Variables	122.81 (.86E-16)	54.53 (.48E-04)	70.49 (.15E-06)	54.53 (.48E-04)	126.04 (.22E-16)	77.64 (.98E-08)	125.66 (.25E-16)	55.39 (.36E-04)

TABLE 3: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results: In Sample   NUMBER OF FACTORS: 10

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
30	Maximum Likelihood	54.59 (.62E-01)	35.20 (.69)	50.17 (.13)	35.20 (.69)	53.63 (.73E-01)	46.85 (.21)	-	-
	Restricted Maximum Likelihood	57.77 (.34E-01)	37.70 (.57)	52.41 (.90E-01)	37.70 (.57)	25.79 (.69)	19.66 (.93)	-	-
250	Maximum Likelihood	129.94 (.19E-10)	81.14 (.13E-03)	102.78 (.19E-06)	81.14 (.13E-03)	103.64 (.15E-06)	87.96 (.19E-04)	105.96 (.71E-07)	63.49 (.10E-01)
	Restricted Maximum Likelihood	154.72 (.21E-14)	105.57 (.81E-07)	124.81 (.12E-09)	105.57 (.81E-07)	113.10 (.67E-08)	95.79 (.18E-05)	124.08 (.15E-09)	81.28 (.12E-03)
750	Maximum Likelihood	201.21 (.23E-22)	116.88 (.19E-08)	144.94 (.82E-13)	116.88 (.19E-08)	133.84 (.48E-11)	101.76 (.27E-06)	201.68 (.19E-22)	115.83 (.27E-08)
	Restricted Maximum Likelihood	225.06 (.13E-26)	143.92 (.12E-12)	171.22 (.36E-17)	143.92 (.12E-12)	140.36 (.45E-12)	112.74 (.75E-08)	225.40 (.11E-26)	142.58 (.20E-12)
	Principal Components	143.62 (.13E-12)	97.88 (.93E-06)	124.14 (.15E-09)	97.88 (.93E-06)	123.07 (.22E-09)	74.52 (.75E-03)	143.64 (.13E-12)	97.73 (.98E-06)
	Instrumental Variables	160.28 (.25E-11)	82.94 (.78E-04)	102.36 (.23E-06)	82.94 (.78E-04)	139.14 (.70E-12)	97.84 (.94E-06)	162.13 (.12E-15)	85.61 (.37E-04)

TABLE 4: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results: In Sample NUMBER OF FACTORS: 15

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
250	Maximum Likelihood	136.61 (.66E-07)	90.57 (.66E-02)	114.23 (.31E-04)	90.57 (.66E-02)	95.41 (.25E-02)	78.08 (.58E-01)	115.33 (.23E-04)	73.55 (.11)
	Restricted Maximum Likelihood	162.99 (.19E-10)	116.21 (.19E-04)	138.54 (.38E-07)	116.21 (.19E-04)	97.66 (.15E-02)	78.55 (.54E-01)	134.88 (.11E-06)	93.32 (.38E-02)
750	Maximum Likelihood	227.77 (.23E-20)	137.59 (.50E-07)	168.24 (.33E-11)	137.59 (.50E-07)	142.48 (.12E-07)	112.92 (.43E-04)	228.59 (.17E-20)	136.25 (.74E-07)
	Restricted Maximum Likelihood	261.81 (.50E-26)	174.62 (.40E-12)	204.05 (.14E-16)	174.62 (.40E-12)	144.82 (.57E-08)	116.15 (.19E-04)	261.14 (.65E-26)	172.99 (.68E-12)
	Principal Components	169.59 (.21E-11)	112.50 (.48E-04)	139.73 (.26E-07)	112.50 (.48E-04)	139.94 (.25E-07)	92.04 (.49E-02)	168.96 (.26E-11)	112.36 (.49E-04)
	Instrumental Variables	168.29 (.33E-11)	106.75 (.19E-03)	135.14 (.10E-06)	106.75 (.19E-03)	149.52 (.36E-12)	109.47 (.10E-03)	174.90 (.36E-12)	109.25 (.11E-03)

TABLE 5: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results: Out of Sample NUMBER OF FACTORS: 5

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
250	Maximum Likelihood	40.98 (.32E-03)	17.95 (.27)	44.24 (.10E-03)	17.95 (.27)	32.86 (.49E-02)	19.04 (.21)	37.60 (.10E-02)	14.11 (.52)
	Restricted Maximum Likelihood	41.52 (.27E-03)	18.37 (.24)	44.61 (.88E-04)	18.37 (.24)	31.82 (.68E-02)	18.45 (.24)	37.81 (.96E-03)	14.18 (.51)
750	Maximum Likelihood	60.29 (.22E-06)	33.03 (.46E-02)	63.32 (.67E-07)	33.03 (.46E-02)	35.66 (.20E-02)	27.03 (.28E-01)	61.55 (.14E-06)	33.13 (.45E-02)
	Restricted Maximum Likelihood	60.79 (.18E-06)	33.27 (.43E-02)	63.61 (.60E-07)	33.27 (.43E-02)	35.83 (.19E-02)	27.14 (.28E-01)	62.02 (.11E-06)	33.36 (.42E-02)
	Principal Components	57.70 (.62E-06)	24.96 (.50E-01)	41.82 (.24E-03)	24.96 (.50E-01)	36.24 (.16E-02)	16.31 (.36)	57.69 (.63E-06)	24.68 (.54E-01)
	Instrumental Variables	49.62 (.14E-04)	20.90 (.14)	50.26 (.11E-04)	20.90 (.14)	34.93 (.25E-02)	15.86 (.39)	49.77 (.13E-04)	20.76 (.14)

TABLE 6: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results: Out of Sample NUMBER OF FACTORS: 10

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
250	Maximum Likelihood	56.83 (.22E-02)	31.11 (.41)	58.57 (.14E-02)	31.11 (.41)	34.61 (.26)	33.23 (.31)	46.04 (.31E-01)	21.24 (.88)
	Restricted Maximum Likelihood	58.27 (.15E-02)	31.83 (.38)	59.07 (.12E-02)	31.83 (.38)	34.23 (.27)	29.15 (.51)	46.08 (.31E-01)	23.07 (.81)
750	Maximum Likelihood	80.67 (.16E-05)	53.50 (.52E-02)	85.09 (.36E-06)	53.50 (.52E-02)	45.83 (.32E-01)	38.83 (.13)	81.49 (.12E-05)	53.82 (.48E-02)
	Restricted Maximum Likelihood	80.90 (.15E-05)	53.82 (.48E-02)	85.41 (.32E-06)	53.82 (.48E-02)	44.73 (.41E-01)	37.90 (.15)	81.69 (.11E-05)	54.14 (.44E-02)
	Principal Components	70.36 (.43E-04)	37.13 (.17)	60.88 (.72E-03)	37.13 (.17)	43.33 (.55E-01)	28.18 (.56)	71.22 (.33E-04)	37.29 (.17)
	Instrumental Variables	65.83 (.17E-03)	39.07 (.12)	69.75 (.52E-04)	39.07 (.12)	45.44 (.35E-01)	37.13 (.17)	72.15 (.25E-04)	40.07 (.10)

TABLE 7: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Daily Results: Out of Sample NUMBER OF FACTORS: 15

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
250	Maximum Likelihood	65.45 (.25E-01)	40.44 (.67)	66.93 (.19E-01)	40.44 (.67)	42.61 (.57)	38.51 (.74)	57.51 (.10)	31.25 (.94)
	Restricted Maximum Likelihood	64.59 (.29E-01)	40.58 (.66)	67.18 (.18E-01)	40.58 (.66)	40.95 (.64)	39.15 (.72)	56.84 (.11)	30.87 (.95)
750	Maximum Likelihood	101.63 (.29E-05)	66.13 (.22E-01)	99.86 (.49E-05)	66.13 (.22E-01)	57.96 (.93E-01)	53.17 (.19)	99.26 (.58E-05)	64.77 (.28E-01)
	Restricted Maximum Likelihood	101.42 (.31E-05)	66.62 (.20E-01)	100.36 (.42E-05)	66.62 (.20E-01)	57.82 (.95E-01)	54.54 (.16)	101.22 (.33E-05)	65.29 (.26E-01)
	Principal Components	79.33 (.12E-02)	46.25 (.42)	73.68 (.45E-02)	46.25 (.42)	47.08 (.39)	34.31 (.88)	80.49 (.90E-03)	46.44 (.41)
	Instrumental Variables	91.62 (.50E-04)	53.26 (.19)	88.85 (.11E-03)	53.26 (.19)	40.06 (.68)	26.46 (.99)	90.45 (.69E-04)	53.11 (.19)

TABLE 8: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Weekly Results: In Sample NUMBER OF FACTORS: 5

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
30	Maximum Likelihood	34.95 (.20E-01)	26.51 (.15)	31.47 (.49E-01)	26.51 (.15)	31.55 (.49E-01)	26.61 (.15)	-	-
	Restricted Maximum Likelihood	37.24 (.11E-01)	29.62 (.76E-01)	33.89 (.27E-01)	29.62 (.76E-01)	30.77 (.58E-01)	26.35 (.15)	-	-
250	Maximum Likelihood	65.69 (.91E-06)	42.12 (.27E-02)	53.50 (.69E-04)	42.12 (.27E-02)	44.71 (.12E-02)	37.02 (.12E-01)	55.08 (.40E-04)	36.01 (.15E-01)
	Restricted Maximum Likelihood	69.89 (.19E-06)	47.45 (.51E-03)	58.02 (.14E-04)	47.45 (.51E-03)	47.85 (.45E-03)	40.58 (.42E-02)	58.73 (.11E-04)	41.02 (.37E-02)
750	Maximum Likelihood	77.69 (.96E-08)	42.20 (.26E-02)	55.48 (.35E-04)	42.20 (.26E-02)	56.83 (.22E-04)	55.80 (.31E-04)	78.09 (.83E-08)	42.31 (.25E-02)
	Restricted Maximum Likelihood	82.27 (.16E-08)	47.76 (.46E-03)	59.97 (.72E-05)	47.76 (.46E-03)	57.64 (.16E-04)	56.00 (.29E-04)	82.79 (.13E-08)	47.98 (.43E-03)
	Principal Components	49.69 (.25E-03)	28.09 (.11)	40.95 (.38E-02)	28.09 (.11)	37.43 (.10)	20.03 (.46)	49.56 (.26E-03)	28.00 (.11)
	Instrumental Variables	71.49 (.10E-06)	37.87 (.92E-02)	42.56 (.23E-02)	37.87 (.92E-02)	51.41 (.14E-03)	45.35 (.99E-03)	71.94 (.88E-07)	37.93 (.90E-02)

TABLE 9: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Weekly Results: In Sample NUMBER OF FACTORS: 10

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
30	Maximum Likelihood	43.59 (.32)	34.94 (.70)	39.72 (.48)	34.94 (.70)	43.65 (.32)	40.91 (.43)	-	-
	Restricted Maximum Likelihood	45.81 (.24)	37.32 (.59)	41.69 (.40)	37.32 (.59)	23.41 (.80)	19.84 (.92)	-	-
250	Maximum Likelihood	87.22 (.23E-04)	58.69 (.28E-01)	71.11 (.18E-02)	58.69 (.28E-01)	67.18 (.45E-02)	61.19 (.17E-01)	71.87 (.15E-02)	50.68 (.12)
	Restricted Maximum Likelihood	105.24 (.90E-07)	75.51 (.58E-03)	87.32 (.22E-04)	75.51 (.58E-03)	70.63 (.20E-02)	65.57 (.66E-02)	84.44 (.51E-04)	63.90 (.95E-02)
750	Maximum Likelihood	121.29 (.41E-09)	73.83 (.90E-03)	88.69 (.15E-04)	73.83 (.90E-03)	84.07 (.57E-04)	73.72 (.92E-03)	122.63 (.25E-09)	74.01 (.86E-03)
	Restricted Maximum Likelihood	137.12 (.15E-11)	90.33 (.93E-05)	105.22 (.90E-07)	90.33 (.93E-05)	92.54 (.48E-05)	84.43 (.52E-04)	138.51 (.88E-12)	90.39 (.91E-05)
	Principal Components	76.37 (.47E-03)	53.08 (.81E-01)	66.00 (.60E-02)	53.08 (.81E-01)	61.55 (.16E-01)	49.45 (.15)	76.40 (.46E-03)	52.98 (.82E-01)
	Instrumental Variables	104.51 (.11E-06)	62.48 (.13E-01)	72.03 (.14E-02)	62.48 (.13E-01)	85.23 (.41E-04)	78.25 (.28E-03)	104.88 (.10E-06)	65.63 (.65E-02)

TABLE 10: BASIS PORTFOLIO COMPARISONS: χ^2 STATISTICS
[p-values in parentheses]

DATA SET: Aggregate Weekly Results: In Sample

NUMBER OF FACTORS: 15

PORTFOLIO FORMATION PROCEDURE

NUMBER OF SECURITIES	ESTIMATION METHOD	MINIMUM IDIOSYNCRATIC RISK		FAMA-MACBETH		QUADRATIC PROGRAMMING: Non-Negative Weights		QUADRATIC PROGRAMMING: Well Diversified	
		Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns	Raw Returns	Excess Returns
250	Maximum Likelihood	94.95 (.27E-02)	68.77 (.20)	82.18 (.30E-01)	68.77 (.20)	64.04 (.34)	54.28 (.68)	82.97 (.26E-01)	59.71 (.49)
	Restricted Maximum Likelihood	114.98 (.25E-04)	87.19 (.12E-01)	100.20 (.88E-03)	87.19 (.12E-01)	62.89 (.37)	52.51 (.74)	101.14 (.71E-03)	77.10 (.68E-01)
750	Maximum Likelihood	149.60 (.13E-08)	94.10 (.32E-02)	109.57 (.98E-04)	94.10 (.32E-02)	94.55 (.30E-02)	90.94 (.61E-02)	151.71 (.68E-09)	93.91 (.34E-02)
	Restricted Maximum Likelihood	174.71 (.38E-12)	119.72 (.74E-05)	134.83 (.11E-06)	119.72 (.74E-05)	100.82 (.76E-03)	98.09 (.14E-02)	175.89 (.26E-12)	119.34 (.82E-05)
	Principal Components	94.73 (.28E-02)	63.84 (.34)	79.67 (.46E-01)	63.84 (.34)	80.73 (.38E-01)	67.64 (.23)	94.68 (.29E-02)	63.93 (.34)
	Instrumental Variables	112.48 (.48E-04)	81.06 (.36E-01)	97.52 (.16E-02)	81.06 (.36E-01)	98.26 (.13E-02)	96.63 (.19E-02)	119.52 (.78E-05)	84.65 (.20E-01)

References

BANZ, ROLF [1981], "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, Vol. 9, pp. 3-18.

BASU, S. [1977], "Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of Market Efficiency," *Journal of Finance*, Vol. 32, pp. 663-682.

BLUME, MARSHALL E. AND ROBERT F. STAMBAUGH [1983], "Biases in Computed Returns: An Application to the Size Effect," *Journal of Financial Economics*, Vol. 12, pp. 387-404.

CANNISTRARO, RICHARD S. [1973], "The Predictive Ability of Price/Earnings Ratios for Determining Risk-Adjusted Performance of Common Stocks," Unpublished, Sloan School of Management, MIT.

CHAMBERLAIN, GARY AND MICHAEL ROTHSCHILD [1983], "Arbitrage, Factor Structure and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, Vol. 51, pp. 1281-1304.

CHEN, NAI-FU [1983], "Some Empirical Tests of the Theory of Arbitrage Pricing," *Journal of Finance*, Vol. 38, pp. 1393-1414.

CHEN, NAI-FU AND JONATHAN E. INGERSOLL JR. [1983], "Exact Pricing in Linear Factor Models with Infinitely Many Assets: A Note," *Journal of Finance*, Vol. 38, pp. 985-988.

CONNOR, GREGORY [1984], "A Unified Beta Pricing Theory," *Journal of Economic Theory*, Vol. 34, pp. 13-31.

CONNOR, GREGORY AND ROBERT A. KORAJCZYK [1984], "Performance Measurement with the Abitrage Pricing Theory: A New Framework for Analysis," Unpublished, Kellogg Graduate School of Management, Northwestern University.

DEMPSTER, ARTHUR P., N.M. LAIRD AND DONALD B. RUBIN [1977], "Maximum Likelihood from Incomplete Data via the E-M Algorithm," *Journal of the Royal Statistical Society*, Vol. Series B 39, pp. 1-38.

DYBVIG, PHILIP H. [1983], "An Explicit Bound on Deviations from APT Pricing in a Finite Economy," *Journal of Financial Economics*, Vol. 12, pp. 483-496.

FAMA, EUGENE F. AND JAMES D. MACBETH [1973], "Risk, Return, and Equilibrium: Some Empirical Tests," *Journal of Political Economy*, Vol. 81, pp. 607-636.

GIBBONS, MICHAEL R. [1982], "Multivariate Tests of Financial Models: A New Approach," *Journal of Financial Economics*, Vol. 10, pp. 3-27.

GRINBLATT, MARK AND SHERIDAN TITMAN [1983], "Factor Pricing in a Finite Economy," *Journal of Financial Economics*, Vol. 12, pp. 495-507.

HAGGLUND, GOSTA [1982], "Factor Analysis by Instrumental Variables Methods," *Psychometrika*, Vol. 47, pp. 209-221.

INGERSOLL JR., JONATHAN E. [1984], "Some Results in the Theory of Arbitrage Pricing," *Journal of Finance*, Vol. 39, pp. 1021-1039.

JOBSON, J.D. AND BOB M. KORKIE [1982], "Potential Performance and Tests of Portfolio Efficiency," *Journal of Finincial Economics*, Vol. 10, pp. 433-466.

LAWLEY, D.N. AND A.E. MAXWELL [1971], *Factor Analysis as a Statistical Method*, New York: Elsevier.

LEHMANN, BRUCE N. AND DAVID M. MODEST [1985a], "The Empirical Foundations of the Arbitrage Pricing Theory I: The Empirical Tests," Unpublished, Graduate School of Business, Columbia University.

——————[1985b], "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons," Unpublished, Graduate School of Business, Columbia University.

——————[1985c], "Temporal Aggregation and the Optimal Construction of Basis Portfolios," Unpublished, Graduate School of Business, Columbia University.

LITZENBERGER, ROBERT H. AND KRISHNA RAMASWAMY [1979], "The Effects of Dividends on Common Stock Prices: Theory and Empirical Evidence," *Journal of Financial Economics*, Vol. 7, pp. 163-195.

MADANSKY, ALBERT [1964], "Instrumental Variables in Factor Analysis," *Psychometrika*, Vol. 29, pp. 105-113.

MADANSKY, ALBERT AND TERRY MARSH [1985], "The Factor Structure of Asset Returns: How Many Factors Are There and Which Ones Are Priced," Unpublished, Sloan School of Management, M.I.T.

REINGANUM, MARC R. [1981], "Misspecification of Capital Asset Pricing: Empirical Anomolies Based on Earnings Yields and Market Values," *Journal of Financial Economics*, Vol. 9, pp. 19-46.

ROLL, RICHARD W. [1977], "A Critique of the Asset Pricing Theory's Tests–Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics*, Vol. 4, pp. 129-176.

——————[1983], "On Computing Mean Returns and the Small Firm Premium," *Journal of Financial Economics*, Vol. 12, pp. 371-386.

ROLL, RICHARD W. AND STEPHEN A. ROSS [1980], "An Empirical Investigation of the Arbitrage Pricing Theory," *Journal of Finance*, Vol. 35, pp. 1073-1103.

ROSENBERG, BARR AND VINAY MARATHE [1979], "Tests of Capital Asset Pricing Hypotheses." In , *Research in Finance: A Research Annual*, Greenwich, Connecticut: JAI Press.

ROSS, STEPHEN A. [1976], "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, Vol. 13, pp. 341-60.

——————[1977], "Risk.Return, and Arbitrage." In Irwin Friend and James L. Bicksler, *Risk Return in Finance*, Cambridge, Massachusetts: Ballinger.

SHANKEN, JAY [1983], "Factor Pricing Models: Synthesis and Extensions," Unpublished, Graduate School of Business Administration, University of California Berkeley.

———[1984], "Multivariate Tests of the Zero-Beta CAPM," Unpublished, Graduate School of Business Administration, University of California Berkeley.

STAMBAUGH, ROBERT F. [1983], "Arbitrage Pricing With Information," *Journal of Financial Economics*. Vol. 12, pp. 357-369.