THE JOURNAL OF FINANCE • VOL. XXXV, NO. 5 • DECEMBER 1980

The Journal of FINANCE

VOL. XXXV DECEMBER 1980 No. 5

An Empirical Investigation of the Arbitrage Pricing Theory

RICHARD ROLL and STEPHEN A. ROSS

ABSTRACT

Empirical tests are reported for Ross' [48] arbitrage theory of asset pricing. Using data for individual equities during the 1962–72 period, at least three and probably four "priced" factors are found in the generating process of returns. The theory is supported in that estimated expected returns depend on estimated factor loadings, and variables such as the "own" standard deviation, though highly correlated (simply) with estimated expected returns, do not add any further explanatory power to that of the factor loadings.

THE ARBITRAGE PRICING THEORY (APT) formulated by Ross [48] offers a testable alternative to the well-known capital asset pricing model (CAPM) introduced by Sharpe [51], Lintner [30] and Mossin [38]. Although the CAPM has been predominant in empirical work over the past fifteen years and is the basis of modern portfolio theory, accumulating research has increasingly cast doubt on its ability to explain the empirical constellation of asset returns.

More than a modest level of disenchantment with the CAPM is evidenced by the number of related but different theories, e.g., Hakansson [18], Mayers [34], Merton [35], Kraus and Litzenberger [23]; by anomalous empirical evidence, e.g., Ball [2], Basu [4], Reinganum [40]; and by questioning of the CAPM's viability as a scientific theory, e.g., Roll [41]. Nonetheless, the CAPM retains a central place in the thoughts of academic scholars and of finance practitioners such as portfolio managers, investment advisors, and security analysts.

There is good reason for its durability: it is compatible with the single most widely-acknowledged empirical regularity in asset returns, their common variability. Apparently, intuition readily ascribes such common variation to a single factor which, with a random disturbance, generates returns for each individual asset via some (linear) functional relationship. Oddly, though, this intuition is wholly divorced from the formal CAPM theory. To the contrary, elegant deriva-

Graduate School of Management, University of California, Los Angeles, and School of Organization and Management, Yale University, respectively.

tions of the CAPM equation have been concocted beginning from the first principles of utility theory; but the model's popularity is not due to such analyses, for they make all too obvious the assumptions required for the CAPM's validity and make no use of the common variability of returns. A review of recent finance texts (e.g., Van Horne, [54, pp. 57–63]) reveals that rationalizations of the CAPM are based instead on the dichotomy between diversifiable and non-diversifiable risk, a distinction which refers to a linear generating process, not to the CAPM derived from utility theory.

The APT is a particularly appropriate alternative because it agrees perfectly with what appears to be the intuition behind the CAPM. Indeed, the APT is based on a linear return generating process *as a first principle*, and requires no utility assumptions beyond monotonicity and concavity. Nor is it restricted to a single period; it will hold in both the multiperiod and single period cases. Though consistent with every conceivable prescription for portfolio diversification, no particular portfolio plays a role in the APT. Unlike the CAPM, there is no requirement that the market portfolio be mean variance efficient.

There are two major differences between the APT and the original Sharpe [50] "diagonal" model, a single factor generating model which we believe is the intuitive grey eminence behind the CAPM. First, and most simply, the APT allows more than just one generating factor. Second, the APT demonstrates that since any market equilibrium must be consistent with no arbitrage profits, every equilibrium will be characterized by a linear relationship between each asset's expected return and its return's response amplitudes, or loadings, on the common factors. With minor caveats, given the factor generating model, the absence of riskless arbitrage profits—an easy enough condition to accept a priori—leads immediately to the APT. Its modest assumptions and its pleasing implications surely render the APT worthy of being the object of empirical testing.

To our knowledge, though, there has so far been just one published empirical study of the APT, by Gehr [17]. He began with a procedure similar to the one reported here. We can claim to have extended Gehr's analysis with a more comprehensive set of data (he used 24 industry indices and 41 individual stocks) and to have carried the analysis farther—to a stage actually required if the tests are to be definitive. Nonetheless, Gehr's paper is well worth reading and it must be given precedence as the first empirical work directly on this subject.

Another empirical study related to the APT is an early paper by Brennan [6], which is unfortunately still unpublished. Brennan's approach was to decompose the residuals from a market model regression. He found two factors present in the residuals and concluded that "the true return generating process must be represented by at least a two factor model rather than by the single factor diagonal model" (p. 30). Writing before the APT, Brennan saw clearly that "it is not possible to devise cross-sectional tests of the Capital Asset Pricing Model, since only in the case of a single factor model is it possible to relate *ex ante* and *ex post* returns" (p. 34). Of course, the APT's empirical usefulness rests precisely in its ability to permit such cross-sectional tests whether there is one factor or many.

The possibility of multiple generating factors was recognized long ago. Farrar

[15] and King [22], for example, employed factor analytic methods. Their work focused on industry influences and was pure (and very worthwhile) empiricism. Since the APT was not available to predict the cross-sectional effects of industry factors on expected returns, no tests were conducted for the presence of such effects.

More recently, Rosenberg and Marathe [44] have analyzed what they term "extra-market" components of return. They find unequivocal empirical support for the presence of such components. Rosenberg and Marathe's work employs extraneous "descriptor variables" to predict intertemporal changes in the CAPM's parameters. They state that "the appropriateness of the multiple-factor model of security returns, with loadings equal to predetermined descriptors, as opposed to a single-factor or market model, is conclusively demonstrated" (p. 100). But, they do not ascertain the *separate* influences of these multiple factors on individual expected returns, and focus instead on a combined influence working through the market portfolio. In other words, they assume the CAPM and decompose the single market beta into its constitutent parts.

Regarding the market portfolio as a construct which captures the influences of many factors follows the theoretical ideas in Rosenberg [45] and Sharpe [52]. Thus, Rosenberg and Marathe's work does not provide a definitive test of the APT.

There are a number of other recent papers which are more or less related to this one. In particular, Langetieg [25], Lee and Vinso [28], and Meyers [36] contain evidence of more than just a single market factor influencing returns. In contrast, Kryzanowski and To [24] give a formal test for the presence of additional factors but find "that only the first factor is non-trivial" (p. 23).

Nevertheless, there seems to be enough evidence in past empirical work to conclude that there may exist multiple factors in the returns generating processes of assets. The APT provides a solid theoretical framework for ascertaining whether those factors, if they exist, are "priced," i.e., are associated with risk premia. The purpose of our paper is to use the APT framework to investigate both the existence and the pricing questions.

In the following section, (I), a more complete discussion of the unique testable features of the APT is provided. Then section II gives our basic tests. It concludes that three factors are definitely present in the "prices" (actually in the expected returns) of equities traded on the New York and American Exchanges. A fourth factor may be present also but the evidence there is less conclusive.

Sections III and IV present two additional tests of the APT. The most important and powerful is in section III, where the APT is compared against a specific alternative hypothesis that "own" variance influences expected returns. If the APT is true, the "own" variance should not be important, even though its sample value is known to be highly correlated cross-sectionally with sample mean returns. We find that the "own" variance's sample influence arises spuriously from skewness in the returns distribution.

In section IV, we present a test of the consistency of the APT across groups of assets. Although the power of this test is probably weak, it gives no indication whatsoever of differences among groups.

Our conclusion is that the APT performs well under empirical scrutiny and that it should be considered a reasonable model for explaining the cross-sectional variation in average asset returns.

I. The APT and its Testability

A. The APT

This section outlines the APT in a fashion that makes it suitable for empirical work. A detailed development of theory is presented in Ross [47, 48] and the intent here is to highlight those conclusions of the theory which are tested in subsequent sections.

The theory begins with the traditional neoclassical assumptions of perfectly competitive and frictionless asset markets. Just as the CAPM is derived from the assumption that random asset returns follow a multivariate normal distribution, the APT also begins with an assumption on the return generating process. Individuals are assumed to believe (homogeneously) that the random returns on the set of assets being considered are governed by a k-factor generating model of the form:

$$\tilde{r}_i = E_i + b_{i1}\tilde{\delta}_1 + \cdots + b_{ik}\tilde{\delta}_k + \tilde{\epsilon}_i,$$
$$i = 1, \cdots, n. \tag{1}$$

The first term in (1), E_i, is the expected return on the i^{th} asset. The next k terms are of the form $b_{ij}\tilde{\delta}_j$ where $\tilde{\delta}_j$ denotes the mean zero j^{th} factor common to the returns of all assets under consideration. The coefficient b_{ij} quantifies the sensitivity of asset i's returns to the movements in the common factor $\tilde{\delta}_j$. The common factors capture the systematic components of risk in the model. The final term, $\tilde{\epsilon}_i$, is a noise term, i.e., an unsystematic risk component, idiosyncratic to the i^{th} asset. It is assumed to reflect the random influence of information that is unrelated to other assets. In keeping with this assumption, we also have that

$$E\{\tilde{\epsilon}_i \mid \tilde{\delta}_j\} = 0,$$

and that $\tilde{\epsilon}_i$ is (quite) independent of $\tilde{\epsilon}_j$ for all i and j. Too strong a dependence in the $\tilde{\epsilon}_i$'s would be like saying that there are more than simply the k hypothesized common factors. Finally, we assume for the set of n assets under consideration, that n is much greater than the number of factors, k.

Before developing the theory, it is worth pausing to examine (1) in a bit more detail. The assumption of a k-factor generating model is very similar in spirit to a restriction on the Arrow-Debreu tableau that displays the returns on the assets in different states of nature. If the $\tilde{\epsilon}_i$ terms were omitted, then (1) would say that each asset i has returns r_i that are an exact linear combination of the returns on a riskless asset (with identical return in each state) and the returns on k other factors or assets or column vectors, $\delta_1, \cdots, \delta_k$. In such a setting, the riskless return and each of the k factors can be expressed as a linear combination of k + 1 other returns, say r_1 through r_{k+1}. Any other asset's return, since it is a linear combination of the factors, must also be a linear combination of the first k + 1 assets' returns. And thus, portfolios of the first $k+1$ assets are perfect substitutes for all other assets in the market. Since perfect substitutes must be priced equally, there must be restrictions on the individual returns generated by the model. This is the core of the APT: there are only a few systematic components of risk existing in nature. As a consequence, many portfolios are close substitutes and as such, they must have the same value.

What are the common or systematic factors? This question is equivalent to asking what causes the particular values of covariance terms in the CAPM. If there are only a few systematic components of risk, one would expect these to be related to fundamental economic aggregates, such as GNP, or to interest rates or weather (although no causality is implied by such relations). The factor model formalism suggests that a whole theoretical and empirical structure must be explored to better understand what economic forces actually affect returns systematically. But in testing the APT, it is no more appropriate for us to examine this issue than it would be for tests of the CAPM to examine what, if anything, causes returns to be multivariate normal. In both instances, the return generating process is taken as one of the primitive assumptions of the theory. We do consider the basic underlying causes of the generating process of returns to be a potentially important area of research, but we think it is an area that can be investigated separately from testing asset pricing theories.

Now let us develop the APT itself from the return generating process (1). Consider an individual who is currently holding a portfolio and is contemplating an alteration of his portfolio. Any new portfolio will differ from the old portfolio by investment proportions x_i $(i = 1, \cdots, m)$, which is the dollar amount purchased or sold of asset i as a fraction of total invested wealth. The sum of the x_i proportions,

$$\sum_i x_i = 0,$$

since the new portfolio and the old portfolio put the same wealth into the n assets. In other words, additional purchases of assets must be financed by sales of others. Portfolios that use no wealth such as $\underline{x} \equiv (x_1, \cdots x_n)$[1] are called arbitrage portfolios.

In deciding whether or not to alter his current holdings, an individual will examine all the available arbitrage portfolios. The additional return obtainable from altering the current portfolio by n is given by

$$\begin{aligned} \underline{x\tilde{r}} &\equiv \sum_i x_i\tilde{r}_i \\ &= (\sum_i x_iE_i) + (\sum_i x_ib_{i1})\tilde{\delta}_1 + \cdots + (\sum_i x_ib_{ik})\tilde{\delta}_k + \sum_i x_i\tilde{\epsilon}_i \\ &\equiv \underline{xE} + (\underline{xb}_1)\tilde{\delta}_1 + \cdots + (\underline{xb}_k)\tilde{\delta}_k + \underline{x\tilde{\epsilon}}. \end{aligned}$$

Consider the arbitrage portfolio chosen in the following fashion. First, we will keep each element, x_i, of order $1/n$ in size; i.e., we will choose the arbitrage portfolio $\underline{x}$ to be well diversified. Second, we will choose $\underline{x}$ in such a way that it

[1] An underscored symbol indicates a vector or matrix.

has no systematic risk; i.e., for each j

$$\underline{xb_j} \equiv \sum_i x_i b_{ij} = 0.$$

Any such arbitrage portfolio, $\underline{x}$, will have returns of

$$\begin{aligned} \underline{x\tilde{r}} &= (\underline{xE}) + (\underline{xb_1})\tilde{\delta}_1 + \cdots + (\underline{xb_k})\tilde{\delta}_k + (\underline{x\tilde{\epsilon}}\,) \\ &\approx \underline{xE} + (\underline{xb_1})\tilde{\delta}_1 + \cdots + (\underline{xb_k})\tilde{\delta}_k \\ &= \underline{xE}. \end{aligned}$$

The term $(\underline{x\tilde{\epsilon}})$ is (approximately) eliminated by applying the law of large numbers. For example, if σ^2 denotes the average variance of the $\tilde{\epsilon}_i$ terms, and if, for simplicity, each x_i exactly equals $\pm 1/n$, then

$$\begin{aligned} \operatorname{Var}(\underline{x\tilde{\epsilon}}) &= \operatorname{Var}(1/n \textstyle\sum_i \tilde{\epsilon}_i) \\ &= [\operatorname{Var}(\tilde{\epsilon}_i)]/n^2 \\ &= \sigma^2/n, \end{aligned}$$

where we have assumed that the ϵ_i are mutually independent. It follows that for large numbers of assets, the variance of $x\tilde{\epsilon}$ will be negligible, and we can diversify away the unsystematic risk.

Recapitulating, we have shown that it is possible to choose arbitrage portfolios with neither systematic *nor* unsystematic risk terms! If the individual is in equilibrium and is content with his current portfolio, we must also have $\underline{XE} = 0$. No portfolio is an equilibrium (held) portfolio if it can be improved upon without incurring additional risk or committing additional resources.

To put the matter somewhat differently, in equilibrium all portfolios of these n assets which satisfy the conditions of using no wealth and having no risk must also earn no return on average.

The above conditions are really statements in linear algebra. Any vector, $\underline{x}$, which is orthogonal to the constant vector and to each of the coefficient vectors, $\underline{b_j}$ $(j = 1, \cdots, k)$, must also be orthogonal to the vector of expected returns. An algebraic consequence of this statement is that the expected return vector, $\underline{E}$, must be a linear combination of the constant vector and the $\underline{b_j}$ vectors. In algebraic terms, there exist $k + 1$ weights, $\lambda_0, \lambda_1, \cdots, \lambda_k$ such that

$$E_i = \lambda_0 + \lambda_1 b_{i1} + \cdots + \lambda_k b_{ik}, \qquad \text{for all} \quad i. \tag{2}$$

If there is a riskless asset with return, E_0, then $b_{0j} = 0$ and

$$E_0 = \lambda_0,$$

hence we will write

$$E_i - E_0 = \lambda_1 b_{i1} + \cdots + \lambda_k b_{ik},$$

with the understanding that E_0 is the riskless rate of return if such an asset exists, and is the common return on all "zero-beta" assets, i.e., assets with $b_{ij} = 0$, for all j, whether or not a riskless asset exists.

If there is a single factor, then the APT pricing relationship is a line in expected return, E_i, systematic risk, b_i, space:

$$E_i - E_0 = \lambda b_i.$$

Figure 1 can be used to illustrate our argument geometrically. Suppose, for example, that assets 1, 2, and 3 are presently held in positive amounts in some portfolio and that asset 2 is above the line connecting assets 1 and 3. Then a portfolio of 1 and 3 could be constructed with the same systematic risk as asset 2, but with a lower expected return. By selling assets 1 and 3 in the proportions they represent of the initial portfolio and buying more of asset 2 with the proceeds, a new position would be created with the same overall risk and a greater return. Such arbitrage opportunities will be unavailable *only* when assets lie along a line. Notice that the intercept on the expected return axis would be E_0 when no arbitrage opportunities are present.

The pricing relationship (2) is the central conclusion of the APT and it will be the cornerstone of our empirical testing, but it is natural to ask what interpretation can be given to the λ_j factor risk premia. By forming portfolios with unit systematic risk on each factor and no risk on other factors, each λ_j can be interpreted as

$$\lambda_j = E^j - E_0,$$

the excess return or market risk premium on portfolios with only systematic factor j risk. Then (2) can be rewritten as,

$$E_i - E_0 = (E^1 - E_0)b_{i1} + \cdots + (E^k - E_0)b_{ik}. \tag{3}$$

Is the "market portfolio" one such systematic risk factor? As a well diversified portfolio, indeed a convex combination of diversified portfolios, the market

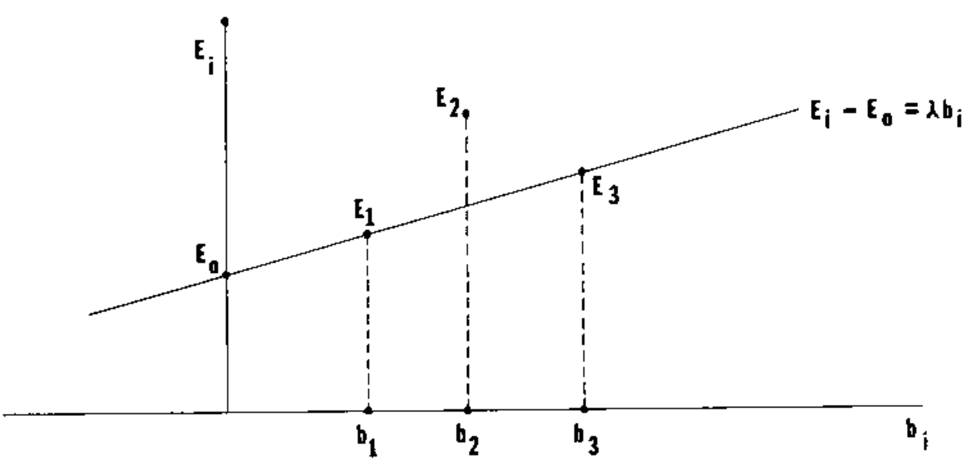

Figure 1.

portfolio probably should not possess much idiosyncratic risk. Thus, it might serve as a substitute for one of the factors. Furthermore, individual asset b's calculated against the market portfolio would enter the pricing relationship and the excess return on the market would be the weight on these b's. But, it is important to understand that *any* well-diversified portfolio could serve the same function and that, in general, k well-diversified portfolios could be found that approximate the k factors *better* than any single market index. In general, the market portfolio plays no special role whatsoever in the APT, unlike its pivotal role in the CAPM, (Cf. Roll [41, 42] and Ross [49]).

The lack of a special role in the APT for the market portfolios is particularly important. As we have seen, the APT pricing relationship was derived by considering any set of n assets which followed the generating process (1). In the CAPM, it is crucial to both the theory and the testing that all of the universe of available assets be included in the measured market portfolio. By contrast, the APT yields a statement of relative pricing on subsets of the universe of assets. As a consequence, the APT can, in principle, be tested by examining only subsets of the set of all returns. We think that in many discussions of the CAPM, scholars were actually thinking intuitively of the APT and of process (1) with just a single factor. Problems of identifying that factor and testing for others were not considered important.

To obtain a more precise understanding of the factor risk premia, $E^j - E_0$, in (3), it is useful to specialize the APT theory to an explicit stochastic environment within which individual equilibrium is achieved. Since the APT is valid in intertemporal as well as static settings and in discrete as well as in continuous time, the choice of stochastic models is one of convenience alone. The only critical assumption is the returns be generated by (1) over the shortest trading period.

A particularly convenient specialization is to a rational anticipations intertemporal diffusion model. (See Cox, Ingersoll and Ross [8] for a more elaborate version of such a model and for the relevant literature references.) Suppose there are k exogenous, independent (without loss of generality) factors, s^j, which follow a multivariate diffusion process and whose current values are sufficient statistics to determine the current state of the economy. As a consequence, the current price, p_i, of each asset i will be a function only of $\underset{\sim}{s} = (s^1, \cdots, s^k)$ and the particular fixed contractual conditions which define that asset in the next differential time unit. Similarly the random return, dr_i, on asset i will depend on the random movements of the factors. By the diffusion assumption we can write

$$d\tilde{r}_i = E_i\, dt + b_{i1}\, d\tilde{s}^1 + \cdots + b_{ik}\, d\tilde{s}^k. \tag{4}$$

It follows immediately that the conditions of the APT are satisfied exactly—with $d\tilde{\epsilon}_i = 0$ and the APT pricing relationship (3) must hold exactly to prevent arbitrage. In this setting, however, we can go further and examine the premia, $E^j - E_0$, themselves.

If individuals in this economy are solving consumption withdrawal problems, then the current utility of future consumption, e.g., the discounted expected value of the utility of future consumption, V, will be a function only of the individual's current wealth, w, and the current state of nature, $\underset{\sim}{s}$. The individual will optimize by choosing a consumption withdrawal plan, c, and an optimal portfolio choice, $\underline{x}$, so as to maximize the expected increment in V; i.e.,

$$\max_{\underline{x}, c} E\{dV\}.$$

At an optimum, consumption will be withdrawn to the point where its marginal utility equals the marginal utility of wealth,

$$u'(c) = V_w.$$

The individual portfolio choice will result from the optimization of a locally quadratic form exactly as in the static CAPM theory with the additional feature that covariances of the change in wealth, $d\tilde{w}$, with the changes in state variables, $d\tilde{s}^i$, will now be influenced by portfolio choice and will, in general, alter the optimal portfolio. By solving this optimization problem and using the marginal utility condition, $u'(c) = V_w$, the individual equilibrium sets factor risk premia equal to

$$E^j - E_0 = (R/c)(\partial c/\partial s^j)\sigma_j^2;$$

where $R = -(wV_{ww})/V_w$, the individual coefficient of relative risk aversion and σ_j^2 is the local variance of (independent) factor s_j. (The interested reader is referred to Cox, Ingersoll and Ross [8] for details.) Notice that the premia $E^j - E_0$ can be negative if consumption moves counter to the state variable. In this case portfolios which bear positive factor s^j risk hedge against adverse movements in consumption, but too much can be made of this, since by simply redefining s^j to be $-s^j$ the sign can be reversed. The sign, therefore, is somewhat arbitrary and we will assume it is normalized to be positive. Aggregating over individuals yields (3).

One special case of particular interest occurs when state dependencies can be ignored. In the log case, $R = 1$, for example, or any case with a relative wealth criteria (see Ross [48]) the risk premia take the special form

$$E^j - E_0 = R(\textstyle\sum_i x_i b_{ij})\sigma_j^2$$

where $\underline{x}$ is the individual optimal portfolio. This form emphasizes the general relationship between b_j and σ_j^2. Normalizing $\sum_i x_i b_{ij}$ to unity by scaling s^j, we have

$$E^j - E_0 = R\sigma_j^2.$$

The risk premium of factor j is proportional to its variance and the constant of proportionality is a measure of relative risk aversion.

For other utility functions, individual consumption vectors can be expressed in terms of portfolios of returns and similar expressions can be obtained. In effect, since the weighted state consumption elasticities for all individuals satisfy the APT pricing relationships, they must all be proportional.[2]

[2] Breeden [5] has developed the observation that homogenous beliefs about E's and b's imply perfect correlation between individual random consumption changes. His results depend on the assumption, made also by APT, that $k < N$.

The risk premium can be written in general as

$$E^j - E_0 = \left[\sum_l w_l R_l \left(\frac{1}{c_l} \right) \frac{\partial c_l}{\partial s_j} \right] \sigma_j^2$$

where l indexes individual agents, w_l is the proportion of total wealth held by agent l, R_l is his coefficient of relative risk aversion, $\frac{1}{c_l}\frac{\partial c_l}{\partial s_j}$ is the partial elasticity of his consumption with respect to changes in the jth factor, and σ_j^2 is the variance of the jth factor. Not very much is known about the term in parentheses and, all other things being equal, about all we can conclude is that risk premia should be larger, the larger the own variance of the factor. We would not expect this result to be specialized to the diffusion model and, in general, we would expect, with beta weights appropriately normalized, that factors with larger own variances would have larger associated risk premia.[3]

Let us return now to the general APT model and aggregate it to a testable market relationship. The key point in aggregation is to make strong enough assumptions on the homogeneity of individual anticipations to produce a testable theory. To do so with the APT we need to assume that individuals agree on both the factor coefficients, b_{ij}, and the expected returns, E_i. It now follows that the pricing relationship (2) which holds for each individual holds at the market level as well. Notice that individual and aggregate risk premia must coincide when there are homogenous beliefs on the expected returns and the factor coefficients.

As with the CAPM, the purpose of assuming homogenous anticipations is not to facilitate the algebra of aggregation. Rather, it is to take the final step to a testable theory. We can now make the rational anticipations assumption that (1) not only describes the *ex ante* individual perceptions of the returns process but also that *ex post* returns are described by the same equation. This fundamental intertemporal rationality assumption permits the *ex ante* theory to be tested by examining *ex post* data. In the next section we will discuss the possibilities for empirical testing which derive from this assumption.

B. Testing the APT

Our empirical tests of the APT will follow a two step procedure. In the first step, the expected returns and the factor coefficients are estimated from time series data on individual asset returns. The second step uses these estimates to test the basic cross-sectional pricing conclusion, (2), of the APT. This procedure is analogous to familiar CAPM empirical work in which time series analysis is used to obtain market betas, and cross-sectional regressions are then run of expected returns, estimated for various time periods, on the estimated betas. While flawed in some respects, the two step procedure is free of some major conceptual difficulties in CAPM tests. In particular, the APT applies to subsets

[3] We have not, of course, developed a complete rational anticipations model in diffusion setting, but it should be clear from this outline that the APT is compatible with the more specific results of Merton [35], Lucas [31], Cox, Ingersoll, and Ross [8], and Ross [48].

of the universe of assets; this eliminates the need to justify a particular choice of a surrogate for the market portfolio.

If we assume that returns are generated by (1), then the basic hypothesis we wish to test is the pricing relationship,

H_0: There exist non-zero constants, $(E_0, \lambda_1, \cdots, \lambda_k)$

such that

$$E_i - E_0 = \lambda_1 b_{i1} + \cdots + \lambda_k b_{ik}, \quad \text{for all } i.$$

The theory should be tested by its conclusions, not by its assumptions. One should not reject the APT hypothesis that assets were priced as if (2) held by merely observing that returns do not exactly fit a k-factor linear process. The theory says nothing about how close the assumptions must fit. Rejection is justified only if the conclusions are inconsistent with the observed data.[4]

To estimate the b coefficients, we appeal to the statistical technique of factor analysis. In factor analysis, these coefficients are called factor loadings and they are inferred from the sample covariance matrix, $\underline{\hat{V}}$. From (1), the population variance, $\underline{V}$, is decomposed into

$$\underline{V} = \underline{B\Lambda B'} + \underline{D}\,, \tag{5}$$

where $\underline{B} = [b_{ij}]$ is the matrix of factor loadings, $\underline{\Lambda}$ is the matrix of factor covariances, and $\underline{D}$ is the diagonal matrix of own asset variances, $\sigma_i^2 = E\{\tilde{\epsilon}_i^2\}$.

From (5), $\underline{V}$ will be unaltered by any transformation which leaves $\overline{B\Lambda B'}$ unaltered. In particular, if $\underline{G}$ is an orthogonal transformation matrix, $\underline{GG'} = \mathrm{I}$, then

$$\begin{aligned}\underline{V} &= \underline{B\Lambda B'} + \underline{D}\\ &= \underline{BGG'\Lambda GG'B'} + \underline{D}\\ &= (\underline{BG})(\underline{G'\Lambda G})(\underline{BG})' + \underline{D}\end{aligned}$$

If $\underline{B}$ is to be estimated from $\underline{\hat{V}}$, then all transforms $\underline{BG}$ will be equivalent. For example, it clearly makes no difference in (1) if the first two factors switch places. More importantly, we could obviously scale up factor j's loadings and scale down factor j by the same constant g and since $bz_{ij}\hat{\delta}_j = gb_{ij}\left(\frac{1}{g}\hat{\delta}_j\right)$ the distributions of returns would be unaltered. To some extent we can eliminate ambiguity by restricting the factors to be orthonormal so they are independent and have unit variance. Alternatively, we could maintain the independence of the factors and construct the loadings for each factor to have a particular norm value, e.g., to

[4] This is a strongly positive view. Testing the APT involves testing H_0 and *not* testing the k-factor model. The latter tests may be of interest in their own right just as any examination of the distribution of returns is of interest, but it is irrelevant for the APT. As Friedman [16, pp. 19–20] points out: one would not be inclined to reject the hypothesis that the leaves on a tree arranged themselves so as to maximize the amount of sunlight they received by observing that trees did not have free will. Similarly, one should not reject the conclusions derived from firm profit maximization on the basis of sample surveys in which managers claim that they trade off profit for social good.

sum to 1 (or −1) and let the factor variances vary. From a theoretical viewpoint these are all equivalent constraints. While they alter the form of the APT null hypotheses, H_0, the statistical rejection region is unaffected.

To see this note that if

$$E_i - E_0 = \underline{b}_i \cdot \lambda,$$

or, in matrix form,

$$\underline{E} - E_0 = \underline{B}\lambda,$$

then

$$\underline{E} - E_0 = (\underline{BG})(\underline{G}'\underline{\lambda})$$

and the linear hypothesis remains true with the exact weights altered by the orthogonal transform.[5] This is a very sensible result. The APT concludes that excess expected returns lie in the space spanned by the factor loadings. Orthogonal transforms leave that space unchanged, altering only the directions of the defining basis vectors, the column vectors of the loadings. As a consequence, we will adopt a statistically convenient restriction to estimate $\underline{B}$, keeping the arbitrariness of the procedure in mind. Notice that this is quite different from the ordinary uses of factor analysis. We are not "rotating" the factors in an arbitrary fashion to try to "interpret" them. Rather, our results are independent of the rotation chosen.

Once the expected returns, E_i, and the loadings, $\underline{B}$, have been estimated, we can then move to the test of H_0. The general procedure is to examine cross-sectional regressions of the form

$$E_i = E_0 + \lambda_1 \hat{b}_{i1} + \cdots + \lambda_k \hat{b}_{ik},$$

where E_0 and $\lambda_1, \cdots, \lambda_k$ are to be estimated. The theory will not be rejected if the joint hypothesis that $\lambda_1 = \cdots = \lambda_k = 0$, is rejected. This is the usual state of statistical testing; we cannot "prove" that a theory is true against an unspecified alternative. We can only fail to reject it.

In Section III a specific alternative will be proposed, namely that the "own" variances, $\sigma^2_{\epsilon_i}$, affect excess returns, and the APT will be tested against this alternative. (This is probably the standard structure which most tests of the APT will take. A specific alternative will be proposed in which some idiosyncratic feature of the assets not reflected in their loadings is hypothesized to explain returns.)

We deal with the specifics of the above tests below, but for the present point out some of the major deficiencies of the procedure. The estimates of b_{ij} found in

[5] Notice, that if we knew the λ_i weights, we could obviously use them to aggregate the factors into a single factor which "explains" excess returns. In this trivial sense the number of factors does not matter. Without further assumptions, though, this begs the question since the λ_i weights must first be estimated to find the proper combination of the factors. For example, if we chose G such that its first column is proportional to λ, then $G'\lambda$ will be a vector with only the first entry non-zero. Under this rotation only a single factor is used to explain excess returns, but as noted above, the result has no empirical content.

the first step are, of course, just estimates and, as such, are subject to sampling error. Let $\tilde{e}_i$ and $\tilde{\beta}_{ij}$ denote the respective sample errors,

$$\hat{E}_i = E_i + \tilde{e}_i$$

and

$$\hat{b}_{ij} = b_{ij} + \tilde{\beta}_{ij}.$$

Under the null hypothesis, then, the cross-sectional regression for any period will be of the form

$$\begin{aligned}\hat{E}_i &= E_i + \tilde{e}_i \\ &= E_0 + \lambda_1 b_{i1} + \cdots + \lambda_k b_{ik} + \tilde{e}_i \\ &= E_0 + \lambda_1 \hat{b}_{i1} + \cdots + \lambda_k \hat{b}_{ik} + \tilde{\xi}_i,\end{aligned}$$

where the regression error

$$\tilde{\xi}_i \equiv \tilde{e}_i - \lambda_1 \tilde{\beta}_{i1} - \cdots - \lambda_k \tilde{\beta}_{ik}.$$

Since the factor analytic estimation procedure to be employed is a maximum likelihood procedure, in a multivariate normal world the estimates will be asymptotically consistent; but very little is known about their small sample properties. In general, we expect $\tilde{\xi}_i$ to be correlated with $\hat{b}_i$ and the cross-sectional regression to suffer from the usual errors-in-variables problems. Clearly, there is a considerable amount of statistical analysis to be carried out before one can feel comfortable with this approach. As a consequence, we stress the tentative and "first try" nature of the empirical work which follows.

II. Empirical Results

A. Data

The data are described in Table I. In selecting them, several more or less arbitrary choices were necessary. For instance, although daily data were available through 1977, the calculations reported in this paper used data only through 1972. The motivation was to secure a calibration or "holdout" sample without sacrificing the advantages of a large estimation sample, large enough for some statistical reliability even after aggregating the basic daily returns into monthly returns. The calibration sample is thereby reserved for later replication and for investigation of problems such as non-stationarity. The cutoff data of 31 December 1972 was selected also to correspond with other published studies of asset pricing, most of which used a pre-1973 period. This should facilitate a comparison of the results.

In our empirical analysis, estimated covariance matrices of returns were computed for groups of individual assets. Calculation of covariances necessitates simultaneous observations—so the beginning and ending dates were specified in order to exclude exceedingly short-lived securities. Although this assured a reasonably large time series sample for every group, there remained some variation across groups in number of observations. This was due evidently to suspen-

Table I

Data Description

Source:	Center for Research in Security Prices Graduate School of Business University of Chicago Daily Returns File
Selection Criterion:	By alphabetical order into groups of 30 individual securities from those listed on the New York or American Exchanges on *both* 3 July 1962 and 31 December 1972. The (alphabetically) last 24 such securities were not used since complete groups of 30 were required.
Basic Data Unit:	Return adjusted for all capital changes and including dividends, if any, between adjacent trading days; i.e., $[(p_{j,t} + d_{j,t})/p_{j,t-1}] - 1$, where p = price, d = dividend, j = security index, t = trading day index.
Maximum Sample Size per Security:	2619 daily returns
Number of Selected Securities	1260, (42 groups of 30 each)

sion of trading, temporary delisting, or simply to missing data for individual securities. None of the 42 groups contained data for all 2619 trading days. The minimum sample size was still 1445 days, however, and only three groups had less than 2000 days. Thirty-six groups (86%) had at least 2400 observations.

The group size of 30 individual securities was a compromise. For some purposes, such as estimating the number of return generating factors present in the economy, the best group size would have included *all* individual assets; but this would have dictated a covariance matrix larger than the processing capacity of the computer. For other purposes, such as comparing covariance structures across groups, statistical power increases with the number of groups, *cet. par.* Unfortunately, the *ceteris* are not *paribus*; for the number of securities *per* group also improves power and the reliability of estimates. We guessed that 30 securities per group would confer reasonable precision for all of the tests envisaged initially and we stuck with 30 as the work proceeded.

B. Estimating the Factor Model

The analysis proceeds in the following stages:

1) For a group of individual assets, (in this case, a group of 30 selected alphabetically), a sample product-moment covariance matrix is computed from a time series of returns, (of New York and American Exchange listed stocks from July 1962 through December 1972).
2) A maximum-likelihood factor analysis is performed on the covariance matrix. This estimates the number of factors and the matrix of loadings.
3) The individual-asset factor loading estimates from the previous step are used to explain the cross-sectional variation of individual estimated expected returns. The procedure here is similar to a cross-sectional generalized least squares regression.
4) Estimates from the cross-sectional model are used to measure the size and statistical significance of risk premia associated with the estimated factors.

This procedure is similar to estimating the size and significance of factor "scores."

5) Steps (1) through (4) are repeated for all groups and the results are tabulated.

The first stage is straightforward and should require no further explanation. There was only one curiosity: every element in the covariance matrix was divided by one-half the largest of the 30 individual variances. This was done to prevent rounding error in the factor analysis and it has no effect whatever on the results since factor analysis is scale free.

In the second stage, an optimization technique suggested by Jöreskog [20] was employed in the form of a program described by Jöreskog and Sörbom [21]. There are several available choices of types of factor analysis. In addition to the maximum likelihood method, there are generalized least squares, unweighted least squares, and approximate methods, among others. The maximum-likelihood method is usually preferable since more is known about its statistical properties, (Cf. Lawley and Maxwell [26]). As we shall see later, however, there may be some problems attendant to the M.L.E. method because the likelihood function involved is that of a multivariate gaussian distribution. To the extent that the data have been generated by a non-gaussian probability law, unknown biases and inconsistencies may be introduced.

Assuming away these problems for the moment, the M.L.E. method provides the capability of estimating the number of factors. This can be accomplished by specifying an arbitrary number of factors, say k, then solving for the maximum likelihood conditional on a covariance matrix generated by exactly k factors. Of course k is set less than the number of securities in the group of 30. A second value of the likelihood function is also found; this one being conditional on the observed sample covariance matrix without any restriction as to number of factors. Then a likelihood ratio, (first likelihood value divided by second), is computed. Under the null hypothesis of exactly k factors, twice the natural logarithm of the likelihood ratio is distributed asymptotically as chi-square with $\frac{1}{2}[(n - k)^2 - (n + k)]$ degrees of freedom. Thus, if the computed chi-square statistic is large (small), then more (fewer) than k factors are required to explain the structure of the generating process. So $k + 1$ $(k - 1)$ factors are specified and another chi-square statistic is computed. The process terminates when the chi-square statistic indicates a pre-selected level, (usually 50%), that an additional factor is required.

We used the alphabetically first group of 30 securities to estimate the number of factors in the way just described, but with the added intention of retaining more factors than a 50% probability level would dictate. We could afford these extra, perhaps superfluous, factors since the third stage of our procedure provides a direct check on the true number of factors in the underlying generating process. An estimated factor introduced spuriously at the factor analysis stage would not be "priced" in the cross-sectional regression; its estimated coefficient should not differ significantly from zero. We wanted to allow the possibility of spurious factors because the same number of true (priced) factors should be present in every group *and* the first group might have been unrepresentative. Fewer than the true number of common factors could have been estimated for group one because of sampling variation. The third stage protects against too many factors estimated at stage two but it does not protect against too few.

For five factors using daily returns over the entire sample period, the chi-square statistic computed from the first group was 246.1. The number of degrees of freedom was 295 and the probability level (.980) implied only two chances in 100 that at least six factors were present in the data. Thus, we specified five factors, retaining this same number in the factor analysis computation for all 42 groups. Table II presents frequencies of the chi-square statistic for the 42 groups of daily returns. The monthly returns used later display a similar pattern.

As the table shows, in 38.1% of the groups, (16 of 42), the likelihood ratio test implied more than a 90% chance that five factors were sufficient. Over three-quarters of the groups had at least an even chance that five were enough. Some sampling variation in the estimated number of factors is inevitable; but the results indicate clearly that five is conservative in the sense of including, with high probability, at least as many estimated factors as there are true factors. Note, however, that a formal goodness-of-fit test using the results in Table II would not quite be legitimate. Since the original covariance matrices were computed over the same time period for all groups, there is probably some statistical dependence across the groups. Thus, the cross-group sample of any statistic is not likely to be a random sample. Since there is positive cross-sectional dependence among the returns, there is also likely to be positive cross group dependence in any statistic calculated from their returns.

With five factors, the model envisaged for each security can be written

$$\tilde{r}_{jt} \equiv \tilde{R}_{jt} - E_j = b_{1j}\tilde{\delta}_{1t} + \cdots + b_{5j}\tilde{\delta}_{5t} + \tilde{\epsilon}_{jt} \tag{6}$$

where R_{jt} is the daily return for day t and security j, E_j is the expected return for j, the b_j's are factor coefficients, the $\tilde{\delta}$'s are the true common factors, and $\tilde{\epsilon}_{jt}$ is a random disturbance completely unrelated to anything else including its own values in other periods. In matrix notation, a group of n individual securities whose returns conform to (6) can be expressed as

$$\tilde{\underline{r}}_t = \underline{B}\tilde{\underline{\delta}}_t + \epsilon_t$$

where $\tilde{\underline{r}}_t$ and $\tilde{\underline{\epsilon}}_t$ are $(n \times 1)$ column vectors, $\underline{B}$ is an $(n \times 5)$ matrix and $\tilde{\underline{\delta}}_t$ is a (5×1) vector. Without loss of generality, the factors can be assumed orthogonal and scaled to have unit variance. Then the null hypothesis represented by equation (6) implies that the covariance matrix of returns takes the form

$$\underline{V} = \underline{BB'} + \underline{D}$$

Table II

Probability that no more than five factors are needed to explain returns	.9	.8	.7	.6	.5	.4	.3	.2	.1	0
Frequency (%)	38.1	16.7	7.14	2.38	11.9	2.38	4.76	4.76	9.52	2.38

Cross-sectional distribution of the Chi-square statistic from a likelihood ratio test that no more than five factors are necessary to explain daily returns, 42 covariance matrices of 30 securities each, NYSE and AMEX listed securities, 1962–72.

where $\underline{D}$ is a (diagonal) matrix whose j^{th} diagonal element is the variance of $\tilde{\epsilon}_{jt}$.

As noted in Section I, although maximum likelihood factor analysis provides a unique estimate of $\underline{V}$, this estimate is compatible with an infinity of estimates for $\underline{B}$, "all equally good from a statistical point of view. In this situation, all the statistician can do is to select a particular solution, one which is convenient to find, and leave the experimenter to apply whatever rotation he thinks desirable" (Lawley and Maxwell [26, p. 11]).

Our program chooses an estimate $\hat{\underline{B}}$ of $\underline{B}$ such that the matrix $\hat{\underline{B}}'\hat{\underline{D}}^{-1}\hat{\underline{B}}$ is diagonal and arranged with its diagonal elements in descending order of magnitude. This constitutes a restriction that guarantees uniqueness, except that $-\hat{\underline{B}}$ is statistically equivalent and, in fact, any column of $\hat{\underline{B}}$ can be reversed in sign. The problem of sign reversal is solved quite easily for the restricted estimates, (see below), but the general non-uniqueness of factor loadings is very troublesome. Essentially, one cannot ascertain with certainty that the first factor in one group of securities is the same as the first factor in another group. For instance, factor number one in group A could conceivably correspond to factor number three in group $K (K \neq A)$. Thus, when the cross-sectional distributions of the loading coefficients are tabulated, there could be a mixing of estimates which apply to different "true" factors.

C. A First Test of the APT

The factor model can be written as

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

and the arbitrage pricing theory requires

$$\underline{E} = \lambda_0 + \underline{B}\underline{\lambda}.$$

Combining the two gives the basic factor process under the null hypothesis that the APT is true,

$$\tilde{\underline{r}}_t = \tilde{\underline{R}}_t - \lambda_0 = \underline{B}\underline{\lambda} + (\underline{B}\tilde{\underline{\delta}}_t + \tilde{\underline{\epsilon}}_t), \tag{7}$$

or, more compactly,

$$\tilde{\underline{r}}_t = \underline{B}\underline{\lambda} + \underline{\xi}_t, \tag{8}$$

where $\underline{\xi}_t$ is the mean zero disturbance at date t caused by intertemporal variation in the factors $\tilde{\underline{\delta}}_t$ and in the diversifiable component $\tilde{\underline{\epsilon}}_t$.

It might seem natural to test the APT via (8) by first estimating the factor loadings, $\underline{B}$, and the mean return vector $\bar{\underline{r}} = \Sigma \underline{r}_t/T$ *from time series, and then running a simple OLS cross-sectional regression analogous to* (8),

$$\bar{\underline{r}} = \hat{\underline{B}}\hat{\underline{\lambda}} + \hat{\underline{\xi}} \tag{9}$$

where $\hat{\underline{\lambda}}$ the OLS regression coefficients, would be the estimated risk premia. A closer examination of (7), however, reveals that this procedure would be biased toward finding risk premia for "priced" factors, even when their true prices are actually zero. To see why, notice that the mean value of $\underline{\delta}_t$, say $\bar{\underline{\delta}} = \Sigma \underline{\delta}_t/T$, must, with probability one, not be exactly zero in any sample. Thus, the cross-sectional regression (9) actually should be written

$$\bar{\underline{r}} = \hat{\underline{B}}(\underline{\lambda} + \bar{\underline{\delta}}) + \bar{\underline{\epsilon}}$$

so that $E(\hat{\underline{\lambda}}) = \underline{\lambda} + \bar{\underline{\delta}}$ will be biased by the time series *sample* mean of the factors, $\bar{\underline{\delta}}$. Of course, the bias should decrease with larger time series sample sizes, but since $\bar{\underline{\delta}}$ will not be exactly zero, however large the time series, $E(\hat{\underline{\lambda}}) \neq 0$ even when $\underline{\lambda} = 0$.

To correct this problem, we have employed a method analagous to that of Fama and MacBeth [14] but adapted to the factor analytic framework. The Fama-MacBeth procedure calculates a cross-sectional regression like (9) *for every time period t*,

$$\underline{r}_t = \hat{\underline{B}}\hat{\underline{\lambda}}_t + \hat{\underline{\xi}}_t$$

and then uses the time series of $\hat{\underline{\lambda}}_t$ to estimate the standard error of the average value of $\hat{\underline{\lambda}}$. This yields an inference about whether the true $\underline{\lambda}$ is non-zero.

A more efficient procedure exploits the factor analysis already conducted with the time series during the estimation of $\underline{B}$. The factor loadings $\hat{\underline{B}}$ are chosen such that $\hat{\underline{V}} \equiv \hat{\underline{B}}\,\hat{\underline{B}}' + \underline{D}$ is the estimated covariance matrix of $\underline{B}\,\hat{\underline{\delta}}_t + \hat{\underline{\epsilon}}_t$, the disturbance term in (7). Thus, a natural generalized least squares cross-sectional regression for each day t is

$$\hat{\underline{\lambda}}_t = (\hat{\underline{B}}'\hat{\underline{V}}^{-1}\hat{\underline{B}})^{-1}\hat{\underline{B}}'\hat{\underline{V}}^{-1}\underline{r}_t \equiv \underline{\Gamma}\underline{r}_t \tag{10}$$

which yields GLS estimates of the risk premia. Furthermore, it can be proven (Lawley and Maxwell [26, pp. 88–89]) that the covariance matrix of the estimates $\hat{\underline{\lambda}}_t$ from (10) is given by

$$\underline{B}'\underline{V}^{-1}\underline{B}. \tag{11}$$

This matrix is particularly convenient since it is constrained to be diagonal by the factor analysis. As a consequence, the estimated risk premia are mutually independent and admit simple t-tests of significance.

For instance, we will report below significance tests for

$$\bar{\hat{\underline{\lambda}}} \equiv \underline{\Gamma}\bar{\underline{r}} \tag{12}$$

whose covariance matrix is

$$\frac{1}{T}\underline{B}'\underline{V}^{-1}\underline{B}, \tag{13}$$

provided the returns are independent over time. Notice that the time series behavior of the estimated factor "scores," the $\hat{\underline{\delta}}$'s, is accounted for by the matrix $\underline{V}$, thereby eliminating the problem created by non-zero $\bar{\underline{\delta}}$ in the simple OLS cross-section (9).

There remain, however, some tricky econometric problems in this procedure. First, equation (11) ignores any estimation errors present in $\hat{\underline{B}}$. This means essentially that the significance tests for $\hat{\underline{\lambda}}$ are only asymptotically correct. There could be an understatement or an overstatement of significance for small samples. We have no way to ascertain the extent of this problem, but we doubt that it introduces a serious error because our sample sizes are "large" by usual statistical standards.

A second difficulty concerns the signs of $\hat{\underline{\lambda}}$. Since the factor loadings ($\underline{B}$) are not unique with respect to sign, neither are their coefficients $\underline{\lambda}$ in (7). Any rotated set of factors would have produced just as adequate a set of loadings. This implies that no importance can be ascribed to the numerical values of $\hat{\underline{\lambda}}$; only their statistical significance is relevant.

Finally, in the cross-sectional models (10) and (12), a value for the zero-beta or risk-free coefficient, λ_0 in (7), must be assumed. It might be thought that $\hat{\underline{\lambda}}_0$ could be obtained easily by adding a column of 1's to $\hat{\underline{B}}$ and computing regression (10) with an augmented matrix of loadings, $[\underline{1}\vdots\underline{B}]$ an augmented $\underline{\Gamma}$ and the total return R_t, in place of the excess return $\underline{r}_t$, as

$$\hat{\underline{\lambda}}_t = \underline{\Gamma}\underline{R}_t,$$

where $\hat{\underline{\lambda}}_t$ now contains an estimate for λ_0 as its first element. Unfortunately, although we report the result of this regression below, it is less satisfactory because the augmented covariance matrix of the estimated risk premia is

$$[\underline{1}\vdots\underline{B}]'\underline{V}^{-1}[\underline{1}\vdots\underline{B}]$$

which is not diagonal except in the fortuitous case when the constant vector is orthogonal to the loadings.

The trade-off, then, is between using a rather arbitrary value of λ_0 in the cross-sectional excess return regression (10) or allowing the data to determine $\hat{\lambda}_0$ but bearing the consequence that the estimates $\hat{\underline{\lambda}}$ are no longer statistically independent. In many applications, mutual independence is merely a nicety since F-tests can be used when dependence among the coefficients is present. In our case, however, constraining the sample design to the independent case is especially important because the $\hat{\underline{\lambda}}$'s at best are some unknown linear combinations of the true $\underline{\lambda}$'s and testing for the number of priced factors or non-zero λ_j's, is thereby reduced to a simple t-test.

Perhaps this will be clarified by considering the results in Table III. The top panel assumes a λ_0 of 6% per annum during the sample period, July 1962 through December 1972. The first results in Table III give the percentage of the groups in which more than a specified number of factors were associated with statistically significant risk premia, $\hat{\underline{\lambda}}$ estimated by (12) and (13). With daily data, 88.1% of the groups had at least one significant factor risk premium, 57.1% had two or more significant factors and in one-third of the groups at least three risk premia were significant. These percentages are far in excess of what would be expected by chance alone under the null hypothesis of no effect. The next row of Table III gives the relevant percentages which would be expected under this null hypothesis. If $\underline{\lambda} = 0$, the chance of observing at least a given number of $\hat{\lambda}_j$'s significant at the 95% level is the upper tail of the binomial distribution with probability of success $p = .05$. For example, the probability of observing at least two significant

Table III

Cross-sectional generalized least squares regressions of arithmetic mean sample returns on factor loadings, (42 groups of 30 individual securities per group, 1962–72 daily returns, standard errors of risk premia (λ) computed from time series)

1 FACTOR	2 FACTORS	3 FACTORS	4 FACTORS	5 FACTORS
I.	$\bar{R}_j - 6\% = \hat{\lambda}_1 \hat{b}_{j1} + \cdots + \hat{\lambda}_5 \hat{b}_{j5}$ (λ_0 assumed at 6%)			
Percentage of groups with at least this many factor risk premia significant at the 95% level				
88.1	57.1	33.3	16.7	4.8
Expected Percentage of groups with at least this many risk premia significant at the 95% level given no true risk premia ($\underline{\lambda} = \underline{0}$)				
22.6	2.26	.115	.003	.00003
Percentage of groups with factor's risk premium significant at the 95% level in natural order from factor analysis				
76.2	50.0	28.6	23.8	21.4
II.	$\bar{R}_j = \hat{\lambda}_0 + \hat{\lambda}_1 \hat{b}_{j1} + \cdots + \hat{\lambda}_5 \hat{b}_{j5}$ (λ_0 estimated)			
Percentage of groups with at least this many factor risk premia significant at the 95% level				
69.0	47.6	7.1	4.8	0
Percentage of groups with this factor's risk premium significant at the 95% level in natural order from factor analysis				
35.7	31.0	23.8	21.4	16.7

$\hat{\lambda}$'s, given $\underline{\lambda} = 0$, is $1 - (.95)^5 - 5(.05)(.95)^4 = .0226$. Notice that this calculation requires zero correlation among the $\hat{\lambda}_j$'s .

If, in fact, four factors are truly significant, then the 4.8 observed significance percentage for five factors (see line 1 of Table III), is almost precisely what one would expect at the 95% level. Similarly, if three are truly significant, the 16.7% of the groups in which at least four are found to be significant exceeds the 9.75% which would occur by chance alone. The disparity is much greater if less than three factors are significant. We can conclude then, that at least three factors are important for pricing, but that it is unlikely that more than four are present.

The second set of results, still with λ_0 assumed equal to 6%, report the percentage of groups in which the first, second, and remaining factors produced by the factor analysis have significant associated risk premia. As noted above, the first factor is selected as the one with the largest diagonal element in $\hat{B}'D^{-1}\hat{B}$, the second has the second largest diagonal element, and so forth, but there is no assurance that corresponding factors agree across different groups. Nevertheless, it is of some interest to examine the significance of the ordered factors and this is reported in the third line of Table III. As can be seen, all factors are significantly greater than the chance level (5%) with particularly heavy weight on the first two. The remaining three are significant, but this may be more a consequence of mixing the order of factors across the groups than of anything important.

The second part of Table III reports similar statistics but with the constant λ_0 estimated instead of assumed. Now the t statistics are no longer independent across the factors and we cannot apply the simple analysis above. But, the statistical results seem to conform well with the previous findings. Perhaps most striking is that at least two factors are significant in 47.6% of the groups while in only 7.1% are three or more significant. This suggests that the three significant factors obtained with λ_0 set equal to 6% may be an over-estimate due to the incorrect choice of the zero-beta return λ_0. When the intercept is estimated, two factors emerge as significant for pricing. However, because the $\hat{\lambda}_j$'s are not mutually independent, there is no standard of comparison for these percentages. As is to be expected, the results for the ordered factors are less significant than those for the λ_0 equal to 6% case, at least for the first and second factors produced by the factor analysis.

The next section (III) tests the APT against a specific alternative. Section IV presents a test for the equivalence of factor structure across the 42 groups.

III. Tests of the APT Against a Specific Alternative

In the previous section, we presented evidence that equity returns seem to depend on several common factors, perhaps as many as four. This many seem to be "priced", i.e., associated with non-zero risk premia which compensate for undiversifiable variation present in the generating process. Although these results are reassuring for the APT, there remains a possibility that other variables also are "priced" even though they are not related to undiversifiable risk. According to the theory, such variables should not explain expected returns; so if some were found to be empirically important, the APT would be rejected.

In this section, we report an investigation of one particular variable, the total variance of individual returns, or the "own" variance. The total variance would not affect expected returns if the APT is valid because its diversifiable component would be eliminated by portfolio formation and its non-diversifiable part would depend only upon the factor loadings and factor variances. It is a particularly good choice to use in an attempt to reject the APT because of its long-documented high positive correlation with sample mean returns.[6] If this sample correlation arises either from statistical estimation errors or else from its relation to factor loadings, the APT would enjoy an additional element of empirical support. If the correlation cannot be ascribed to these causes, however, then this would constitute evidence against the theory.

The procedure of this section is relatively straightforward: cross-sectionally (across individual assets), we regress estimates of expected returns on the five factor loading estimates described in the previous section *and* on

$$s_j = [\textstyle\sum_{t=1}^{T} (R_{jt} - R_j)/T]^{1/2}, \qquad j = 1, \cdots, N$$

the standard deviation of individual returns. This test is less efficient for detecting

[6] See, e.g., Douglas [10] and Lintner [30]. The "own" variance received very careful scrutiny in Miller and Scholes [37], and has been the object of recent theoretical inquiry in Levy [29].

"priced" factors than the factor analysis based test reported previously. Now, however, there is no alternative to using an ordinary regression approach since the extra variable s_j is not a factor loading and is not produced by the factor analysis.

Some evidence on the apparent explanatory power of the own standard deviation, s_j, is presented in Table IV. On average over the 42 groups of securities, the t-statistic (coefficient/standard error of coefficient) was 2.17 for s_j. 45.2% of the groups displayed statistically significant effects of s_j on mean sample returns at the 95% level of significance. In contrast, the F-test that at least some (one or more) factor loading had an effect on the mean return was significant at the 95% level for only 28.6% of the 42 groups.

A caution mentioned earlier in connection with all of our results should be reiterated: there was probably some positive dependence across groups, so the percentage of groups whose statistics exceed a critical value may overstate the actual significance of the relation between explanatory variables and expected returns. Nevertheless, the magnitude of the numbers would certainly appear to support a conclusion that the relation is statistically significant. The "explained"

Table IV

Cross-sectional regression[a] of estimated expected returns on factor loadings and individual total standard deviations of return (summary for 42 groups of 30 individual securities per group, 1962–72 daily returns)

Arithmetic Mean	Standard Error of Mean	Percentage of Groups Whose Statistic Exceeds 95% Critical Level[b]
Across 42 Groups		
t-statistic, test for most significant factor loading having no effect on expected return.		
2.19	.162	47.6
t-statistic: test for individual total standard deviation having no effect on expected return.		
2.17	.303	45.2
F-statistic: test for no effect by any factor loading on expected return (in addition to the effect of standard deviation).		
2.21	.295	28.6

[a] The regression equation for group g is

$$\bar{R}_j = \hat{\lambda}_{0g} + \hat{\lambda}_{1g}\hat{b}_{1j} + \cdots + \hat{\lambda}_{5g}\hat{b}_{5j} + \hat{\lambda}_{6g}s_j + \xi \quad j = 1, \cdots, 30$$

where $\bar{R}_j$ is the sample arithmetic mean return for security j, $\hat{b}_{kj}$ is security j's loading on factor k, the $\hat{\lambda}$'s are regression coefficients, s_j is individual asset j's total standard deviation of daily returns during the sample period and ξ_j is a residual.

[b] With 30 observations per group and six explanatory variables, the 95% critical value is 2.06 for the t-statistic and 2.64 for the F-statistic.

variation is quite high: the coefficient of multiple determination (adjusted R^2) is .743 on average over the 42 groups. Even the group with lowest explained variation has an R^2 of .561 (and recall that these are individual assets!). Without s included, the average adjusted R^2 is .563 and the minimum R^2 over the 42 groups is .166.

The apparently significant explanatory power of the "own" standard deviation (s) suggests that the arbitrage pricing theory may be false. Since arbitrageurs should be able to diversify away the non-common part of s, it should not be priced. There is reason, however, for a closer examination before rejecting the APT entirely.

A possible source of a spurious effect of the own variance on expected return is skewness in the distribution of individual returns. Positive skewness can create positive dependence between the *sample* mean and *sample* standard deviation (and vice versa for negative skewness). Miller and Scholes [37] argued convincingly that skewness could explain the sample mean's dependence on "own" variance. Our results below tend to support the Miller-Scholes argument within the APT context.

The distribution of individual daily returns are indeed highly skewed. Table V gives some sample results. As indicated there, 1213 out of 1260 individual assets, (96.3%), had positive estimated measures of skewness. There was considerable variation across assets, too. Although the sampling distribution of the skewness measure SK is not known and is difficult to tabulate even under the assumption of lognormality, there appears to be too much cross-sectional variation in SK to be ascribed to chance alone. Thus, individual assets probably differ in their population skewness. Note that intertemporal aggregation to monthly returns reduces the skewness only slightly.

Skewness is cross-sectionally correlated positively with the mean return and even more strongly with the standard deviation. Some part of this correlation may itself arise from sampling variation and some part too could be present in the population parameters. There is really no way to sort this out definitively. The strong cross-sectional regressions in the last panels of Table V suggest that attempts to expunge the spurious sampling dependence between sample mean return and standard deviation by exploiting the measured sample skewness, either as an additional variable in the cross-sectional regression or as a basis for skewness-sorted groups which might have less remaining spurious dependence, are probably doomed to weak and ambiguous results.[7] Also, such methods would be biased against finding a true effect of standard deviation, if one exists.

A procedure[8] which is charming in its simplicity and seems to resolve many of the statistical problems occasioned by skewness can be used if the observations are not too serially dependent: simply estimate each parameter from a different set of observations. In the present application, for example, we are concerned with sampling dependencies among estimates of all three parameters, expected return, factor loadings, and "own" standard deviation. If the time-series obser-

[7] As Martin [33] shows, using sample skewness and standard deviation both as additional explanatory variables causes severe econometric problems.

[8] We are grateful to Richard McEnally for suggesting this procedure.

Table V

Information About Skewness for Daily and Monthly Returns 1260 New York and American Listed Assets, 1962–72

Data Interval	Percent Positive	Mean	Standard Deviation	Smallest	Largest	Correlation with $\bar{R}$	Correlation with $\mathrm{Log}_e(s_j)$
			Product—Moment Skewness Measure, SK				
Daily	96.3	.681	.551	−2.06	4.56	.211	.452
Monthly	90.0	.654	.688	−1.04	5.94	.212	.520

Cross-Asset Regressions

Data Interval	$\hat{b}_1$	$t_{\hat{b}_1}$	$\hat{b}_2$	$t_{\hat{b}_2}$	Adjusted R^2
		$(\bar{R}_j - \bar{R}_{5,j})/\log_e(s_j) = \hat{b}_0 + \hat{b}_1 SK_j \quad j = 1, \cdots, 1260$			
Daily	1.82	20.6	—	—	.251
Monthly	1.14	36.6	—	—	.515
		$\bar{R}_j - \bar{R}_{5,j} = \hat{b}_0 + \hat{b}_1 SK_j + \hat{b}_2 \log_e(s_j) \quad j = 1, \cdots, 1260$			
Daily	5.17	7.87	29.2	68.0	.838
Monthly	7.33	24.6	8.17	32.1	.728

Definitions:

R_{jt} = Return for asset j in interval t

T = Total number of intervals in sample

$\bar{R}_j = \sum_t R_{jt}/T$

$\bar{R}_{5,j}$ = Sample mean after excluding the 25% smallest and 25% largest values of R_{jt}

$s_j = [\sum_t (R_{jt} - \bar{R}_j)^2/T]^{1/2}$

$SK_j = [\sum_t (R_{jt} - \bar{R}_j)^3/T\,]/s_j^3$

vations are temporarily uncorrelated such dependencies could be removed by using observations 1, 4, 7, 10, . . . to estimate the expected return, observations 2, 5, 8, 11, . . . to estimate the factor loadings, and 3, 6, 9, 12, . . . to estimate the standard deviation of returns. With complete intertemporal independence, there would be no sampling covariation among the estimates and only the cross-asset population relationships would remain.

The daily returns for each asset are indeed close to independent over time. There may be some slight negative dependence but it has a low order of magnitude. Unfortunately, this is not true for the squared returns. There is positive intertemporal dependence in absolute price changes or in squared changes. This implies that the standard deviation of returns and the factor loadings estimated from non-overlapping adjacent days would still retain some sampling dependence. But since there are so many available time series observations (2619), we have the luxury of skipping days and estimating the parameters from non-overlapping observations "insulated" be at least one day. Table VI summarizes the results obtained with daily observations, using days 1, 7, 13, ... for the estimated expected returns, observations 3, 9, 15, ... for the factor loadings, and 5, 11, 17, ... for the standard deviations. This has had the effect of reducing the number of time series observations used in the estimation of each parameter from 2619 to 436. Note that the factor loadings were estimated in the usual way but for covariance matrices computed only with observations 3, 9, 15, ...

These results are to be compared with those reported in Table IV where all estimates were computed from the same sample observations. Only nine of the 42 groups now display a significant t-statistic for s. Given the possibility of cross-group interdependence, this is only the weakest conceivable evidence for an effect by s on expected returns. The remaining effect drops even further when more "insulating" days are inserted between observations used to estimate the parameters. When three days are skipped rather than just one day, only seven groups out of 42 (16.7%) display significant effects for s at the 95% level. This supports

Table VI

Cross-sectional regressions of estimated expected returns on factor loadings and individual total standard deviations of return (summary for 42 groups of 30 individual securities per groups, 1962–72 daily observations with estimators taken from non-overlapping subsamples)[a]

Arithmetic Mean	Standard Error of Mean	Percentage of Groups Whose Statistic Exceeds 95% Critical Level[b]
Across 42 Groups		
t-statistic: test for most significant factor loading having no effect on expected return.		
2.27	.111	57.1
t-statistic, test for individual total standard deviation of return having no effect on expected return.		
.941	.204	21.4
F-statistic, test for no effect by any factor loading on expected return (in addition to the effect of standard deviation).		
2.24	.183	31.0

[a] The estimated returns are obtained from daily observations 1, 7, 13, ···2617; the factor loadings from observations 3, 9, ···2619; the standard deviations from observations 5, 11, ···2615.

[b] The regression equation and 95% critical values are given in nn. a and b of Table IV.

the argument that serial dependence in *squared* returns may be responsible for the small remaining effect of s shown in Table VI.

In contrast to the reduced impact of standard deviation, the estimated influence of the factor loadings have increased, (though admittedly only by a small amount). For example, the most significant factor loading now has a t-statistic of at least 2.06 in 57.1% of the groups.

Again, the groups are not independent; so caution should be exercised when interpreting the results. The results are not definitive but they *are* consistent with most of the frequently-observed sampling dependence between standard deviation and mean return being attributable to effects working through factor loadings and to spurious effects due to skewness.

As a final test, we conducted an experiment similar to that developed by Fama and MacBeth [14]. Here is an outline of the procedure:

a) Using daily observations 3, 9, 15, ... the five factor loadings $\hat{b}_{1j}, \ldots \hat{b}_{5j}$ are estimated for each asset in each of the 42 groups of 30 assets.
b) Using daily observations 5, 11, 17, ... the "own" standard deviation of return s_j is computed for each asset.
c) Using observations 1, 7, 13, ... the following cross-sectional regression is computed for each group, g. $R_{jt} = \hat{\lambda}_{0gt} + \hat{\lambda}_{1gt}b_{1j} + \ldots + \hat{\lambda}_{5gt}b_{5j} + \hat{\lambda}_{6gt}s_j + \xi_{jt}$, $j = 1, \ldots 30$ within each group and for all groups $g = 1, \ldots 42$. This yields 42 time series of vectors, $\hat{\lambda}_{gt} = \hat{\lambda}_{0gt}, \ldots \hat{\lambda}_{6gt}$, of estimated factors $\hat{\lambda}_{1gt}$ through $\hat{\lambda}_{5gt}$, of the riskless interest rate and of the effect of "own" standard deviation $\hat{\lambda}_{6gt}$.
d) The time series of $\hat{\lambda}_{6gt}$ is used to compute a standard error for the mean value, i.e., for $\bar{\lambda}_{6g} = \Sigma_t \hat{\lambda}_{6gt}/T$, in order to test for the significant presence of an "own" variance effect.

The results indicate that just three of the 42 groups (7.1%) display a significant effect of s on expected return at the 95% level. Since just one less group, two out of 42, would be fewer than the number to be expected by pure chance, there seems to be little remaining reason to reject the hypothesis that individual expected returns are unaffected by the "own" variance of returns.

This procedure also could be used to estimate the significance of different factors. However, due to the factor identification problem, the time series of factor values from one group will probably not be the same as the factor values for a different group. Furthermore, the resulting tests are less powerful than the factor-analysis based tests reported in Section II. They do indicate, however, that 17 groups (40.5%) have at least one significant factor and ten groups (23.8%) have at least two significant. This is an indication of fewer significant factors than in the factor-analysis tests but such a result is to be anticipated with a less powerful method.[9]

[9] Following Fama-MacBeth [14], a test of market efficiency can be conducted by regarding that the $\hat{\lambda}$ as excess returns on portfolios. (They can be interpreted as portfolios that load exclusively on a given factor). The returns should be serially uncorrelated in an efficient market. We found the first ten lagged autocorrelations, each subsuming six trading days, to be insignificantly different from zero. For example, the ten lagged serial correlation coefficients of $\hat{\lambda}_{11}$, (the first factor of the first group), are .00726, .0432, −.0197, −.0123, −.0201, −.112, −.0412, −.000979, −.0624, .0739. The sample size is 430 ± 5, (depending on the lag).

IV. A Test for the Equivalence of Factor Structure Across Groups

One of the most troubling econometric problems in the two preceding sections was due to the technological necessity of splitting assets into groups. Since the calculations were made for each group separately, but over the same time interval, the results are potentially susceptible to spurious sampling dependence among the groups. Also, due to the factor identification problem, there is no good way to ascertain whether the *same* three (or four) factors generate the returns in every group. It's conceivable, (but we think unlikely) that each of the 42 groups displays three *different* factors. This would imply that the actual number of common factors is 3 × 42, or at least some number considerably larger than three.

Even if the APT is true, the same underlying common factors can be "rotated" differently in each group. However, there is one parameter, the intercept term (λ_0 in eq.(2)) which should be identical across groups, whatever the sample rotation of the generating factors. Recall that λ_0 should be the expected return on either the riskless rate of interest or on an asset with no sensitivity to the common factors. This suggests that a simple test of the APT and of the cross-group consistency of factor structure can involve ascertaining whether the λ_0's estimated for the 42 groups are significantly different.

Since the test must also correct for inter-group dependence, a reasonable approach would use the time series estimated intercepts from the Fama-MacBeth type cross-sectional regressions computed in the last part of Section III; (Cf. pp. 46–48). For each group g, $\hat{\lambda}_{0gt}$ is the cross-sectional intercept for day t, from a cross-sectional regression on the factor landings ($\hat{b}$'s) and the "own" standard deviation (s_i) estimated from different but interleaved observations. Since each group has a time series $\hat{\lambda}_{0g}$ whose members are possibly correlated across groups, the appropriate test is Hotelling's T^2 for differences in adjacent groups.[10] That is, let

$$Z_{g/2,t} = \hat{\lambda}_{0,g-1,t} - \hat{\lambda}_{0,g,t} \quad g = 2, 4, 6, \ldots.$$

be computed for each naturally-ordered pair of groups with a "sufficient" number of observations. We assumed that 400 was a sufficient number. There were 38 groups with at least 400 observations from the calendar observations used in the regressions (i.e., from observation 1, 7, 13, ...). Thus, there were 19 time-series for $Z_{g/2}$, ($g = 2, \ldots 38$).

The composite null hypothesis to be tested is

$$H_0\colon E(Z_{g/2}) = 0, g = 2, 4, \ldots 38$$

and Hotelling's T^2 conducts this test by using the quadratic form

$$T^2 = N \bar{\underset{\sim}{Z}}' \underline{\Sigma}^{-1} \bar{\underset{\sim}{Z}}$$

where $\bar{\underset{\sim}{Z}}$ is the vector of sample means of the $\bar{Z}_{g/2,t}$'s and $\underline{\Sigma}$ is their sample covariance matrix. The sample size is N. Since simultaneous observations are required to compute the covariance matrix, if any stock in any of the 38 groups

[10] See Press [39, ch. 6] for a general explanation of Hotelling's T^2.

had a missing observation, that observation could not be used. This resulted in a further reduction to 188 simultaneous observations.

Hotelling's T^2 value for these observations was 16.9 and the corresponding F statistic with 19 and 169 degrees of freedom was located at the .298 fractile of the null distribution. Thus, there is absolutely no evidence that the intercept terms were different across groups.

However, we do admit that this test is probably quite weak. There is a very low degree of explanatory power in the daily cross-sectional regressions and thus the sampling variation of each $\hat{\lambda}_{0gt}$ is quite large. Furthermore, Hotelling's test in small samples requires multi-variate normality. It is known, however, to be asymptotically robust and in the bivariate case is robust even for quite modest sample sizes much smaller than ours; (Cf. Chase and Bulgren [7]).

V. Conclusion

The empirical data support the APT against both an unspecified alternative—a very weak test—and the specific alternative that own variance has an independent explanatory effect on excess returns. But, as we have emphasized, these tests are only the beginning and should be viewed in that light.

A number of the empirical anomalies in the recent literature could be re-examined in the context of these results. For example, the APT would predict that insofar as price-earnings ratios have explanatory power for excess returns, they must be surrogates for the factor loadings. This provides the basis for an alternative test of the APT. On the longer term agenda, the statistical underpinnings of our analysis must be shored. Work on the small sample properties of factor analysis is scarce, and for nonnormal distributions, results appear to be nonexistent.

Lastly, of course, an effort should be directed at identifying a more meaningful set of sufficient statistics for the underlying factors. While this is not a necessary component of tests of the APT, it is an interesting and worthwhile pursuit of its own.

The issue in all of this, of course, is not whether the APT is true or false. Like all the theories that are not empty, it is false that some degree of precision in the testing: if we test long enough, all interesting theories are rejected. Rather, the question is what we will learn from these tests on how well the theory performs in competition with specific alternatives. At stake is the basic intuition of the APT that systematic variability alone affects expected returns, and this is the central theme of modern asset pricing theory.

Acknowledgement

The comments and suggestions of listeners to our preliminary results on this subject have been invaluable in improving many half-baked procedures. To the extent that our soufflé is finally ready to come out of the oven, we owe a particular debt to participants in seminars at Berkeley/Stanford, Laval, Karlsruhe, and Southern California, and to the participants in the conference on "new issues in the asset pricing model" held at Coeur d'Alene, Idaho, and sponsored by Washington State University. This work originated while the authors were Leslie Wong summer fellows at the University of British Columbia in 1977. Comments by Michael Brennan, Thomas Copeland and Richard McEnally have been especially helpful. Of course, no one but us can be held responsible for remaining errors.

REFERENCES

1. J. Aitcheson and J.A.G. Brown. *The Lognormal Distribution.* Cambridge: Cambridge University Press, 1957.
2. Ray Ball. "Anomalies in Relationships Between Securities' Yields and Yield—Surrogates." *Journal of Financial Economics* (June/September 1978), 103–26.
3. Fischer Black. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business* 45 (July 1972), 444–54.
4. S. Basu. "Investment Performance of Common Stock in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance* 23 (June 1977), p. 663–82.
5. Douglas T. Breeden. "An Inter-Temporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." Unpublished manuscript, University of Chicago, Graduate School of Business, August 1978.
6. M. J. Brennan. "Capital Asset Pricing and the Structure of Security Returns." Unpublished manuscript, University of British Columbia, May 1971.
7. G. R. Chase and William G. Bulgren. "A Monte Carlo Investigation of the Robustness of T^2." *Journal of the American Statistical Association* 66 (September 1971), 499–502.
8. John C. Cox, Jonathan E. Ingersoll, Jr. and Stephen A. Ross. "A Theory of the Term Structure." *Econometrica*. (forthcoming).
9. Peter K. Clark. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica* 41 (January 1973), 135–55.
10. George W. Douglas. "Risk in the Equity Market: An Empirical Appraisal of Market Efficiency." *Yale Economic Essays* 9 (Spring 1969), 3–45.
11. Eugene F. Fama. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (May 1970), 383–417.
12. Eugene F. Fama. *Foundations of Finance.* New York: Basic Books, 1976.
13. Eugene F. Fama. "The Behavior of Stock Market Prices." *Journal of Business* 38 (January 1965), 34–105.
14. Eugene F. Fama and James D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy* 38 (May 1973) 607–36.
15. D. E. Farrar. *The Investment Decision Under Uncertainty.* Englewood Cliffs, N.J.: Prentice-Hall dissertation series, 1962.
16. Milton Friedman. "The Methodology of Positive Economics," in *Essays in Positive Economics.* Chicago: University of Chicago Press, 1953.
17. Adam Gehr, Jr. "Some Tests of the Arbitrage Pricing Theory." *Journal of the Midwest Finance Association* (1975), 91–105.
18. Nils H. Hakansson. "Capital Growth and the Mean-Variance Approach to Portfolio Selection." *Journal of Financial and Quantitative Analysis* 6 (January 1971), 517–557.
19. Der-Ann Hsu, Robert B. Miller, and Dean W. Wichern. "On the Stable Paretian Behavior of Stock Market Prices." *Journal of the American Statistical Association* 69 (March 1974), 108–13.
20. Karl G. Jöreskog. "Some Contributions to Maximum Likelihood Factor Analysis." *Psychometrika* 32 (December 1967), 443–82.
21. Karl G. Jöreskog and Dag Sörbom. *EFAP, Exploratory Factor Analysis Program.* National Educational Resources, Inc., 1976.
22. Benjamin F. King. "Market and Industry Factors in Stock Price Behavior." *Journal of Business* 39 (January 1966, supp.), 139–90.

23. Alan Kraus and Robert Litzenberger. "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance* 31 (September 1976), 1085–1100.
24. Lawrence Kryzanowski and To Minh Chan. "General Factor Models and the Structure of Security Returns." Working paper, Concordia University, Faculty of Commerce and Administration, (September 1979).
25. Terence C. Langetieg. "An Application of a Three-Factor Performance Index to Measure Stockholder Gains from Merger." *Journal of Financial Economics* 6 (December 1978), 365–83.
26. D. N. Lawley and A. E. Maxwell. *Factor Analysis as a Statistical Method.* London: Butterworths, 1963.
27. Cheng F. Lee. "Functional Form, Skewness Effect, and the Risk-Return Relationship." Abstracted in *Journal of Financial and Quantitative Analysis.* (March 1977), 55–72.
28. Cheng F. Lee and Joseph D. Vinso. "Single vs. Simultaneous Equation Models in Capital Asset Pricing: The Role of Firm-Related Variables." *Journal of Business Research* (1980), 65–80.
29. Haim Levy. "Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio." *American Economic Review* (September 1978).
30. John Lintner. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics* 47 (February 1965), 13–37.
31. Robert E. Lucas, Jr. "Asset Prices in an Exchange Economy." *Econometrica* 46 (November 1978), 1429–45.
32. Benoit Mandelbrot. "The Variation of Certain Speculative Prices." *Journal of Business* 36 (October 1963), 394–419.
33. Charles G. Martin. "Ridge Regression Estimates of the *Ex Post* Risk-Return Tradeoff on Common Stock." *Review of Business and Economic Research* 13 (Spring 1978), 1–15.
34. David Mayers. "Non-Marketable Assets and Capital Market Equilibrium Under Uncertainty." 223–248 in Michael C. Jensen, ed., *Studies in the Theory of Capital Markets.* New York: Praeger, 1972.
35. Robert C. Merton. "An Inter-Temporal Capital Asset Pricing Model." *Econometrica* 41 (September 1973), 867–87.
36. Stephen L. Meyers. "A Re-examination of Market and Industry Factors in Stock Price Behaviors." *Journal of Finance* 28 (June 1973), 695–706.
37. Merton H. Miller and Myron Scholes. "Rates of Return in Relation to Risk: A Re-examination of Some Recent Findings." 47–48 in Michael C. Jensen, ed., *Studies in the Theory of Capital Markets.* New York: Praeger, 1972.
38. Jan Mossin. "Equilibrium in a Capital Asset Market." *Econometrica* 34 (October 1966), 768–83.
39. S. James Press. *Applied Multivariate Analysis.* New York: Holt, Rinehart, and Winston, Inc., 1972.
40. Marc R. Reinganum. "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings Yields and Forecasts." Unpublished manuscript, Graduate School of Business, University of Chicago, 1978.
41. Richard Roll. "A Critique of the Asset Pricing Theory's Tests." *Journal of Financial Economics* 4 (May 1977), 129–76.
42. Richard Roll. "Ambiguity When Performance is Measured by the Securities Market Line." *Journal of Finance* 33 (September 1978), 1051–69.
43. Richard Roll. "Testing a Portfolio for Ex Ante Mean/Variance Efficiency." In E. Elton and M. Gruber, eds., *TIMS Studies in the Management Sciences.* Amsterdam: North-Holland, 1979.
44. Barr Rosenberg and Vinay Marathe. "Tests of Capital Asset Pricing Hypotheses." Unpublished manuscript, University of California, Berkeley, 1977.
45. Barr Rosenberg. "Extra-Market Components of Covariance in Security Returns." *Journal of Financial and Quantitative Analysis* 9 (March 1974), 263–74.
46. Barr Rosenberg and Michael Houglet. "Error Rates in CRSP and Compustat Data Bases and Their Implications." *Journal of Finance* 29 (September 1974), 1303–10.
47. Stephen A. Ross. "Return Risk, and Arbitrage." In Irwin Friend, and James L. Bicksler, eds., *Risk and Return in Finance*, I, 189–218. Cambridge, Mass.: Ballinger, 1977.

48. Stephen A. Ross. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13 (December 1976), 341–60.
49. Stephen A. Ross. "The Current Status of the Capital Asset Pricing Model (CAPM)." *Journal of Finance* 33 (June 1978), 885–90.
50. William F. Sharpe. "A Simplified Model for Portfolio Analysis." *Management Science* 9 (January 1963), 277–93.
51. William F. Sharpe. "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk." *Journal of Finance* 19 (September 1964), 425–42.
52. William F. Sharpe. "The Capital Asset Pricing Model: A 'Multi-Beta' Interpretation." In H. Levy and M. Sarnat, eds., *Financial Decision Making Under Uncertainty.* New York: Academic Press, 1977.
53. Ledyard R. Tucker and Charles Lewis. "A Reliability Coefficient for Maximum Likelihood Factor Analysis." *Psychometrika* 38 (March 1973), 1–10.
54. James C. Van Horne. *Financial Management and Policy,* 4th ed. Englewood Cliffs, N.J. Prentice-Hall, 1977.