# The Arbitrage Pricing Theory
# and Multifactor Models of Asset Returns*

## Gregory Connor
London School of Economics and Political Science

and

## Robert A. Korajczyk
Northwestern University

Working paper #139

September 1992
Revised:  September 30, 1993

Comments Welcome
Forthcoming in *Finance Handbook*,
edited by Robert Jarrow, Vojislav Maksimovic, and William Ziemba

Robert A. Korajczyk
Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2006
**USA**
Telephone: (708) 491-8336
FAX: (708) 491-5719
Internet: r-korajczyk@nwu.edu

Gregory Connor
Financial Markets Group
London School of Economics
Houghton Street
London WC2A 2AE
**UNITED KINGDOM**
Telephone: 44-071-955-7320
FAX: 44-071-242-1006
E-mail: connorg@vax.lse.ac.uk

## I.    Introduction

The Arbitrage Pricing Theory (APT) of Ross (1976, 1977), and extensions of that theory, constitute an important branch of asset pricing theory and one of the primary alternatives to the Capital Asset Pricing Model (CAPM).  In this chapter we survey the theoretical underpinnings, econometric testing, and applications of the APT.  We aim for variety in viewpoint without attempting to be all-inclusive.  Where necessary, we refer the reader to the primary literature for more complete treatments of the various research areas we discuss.

In Section II we discuss factor modelling of asset returns.  The APT relies fundamentally on a factor model of asset returns.  Section III describes theoretical derivations of the APT pricing restriction.  Section IV surveys the evidence from estimates and tests of the APT.  In Section V we discuss several additional empirical topics in applying multifactor models of asset returns.  We survey applications of the APT to problems in investments and corporate finance in Section VI.  We provide some concluding comments in Section VII.

## II.    Strict and Approximate Factor Models

Stock and bond returns are characterized by a very large cross-sectional sample (in excess of 10,000 simultaneous return observations in some studies) with strong co-movements.  The fundamental sources of these co-movements are not always obvious and are not easily measured.  Such a statistical system, where a few unobservable sources of system-wide variation affect many random variables, lends itself naturally to factor modelling.  The APT begins by assuming that asset returns follow a factor model.

In a factor model, the random return of each security is a linear combination of a small number of common, or pervasive, factors, plus an asset-specific random variable.  Let n denote the number of assets and k the number of factors.  Let f denote the k-vector of random factors, B the n×k matrix of linear coefficients representing assets' sensitivities to movements in the factors (called factor betas or factor loadings), and ε the n-vector of asset-specific random variables (called the idiosyncratic returns).  We can write the n-vector of returns, r, as expected returns

plus the sum of two sources of random return, factor return and idiosyncratic return:

$$r = E[r] + Bf + \varepsilon, \tag{1}$$

where $E[f] = 0$, $E[\varepsilon] = 0$, and $E[f\varepsilon'] = 0$. The beta matrix, B, is defined by the standard linear projection, $B = E[(r - E[r])f'](E[ff'])^{-1}$. Given a vector of returns, r, and a vector of zero-mean variates, f, the standard linear projection divides the returns into expected returns, k linear components correlated with f, and zero-mean idiosyncratic returns uncorrelated with f. The standard linear projection imposes no structure on the returns or factors besides requiring that the variances and expected returns exist. In Sections II.1 and II.3 below, we provide enough additional structure on (1) so that the idiosyncratic returns are diversifiable risk and the factor risks are not.

## II.1    Strict Factor Models

Since the factors and idiosyncratic risks in (1) are uncorrelated, the covariance matrix of asset returns, $\Sigma = E[(r - E[r])(r - E[r])']$, can be written as the sum of two matrices, the covariance matrix of each security's factor risk and the covariance matrix of idiosyncratic risks:

$$\Sigma = BE[ff']B' + V \tag{2}$$

where $V = E[\varepsilon\varepsilon']$. In a strict factor model, the idiosyncratic returns are assumed to be uncorrelated with one another. This means that the covariance matrix of idiosyncratic risks, V, is a diagonal matrix. This captures the essential feature of a strict factor model: the covariance matrix of securities can be decomposed as the sum of a matrix of rank k and a diagonal matrix of rank n. This imposes restrictions on the covariance matrix as long as k is less than n.

A strict factor model divides a vector random process into k common sources of randomness (each with linear impact across assets) and n asset-specific sources of randomness. One of Ross's insights was to see that this model could be employed to separate the nondiversifiable and diversifiable components of portfolio risk. Suppose that there are many available assets (i.e., n is large). The idiosyncratic variance of a portfolio with portfolio

proportions equal to $\omega$ is:

$$\omega' V \omega \; = \; \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 \; \leq \; (\max_i \sigma_i^2) \sum_{i=1}^{n} \omega_i^2 \; .$$

Since the portfolio weights sum to one, the average portfolio weight is 1/n. If the holdings are spread widely over the n assets (so that all the portfolio weights are close to 1/n) then the sum of squared portfolio weights approaches zero as n goes to infinity. As long as there is an upper bound on the idiosyncratic variances of the individual assets, the idiosyncratic variance of any well-spread portfolio will be near zero. Therefore, given a strict factor model and many assets, the idiosyncratic returns contain only diversifiable risk.

**II.2    Choice of Rotation**

There is a *rotational indeterminacy* in the definition of the factors and the betas in equation (1). Given B and f, consider any nonsingular k×k matrix L and construct $B^* = BL$ and $f^* = L^{-1}f$. Replacing B and f with B* and f* yields an observationally equivalent return generating model. There are various approaches to choosing which of the infinite set of (B, f) pairs to use. Often the analyst chooses to simplify (2), without loss of generality, by letting $E[ff'] = I_k$. Another common choice of rotation is the eigenvector decomposition. Given a strict factor model, define the square root inverse matrix of V, $V^{-1/2}$, in the obvious way: $(V^{-1/2})_{ii} = (V_{ii})^{-1/2}$ and $(V^{-1/2})_{ij} = 0$ for i≠j. Scale the covariance matrix of returns by pre- and post-multiplying by $V^{-1/2}$: $\Sigma^* = V^{-1/2} \Sigma V^{-1/2}$. (Note that if we scale each asset return by its idiosyncratic standard deviation then Σ* is the covariance matrix of the rescaled returns). Using (2) we can write:

$$\Sigma^* = J \Lambda J' + I_k.$$

where J is the n×k matrix of the first[1] k eigenvectors of $\Sigma^*$ and $\Lambda$ is a k×k diagonal matrix of the associated eigenvalues squared [see Chamberlain and Rothschild (1983)]. One choice of rotation is to set B = J. This choice is often used in econometric work since there are well-known techniques for calculating the dominant eigenvectors of a matrix.

The factors underlying the co-movements in security returns presumably come from economy-wide shocks to expected cash flows and required returns. Suppose that we can exactly identify the economic shocks giving rise to the co-movements; let g denote the k-vector of these observable economic shocks. The statistical factors, f, and economic shocks, g, are equivalent if $g = L^{-1}f$ for a nonrandom k×k matrix L. In this case, the obvious choice of rotation is $f^* = g$. More realistically, the statistical factors in security returns and any set of observed economic shocks will be imperfectly correlated. There are various statistical techniques used to rotate the factors to be "as close as possible" to the observed economic shocks [see, for example, Burmeister and McElroy (1988)].

Suppose for now that the economic shocks and statistical factors are equivalent and consider the obvious rotation $f^* = g$. Rewriting (1) using this rotation gives:

$$r = E[r] + B^*g + \varepsilon. \qquad (3)$$

A model using economically interpretable factors, as in (3), has notable advantages over (1). Since the factors are observed economic shocks, we can interpret the beta coefficients $B^*$ in economically meaningful ways. After estimation, we can make statements such as "asset i has a high inflation risk." Contrast this with the betas estimated using the eigenvector rotation. Here we can only make statements such as "asset i has a high sensitivity to eigenvector 2 risk." Since the eigenvectors are statistical artifacts, the betas from them provide little interpretable information. Most APT researchers would agree that, other things being equal, an economically meaningful rotation, as in (3), is preferable to (1). From an empirical viewpoint, other things are not equal. Models with statistically generated factors fit the returns data better than ones with economic shocks as proxies for the factors, in the sense that the fraction of time series variance

explained is higher.  This is distinct from the question of which set of factors perform better in terms of explaining cross-sectional differences in asset returns.

As the number of securities grows large, the de-meaned returns to well-diversified portfolio returns approximate a linear combination of the factors.  That is, any portfolio $\omega$ such that $\omega'\varepsilon \approx 0$ has de-meaned return [from equation (1)] approximately equal to a linear combination of the factors.  That is, $r_\omega - E[r_\omega] \approx b_\omega\, f$, where $b_\omega = \omega'B$ is the $1 \times k$ vector of factor betas for this portfolio.  Thus, any set of k well-diversified and linearly independent portfolios has de-meaned returns approximately equivalent to a rotation of the factors [see Admati and Pfleiderer (1985)].  We will call such portfolios factor representing or factor-mimicking portfolios.

## II.3   Approximate Factor Models

In order for a strict factor model to have empirical content, k must be less than n.  For stock market return data, k is usually taken to be *much* less than n.  A typical empirical study with U.S. equity returns will have k in the range of one to fifteen, whereas n, the number of available U.S. equity returns, is from one thousand to six thousand (depending upon the selection criteria).  A strict factor model imposes too severe a restriction on the covariance matrix of returns when n/k is this large.

Ross uses the factor model assumption to show that idiosyncratic risks can be diversified away in a many-asset portfolio.  The strict diagonality of V is sufficient for Ross's diversification argument, but not necessary.  Chamberlain (1983) and Chamberlain and Rothschild (1983) develop an asymptotic statistical model for asset returns data called an *approximate factor model*.  This model preserves the diversifiability of idiosyncratic returns but weakens the diagonality condition on V.  It also imposes a condition which ensures that the factor risks are *not* diversifiable.

An approximate factor model relies on limiting conditions as the number of assets grows

large. We start with an infinite sequence of random asset returns, $r_i$, $i = 1, 2, ...$, with finite means and variances. We treat the observed assets as the first n assets from this infinite sequence, and impose limiting conditions as n goes to infinity.

Let f denote a k-vector of mean-zero random variates with finite variances. We can always express asset returns using the standard linear projection (1): expected returns plus a beta matrix multiplied by the factors plus idiosyncratic return where the factors and idiosyncratic returns are uncorrelated. Therefore, we can always describe the covariance matrix of asset returns using (2): $\Sigma = BE[ff']B' + V$. We seek conditions on $\Sigma$ to ensure that the idiosyncratic returns are diversifiable and the factor risks are not. We say that $\varepsilon$ is *diversifiable risk* if $\lim_{n \to \infty} \omega^{n\prime}\omega^n = 0$ implies $\lim_{n \to \infty} E[(\omega^{n\prime}\varepsilon)^2] = 0$, where $\omega^n$ is the $n \times 1$ vector of portfolio weights for a portfolio formed from the first n assets. This means that all well-diversified portfolios have idiosyncratic variance near zero. A converse condition is imposed on factor risks. Let $z^j$ denote an n-vector with a one in the $j^{th}$ component and zeros elsewhere. The factors, f, are *pervasive risk* if for each $z^j$, $j=1,...,k$, there exists an $\omega^n$ such that $\lim_{n \to \infty} \omega^{n\prime}\omega^n = 0$ and $\omega^{n\prime}B = z^j$ for all n. This condition guarantees that each factor risk affects many assets in the economy.

Chamberlain and Rothschild define an *approximate factor structure* as a factor decomposition where the $\varepsilon$'s are diversifiable and the f's are pervasive. They show that the diversifiable risk condition is equivalent to a finite upper bound on the maximum eigenvalue of $V^n$ as n goes to infinity and that the pervasiveness condition is equivalent to the minimum eigenvalue of $B^{n\prime}B^n$ going to infinity with n.

Note that the covariance matrix of returns is the sum of two components, $B^nB^{n\prime}$ and $V^n$. In an approximate factor model, $B^nB^{n\prime}$ has all of its eigenvalues going to infinity, whereas $V^n$ has a bound on all its eigenvalues. Chamberlain and Rothschild show that these bounds carry over to the covariance matrix. In an approximate factor model, the k largest eigenvalues of the covariance matrix go to infinity with n, and the $k+1^{st}$ largest eigenvalue is bounded for all n. They prove that this is a sufficient condition as well. Consider a countable infinity of assets

whose sequence of covariance matrices has exactly k unbounded eigenvalues. Then the asset returns necessarily follow an approximate k-factor model. So the conditions on the sequence of eigenvalues (k[th] unbounded, k+1st bounded) characterize an approximate k-factor model.

An intuitive example of an approximate factor model is a sector and industry model of risk. Suppose that there is a large number (n) of assets each representing the common shares of one firm. Each firm belongs to one of a large number (m) of industries each with a small number (h, with h approximately equal to n/m) of firms. Idiosyncratic returns are correlated within industries but uncorrelated across industries. In this case, the covariance matrix of idiosyncratic returns consists of a series of h×h sub-matrices along the diagonal and zeros elsewhere, when assets are ordered by industry grouping. The sub-matrices are the within-industry covariance matrices. Holding h constant and letting n and m increase, this series of covariance matrices has bounded eigenvalues.[2]

On the other hand, suppose that there is a small number, k, of sectors, each containing n/k firms. All firms within sector j are subject to sector shock $f_j$ with unit betas (for simplicity). Firms in sector j are unaffected by the shocks of other sectors. Given these assumptions, the sector shocks constitute pervasive risk.[3] Note the clear distinction between industries (a small proportion of the firms are in each industry) versus sectors (a substantial proportion of the firms are in each sector).

Connor and Korajczyk (1993) suggest that, for econometric work, imposing a mixing condition on the sequence of idiosyncratic returns is more useful than the bounded eigenvalue condition alone. The cross-sectional sequence of idiosyncratic returns, $\varepsilon_i$, i=1,..., is called a mixing process if the probability distribution of $\varepsilon_{i+m}$, conditional on $\varepsilon_i$, approaches its unconditional distribution as m goes to infinity. [See White and Domowitz (1984) for a discussion of mixing processes and their applications]. The idiosyncratic return of an asset may be strongly related to those of a few other "close" assets, but it must have asymptotically zero relationship to most assets. The mixing condition differs from the bound on eigenvalues in that

it restricts the entire conditional probability distribution rather than only the covariance matrix. In many estimation problems, restrictions on the covariances alone are not enough to derive asymptotic distributions of test statistics.

## II.4   Conditional Factor Models

There is clear empirical support for time-varying means and variances in asset returns, and this has led to some recent work on time-varying (or dynamic) factor models of returns. Dropping the assumption that returns are independently and identically distributed through time, and rewriting (1) with an explicit time subscript, gives:

$$r_t = E_{t-1}[r_t] + B_{t-1}f_t + \varepsilon_t.$$

Let $B_{t-1}$ be chosen by the conditional projection of $r_t$ on $f_t$ [i.e., $B_{t-1} = E_{t-1}[(r_t - E_{t-1}[r_t])f_t'](E_{t-1}[f_tf_t'])^{-1}$] so that $E_{t-1}[\varepsilon_t'f_t] = 0$. The conditional covariance matrix can be written as:

$$\Sigma_{t-1} = B_{t-1}\Omega_{t-1}B_{t-1}' + V_{t-1},$$

where $\Omega_{t-1} = E_{t-1}[f_tf_t']$ and $V_{t-1} = E_{t-1}[\varepsilon_t\varepsilon_t']$. Even if we impose that $V_{t-1}$ is diagonal for all t, the system is not statistically identified in this general form. Suppose that we observe the returns on n securities for T periods. For each date, t, we must estimate the n elements of $V_{t-1}$, the nk elements of $B_{t-1}$, and the $k^2$ elements of $\Omega_{t-1}$. This gives a total of $T(n+nk+k^2)$ parameters to be estimated from nT return observations. Obviously we must impose more structure to get an identified model.

Moving to dynamic models can eliminate some rotational indeterminacies of static models but some indeterminacies remain. For example, time variation in $B_{t-1}\Omega_{t-1}B_{t-1}'$ could be caused by time variation in the factor betas, $B_{t-1}$, with homoscedastic factors (i.e., $\Omega_{t-1} = \Omega$); by heteroscedastic factors with constant factor betas; or by time variation in both $B_{t-1}$ and $\Omega_{t-1}$. The structure imposed on $B_{t-1}$ and $\Omega_{t-1}$ for identifiability is related to the assumed nature of the dynamic influence. Suppose that the analyst assumed that $\Omega_{t-1} = I_k$ for all t, so that the factors are homoscedastic. Then, all of the dynamics are due to time variation in $B_t$. Alternatively the

analyst can assume that $B_{t-1} = B$ for all t, in which case any time-variation in the factor model appears in $\Omega_{t-1}$ [for example, see Engle, Ng, and Rothschild (1990)].

Conditional on the assumed source of time variation in the factor model ($B_{t-1}$ or $\Omega_{t-1}$) and the assumed time-series properties of the dynamics [e.g., Engle, Ng, and Rothschild (1990) assume that the factors follow GARCH[4] processes], the dynamic structure can eliminate some of the standard rotational indeterminacy found in static factor models [Sentana (1992)].

Some recent papers in this area also allow for time variation in expected returns, so that $E_{t-1}[r_t]$ is not constant [see, for example, Engle, Ng, and Rothschild (1990)].

## III.    Derivation of the Pricing Restriction

Now we will use the factor model of returns to derive the APT pricing result:

$$E[r] \approx \iota^n \lambda_0 + B\lambda \tag{4}$$

where $\lambda_0$ is a constant, $\lambda$ is a k-vector of factor risk premia, and $\iota^n$ is an n-vector of ones.  The approximate equality sign " $\approx$ " in (4) reflects the fact that the APT holds only approximately, requiring that the economy has a large number of traded assets in order to be an accurate pricing model, on average.

## III.1    Exact Pricing in a Noiseless Factor Model

We begin with a *noiseless* factor model (one with no idiosyncratic risk), where $r = E[r] + Bf$.  This is much too strong a restriction on asset returns but is useful for the intuition it provides.  In this case, an exact arbitrage argument is sufficient for the APT.  Here we do not need a large number of assets, and there is no approximation error in the APT pricing restriction. The result comes from Ross (1977).  To derive the APT in this case, project $E[r]$ on $\iota^n$ and B to get projection coefficients $\lambda_0$ and $\lambda$ and a projection residual vector $\eta$:

$$E[r] = \lambda_0 \iota^n + B\lambda + \eta. \tag{5}$$

By the property of projection residuals we have $\eta'B = 0$ and $\eta'\iota^n = 0$.  Consider the n-vector, $\eta$,

viewed as a portfolio of asset purchases and sales. This portfolio has zero cost since $\eta'\iota^n = 0$ and no randomness since $\eta'B = 0$. If this portfolio had a positive expected return, it would represent an arbitrage opportunity, that is, a zero-cost portfolio with a strictly positive expected payoff and no chance of a negative payoff. The existence of an arbitrage opportunity is inconsistent with even the weakest type of pricing equilibrium. The expected payoff of the portfolio under consideration is $\eta'E[r] = \eta'\eta$. This can only be zero if $\eta = 0$. So the APT pricing model (4) holds with equality in the absence of arbitrage opportunities.

We can combine the noiseless factor model, $r = E[r] + Bf$, with the APT pricing result, $E[r] = \iota^n\lambda_0 + B\lambda$, to show that $r = \iota^n\lambda_0 + B(f+\lambda)$. A unit-cost portfolio is any collection of assets such that $\omega'\iota^n = 1$. The payoff to a unit-cost portfolio is a portfolio return. A unit cost portfolio with $\omega'B = 0$ (no factor risk) has a risk-free return equal to $\lambda_0$. As long as the $(k+1)\times n$ matrix $[\iota^n, B]$ has rank $k+1$, we can construct such a portfolio,[5] and identify $\lambda_0$ as the riskfree return. A unit cost portfolio with a unit sensitivity to factor j and zero sensitivity to the other factors has expected return $\lambda_0 + \lambda_j$. Hence, the k-vector $\lambda$ measures the risk premia (expected returns above the risk-free return) per beta-unit of each factor risk. These risk premia are dependent on the factor rotation, which affects the scales of the betas.

**III.2   Approximate Non-Arbitrage**

The argument used for the noiseless factor model can be extended to a strict or approximate factor model. In this case, we get a pricing relation which holds approximately in an economy with many assets. For generality we work with the case of an approximate factor model. We combine the original formulation of Ross (1976) with some refinements of Huberman (1982). Consider the orthogonal price deviations $\eta$ defined by (5), as in the noiseless case. Define a sequence of portfolios as follows: the $n^{th}$ portfolio consists of holdings of the first n assets in proportion to their price deviations, scaled by the sum of squares of these deviations:

$$\omega^n = \eta^n/(\eta^{n'}\eta^n).$$

Using the same steps as in the noiseless case, one can show that the cost of each of these portfolios is zero, the expected payoff of each is 1, and the variance is $(\eta^{n\prime}\eta^n)^{-2}\eta^{n\prime}V^n\eta^n$. Using the property of the maximum eigenvalue, we have $\eta^{n\prime}V^n\eta^n \leq (\eta^{n\prime}\eta^n)\|V^n\|$, where $\|V^n\|$ denotes the maximum eigenvalue of $V^n$. Therefore the portfolio variance is less than or equal to $(\eta^{n\prime}\eta^n)^{-1}\|V^n\|$. Since $\|V^n\|$ is bounded (by the definition of an approximate factor model), the variance of this sequence of portfolios goes to zero as n increases if $\eta^{n\prime}\eta^n$ is not bounded above. This would constitute a sequence of "approximate arbitrage portfolios." These portfolios have zero cost, unit expected payoff, and variance approaching zero, as the number of assets in the economy increases. Ross (1976), Huberman (1982), Ingersoll (1984), and Jarrow (1988) show that approximate arbitrage portfolios will not exist in a well-functioning capital market. If we rule out approximate arbitrage portfolios, then $\eta^{n\prime}\eta^n$ must be bounded for all n.

The bound on the sum of squared pricing errors has the following interpretation. Although the APT can substantially misprice any one asset (or any limited collection of assets), the prices of most assets in a many-asset economy must be closely approximated. Let c denote the upper bound on $\eta^{n\prime}\eta^n$. The average squared pricing error is less than c/n and, for any $\zeta > 0$, only $c/\zeta$ assets have squared pricing errors greater than or equal to $\zeta$. The proportion of assets with squared pricing errors greater than $\zeta$ goes to zero as n increases.

The approximate nature of the APT pricing relation in (4) causes important problems for tests of the APT. With a finite set of assets, the sum of squared pricing errors must be finite, so we cannot directly test whether $\eta^{n\prime}\eta^n$ is bounded. Shanken (1982) argues that the weakness of this price approximation renders the APT untestable. He shows that this pricing bound is not invariant to "re-packaging" the assets into an equivalent set of n unit-cost portfolios [Gilles and LeRoy (1991) make a similar argument]. Shanken argues that only equilibrium-based derivations of the APT (which can provide an exact pricing approximation) are truly testable. The equilibrium-based derivations involve additional assumptions besides those needed to derive (4), and are discussed below. Ingersoll (1984) notes that the APT pricing approximation will be

close for all well-diversified portfolios (since the pricing errors diversify away). He argues that the pricing of these portfolios should be of more concern to the economist than the pricing of individual assets, and therefore the weakness of the pricing approximation for individual assets is not crucial. A well-diversified factor-mimicking portfolio will have an expected excess return close to the factor risk premium. Heston (1991) builds on Ingersoll's analysis to show that the weakness of the pricing approximation does not affect some statistical tests based on large cross-sections of assets.

Reisman (1992b) expands on Shanken's argument. He proves that the approximate-arbitrage pricing bound is unaffected by measurement error in the factors. If there are k true factors, then *any* set of k or more random variables which are correlated with the factors can be used as proxies. For example, almost *any* set of k or more individual assets returns (as long as they have differing beta coefficients) can be used as factors. The finite bound on the sum of squared APT pricing errors absorbs the additional pricing error generated by any mis-measurement of the factors or overestimate of the number of factors [also see Shanken (1992b)].

So far we have considered an economy with a large, but finite, number of assets. Chamberlain (1983) extends the APT to an economy with an infinite number of assets. To accomplish this, he expands the space of portfolio returns to include infinite-dimensional linear combinations of asset returns. Let $r^n$ denote the n-vector of the first n of the infinite set of assets. We define a limit portfolio return as the limit of the returns to n-asset portfolios as n goes to infinity:

$$r_\omega = \lim_{n \to \infty} \omega^{n\prime} r^n.$$

$$(6)$$

The limit in (6) is usually taken with respect to the second-moment norm $\|r_\omega\| = E[r_\omega^2]$. A simple example of a convergent sequence of portfolios is $\omega^n = (1/n, 1/n, ..., 1/n)$. Note that, element-by-

element, this sequence of portfolio weights converges to a zero vector. Yet the limit portfolio of this sequence has a well-defined, non-zero return in most cases.[6] Limit portfolio returns can be perfectly diversified, that is, have idiosyncratic variance of exactly zero.

Ross (1978b) and Kreps (1981) develop an exact non-arbitrage pricing theory (this is not the same as the APT). In the absence of exact arbitrage opportunities, there must exist a positive, linear pricing operator over state-contingent payoffs. Chamberlain and Rothschild (1983) show that in an infinite-asset model the approximate-arbitrage APT is an extension of the Ross-Kreps exact non-arbitrage pricing theory. In the absence of approximate arbitrage, the positive linear pricing operator defined by Ross and Kreps must be continuous with respect to the second moment norm. Given an approximate factor model for asset returns, this continuity condition implies the same bound on APT pricing errors described above. Reisman (1988) extends the Chamberlain-Rothschild result to general normed linear spaces. He shows that the APT can be reduced to an application of the Hahn-Banach theorem using two assumptions: one, the non-existence of approximate arbitrage opportunities for limit portfolios, and two, the approximate factor model assumption on the countably infinite set of asset returns.

Stambaugh (1983) extends the APT to an economy in which investors have heterogeneous information and/or the econometrician has less information than investors. Unconditional asset returns must obey a factor model, but the conditional asset returns (as perceived by an investor with special information) need not obey a factor model. In the absence of approximate arbitrage (for an informed or uninformed investor, or both) the APT pricing restriction holds using the unconditional betas.

## III.3    Competitive Equilibrium Derivations of the APT

There are advantages to the approximate-arbitrage proof of the APT, since the nonexistence of approximate arbitrage opportunities is such a weak assumption. One drawback is the weakness of the pricing approximation. As an alternative to the approximate arbitrage

approach, one can derive the APT by imposing competitive equilibrium. This gives a stronger

pricing approximation, and links the APT with other equilibrium-based pricing models.

Consider an investor with a risk-averse utility function u($\cdot$) for end-of-period wealth.

Suppose that returns obey an approximate factor model, with the additional assumption that

idiosyncratic risks are conditionally mean zero given the factors:

$$E[\varepsilon \mid f] = 0.$$

Let $W_0$ denote the investor's wealth at time 0. In competitive equilibrium, a first-order portfolio

optimization condition must hold for every investor:

$$E[u'(W_0\omega'r)r] = \iota^n\gamma, \tag{7}$$

for some positive scalar $\gamma$. For notational simplicity, let $W_0 = 1$. Inserting (1) into (7),

separating the three additive terms and bringing constants outside the expectations operator

gives:

$$E'[r] = \iota^n\lambda_0 + B\lambda + \frac{E[u'(\omega'r)\varepsilon]}{E[u'(\omega'r)]} \tag{8}$$

where $\lambda_0 = \gamma/E[u'(\omega'r)]$ and $\lambda = -E[u'(\omega'r)f]/E[u'(\omega'r)]$. The competitive equilibrium derivations

of the APT assume a sufficient set of conditions so that the last term in (8) is approximately a

vector of zeros. Note that this last term is the vector of risk premia the investor assigns to the

idiosyncratic returns. So proving the equilibrium APT amounts to showing that, in competitive

equilibrium, investors will assign a zero or near-zero risk premium to each idiosyncratic return.

Chen and Ingersoll (1983) assume that in competitive equilibrium some investor has a

portfolio return with no idiosyncratic risk. Let $r_N$ denote this portfolio return where $r_N =$

$E[r_N] + bf$ for some k-vector b. Using $E[\varepsilon \mid f] = 0$, we have $E[u'(r_N)\varepsilon] = E[E[u'(r_N)\varepsilon \mid f]] = 0$. So

in the Chen and Ingersoll (1983) model, the APT holds exactly.

Consider again the optimality condition (7), but assume that the chosen portfolio is well-

diversified (with idiosyncratic variance near zero) but not perfectly diversified.  Consider an

exact first-order Taylor expansion of $u'(\omega'r) = u'(\omega'E[r] + \omega'Bf + \omega'\varepsilon)$ around  $\omega'E[r] + \omega'Bf$:

$$u'(\omega'r) = u'(\omega'E[r] + \omega'Bf) + (\omega'\varepsilon)u''(\omega'E[r] + \omega'Bf + \delta),$$

where $\delta$ is the Taylor residual term.  Therefore,

$$E[\varepsilon u'(\omega'r)] = E[\varepsilon u'(\omega'E[r] + \omega'Bf)] + E[\varepsilon(\omega'\varepsilon)u''(\omega'E[r] + \omega'Bf + \delta)]. \tag{9}$$

The first vector term of (9) is exactly zero, as noted above.  Under reasonable assumptions, every

component of the second vector term is near zero if the chosen portfolio is well-diversified.  For

simplicity, suppose the investor has quadratic utility, so that $u''$ is a constant.  Then

$E[\varepsilon(\omega'\varepsilon)u''(\omega'Bf + \delta)] = E[\varepsilon\omega\varepsilon']u'' = E[\varepsilon\varepsilon']\omega u''$ .  Consider an arbitrary term of this n-vector

(the $i^{th}$ term) and note that $(E[\varepsilon\varepsilon']\omega u'')_i \leq \|V\|\omega'\omega u''$, which goes to zero as $\omega'\omega$ goes to zero.

With non-quadratic utility, the proof that this term approaches zero is messier than, but not

fundamentally different from, the quadratic utility case [see, for example, Dybvig (1983) or

Grinblatt and Titman (1983)].

The model above has the shortcoming that it assumes a particular form for the

equilibrium portfolio returns of investors.  It is preferable in economic modelling to derive the

properties of endogenous equilibrium variables (such as portfolio returns) rather than to impose

assumptions on them.  Dybvig (1983) develops a simple and elegant equilibrium version of the

APT which accomplishes this.  Dybvig assumes that all investors have constant relative risk

aversion and that the security market is effectively complete.  That is, all welfare-increasing

trading opportunities are available [see Ingersoll (1987, chapter 8) for a discussion of effectively

complete markets].  When the security market is effectively complete, one can construct a

representative investor for the economy.  By definition, the representative investor finds it

optimal, conditional on budget constraints, to hold the market portfolio.  Dybvig assumes that

the market portfolio is well-diversified (which is an assumption common to all equilibrium

derivations of the APT[7]). Dybvig considers the optimality condition (7) for the representative

investor who holds the market portfolio, and derives the utility function for this investor (it is a

linear combination of the constant relative risk aversion functions of the investors). He shows that the Taylor residual in expression (9) converges to zero for each asset, given this utility function. Connor (1982) and Grinblatt and Titman (1983) develop models broadly similar to Dybvig's, though differing in details.

The equilibrium version of the APT can also be derived using Chamberlain's infinite-asset techniques. Connor (1984) requires that the market portfolio return is a perfectly diversified limit portfolio return. He allows investors to hold limit portfolios in equilibrium. He then shows [along the lines of Chen and Ingersoll (1983) discussed above] that in competitive equilibrium all investors choose to hold perfectly diversified portfolios and the APT pricing relation holds exactly for every asset.

Milne (1988) adds a real investment side to the equilibrium APT. Each corporation owns a capital investment function which produces random profits. The firms are purely equity financed. The model is static; each firm issues equity and invests the proceeds in its investment technology, which produces a random profit at the end of the period. Recall that the equilibrium version of the APT requires that the market portfolio is well-diversified. With production, the relative supplies of the various assets are endogenous to the model since the issuance of equity depends upon the capital investment decisions of firms. The pricing theory requires that the capital investment plans chosen by corporations must be such that the market portfolio is well-diversified after the firms make their decisions [also see Brock (1982)].

### III.4   Mean-Variance Efficiency and Exact Factor Pricing

Mean-variance efficiency mathematics can be employed to restate the APT pricing restriction. This restatement is particularly useful for econometric modelling.

Consider the set of unit-cost portfolios obtainable as linear combinations of an n-vector of asset returns, r. Note that for any portfolio return, $r_\omega$, we can define the one-factor projection equation:

$$r = E[r] + bf + \varepsilon \tag{10}$$

where $f = r_\omega - E[r_\omega]$. A single-beta pricing model holds with respect to (10) if:

$$E[r] = \iota^n \lambda_0 + b\lambda \tag{11}$$

for some scalars $\lambda_0$ and $\lambda$. Define a mean-variance efficient portfolio as a unit-cost portfolio which minimizes variance subject to $E[\omega'r] = c$ for some c. One can show[8] that (11) is the necessary and sufficient condition for the mean-variance efficiency of $\omega$. Therefore, proving that (11) holds is equivalent to proving that $\omega$ is mean-variance efficient. Note that this is *not* a pricing theory; it is a mathematical equivalence between the pricing restriction (11) and the mean-variance efficiency of $\omega$. If $\omega$ is the market portfolio, then (11) is the conventional statement of the CAPM. We can equivalently re-state the CAPM as "the market portfolio is mean-variance efficient."

The relationship between mean-variance efficiency and beta pricing carries over to a multi-beta model. Given an n-vector of returns r, consider any set of k portfolio returns $r_{\omega 1}$, $r_{\omega 2}$, ..., $r_{\omega k}$ and the projection equation:

$$r = E[r] + Bf + \varepsilon,$$

where $f_j = r_{\omega j} - E[r_{\omega j}]$ for j=1, ..., k. Hypothesize linear pricing with respect to these factor-mimicking portfolios as in equation (11)

$$E[r] = \iota^n \lambda_0 + B\lambda. \tag{12}$$

Grinblatt and Titman (1987) show that (12) holds if and only if some linear combination of the portfolios $\omega_1$, ..., $\omega_k$ is mean-variance efficient. Chamberlain (1983) derives this same result for large-n and infinite-n models. Chamberlain shows that if a linear combination of factor portfolios converges to a mean-variance efficient portfolio as n goes to infinity, then the deviations from APT pricing go to zero. He gives explicit bounds on the speed of convergence of the sum of squared pricing errors to zero. In an infinite-asset economy, if a linear combination of factor portfolios is mean-variance efficient, then the APT holds exactly.

The Grinblatt-Titman and Chamberlain analysis is not an independent pricing theory, but

rather a useful reinterpretation of the APT pricing formula. The mean-variance efficiency criteria restates the mathematical relationship between expected returns and betas given by (12). That is, we can re-state the APT pricing restriction as "a linear combination of factor portfolios is mean-variance efficient." This alternative characterization proves very useful for econometric modelling; see Shanken (1987a) and Kandel and Stambaugh (1989) for the derivation of APT test statistics based on this approach. Most econometric analyses of the APT along these lines have relied on the exact finite-n model of Grinblatt and Titman. Given the interesting cross-sectional asymptotic analysis of Heston (1991), Reisman (1992b), and Mei (1993), it might be useful to extend this econometric framework to encompass the large-n asymptotic mean-variance efficiency described by Chamberlain (1983).

The equivalence between the mean-variance efficiency of factor portfolios and exact APT pricing also sheds light on the relationship between the CAPM and APT. Assume that the market portfolio is perfectly diversified (it has zero idiosyncratic variance). Some variation on this assumption is necessary if we are to derive the APT using an equilibrium argument, and it is widely accepted as a natural assumption even when the model is derived via approximate arbitrage [see, e.g., Ingersoll (1984) and Dybvig and Ross (1985)]. This assumption implies that the return to the market portfolio is a linear combination of factor portfolio returns. The APT holds if any linear combination of factor portfolios is mean-variance efficient. The CAPM holds if the market portfolio (a *particular* linear combination of factor portfolios) is mean-variance efficient. Note that the CAPM requires observation of the market portfolio returns whereas the APT needs observations of the factors or factor-mimicking portfolios. Analysts differ on which is easier to observe [e.g., Shanken (1982, 1985) and Dybvig and Ross (1985)].

Wei (1988) constructs a model which combines features of the CAPM and APT. He assumes that asset returns obey an approximate factor model, and that the idiosyncratic returns obey a mutual fund separating condition. (The simplest case is that the idiosyncratic returns are independent of the factors and multivariate normal). He shows that in competitive equilibrium,

an exact k+1-factor pricing model holds.  Consider the projection equation linking the market portfolio return and the factors:

$$r_q = E[r_q] + h'f + \varepsilon_q.$$

Wei calls the random variable $\varepsilon_q$ the "residual market factor."  He shows that there exists a k-vector $\lambda$ and scalar $\lambda_q$ such that

$$E[r] = \iota^n\lambda_0 + B\lambda + \beta\lambda_q,$$

where $\beta = cov(r,\varepsilon_q)/var(\varepsilon_q)$.  If the market portfolio is well-diversified, then the $k+1^{st}$ factor premium is redundant (since $\beta$ is a linear combination of B) and the pricing equation holds with only k factors.

## III.5   Pricing Dynamics

Dynamic versions of the APT generally specify an exogenous factor model for the cash flows (dividends) paid by firms and derive the factor model for the prices of securities endogenously.  Discrete time dynamic models are derived in Jagannathan and Viswanathan (1988), Bossaerts and Green (1989), Connor and Korajczyk (1989), and Hollifield (1993).  Even if the factor loadings (betas) for the cash flow process are time invariant, the betas of asset returns (relative to factor-mimicking portfolios) will be functions of the current information set. In the static APT we can replicate the priced payoff from a security with the riskless asset and factor-mimicking portfolios.  Jagannathan and Viswanathan (1988) show that in a multiperiod economy there is, in general, a different riskless asset for every maturity (i.e., a discount bond with that maturity).  Thus, even though the risky components of assets' payoffs are driven by, for example, a one-factor model, asset returns may follow an infinite factor structure (corresponding to a portfolio mimicking the single factor, plus the returns on discount bonds for every maturity).

Connor and Korajczyk (1989) develop a multiperiod model in which they assume that per-share dividends, rather than asset returns, obey an approximate factor model.  They show that expected returns obey the exact APT pricing restriction at each date.  However, the general

version of their model is not statistically identified since the beta coefficients and factor risk premia vary through time. They describe additional conditions on preferences and the stochastic process for dividends which give a statistically identified model. The stochastic process for dividends is such that the infinite number of term structure factors in the general formulation of Jagannathan and Viswanathan (1988) are replaced by the return on a single consol bond.

Bossaerts and Green (1989) develop an alternative with a more explicit description of the time-varying risk premia. They treat the special case of a one-factor model for dividends, but it is straightforward to generalize much of their analysis to a multi-factor model. The return on a consol bond plays an important role in their model. They give explicit, testable expressions for the time-variation in asset betas and the factor risk premia by substituting observable quantities (returns on a reference portfolio and the relative prices of assets) for the unobservable return on the consol bond.

Engle, Ng, and Rothschild (1990) also develop a multi-period equilibrium version of the APT. They begin along the lines of Chen and Ingersoll (1983) by assuming that the marginal utility of consumption for a representative investor can be described as a function of k random factors. They also assume that returns at each date follow an approximate factor model with conditionally mean zero idiosyncratic returns. The standard first-order condition for a budget-constrained optimal portfolio [i.e., equation (7)] gives an exact version of the APT.

Bansal and Viswanathan (1993) also rely on an assumption that the marginal utility of consumption of a representative investor can be described by a (potentially non-linear) function of k random factors. They do not assume that all assets have returns given by an approximate factor model. In any competitive equilibrium, all assets, even those not obeying an approximate factor model, must have expected returns given by the Ross-Kreps positive linear state space pricing function. The contribution of Bansal and Viswanathan is to note that this state pricing function can be described as a function of the k random factors which explain the representative investor's marginal utility. This gives rise to a *nonlinear* k-factor pricing model [see Latham

(1989) for a related model].

Reisman (1992a), building on earlier work by Ohlson and Garman (1980), extends the APT to a continuous-time economy. He assumes that there are a large number of assets, each paying a liquidating dividend at the terminal date T. From time 0 to T asset prices are continuously set so as to exclude approximate arbitrage. He assumes that the continuous-time information flow about the vector of terminal dividends follows a continuous-time approximate factor model. He shows that instantaneous expected returns obey the APT with bounded pricing errors, almost surely.

Chamberlain (1988) develops an intertemporal equilibrium asset pricing model which integrates the APT with Merton's (1973) Intertemporal Capital Asset Pricing Model (ICAPM). In Chamberlain's model, trading lasts from 0 to T and investors can trade continuously during that time. There exists a countably infinite set of assets; the vector of random asset payoffs at time T, conditioned at any time, t, between 0 and T, follows a continuous-time approximate factor model. Chamberlain assumes that the market portfolio is well-diversified. He proves that, at each time t, asset prices obey the APT formula, *and* that this formula is identical to the CAPM pricing formula (if k equals one) or the ICAPM formula (if k is greater than or equal to one). Constantinides (1989) gives an alternative proof in a slightly different framework.

Chamberlain's model is an important contribution for the way it rigorously unifies the APT and ICAPM. In his framework, these two pricing models are not testably distinct. Connor and Korajczyk (1989) argue that the APT and ICAPM should be separated by econometric and empirical considerations rather than theoretical ones. The ICAPM stresses the role of state variables as the fundamental determinants of asset risk premia, whereas the APT stresses the pervasive factors in random returns as the key determinants. Chamberlain's model shows that these two categories are not always distinct: the set of state variables of the ICAPM can be identical to the set of pervasive factors of the APT.

**IV.    Empirical Analysis of the APT**

Analyses of the factor structure of asset returns actually predate the APT.  Rather than being motivated by the pricing implications of the APT, this strand of the literature was primarily motivated by a desire to describe, in a parsimonious manner, the covariance structure of asset returns, $\Sigma$.  The covariance matrix of asset returns is, of course, a major component of a portfolio optimization problem.  Estimation of the unrestricted covariance matrix of n securities requires the estimation of $n \times (n + 1)/2$ distinct elements.  The single index, or diagonal, model of Sharpe (1963) postulated that all of the common elements of returns were due to assets' relations with the index.  Thus, only $3 \times n$ parameters needed to be estimated:  n "betas" relative to the index, n unique variances, and n intercept terms.  This approach reduced much of the noise in the estimate of $\Sigma$.  One could view the single-index model as a strict one-factor model with a prespecified factor.  In practice, the single index did not describe all of the common movements across assets (i.e., the residual matrix is not diagonal) so there seemed to be some additional benefit from using a multifactor model.  With k factors there are still only $n \times (k + 2)$ parameters to estimate ($n \times k$ betas, n intercepts or means, and n unique variances).  Some studies in this area are Farrar (1962), King (1966), Cohen and Pogue (1967), and Elton and Gruber (1973).

Our primary interest, however, is the evidence regarding the pricing implications of the APT.  As discussed in Section III, the main implication of the APT is that expected returns on assets are approximately linear in their sensitivities to the factors [equation (4)]:

$$E[r] \approx \iota^n \lambda_0 + B\lambda.$$

With additional restrictions used in some competitive equilibrium derivations of the APT (Section III.3), it is possible to obtain the pricing relation as an equality.  Since standard statistical methods are not amenable to testing approximations, most empirical tests actually evaluate whether (4) holds as an equality.  Thus the tests are joint tests of the APT plus any ancillary assumptions required to obtain the exact pricing relation [Shanken (1985)]:

$$E[r] = \iota^n \lambda_0 + B\lambda. \tag{13}$$

Once the relevant factors have been identified or estimated, approaches to analyzing and testing the APT have, to a large extent, mirrored developments in analyzing and testing other asset pricing models, such as the CAPM [see Ferson (1993) for a review of tests of asset pricing models]. Various aspects of (13) have been investigated. Some authors have focussed on evidence regarding the size and significance of the factor risk premia vector, $\lambda$. One testable restriction of the model is that the implied risk premia are the same across subsets of assets. That is, if we partition the return vector, r, into components, $r^1$, $r^2$, ..., $r^s$, with $B^i$ representing the same partitioning of B, and investigate the subset pricing relations

$$E[r^i] = \iota^n \lambda_0^i + B^i \lambda^i \quad i = 1, 2, ..., s, \tag{14}$$

then $\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ for all i. Another restriction implied by the pricing model is that variables in the agents' information set should not allow us to predict expected returns which differ from the relation in (13). These restrictions form the basis for testing the APT.

The exact pricing relation (13), along with the factor model for the return generating process (1), imply that the n-vector of returns at time t, $r_t$, is given by:

$$r_t = \iota^n \lambda_{0,t-1} + B(\lambda_{t-1} + f_t) + \varepsilon_t. \tag{15}$$

The riskless rate of return, $\lambda_{0,t-1}$, and the risk premia, $\lambda_{t-1}$, have a time t-1 subscript since they are determined by expectations conditional on information at time t-1. If we observe the return on the riskless asset, $\lambda_{0,t-1}$, we get an equivalent relation between returns in *excess* of the riskless rate $R_t = r_t - \iota^n \lambda_{0,t-1}$, B, and the factor returns, $\lambda_{t-1} + f_t$:

$$R_t = B(\lambda_{t-1} + f_t) + \varepsilon_t. \tag{16}$$

All empirical analyses of the APT involve analysis of a panel of asset return data in which we observe a time series of returns (t = 1, 2, ..., T) on a cross-sectional sample of assets or portfolios (the n different assets in $r_t$ or $R_t$). Even though all empirical studies combine cross-sectional and time-series data, it is common to classify them as cross-sectional or time-series studies on the basis of the approach used in the final, testing stage of the analysis. That is, conditional on B, (15) and (16) can be thought of as cross-sectional regressions in which the

parameters being estimated are $\lambda_{0,t-1}$ and $(\lambda_{t-1} + f_t)$.  Conversely, conditional on $\lambda_{0,t-1}$ and $(\lambda_{t-1} + f_t)$, (15) and (16) can be thought of as time-series regressions in which the parameters being estimated are the elements of B.  Some studies, such as McElroy and Burmeister (1988), jointly estimate the model in one step.  We will first consider a sample of cross-sectional tests of the APT and then describe some time-series tests.

## IV.1  Cross-sectional Tests of the APT

For the moment, assume that we observe the $n \times k$ matrix B, representing the assets' sensitivity to the factors.  Then (15) and (16) can be viewed as cross-sectional regressions of $r_t$ and $R_t$, respectively, on a constant and the matrix k factor sensitivities, B.

$$r_t = \iota^n F_{0,t-1} + BF_t + \varepsilon_t \tag{17}$$

$$R_t = \iota^n F_{0,t-1} + BF_t + \varepsilon_t \tag{18}$$

The parameters to be estimated are an intercept, $F_{0,t-1}$, and the k-vector of slope coefficients, $F_t$.  The parameters can be estimated by a variety of methods, including ordinary least squares (OLS), weighted least squares (WLS), and generalized least squares (GLS).  Under standard conditions, the estimates are unbiased and consistent.  That is, as the cross-sectional sample size, n, approaches infinity, $\hat{F}_{0,t-1}$ should be equal to $\lambda_{0,t-1}$ in (17) and zero in (18) and $\hat{F}_t$ should be equal to the vector of factor realizations, $\lambda_{t-1} + f_t$ (where $^\wedge$ denotes the estimate of the parameter).

In a given period, we cannot disentangle from $\hat{F}_t$ the risk premia $\lambda_{t-1}$ and the unexpected factor shocks $f_t$.  However, given a time series of returns $r_t$ (t = 1, 2, ..., T), we can estimate a cross-sectional regression for each period, yielding a time series of estimates $\hat{F}_1, \hat{F}_2, ..., \hat{F}_T$ (as well as $\hat{F}_{0,0}, \hat{F}_{0,1}, ..., \hat{F}_{0,T}$).  Since the unexpected factor shocks are conditionally mean zero (otherwise they would not be unexpected), we can learn about the risk premium vector by investigating the time-series average of the estimates, $\bar{F} = (\hat{F}_1 + \hat{F}_2 + ... + \hat{F}_T)/T$ and $\bar{F}_0 = (\hat{F}_{0,0} + \hat{F}_{0,1} + ... + \hat{F}_{0,T})/T$.  If the risk premium vector is stationary, with mean $\lambda$, then $\bar{F}$ should converge to $\lambda$ since the average of the $f_t$ will converge to zero.[9]  The precision of our estimate of

$\lambda$, $\overline{F}$, can be estimated by the time-series variability of $F_t$.

We can also test the predictions of the exact APT by augmenting the cross-sectional regressions in (17) and (18) with a n × j matrix of firm-specific instruments, $Z_{t-1}$, observable at the beginning of the period:[10]

$$r_t = \iota^n F_{0,t-1} + BF_t + Z_{t-1}\delta + \varepsilon_t \tag{19}$$

$$R_t = \iota^n F_{0,t-1} + BF_t + Z_{t-1}\delta + \varepsilon_t \tag{20}$$

where $\delta$ is an j-vector of parameters. If the model is correct, cross-sectional differences in expected returns should only be due to differences in factor sensitivities, B, and not due to other variables such as the instruments, $Z_{t-1}$. Therefore, values of $\delta$ different from zero are inconsistent with the model.

In the raw return regression (17), the estimate $F_{0,t-1}$ represents a unit investment portfolio with zero exposure to factor risk (or market risk in the CAPM) and should converge to the riskless rate of interest. The estimate $F_t$ represents a set of k zero-investment (arbitrage) portfolios, with portfolio j having a sensitivity of unity to the $j^{th}$ factor and a sensitivity of zero to the remaining factors [see Fama (1976, chapter 9)]. Thus the vector $F_t$ represents a set of excess returns to factor-mimicking portfolios.

This cross-sectional approach is used by Fama and MacBeth (1973) to test the CAPM. Unlike the assumption we made above, however, we are not generally endowed with the true matrix of factor sensitivities, B. Fama and MacBeth (1973) propose using, in an initial stage, time-series regressions of asset returns on a proxy for the market portfolio to obtain estimates of the sensitivities, or betas. The second-stage cross-sectional regressions then use these estimates as the independent variables. Fama and MacBeth (1973) also include as instruments [our $Z_{t-1}$ in (19) and (20)] the squared values of beta and the asset-specific, or residual, risk as measured by the standard deviation of the error from the first-stage time-series regressions.

Given that the cross-sectional regressions use estimates of B instead of the true value, the

regressions suffer from an errors-in-variables (EIV) problem. Since the betas of portfolios are more precisely measured than the betas of individual assets, Fama and MacBeth use portfolios of assets in the cross-sectional regressions instead of individual assets. This reduces the EIV problem. The portfolios are formed in a manner designed to maintain cross-sectional dispersion in the independent variable, beta [see Fama and MacBeth (1973) or Fama (1976, chapter 9) for details]. A multiple-factor analog of this two-pass, cross-sectional regression procedure forms the basis of many tests of the APT.[11] The two-pass procedure is analyzed and extended in Shanken (1992a).

The first step in the Fama-MacBeth procedure is to obtain an estimate of the matrix of asset sensitivities to the factors, B. If we observe the factors, f, directly, then B, E[r], and V in (1) and (2) can be estimated through standard time-series regression procedures as is done in Fama and MacBeth (1973) using the returns on a market portfolio proxy. This approach forces us to choose the factors *ex ante*. An alternative approach to estimating B that relies only on the assumed strict factor model is factor analysis [see, for example, Morrison (1976, ch. 9) or Anderson (1984, ch. 14)]. Let us assume that returns follow a strict factor model, have a multivariate normal distribution cross-sectionally, and are independently and identically distributed through time. Let $\Sigma$ denote the sample covariance matrix of returns, estimated using T time-series observations of n securities, with $T > n$. Under these conditions, the $n \times n$ matrix, $\Sigma$, has a Wishart distribution. The parameters of the distribution are the $n \times k$ matrix of factor betas, B, and the n idiosyncratic variances $V_{ii}$ $i=1,...,n$. (Note that the off-diagonal elements of V equal zero by assumption). The maximum likelihood estimates $\mathbb{B}$ and $\mathbb{V}$ are those which maximize the likelihood of observing $\Sigma$ given $B = \mathbb{B}$ and $V = \mathbb{V}$. Various numerical techniques have been suggested for solving the maximum likelihood problem. The first-order conditions for a maximum can be written as follows:

$$\text{diag}[\mathbb{B}\mathbb{B}' + \mathbb{V}] = \text{diag}[\Sigma]$$

$$\Sigma\mathbb{V}^{-1}\mathbb{B} = \mathbb{B}(I + \mathbb{B}'\mathbb{V}^{-1}\mathbb{B}).$$

The first-order conditions are necessary but not sufficient. They do not encompass the restriction that the diagonal elements of V must be nonnegative.[12] Also, the matrix B is only identified up to an orthogonal transformation. This is known as rotational indeterminacy (see section II.2 above). The computational complexity of this maximum likelihood problem increases dramatically with n. This has led some analysts to use small cross-sectional samples. Alternative computational algorithms have been developed to alleviate these problems. These issues will be discussed in the context of particular empirical studies below.

To our knowledge, the first empirical analysis of the APT is by Gehr (1978), who uses a variant of the cross-sectional approach. This study applies factor analysis to a set of 41 individual company returns (chosen from different industries) to obtain an initial set of factor-mimicking portfolios, $F_t$, in two steps. In step one, factor analysis is applied to the sample covariance matrix in order to obtain an estimate of the assets' matrix of factor sensitivities, B (called factor loadings in the factor analysis literature). In step two, a cross-sectional regression of asset returns on B [as in (17)] gives an initial estimate of the factor-mimicking portfolios, $F_t$ (called factor scores in the factor analysis literature). For a second set of assets (24 industry portfolios), the matrix of betas is then estimated by a time-series regression of asset returns on the returns of either one, two, or three initial factor-mimicking portfolios. Finally, the average premium vector, $\overline{F}$, is estimated from a cross-sectional regression of average returns of the 24 industry portfolios, $\overline{r}$, on their estimated betas.

In our description of the cross-sectional regression approach above, we estimated $F_t$ for each period and then averaged these estimates to get $\overline{F}$. In Gehr (1978) the returns are averaged first and then regressed on the beta matrix. If the beta matrix is held constant over the period, these two approaches will lead to *identical* point estimates. However, the standard errors calculated from the time series of the $F_t$ will be different than the OLS standard errors from the single regression of average returns on betas. The time-series standard errors should be preferable since they incorporate cross-sectional dependence and heteroscedasticity that is

ignored in the OLS standard errors. Shanken (1992a) suggests additional adjustments to the time-series standard errors to account for the EIV problem in the betas.

Gehr (1978) uses 30 years of monthly data to estimate the vector of average risk premia, $\overline{F}$. His focus is on whether the premia are significantly different from zero, and therefore no explicit tests of the model's over-identifying restrictions are performed in the study. Over the 30-year period only one of the three factors, the third factor, has a significant premium. Over the three 10-year subintervals there were one, none, and two factors, respectively, with significant premia.

Roll and Ross (1980) estimate factor risk premia and test the APT restrictions with a sample of daily returns on 1260 firms over the period from July 1962 to December 1972. Due to computational considerations, they divide the cross-sectional sample into 42 groups of thirty firms each and perform an analysis on each group. For a five factor model they use maximum likelihood factor analysis to estimate B, the matrix of assets' sensitivities to the factors. Given this estimate of B, $\hat{B}$, they perform cross-sectional regressions of asset returns on $\hat{B}$, as in (17). They also perform cross-sectional regressions of asset excess returns (i.e., returns in excess of an assumed riskless rate, $\lambda_0$, of 6% per annum) on $\hat{B}$, as in (18). As in Fama and MacBeth (1973), the cross-sectional regressions are estimated each period and the risk premia are measured by the time-series average of the estimates, $\overline{F}$. Roll and Ross (1980) use generalized least squares in the cross-sectional regressions rather than OLS. The relevant covariance matrix for the GLS weighting is obtained from the inputs to the factor analysis step. The results indicate that as many as four factors have significant risk premia.

Roll and Ross (1980) test the APT by including the sample standard deviation of the asset as an instrument in cross-sectional regressions such as (19) and (20). In their tests, the estimate of the standard deviation is not predetermined. In one version of this test (their Table IV), the sample standard deviation, estimated beta matrix, $\hat{B}$, and asset returns are from the same sample. In this case, the test strongly rejects the APT because of the apparent significant relation

between mean returns and standard deviation, even after controlling for factor risk.  As Roll and

Ross (1980) point out, the use of the same sample to estimate the dependent and independent

variables in the regressions may lead to spurious significance of the parameter $\delta$ in (19) and (20).

This could be caused by correlation in the sampling errors of mean returns and sample standard

deviations [this problem is also discussed in Miller and Scholes (1972) and Lehmann (1990)].

To overcome this problem, Roll and Ross (1980) perform the tests using disjoint subsets of the

data to estimate the inputs.  That is, they use observations 3, 9, 15, etc. to get the estimated factor

sensitivities, B; use observations 5, 11, 17, etc. to estimate the standard deviation; and use the

returns for observations 1, 7, 13, etc. to estimate the cross-sectional regression (19).  The use of

disjoint subsets to estimate the inputs should reduce the potential for spurious significance.  In

this case, three of the forty-two groups of assets have a statistically significant value of $\delta$.  They

argue that there is little evidence against the hypothesis that an asset's own standard deviation

has no incremental power over the asset's factor sensitivities in explaining mean returns.

An additional implication of the model, shown in (14), is that the implied zero-beta (or

riskless) return, $\lambda_0$, and the implied risk premia, $\lambda$, should be the same across subsets of assets.

Because of the standard rotational indeterminacy of the estimate, B, from factor analysis, Roll

and Ross (1980) cannot compare $\lambda^i$ to $\lambda^j$ (where i and j denote different subgroups of assets)

because the rotations across the subgroups may be different.  However, they can compare $\lambda_0^i$ and

$\lambda_0^j$.  In a final test they use a Hotelling $T^2$ test to test the equality of the mean zero-beta return

across 38 of the 42 groups (four groups were excluded because of lack of time-series data).

They could not reject the hypothesis that the mean zero-beta returns were the same across

groups.

One of the advantages of using daily data, as in Roll and Ross (1980), is the large number

of time-series observations available for estimation.  This is particularly important when the

sample is to be subdivided to estimate factor sensitivities, standard deviations, and mean returns

over separate observations.  However, the use of daily data causes some problems in terms of

estimating the matrix of factor sensitivities, B. The main input into factor analysis is the sample covariance matrix of asset returns. The standard sample covariance assumes that we have returns that are synchronous (i.e., observed over the same period). In a given observation period (in this case, a day) the returns on one asset are actually measured over a different time interval than the returns on another asset, in general. This is due to the fact that returns are calculated from the percentage change in closing prices (adjusted for any distributions on that day). The closing prices are usually the price of the last trade of the day. This last trade might have occurred at the close of the day for some assets but earlier in the day for others. The usual pairwise sample covariance will tend to underestimate the true covariance because it is only measuring the comovement over the typical daily common observation period across assets. The non-synchroneity also induces lead and lagged cross-correlations. The extent of the bias in the covariance estimates depends on the severity of the non-synchroneity.

This bias is not restricted to daily data; it is present at any observation frequency. However, the bias is a function of the amount of non-synchroneity, as a fraction of the observation period. This will be much larger for daily observations than for monthly observations, for example. The equivalent problem occurs in applications of the CAPM or event studies that need to adjust for cross-sectional differences in sensitivities to a market index. Scholes and Williams (1977), Cohen, Hawawini, Maier, Schwartz, and Whitcomb (1983), and Andersen (1989) propose estimators for beta which correct for the bias in the standard OLS estimate of beta. The estimators consist of the sum of lead, contemporaneous, and lagged betas, adjusted for the serial correlation in the market.

Shanken (1987b) recognizes that the same type of synchroneity problem arises in the use of factor analysis to obtain first stage estimates of B. He proposes a covariance matrix estimator based on the methods of Cohen, *et al* (1983) for use in the factor analysis stage. Shanken (1987b) applies this approach to a set of assets chosen to be comparable to the sample in Roll and Ross (1980). Empirically, he finds that the average estimate of pairwise covariance,

adjusted for non-synchronous trading, is twice as large as the average unadjusted sample covariance. In fact, almost all (97%) of the adjusted estimates are larger than the unadjusted estimates. He also finds that factor-mimicking portfolios constructed from B's adjusted for non-synchroneity have small correlations with portfolios constructed from unadjusted B's. This implies that using unadjusted covariance matrices in the factor analysis stage is not just an innocuous choice of a different rotation of the same factors. The evidence in Shanken (1987b) indicates that non-synchronous trading may induce significant biases when applying factor analysis to high frequency data. Therefore, if one wishes to use daily data in order to increase the size of the time-series sample, some adjustment for non-synchroneity should be considered.

Brown and Weinstein (1983), using a data set and time period chosen to be the same as those chosen by Roll and Ross (1980), test the equality of the risk premia across subgroups of assets [i.e., they test $\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ in (14)].[13] Rather than performing the analysis on 42 groups of thirty stocks each, they use twenty-one groups of sixty stocks each. Each group of sixty assets is divided into two subgroups of thirty assets. For each group of sixty securities, maximum likelihood factor analysis is used to get an estimate, B, of the matrix of factor sensitivities as well as estimates for the two subgroups, $B^1$ and $B^2$. Let $B_u$ be the unrestricted factor beta matrix formed by stacking $B^1$ and $B^2$ (i.e., $B_u' = [B^{1'} : B^{2'}]$). An unrestricted form of the model is estimated by a cross-sectional GLS regression of the form (19) in which returns are regressed on $\iota^{60}$, $B_u$, and Z. The top $30 \times (k + 1)$ submatrix of instruments, Z, is a matrix of zeros and the bottom $30 \times (k + 1)$ submatrix of Z is equal to $[\iota^{30} : B^2]$. A restricted form of the model is estimated by a cross-sectional GLS regression of the form (17) in which returns are regressed on $\iota^{60}$ and B. The test statistic is formed from the diagonal elements of the restricted and unrestricted residual covariance matrices.[14] The test is equivalent to a test for a shift in the regression parameters (sometimes referred to as a Chow test). The law of one price implies that the price of risk should be the same across subgroups. Brown and Weinstein (1983) test the hypothesis of equal price of risk across subgroups for three, five, and seven factor models. They

find that the restrictions are rejected at standard levels of statistical significance but argue that this may be an artifact of the large number of observations available. That is, holding the size of the test (i.e., the probability of type I error) constant, the probability of a type II error approaches zero as the number of observations increases. Brown and Weinstein (1983) propose using a posterior odds ratio approach to alter the size of the test to reflect the large sample. After this adjustment, the tests still reject the hypothesis of equal prices of risk approximately fifty percent of the time.

Early factor-analytic-based empirical analyses of the APT tended to focus on small subgroups of securities (between 24 and 60 assets per group in the studies discussed above) because of the computational problems associated with performing factor analysis of large-scale covariance matrices. Much subsequent research has been devoted to developing methods that can accommodate large cross-sectional samples. One such method is proposed in Chen (1983). He analyzes daily stock return data over the 16-year period from 1963 through 1978, divided into four four-year subperiods. The number of assets analyzed in the subperiods is 1064, 1562, 1580, and 1378, respectively. He chooses the first 180 stocks (alphabetically) in each subperiod and uses factor analysis to estimate the factor sensitivities for a ten factor model. Factor-mimicking portfolios for a five factor model are then formed from these same 180 stocks by a mathematical programming algorithm that imposes a penalty for choosing portfolio weights very different from 1/n and which also disallows short positions. The factor sensitivities of the remaining n - 180 assets are estimated from their covariances with the factor-mimicking portfolios [see Chen (1983, equation A1)]. Cross-sectional regressions of the form (17) are estimated for the five factor APT and the CAPM (where the S&P 500, equal-weighted CRSP portfolio, and value-weighted CRSP portfolio are used as proxies for the market portfolio). The asset returns on even days are used as dependent variables while the factor sensitivities, B, and CAPM betas are estimated with data from odd days. Chen (1983) finds that the vector of average factor risk premia, $\overline{F}$, is significantly different from the zero vector.

Many studies focus only on the question of whether the restrictions implied by the APT can be rejected. A more important question is whether the model outperforms or underperforms alternative asset pricing models. This is a difficult problem because the competing hypotheses (e.g., the APT versus the CAPM) are not nested. That is, one hypothesis is not a restricted version of the other hypothesis. Chen (1983) addresses this issue by applying methods of testing non-nested hypotheses [see Davidson and Mackinnon (1981)]. Let $\hat{r}_{i,t,APT}$ denote the fitted value for $r_{i,t}$ from the regression (17) when the estimated factor sensitivities are used to form B, and let $\hat{r}_{i,t,CAPM}$ denote the fitted value for $r_{i,t}$ from the regression (17) when the estimated market betas are used to form B. Consider the cross-sectional regression

$$r_{i,t} = \alpha_t \hat{r}_{i,t,APT} + (1 - \alpha_t)\hat{r}_{i,t,CAPM} + e_{it}. \tag{21}$$

The time series of $\alpha_t$ can be used to calculate the mean value $\bar{\alpha}$, and the standard error of $\bar{\alpha}$. If the APT is the appropriate model of asset returns then one would expect $\bar{\alpha}$ to equal 1.0, while one would expect $\bar{\alpha}$ to equal zero if the CAPM is the appropriate model. Chen finds that, across the four subperiods and across various market portfolio proxies, he can often reject both the hypothesis that $\alpha = 0$ and the hypothesis that $\alpha = 1$. However, the point estimates are all very close to one. That is, $\bar{\alpha}$ is between 0.938 and 1.006. Also, Chen (1983) finds that the residuals from the CAPM cross-sectional regression (17) can be explained by the factor sensitivities while the residuals from the APT cross-sectional regression are not explained by assets' betas relative to the market portfolio. Thus, the data seem to support the APT as a better model of asset returns.

Chen (1983) also compares the returns on a portfolio of high variance stocks to the returns on a portfolio of low variance stocks constructed to have the same estimated factor sensitivities. If the APT is correct these two portfolios should have the same expected returns (since they have the same factor sensitivities, B). There is no significant difference in returns. The same procedure is applied to portfolios of large capitalization and small capitalization stocks. Chen finds that, while all of the point estimates indicate that large firms had lower

returns than small firms with the same factor risk, the difference is statistically significant in only one of the four subperiods.  He concludes that the size anomaly is explained by differences in factor risk.

Reinganum (1981) uses the same method of factor beta estimation as Chen (1983) to compare ten portfolios formed on the basis of market value of equity.  The returns on these portfolios are compared to control portfolios constructed to have the same sensitivity to the factors.  This is done for three, four, and five factor models.  Unlike Chen (1983), Reinganum (1981) concludes that the size anomaly is not explained by the APT.

The above studies use factor analysis, or some variant, to estimate assets' factor betas. An alternative approach is taken by Chen, Roll, and Ross (1986) who specify, *ex ante*, a set of observable variables as proxies for the systematic "state variables" or factors in the economy. The prespecified factors are (i) the monthly percentage change in industrial production (lead by one period)[15]; (ii) a measure of unexpected inflation; (iii) the change in expected inflation[16]; (iv) the difference in returns on low-grade (Baa and under) corporate bonds and long-term government bonds; and (v) the difference in returns on long-term government bonds and short-term Treasury bills.

Sixty months of time-series observations are used to estimate assets' betas relative to these prespecified factors.  Given these estimates of the factor sensitivities, $B$, cross-sectional regressions of returns on $B$ [as in (17)] are estimated in order to get estimates of the returns on factor-mimicking portfolios, $F_t$.  As in Fama and MacBeth (1973), portfolios rather than individual assets are used in these second-stage regressions in order to reduce the EIV problem caused by the use of $\hat{B}$ rather than $B$.  Chen, Roll, and Ross (1986) form twenty portfolios on the basis of firm size (market capitalization of equity) at the beginning of the particular test period. The average risk premia are estimated for the full sample period, January 1958 to December 1984, as well as three subperiods.

The average factor risk premia, $\bar{F}$, are statistically significant over the entire sample

period for the industrial production, unexpected inflation, and low-grade bond factors, and is marginally significant for the term-spread factor (v). To check how robust the results are to changes in the prespecified factors, Chen, Roll, and Ross (1986) perform the above exercise with the change in industrial production factor replaced by several alternative factors. One can view this as estimating (19) with the extra instruments, $Z_{t-1}$, being the betas on the extra factors. If the specified model is adequate, then $\delta$ should be equal to zero.

In the CAPM, the appropriate measure of risk is an asset's beta with respect to a market portfolio. Therefore, one logical alternative candidate as a factor would be a market portfolio proxy. The above analysis is conducted with the annual industrial production factor replaced by a market portfolio factor (either the equal-weighted or the value-weighted NYSE portfolio). They find that the risk premia on the market factors are not statistically significant when the other factors are included in the regression (17).

Consumption-based asset pricing models [e.g., Lucas (1978) and Breeden (1979)] imply that risk premia are determined by assets' covariance with agents' intertemporal marginal rate of substitution in consumption. This can be approximated by assets' covariance with changes in consumption. The growth rate in per capita real consumption is added as a factor (to replace the market portfolios). This growth rate is actually lead by one period to reflect the fact that there are lags in data collection. The risk premium on the consumption factor is not significant when the other five prespecified factors are included.

The last alternative factor analyzed by Chen, Roll, and Ross (1986) is the percentage change in the price of oil. The same analysis as above is performed with the beta of assets' returns with respect to changes in oil prices replacing the other alternative factors. The estimated risk premium associated with oil price shocks is statistically insignificant for the full period and for two of the three subperiods. The subperiod in which the premium is statistically significant is the 1958-1967 period.

Chen, Roll, and Ross (1986) conclude that the five prespecified factors provide a

reasonable specification of the sources of systematic and priced risk in the economy. This is based largely on their results which suggest that, after controlling for factor risk, other measures of risk (such as market betas or consumption betas) do not seem to be priced.

Chan, Chen, and Hsieh (1985) use the same set of factors as Chen, Roll, and Ross (1986) in order to determine whether cross-sectional differences in factor risk are enough to explain the size anomaly evident in the CAPM literature and in some previous APT studies [e.g., Reinganum (1981)]. For each test year from 1958 to 1977, an estimation period is defined as the previous five year interval (i.e, 1953-1957 is the estimation period for 1958, 1954-1958 is the estimation period for 1959, etc.). The sample consists of all NYSE firms that exist at the beginning of the estimation period and have price data at the end of the estimation period. Firm size is defined as the market capitalization of the firm's equity at the end of the estimation period. Each firm is ranked by firm size and assigned to one of twenty portfolios.

Chan, Chen, and Hsieh (1985) estimate the factor sensitivities of the twenty size-based portfolios relative to the prespecified factors and the equal-weighted NYSE portfolio over the estimation period. In the subsequent test year, cross-sectional regressions, such as (17), of portfolio returns on the estimated factor sensitivities, B, are run each month. This is repeated for each test year and yields a monthly time series of returns on factor-mimicking portfolios from January 1958 to December 1977.

If the risk premia from the factor model explain the size anomaly, then the time-series averages of the residuals from (17) should be zero. Chan, Chen, and Hsieh (1985) use paired $t$ tests and the Hotelling $T^2$ test to determine if the residuals have the same means across different size portfolios.[17] These tests are equivalent to estimating (19) where $Z_{t-1}$ represent various combinations of portfolio dummy variables and testing whether the elements of the vector $\delta$ are equal to each other.

Chan, Chen, and Hsieh (1985) find that the risk premium for the equal-weighted market portfolio is positive in each subperiod, but is not statistically significant. Over the entire period

they find significant premia for the industrial production factor, the unexpected inflation factor, and the low-grade bond spread factor. They find that the average residuals are not significantly different across portfolios and that the difference in the average residuals between the portfolio of smallest firms and the portfolio of largest firms, while positive, is not significantly different from zero. The average difference in monthly returns between these two portfolios is 0.956%; 0.453% is due to the low-grade bond risk premium, 0.352% is due to the NYSE market risk premium, 0.204% is due to the industrial production risk premium, and 0.120% is left unexplained.

Chan, Chen, and Hsieh also run regressions such as (19) in which the instrument, $Z_{t-1}$, is the logarithm of firm size. When the B matrix includes the betas for the prespecified factors and the equal-weighted NYSE portfolio, then the coefficient on firm size, $\delta$, is statistically significant. When B only contains betas for the prespecified factors, then $\delta$ is insignificant. They conclude that the multifactor model explains the size anomaly.

Shanken and Weinstein (1990) reevaluate the evidence on the risk premia associated with the prespecified factors used in Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986). While Shanken and Weinstein (1990) use the same set of five prespecified factors and time periods similar to those in Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986), they make several changes in the procedures. One adjustment is an EIV correction for the time-series standard errors of $\bar{F}$, which is derived in Shanken (1992a). This correction tends to increase the standard errors and, hence, decrease the reported test statistics.

A second change involves the manner in which the size-based portfolios are formed for the estimation of the matrix of factor sensitivities of those portfolios. Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986) form the size-based portfolios on the basis of the market capitalization of the firms at the *end* of the estimation period. For example, betas are estimated by Chan, Chen, and Hsieh (1985) over the period 1953-1957 for twenty size-based portfolios formed on the basis of market capitalization at the end of December 1957. Given these

estimates, cross-sectional regressions are run for the twelve months of 1958.  While this approach does not induce bias in the portfolio returns for 1958, it may induce correlation between the estimation error in betas, $\hat{B} - B$, and the allocation of firms to portfolios.  For example, some of the firms allocated to the small firm portfolios in December 1957 will have had poor performance over the period 1953-1957, while the opposite is true for the firms allocated to the large firm portfolios.  However, if the current beta is related to past performance, then the historical betas calculated over 1953-1957 will systematically misstate the current level of beta.  For example, leverage effects could lead to a negative relation between beta and performance (i.e., increases in beta for poor performers and decreases in beta for good performers).  This type of effect will cause the historical estimate of beta (as an estimate of beta for the *next year*) to be too small for the small firm portfolios and too large for the large firm portfolios.  Shanken and Weinstein (1990) argue that this decrease in dispersion of betas would lead to an upward bias in the estimated risk premia from the cross-sectional regressions (assuming the premia are non-zero in the first place).  This bias could lead to spurious significance in the estimated risk premia.

The alternative portfolio formation procedure used by Shanken and Weinstein (1990) is to form size portfolios at the beginning of each year and use asset returns over the subsequent year to estimate betas.  For example, for the 1953-1957 estimation period, form portfolios at the end of  December 1952 to calculate returns in 1953, form portfolios at the end of December 1953 to calculate returns in 1954, and so on.  This procedure does not induce correlation between beta estimation errors and portfolio groupings since the allocation to groups is chosen *ex ante*.

Shanken and Weinstein (1990) estimate cross-sectional regressions (18) using betas estimated from the prior five-year period as well as betas estimated over the same period as the cross-sectional regressions.  They check the sensitivity of the results to the number of portfolios used by estimating the cross-sectional regressions with 20, 60, and 120 portfolios (using WLS).  They also estimate restricted versions of the cross-sectional regressions that take advantage of

the fact that some of the prespecified factors are excess returns on financial assets. $\overline{F}_j$, the $j^{th}$ element of $\overline{F}$, is the excess return on a portfolio that mimics factor j. If factor j is an asset excess return then it mimics itself without error, so we can impose the restriction that $\overline{F}_j$ is equal to the time-series mean of the factor. The sample period of 1958-1983 is divided into three subperiods, 1958-1967, 1968-1977, and 1978-1983. With a design similar to that used by Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986) (using the prior period betas, 20 portfolios, and without the above restrictions imposed) none of the factor risk premia are statistically significant (at the 5% level) in the three subperiods. Only the industrial production factor premium is significant over the entire sample period. Using a larger number of cross-sections increases the evidence for a significant price of risk for this factor and provides some evidence for a significant risk premium associated with the low-grade bond factor in the first subperiod.

The use of contemporaneously estimated betas does not seem to influence the results greatly. The restricted estimates described above tend to decrease the significance of the low-grade bond factor and increase the significance of the industrial production and term-structure factors.

Similar to Chan, Chen, and Hsieh (1985), Shanken and Weinstein (1990) use the Hotelling $T^2$ statistic[18] to test whether the portfolio residuals from (18) have a mean of zero. The $T^2$ tests do not reject the hypothesis that the residuals have a mean of zero for both the unrestricted and restricted estimators. They also test whether the price of risk is equal across small and large firms. This is done in the framework of (20) where the instrument, $Z_{t-1}$, is the product of B and a dummy variable. The dummy variable is equal to one if the portfolio is one of the first n/2 size-based portfolios (where n is the total number of portfolios) and is equal to zero otherwise. If the price of risk is the same across subgroups, then $\delta$ should be zero. There is little evidence of differential pricing of risk for both the unrestricted and restricted estimators.

As in the previous studies, Shanken and Weinstein (1990) check the specification of the prespecified factor by including betas relative to a market portfolio proxy (the value-weighted

CRSP index) in the cross-sectional regressions. Using the design of Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986), the estimated market risk premium is not significant. Using the restricted model or the unrestricted model with contemporaneous betas, Shanken and Weinstein find that the estimated risk premium on the market proxy is statistically significant.

The results of Shanken and Weinstein (1990) suggest that the previous significance of the prespecified factor risk premia and the ability of those factors to render the market risk premium insignificant may be sensitive to the portfolio formation strategy and to whether or not one uses the EIV adjustment. The results also suggest that the choice of the number of assets or portfolios used in estimating the parameters in the cross-sectional regressions (equations 17-20) may have an important influence on the precision of the estimates.

A related issue regarding the portfolio formation process's influence on the power of statistical tests is raised in Warga (1989). He argues that the manner in which portfolios are chosen will tend to maximize the cross-sectional dispersion of assets' sensitivities to some factors but will yield low dispersion of assets' sensitivities to other factors. Dispersion in betas is important for the precision of the estimates in the cross-sectional regressions. The typical methods will then give precise estimates of the premia for some factors and imprecise estimates for others. He provides evidence that the size-based stratification will yield dispersion in assets' sensitivities to the low-grade bond factor but will yield low dispersion in assets' sensitivities to the market portfolio proxy. This implies low power against the hypothesis that the market risk premium is zero and may be an additional reason why Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986) found that market risk was insignificant. The larger number of portfolios in some of the tests in Shanken and Weinstein (1990) will increase dispersion in the betas and lead to more precise estimates.

Studies which test the APT in ways similar to Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986) include Burmeister and Wall (1986), Berry, Burmeister, and McElroy (1988b), Connor and Uhlaner (1988), Ferson and Harvey (am, 1991b), Wei, Lee, and Chen

(1991), and Cragg and Donald (1992).

## IV.2  Time-Series Tests of the APT

Now, rather than assuming we observe the matrix of factor betas, B, let us assume that

we observe $\lambda_{0,t-1}$, and $\lambda_{t-1} + f_t$, which represent the return on a zero-beta asset and the vector of

excess returns (i.e., returns in excess of the zero-beta return) of k portfolios which are perfectly

correlated with the factors.[19]  We can then view (15) and (16) as restricted versions of time-series

regressions of asset excess returns on the factor portfolio returns ($\lambda_{t-1} + f_t$) in which the

parameters to be estimated are the entries in the factor beta matrix, B.  For example, let $F_t$ denote

$\lambda_{t-1} + f_t$, assume that B is constant over time, and consider the time-series system of regressions:

$$R_t = \alpha + BF_t + \varepsilon_t \tag{22}$$

where $\alpha$ is an $n \times 1$ vector of intercept coefficients.  A testable restriction implied by the pricing

model is that $\alpha = 0$.  This approach to testing the specification of asset pricing models is used by

Black, Jensen, and Scholes (1972) to test the CAPM where $F_t$ represents the excess return on a

market portfolio proxy (the equal-weighted NYSE portfolio in their case).  Jobson (1982)

discusses this approach in an APT context.  A variant of this approach applies when the riskless

or zero-beta return is not observed.  Let $F_t^*$ denote $\lambda_{0,t-1}\iota^k + F_t$, the "raw" returns (i.e., not in excess

of the zero-beta return) on a set of k factor-mimicking portfolios, and consider the time-series

regression:

$$r_t = \alpha + BF_t^* + \varepsilon_t. \tag{23}$$

Under the assumption that $\lambda_{0,t-1}$ is constant through time and equal to $\lambda_0$, the asset pricing model

implies the restriction:

$$\alpha = (\iota^n - B\iota^k)\lambda_0.$$

This approach is used in a CAPM context in Gibbons (1982) with $F_t^*$ being the equal-weighted

NYSE portfolio.

The pricing restrictions that we have seen so far are equivalent to having some linear

combination of factor-mimicking portfolios on the mean/variance efficient frontier of asset returns (as discussed in section III.4).  A stronger condition is that the factor-mimicking portfolios span the entire mean/variance efficient frontier.  Spanning would imply the restrictions that $\alpha = 0$, or equivalently $B\iota^k = \iota^n$, in (23) [see Huberman and Kandel (1987)].

Lehmann and Modest (1988) perform time-series based tests of the APT restriction, $\alpha = 0$ in (22) and (23).  They divide the period from 1963 to 1982 into four five-year subperiods. Firms traded on the NYSE and AMEX that do not have missing daily data over a subperiod comprise the sample.  For each subperiod, 750 of these firms are selected at random and their daily returns are used to estimate the covariance matrix of returns.  Factor analysis is applied to the covariance matrix of returns in order to estimate the factor sensitivities of the assets. Lehmann and Modest (1988) use the *EM* algorithm [see Dempster, Laird, and Rubin (1977)] to factor analyze the full $750 \times 750$ return covariance matrix.[20]  This eliminates the need to analyze many small subsets of data, as was done previously by many authors.  The ability to use large numbers of individual assets to form factor-mimicking portfolios is an important improvement because it allows us to form well-diversified portfolios without inadvertently masking important characteristics of the data.

Given the $n \times k$ matrix of estimated factor sensitivities, $B$, and an estimate of the idiosyncratic covariance matrix, $\mathcal{V}$ (assumed to be diagonal), Lehmann and Modest form $k$ factor-mimicking portfolios and a zero-beta mimicking portfolio by minimizing the idiosyncratic risk of the portfolio subject to the constraint that the portfolio only has sensitivity to one factor. That is, the n-vector of portfolio weights for the $j^{th}$ factor-mimicking portfolio, $w_j$, is chosen to solve:

$$\min_{\omega_j} \quad \omega_j^/ \hat{V} \omega_j$$

$$\text{such that } \omega_j^/ \hat{B}_{.,s} = 0 \quad \text{for all } s \neq j \tag{24}$$

$$\omega_j^/ \iota^n = 1$$

where $B_{.,s}$ denotes the $s^{th}$ column of B. The zero-beta portfolio is formed in the same way except that $\omega_j^/ B_{.,s} = 0$ for all s [see Lehmann and Modest (1985, 1988) for details]. Given these portfolio weights, they calculate weekly returns on factor-mimicking portfolios for models with five, ten, and fifteen factors. Excess returns of these factor-mimicking portfolios are used as $F_t$ in the regressions (22) and raw returns are used as $F_t^*$ in the regressions (23).

Lehmann and Modest (1988) calculate several sets of weekly returns to be used as $R_t$ and $r_t$. All NYSE and AMEX firms that meet the data requirements are allocated to quintile and ventile portfolios. Two sets of sized-based portfolios are formed by ranking firms by market capitalization at the beginning of the test period and forming five and twenty equally-weighted portfolios, respectively. Two sets of dividend yield-based portfolios are formed by ranking firms by dividend yield in the year before the test period. The first portfolio in each set contains all firms with a zero dividend yield. The remaining assets are allocated equally to the other four or nineteen portfolios (depending on whether there are five or twenty portfolios in $R_t$). Finally, two sets of variance-based portfolios are formed by ranking firms by their sample variances in the year before the test period (using daily data) and forming five and twenty equally-weighted portfolios, respectively. The various sets of weekly portfolio returns are regressed on the raw or excess returns on the factor-mimicking portfolios in a standard multivariate regression analysis. Similar regressions are run with single-index market portfolio proxies, the CRSP equal-weighted and value-weighted portfolios.

Using the five size-based quintile portfolios, Lehmann and Modest (1988) reject the hypothesis (at p-values less than 5%) that $\alpha = 0$ in (22) and (23) for both of the CRSP indices

and the 5, 10, and 15 factor models (their Table 1).  Using the twenty size ventile portfolios, the single-index models are rejected while the APT models are generally not rejected.  Given that the models are rejected with the quintile portfolios, Lehmann and Modest (1988) argue that the failure to reject the models with the ventile portfolios may be due to lower power in that specification.

Using the five dividend quintile portfolios the single-index models are rejected while only the single-index model using the equal-weighted portfolio is rejected using the twenty yield portfolios.  The APT models are not rejected using either the quintile or ventile portfolios (their Table 4).  The results for the variance-based portfolios are similar to the results for the dividend yield portfolios (their Table 5).

As discussed above, if the factor portfolios span the mean/variance efficient frontier, then there is a testable restriction on the factor sensitivities, $B\iota^k = \iota^n$ in the regression (23).  Lehmann and Modest (1988, Table 8) test this restriction which is overwhelmingly rejected.

Lehmann and Modest (1988) conclude that, while the APT is rejected on the basis of the regressions with size-based portfolios, its apparent ability to explain the dividend yield and variance effects that are unexplained by the CAPM (with standard proxies for the market portfolio) make it a good alternative model of asset pricing.

Connor and Korajczyk (1988a) also use a large number of individual assets to form factor-mimicking portfolios.  They use the asymptotic principal components procedure derived in Connor and Korajczyk (1986).  The asymptotic principal components procedure provides a computationally feasible method of estimating factor-mimicking portfolios from very large cross-sections.  Let R denote the $n \times T$ matrix of excess returns on assets, assume that asset returns follow an approximate k-factor model, and define $\Omega$ to be equal to $R'R/n$.  Connor and Korajczyk (1986) show that the first k eigenvectors of the matrix $\Omega$ converge to excess returns on factor-mimicking portfolios (subject to the typical rotational indeterminacy).  Note that $\Omega$ is a $T \times T$ matrix so that one only needs to perform eigenvector decompositions of a $T \times T$ matrix,

regardless of the size of the cross-sectional sample.  Factor analytic approaches require the decomposition of an n × n matrix followed by a portfolio formation procedure such as (24) or cross-sectional regressions.  For large n and moderate T the computational burden of asymptotic principal components is much smaller than factor analytic procedures.  Also, the procedure does not require that T be larger than n, only that T be larger than k, the number of factors.  Some studies have used asymptotic principal components with cross-sectional samples in excess of 11,000.

Connor and Korajczyk (1988a) use monthly data on NYSE and AMEX firms over the twenty year period from 1964 to 1983.  The sample period is divided into four 5-year subperiods.  In each subperiod, the asymptotic principal components technique is applied to the returns, in excess of the one-month Treasury bill return, for all firms without any missing monthly returns over the subinterval.  This yields excess returns on factor-mimicking portfolios constructed from samples of 1487, 1720, 1734, and 1745 firms in the respective subperiods.  These portfolio excess returns are used as $F_t$ in (22) to test five-factor and ten-factor versions of the APT.

There are two sets of test assets used as $R_t$ in (22).  The first is a set of ten size-based portfolios.  Firms are ranked on the basis of market capitalization at the beginning of the five year subperiod and are allocated to ten equal-weighted size decile portfolios.  This is similar to the portfolio formation strategy of Lehmann and Modest (1988) except that there are ten rather than five or twenty portfolios.  The second set of test assets is the entire sample of individual assets for each subperiod.  The statistics used to test the hypothesis that $\alpha = 0$ require a decomposition of the idiosyncratic covariance matrix, $V$.  The tests of Lehmann and Modest (1988) and Connor and Korajczyk (1988a) when portfolios are used as $R_t$ do not place any restrictions on the specific form (such as diagonality) of $V$.[21]  However, when using individual assets, an unrestricted $V$ is not feasible (if for no other reason tan that there are more parameters to estimate than observations in the data).  The approach taken by Connor and Korajczyk (1988a) in this case is to assume that V is block diagonal by industry, where industry is defined

by 3-digit SIC codes.  That is, within a 3-digit industry V is unrestricted but $V_{i,j}$ is assumed to be zero if firms i and j are in different industries.  Connor and Korajczyk (1988a) also estimate an alternative regression which includes instruments, $Z_{t-1}$:

$$R_t = \alpha + BF_t + \delta Z_{t-1} + \varepsilon_t \tag{25}$$

where $Z_{t-1}$ is a January dummy variable, equal to 1 if month t is January and zero otherwise.  This is the time-series equivalent of (20) and the asset pricing model implies that $\alpha = 0$ and $\delta = 0$.  The choice of a January dummy variable for Z is motivated by the inability of the CAPM to explain seasonality in asset returns [Keim (1983)].

The test statistics in Connor and Korajczyk (1988a) are modified likelihood ratio statistics [see Rao (1973, pp 554-556)] which have an exact small sample distribution under the null hypothesis that the idiosyncratic returns, $\varepsilon_t$, are multivariate normal.  The modified statistic is used because the standard asymptotic tests seem to have poor small sample properties [Binder (1985) and Shanken (1985)].

Using the size portfolios as test assets, Connor and Korajczyk (1988a) reject (at the 5% level) $\alpha = 0$ in (22) for the value-weighted CAPM as well as the APT with five and ten factors, while the CAPM using the equal-weighted CRSP proxy is not rejected.  Using the seasonal instruments as in (25), the hypothesis that $\delta = 0$ is strongly rejected for the market portfolio proxies but not for the APT models, while the hypothesis that $\alpha = 0$ is rejected for the APT but not for the market proxies.

The test statistics seem to indicate that the APT models do a better job of explaining the seasonality in size portfolio returns but a worse job of explaining the non-seasonal size anomaly, relative to the single index CAPM-like models.  However, given that the models are not nested, a direct comparison of the test statistics can be misleading.  That is, a larger and therefore "more significant" test statistic for one model versus another does not necessarily mean that the former model fits the data less well.  As an analogy, consider testing $\alpha_i = 0$ for a single portfolio or asset i in (22), with $F_t$ either being a vector of five factors or a single market portfolio.  This test is a

simple t-test, defined as the estimate, $\hat{\alpha}_i$, divided by its standard error.  The t-statistic can be larger for a given model either because $\hat{\alpha}_i$ is larger *or* because the standard error is smaller (i.e., $\hat{\alpha}_i$ is measured with less error).  Using multiple factors tends to increase the $R^2$ of the regression and, consequently, the precision of the estimates of $\alpha_i$ increases.  Thus, we can have smaller deviations from the null hypothesis, $\alpha_i = 0$, in an economic sense that are more significant in the statistical sense.  As an informal check for this, Connor and Korajczyk (1988a) plot the estimates of $\alpha_i$ and $\delta_i$ for the size portfolios.  The plots bear out the indication that the APT models perform better in terms of explaining the seasonal effects.  There is a pronounced size pattern in $\delta_i$ for the CAPM models but no pattern for the APT models.  However, in contrast to the impression that might be given by the test statistics, there is no clear-cut difference in the magnitude of $\alpha_i$ between the APT models and the single-index models.  The stronger rejections of the restriction that $\alpha = 0$ in (25) seem to be due to greater precision of the estimate of $\alpha$ for the APT relative to the CAPM.[22]

In Connor and Korajczyk (1988a), the tests using individual assets rather than the size based portfolios do not provide much power to discriminate between models.  For most subperiods and hypotheses [i.e., $\alpha = 0$ in (22), $\alpha = 0$ in (25), and $\delta = 0$ in (25)] the tests either reject all models or fail to reject all models.  For a few of the tests the statistics lead to rejection of the CAPM and fail to reject the APT, while there are no cases of the reverse happening.  Finally, they test whether the estimates $\alpha_i$ and $\delta_i$ are related to market capitalization of the firm using a large-sample approximation to a posterior odds ratio.  The CAPM is rejected in almost every subperiod while the APT models tend to reject the hypothesis that $\alpha$ is not related to size but fail to reject that $\delta$ is not related to size.  This is consistent with the pattern of pricing errors for the size-based portfolios described above.

Just as some authors have specified, *ex ante*, certain macroeconomic series as being the pervasive factors [e.g., Chen, Roll, and Ross (1986)], other authors have specified, *ex ante*, sets of portfolios whose returns are assumed to be maximally correlated with the pervasive factors.

When macroeconomic series are used, a second step is required to form factor-mimicking portfolios (generally through cross-sectional regressions of asset returns on estimated betas). When *ex ante* specified portfolios are used, one can avoid the second step since the factors are asset returns which contain the appropriate risk premia. Huberman and Kandel (1987) specify the factors to be three size-based portfolios. Fama and French (1993) specify the factors to be five portfolio excess returns: (i) the return on a value-weighted market portfolio (in excess of the one-month Treasury bill return); (ii) the difference in returns on a small-firm portfolio and a large firm portfolio; (iii) the difference in returns on a portfolio of firms with high book-to-market equity (i.e., book value of equity relative to market value of equity) and a portfolio of firms with low book-to-market equity; (iv) the difference in the return on a long-term government bond portfolio and the return on the one-month Treasury bill; and (v) the difference in the return on a long-term corporate bond portfolio and the return on a long-term government bond portfolio.

Huberman and Kandel (1987) and Fama and French (1993) find that the multi-factor models do a much better job in explaining asset returns (i.e., values of $\alpha$ close to zero) than do standard single-index models.

McElroy and Burmeister (1988) postulate macroeconomic variables as observable factors and use nonlinear time-series regression to estimate the parameters of the factor model. Their approach allows joint estimation of the parameters of the model in one step rather than the two step procedures common to many of the previous studies. The pricing restrictions of the APT imply cross-equation restrictions on the statistical model. They use monthly returns on 70 individual stocks (from January 1972 through December 1982) as the set of test assets and five prespecified factors that are similar to the factors used by Chen, Roll, and Ross (1986). The five factors are: (i) the difference in returns of long-term corporate bonds and long-term government bonds plus a constant;[23] (ii) the difference in returns on long-term government bonds and short-term Treasury bills; (iii) a measure of unexpected deflation (the negative of unexpected

inflation); (iv) a measure of unexpected growth in sales; and (v) either a return on market index (the S&P 500 portfolio) or a "residual market factor" equal to the residuals from a regression of the market index on the other four factors.

Assuming that the prespecified factors correspond to the factor innovation, $f_t$, that the factor risk premia are constant through time ($\lambda_{t-1} = \lambda$ for all t), and that the exact pricing model holds, we can rewrite (16) as the multivariate time-series regression:

$$R_t = B\lambda + Bf_t + \varepsilon_t \tag{26}$$

where the parameters to be estimated are B and $\lambda$. The n - k nonlinear cross-equation restrictions implied by the model are requirements that the intercept in (26) be equal to $B\lambda$. McElroy and Burmeister (1988) present an error components motivation for including either the return on a well-diversified portfolio or the residuals from a regression of the return on that portfolio on the other macroeconomic factors (the "residual market factor") as one of the factors. In either case the model implies testable restrictions of the same form as above. They estimate (26) using iterated nonlinear seemingly unrelated regression (INLSUR)[24] and find that the estimated risk premia $\lambda$ are significantly different from zero (at the 5% level) for each factor except the unexpected deflation factor. The overidentifying cross-equation restrictions are not rejected, leading McElroy and Burmeister to conclude that the multifactor model used here is a "useful empirical framework" for linking macroeconomic innovations to expected asset returns.

Bossaerts and Green (1989) and Hollified (1993) test dynamic versions of the APT. They find that static, constant parameter models are rejected, while the dynamic models perform well. Bansal and Viswanathan (1993) implement their non-linear APT by noting that the return on the aggregate wealth portfolio and yields on default free bonds are free of idiosyncratic risk. The value-weighted NYSE portfolio is used as an aggregate wealth proxy while the yield on one-month Treasury bills and the yield spread between six and nine month Treasury bills are used as the idiosyncratic risk free yields. Some agent's intertemporal marginal rate of substitution is postulated to be an unknown non-linear function of these three idiosyncratic risk

free variables.  Bansal and Viswanathan (1993) use semi-non-parametric techniques to estimate the intertemporal marginal rate of substitution or state pricing function.  They find that linear versions of the model are rejected in favor of non-linear versions, using their choice of factors. They also find that a one-factor (the NYSE portfolio) non-linear model is rejected in favor of a two-factor (the NYSE portfolio and the one-month Treasury bill) non-linear model.  There is not much support for adding a third factor (the yield spread).  While there is some evidence indicating that the non-linear APT does not completely price assets and dynamic trading strategies, Bansal and Viswanathan (1993) argue that the non-linear models perform better than linear versions of their model.

As noted above, the results of classical significance tests can be difficult to interpret.  For example, the causes or economic implications of rejecting or failing to reject a model are often not addressed [see McCloskey (1985)].  Do we reject a model because it is a poor description of the data or because we have a huge amount of data?  Do we fail to reject a model because it is a good description of the data or because the tests have no power?  What is an economically significant departure from the model?

McCulloch and Rossi (1990, 1991) provide Bayesian analyses of time-series implementations of the APT which explicitly incorporate an evaluation of the informativeness of the data and measures of economic significance, in addition to statistical significance. McCulloch and Rossi (1991) evaluate the performance of the APT by calculating posterior odds ratios.  They use the same sample and factor-mimicking portfolio formation methods as Connor and Korajczyk (1988a) and investigate the null hypothesis that $\alpha = 0$ in (22).  The posterior odds ratio, K, for the null hypothesis versus the alternative that $\alpha \neq 0$ is given by:

$$K = \frac{p(D \mid \alpha = 0)}{p(D \mid \alpha \neq 0)} \times \frac{p(\alpha = 0)}{p(\alpha \neq 0)} \tag{27}$$

where D represents the sample data, $p(\alpha = 0)/p(\alpha \neq 0)$ is the prior odds ratio, and

$p(D \mid \alpha = 0)/p(D \mid \alpha \neq 0)$ is a ratio of predictive densities. The odds ratio explicitly takes into account the informativeness of the data. An odds ratio greater than 1:1 favors the null hypothesis while an odds ratio less than 1:1 favors the alternative hypothesis.

For a one-factor model McCulloch and Rossi (1991) find that the odds ratio favors the alternative hypothesis ($\alpha \neq 0$), except for the case when the prior distribution is relatively uninformative. For a five-factor model they find that the odds ratio favors the null hypothesis ($\alpha = 0$), except for the case when the prior distribution is relatively informative. The sensitivity of the odds ratio to the specification of the prior distribution leads McCulloch and Rossi to conclude that the data are relatively uninformative about the model.

McCulloch and Rossi (1990) derive utility-based metrics to assess the economic significance of deviations from the exact APT pricing restrictions. McCulloch and Rossi (1990) construct weekly returns on all NYSE and AMEX firms from January 1, 1963 to December 31, 1987. They construct weekly excess returns on factor-mimicking portfolios using the asymptotic principal components procedure of Connor and Korajczyk (1988b) and construct weekly returns on ten size-based portfolios with monthly rebalancing. The ten size-based portfolios are the test assets whose vector of pricing errors, $\alpha$, should be zero.

McCulloch and Rossi (1990) begin by evaluating the posterior distribution of $\alpha$ in (22) using a diffuse prior. They find evidence against the APT in the sense that the mass of the posterior distribution of $\alpha$ is often far from the null hypothesis of zero. McCulloch and Rossi (1990) wish to determine whether these deviations from the null hypothesis are economically significant. A reasonable metric is how much utility one would lose by assuming the null hypothesis is true. To determine this they investigate the posterior distribution of the difference in certainty equivalents between two utility-maximizing investors; one choosing portfolios assuming $\alpha \neq 0$ and the other choosing portfolios assuming $\alpha = 0$. A negative exponential utility function is postulated and normality of asset returns is assumed. The hypothetical investors choose to allocate their portfolios across the ten size-based portfolios and the riskless asset.

McCulloch and Rossi (1990) find that the dispersion on the posterior distribution of the certainty equivalents is quite large when the analysis is performed over five-year subintervals, thus confirming the odds ratio results indicating that the data are relatively uninformative. Over the full sample, however, the posterior distribution of the certainty equivalents is much tighter and closer to zero, the value implied by the null hypothesis. The predictive distribution of returns, with and without the restriction that $\alpha = 0$, is used to derive efficient frontiers. McCulloch and Rossi (1990) conclude that there is an economically significant difference between the unrestricted and restricted frontiers, but that the high level of parameter uncertainty makes definitive statements about the validity of the APT difficult.

Geweke and Zhou (1993) evaluate the posterior distribution of the average squared pricing deviations (the cross-sectional average of $\alpha_i^2$, i = 1, 2, ..., n) from the APT. They use industry and size-decile portfolios to estimate the posterior mean of the average squared pricing deviations. They argue that a one factor model explains most of the variation in expected returns with the remaining variation being economically negligible.

## IV.3    Summary of Tests of the APT

The tests often reject the overidentifying restrictions of the APT. However, this by itself is not as useful as a direct comparison of the APT to competing models of asset returns. This type of comparison is made difficult by the fact that the models are not, in general, nested models. In the cases in which the APT is compared to implementations of the CAPM, the APT seems to fare well in the sense that it does a better job of explaining cross-sectional differences in asset returns [e.g., the non-nested hypothesis tests of Chen (1983)], it seems to explain some pricing anomalies relative to the CAPM [e.g., the dividend yield anomaly seems to be eliminated by the APT in Lehmann and Modest (1988) while there are mixed results about the APT's ability to explain the size anomaly], and it generally has smaller pricing errors than the CAPM [e.g., the absolute size of $\alpha$ seems to be smaller for the APT, see Connor and Korajczyk (1988a, Figures 1-

6)].

On the other hand, there is evidence which suggests that the asset pricing models are not providing much information about unconditional cross-sectional differences in expected returns. In standard tests of the models, this is evident through the frequent inability of researchers to find significant risk premia for market risk or factor risk. The lack of information provided by the models is also evident in the sensitivity of the posterior odds ratios to changes in prior distributions [McCulloch and Rossi (1991)] and in the large dispersion in the posterior distributions of the difference in certainty equivalents in the utility-based approach of McCulloch and Rossi (1990). These difficulties are essentially all related to the fact that, given the inherent variability in asset returns, it is difficult to measure unconditional mean return with much precision. This problem is one shared by all models of unconditional asset pricing and is not specific to the APT.

## V.     Other Empirical Topics

The APT does not provide an *a priori* specification of the appropriate number of priced factors. The choice of the appropriate number of factors is complicated by the fact that, with a finite number of assets, alternative rotations of the factors can change the apparent factor structure [Shanken (1982)]. In Section V.1 we survey the literature on testing for the appropriate number of factors. In Section V.2 we discuss alternative methods of forming factor-mimicking portfolios that have not been discussed above and section V.3 contains a survey of international applications of the APT.

### V.1     Tests for the Appropriate Number of Factors

Estimates and tests of the APT require, as a maintained hypothesis, that returns follow a factor model with a pre-specified number of factors. Roll and Ross (1980) use a likelihood ratio test of the hypothesis that k factors are sufficient to characterize U.S. stock market returns. The data set and empirical estimation methodology of their paper have been discussed in Section

IV.1 above. The likelihood ratio test comes from the factor analysis literature [e.g., see Morrison (1976, section 9.5)] and is given by:

$$\left[ T - 1 - \frac{2n + 5}{6} - \frac{2k}{3} \right] \ln\left[ \frac{\hat{\Sigma}_k}{\hat{\Sigma}} \right]$$

where $\Sigma_k$ is the maximum likelihood estimate of the $n \times n$ covariance matrix of returns, $\Sigma$, under the constraint that returns follow a strict k-factor model; $\Sigma$ is the unconstrained maximum likelihood estimate $\Sigma$; T is the size of the time-series sample; and n is the number of assets in the cross-section. If asset returns follow a strict k-factor model and have a multivariate normal distribution, then the test statistic has an asymptotic distribution that is $\chi^2$ with degrees of freedom equal to $[(n - k)^2 - n - k]/2$ (where asymptotic means large T and fixed n). Roll and Ross apply the likelihood ratio test to 42 groups of 30 stocks each (sorted alphabetically). They find that, for most groups, five factors seems sufficient. In 32 of the 42 groups, the *p*-values of the test statistics (for the hypothesis that five factors were sufficient) were less that 0.50. Roll and Ross stress the tentative nature of their statistical tests; their paper is the first full-scale estimation and testing of the APT.

Dhrymes, Friend, and Gultekin (1984) increase the number of securities in each estimation group from 30 [the number used in Roll and Ross (1980)], to 60, 120, and 180. They repeat the likelihood ratio test for the number of factors on these larger cross-sectional sample sizes. They find that as the number of securities covered in the test increases, the number of statistically significant factors also increases. The Dhrymes, Friend, and Gultekin result is confirmed on British stock market returns data by Diacogiannis (1986).

There are at least two reasons why one might find that the number of significant factors increases as the number of assets increases. First, the likelihood ratio statistics are only asymptotically $\chi^2$. Conway and Reinganum (1988) demonstrate that there is a pronounced

tendency to find *too many* factors in small samples (i.e., small time-series samples).  If we hold the size of the time series fixed at T, and increase the number of cross-sections, n, then the effective size of the sample decreases and the small sample bias in favor of finding extra factors increases [also see Raveh (1985)].

Secondly, the likelihood ratio test assumes a strict factor model.  Suppose instead that returns obey an *approximate* factor model with, say, five factors.  In addition to the five pervasive factors, there are within-industry effects and other sources of cross-firm correlations which are not strong enough to qualify as pervasive sources of risk.  Using groups of thirty securities chosen randomly, the analyst is unlikely to identify these second-order sources of correlations as factors.  As the number of securities in the test increases, these "unimportant" factors may become statistically significant [Roll and Ross (1984b)].  The Dhrymes, Friend, and Gultekin (1984) findings highlight the weakness of the exact (as opposed to approximate) factor model assumption for security market returns data.

A separate issue regarding the Roll and Ross (1980) test for the number of factors is related to the adjustments for nonsynchroneity in Shanken (1987b).  As discussed in Section IV.1 above, Shanken adjusts the daily return covariance estimates for the presence of nonsynchronous trading.  He applies the likelihood ratio test to the adjusted covariance matrix, with different results.  Following Roll and Ross (1980) by using alphabetically-sorted groups of 30 securities each, Shanken finds at least a 99% chance of greater than ten factors in all cases.

The work of Chamberlain and Rothschild (1983) on approximate factor models has led to a search for alternative tests for the number of factors that are robust to the existence of an approximate, rather than a strict, factor model.  Recall from Section II.3 above that an approximate k-factor model is equivalent to exactly k eigenvalues of the covariance matrix of returns going to infinity as the number of cross-sections, n, increases to infinity.  If we can observe the sequence of covariance matrices (with increasing n) then we can look for the number of eigenvalues which grow unboundedly with n.  Note that this type of test relies only on an

approximate, not a strict, factor model, a substantial advantage for equity market returns data. Luedecke (1984) and Trzcinka (1986) provide the first statistical analysis along these lines. The problem, as they both note, is that the sampling properties of n-asymptotic (as opposed to T-asymptotic) eigenvalues are unknown, and so their work is exploratory. They find that the first eigenvalue of the sample covariance matrix is much larger than the others, and that all of the eigenvalues increase as n increases. By one possible standard (dominant eigenvalues), the empirical evidence indicates a single-factor model, whereas by another possible standard (increasing eigenvalues with n), the evidence points to a many-factor model.

Brown (1989) analyzes the behavior of the eigenvalues of the sample covariance matrix, $\Sigma$, through simulations. The simulated asset returns follow a four-factor model. Brown (1989) analytically derives the behavior, as n increases, of the eigenvalues of the population covariance matrix, $\Sigma$. The first four population eigenvalues grow with n while the remaining eigenvalues are constant. Brown then investigates the behavior of the sample eigenvalues through simulation. He applies the Luedecke-Trzcinka test to a simulated sample with the same dimensions (n and T) as that of Trzcinka. He finds that the first eigenvalue dominates (as in Luedecke and Trzcinka) and that *all* the other eigenvalues increase with n (again, as in Luedecke and Trzcinka). It is clear from Brown's simulations that we cannot infer the behavior, as n increases, of population eigenvalues, from the behavior, as n increases but with T *held constant*, of the sample eigenvalues. The problem is not the total number of return observations, but the *relative* size of the cross-sectional and time-series samples. This issue is also discussed in Connor and Korajczyk (1993).

Korajczyk and Viallet (1989) suggest a test for the number of factors which relies on the fact that well-diversified portfolios have no idiosyncratic risk (in the limit, as n approaches infinity). Assume that asset returns follow an approximate k-factor model, but that a k+1-factor model is estimated, where the k+1st factor is just picking up some idiosyncratic cross-correlations. In a time-series regression of a well-diversified portfolio's returns on the k+1

factors the coefficients should be statistically significant for the k pervasive factors and zero for factor k+1. Korajczyk and Viallet (1989) use the equal-weighted market portfolio (i.e., the portfolio weights are 1/n) as a proxy for a well-diversified portfolio. They find that this test identifies a large number of significant factors. This might be due to the fact that there are a large number of factors *or* due to the fact that the equal-weighted portfolio is, strictly speaking, only well-diversified when n is equal to infinity. Thus, the test may be finding factors due to the idiosyncratic risk left in the portfolio. This test is generalized in Heston (1991, example 5) to the case where the limiting portfolios are well-diversified, but need not have equal weights.

Connor and Korajczyk (1993) provide a different test for the number of factors in an approximate factor model. They analyze the decrease in cross-sectional average idiosyncratic variance in moving from a k factor model to a k+1 factor model. If returns are generated by a k factor model, then the expected decrease is zero, and Connor and Korajczyk provide a test statistic for a significant decrease. They find that the data suggest between one and six statistically significant factors.

The inferences from alternative tests for the number of factors tend to be bi-modal. There is a group of tests that indicates a very large number of factors and a group of tests that indicates a rather small number of factors. At this stage, there does not seem to be a general consensus on this point. A common approach taken by authors, in the face of this uncertainty about the appropriate number of factors, is to perform their analyses with various numbers of factors to determine whether the results are sensitive to the addition of factors.

**V.2     Alternative Factor-mimicking Portfolio Estimation Methods**

In Section IV we discussed several methods of constructing sets of factor-mimicking portfolios for use in testing the APT and estimating the risk premium associated with factor risk. The most frequently used approach is the cross-sectional regression of asset returns on some estimate of factor sensitivities, B, as in (17) and (18). The estimate of B may come from a time-

series regression of asset returns on prespecified factors or from factor analysis if the factors are not prespecified. An alternative approach is to prespecify the matrix of factor sensitivities directly. That is, assume that certain observable, firm specific, variables are equal to the factor sensitivities (or at least that they are equal to some linear combination of the factor sensitivities).

For example, assume that we can observe k attributes for each of the n firms (such as firm size, earnings/price ratios, etc.). Call the $n \times k$ matrix of attributes X. If we are willing to assume that $X = BL$ where L is some $k \times k$ nonsingular matrix, then cross-sectional regressions of returns on X will yield factor-mimicking portfolios that span the same space as portfolios created by regressing returns on B. The most important assumption is, of course, that $X = BL$. This is not very different from the implicit assumption used in studies that prespecify the factors to be particular macroeconomic innovations (i.e., that the macroeconomic variables are $L^{-1}F$ where L is a $k \times k$ nonsingular matrix and F is the $k \times T$ matrix of true factors).

This type of procedure is discussed by Rosenberg (1974) and used by Kale, Hakansson, and Platt (1991) who chose the firm attributes to include book value-to-price ratios, firm size, dividend yield, fraction of sales in various industries, and several other attributes.

Fama and French (1992) investigate the power of several firm attributes (size, book value/market value, leverage, earnings/price and market beta) to explain cross-sectional differences in asset returns. They use Fama-MacBeth cross-sectional regressions to estimate the excess returns on portfolios with unit average levels of each attribute (and zero average level of the other attributes). Fama and French (1992) find that the attributes of size and book/market ratios absorb the effects of the other attributes and that the market beta has no explanatory power. They conclude that there are multidimensional aspects of risk that are proxied by size and book/market ratios but not by betas relative to a market proxy. One possible interpretation of the results is that a multifactor asset pricing model is being used to price assets and that size and book/market ratios are good proxies for assets' sensitivities to the factors.

Mei (1993) suggests an alternative approach to estimating factor-mimicking portfolios.

He uses cross-sectional regressions to estimate the returns on factor-mimicking portfolios, but instead of using B as the set of independent variables, he uses realized returns from a prior period. The intuition for this can be most easily seen if we consider a noiseless factor model as described in Section III.1 [i.e., $\varepsilon = 0$ in (1)]. In this case the excess returns from the prior period are proportional to B since $R_t = B(\lambda_{t-1} + f_t)$ as in (16). Thus, if B is constant through time, the cross-sectional regression of excess returns on past excess returns is the same as a regression of returns on B [up to a scale transformation which is a function of the prior period factors, $(\lambda_{t-1} + f_t)$]. Mei (1993) suggests an instrumental variable approach to account for the fact that the return generating process does have an idiosyncratic return component.

### V.3    Tests of International Models

The empirical work described above uses data on assets in the United States exclusively. There have been a number of papers that perform the same or similar tests on the assets of other countries individually. There have also been a number of papers that use the APT to analyze asset returns across two or more countries.

Examples of single-economy applications of the APT are Chan and Beenstock (1984) and Abeysekera and Mahajan (1987) for the United Kingdom; Dumontier (1986) for France; Hamao (1988) and Brown and Otsuki (1990) for Japan; Hughes (1984) for Canada; and Winkelmann (1984) for Germany. Generally these papers have yielded similar inferences for these economies as the papers dealing with data from the United States. We will not describe these papers in detail here.

International versions of the APT are derived in Ross and Walsh (1983), Solnik (1983), and Levine (1989). Under the assumption that the exchange rate follows the same factor model as asset returns, Ross and Walsh (1983) and Solnik (1983) show that the same basic linear pricing result holds. If the exchange rate is spanned by the factors (i.e., it has no idiosyncratic risk) then we can change numeraires without changing the factor structure. On the other hand, if

the exchange rate has idiosyncratic risk, then changing numeraires will entail introducing an additional, but unpriced, factor [see, for example, Clyman, Edelson, and Hiller (1991)].

Integration across national markets would require that common sources of risk be priced in a consistent manner across countries. A number of authors have used international versions of the APT to assess the severity of capital controls, or barriers to market integration. Also, the assumption that exchange rates follow the same type of factor structure and are priced in a manner consistent with other assets has implications for the pricing of forward positions in currencies.

Cho, Eun, and Senbet (1986) use a variant of factor analysis, inter-battery factor analysis [see Cho (1984)], to estimate the factor sensitivity matrix, B, for factors common across pairs of countries. Inter-battery factor analysis is computationally less burdensome than standard factor analysis since it estimates factor sensitivities only for common factors. A drawback to the technique is that it cannot estimate country-specific factors, which are not ruled out, *a priori*, by the international APT. Cho, Eun, and Senbet (1986) then test for consistent pricing across countries [as in (14), where a subset is defined as the assets of one country] in a manner similar to that of Brown and Weinstein (1983).

Their sample consists of returns on 349 stocks from eleven countries from January 1973 through December 1983. The tests are performed separately for each possible pair of countries. Three hypotheses are investigated. The first is that $\lambda_0^i = \lambda_0$ in (14), the second is that $\lambda^i = \lambda$ in (14), and the third is that both $\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$. Since inter-battery factor analysis picks out only common factors, the second hypothesis alone is strictly implied by the exact version of the APT. The values of $\lambda_0^i$ may differ across countries since they could incorporate the risk premia for factors specific to that country which are still not globally diversifiable.[25] They reject (at the 5% level) the hypothesis that $\lambda_0^i = \lambda_0$ in three of the 55 country pairs. The hypothesis that $\lambda^i = \lambda$ is rejected in 30 of the 55 cases and the joint hypothesis that $\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ is rejected in 32 of the 55 pairs. Although the tests are not independent, the large fraction of rejections lead Cho,

Eun, and Senbet (1986) to conclude that the second and third hypotheses are not supported by the data. They suggest that this rejection may be due to lack of integration of capital markets or possibly to differential tax effects across countries.

Berges-Lobera (undated) tests for equality of factor risk premia across common stocks traded in the United States, Canada, the United Kingdom, and Spain. Monthly data from 1955 through 1980 are used for 100 firms each in the U.S. and U.K., 82 firms in Canada, and 62 firms in Spain. The hypothesis that pricing across markets is consistent is not rejected for the United States and Canada but is rejected for the United Kingdom/United States and United Kingdom/Canada pairs. The estimated risk premia for Spain are not precise enough to draw firm conclusions.

Korajczyk and Viallet (1989) perform time-series tests, as in (22) and (25), of single-economy and international versions of the CAPM and APT. They use monthly stock return data from France, Japan, the United States, and the United Kingdom over the period from January 1969 to December 1983. The number of firms with return data available ranges from 4211 to 6692. The asymptotic principal components technique is used to estimate the returns on factor-mimicking portfolios, $F_t$. The test assets that make up $R_t$ are sets of size based decile portfolios. For the single-economy versions, the factor portfolios and size portfolios are estimated using assets from one country (e.g., single-economy models for Japan would use Japanese stocks to estimate $F_t$ and form $R_t$). In the international versions, all of the assets are used to estimate $F_t$ and form the size portfolios $R_t$. In the international versions of the model, tests of the restriction $\alpha = 0$ in (22) are implicitly tests of equal prices of risk across countries [$\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ in (14)]. This is due to the fact that the method of forming factor-mimicking portfolios assumes consistent factor pricing across assets. Any differences in the pricing of factor risk across countries is then picked up in the intercept, $\alpha$, of the time-series regression. Over the full sample, the statistical tests provide some evidence against all of the models (CAPM and APT in single-economy and international versions). The APT seems to perform better than the CAPM,

in terms of the magnitudes of $\alpha$. An analysis of the size of the $\alpha$ across models does not yield a clear advantage to either single-economy or international versions of the models.

The sample period used in this study includes several important changes in international capital markets. There is a trend toward the relaxation of capital controls, which should lead to greater integration of markets. Also, the period includes a switch from fixed to floating exchange rates. Korajczyk and Viallet (1989) identify two periods, 1974 and 1979, as being particularly important periods of change. Estimates of $\alpha$ which allow for these periods to be isolated indicate that the rejections of the hypothesis that $\alpha = 0$ seem to be due to the earliest period (before February 1974). Since this corresponds to the period with the most severe barriers to international capital movements, the results are consistent with important pricing effects of capital controls.

Gultekin, Gultekin, and Penati (1989) use the APT to investigate the effect of a particular change in capital controls, a revision of Japan's Foreign Exchange and Foreign Trade Control Law (FEFTCL), which took effect in December 1980 [see Suzuki (1987)]. The revision of the FEFTCL amounted to a change from a regime with many barriers to capital flows to a regime with essentially no barriers to capital flows.

Gultekin, Gultekin, and Penati (1989) argue that while barriers to capital movements before the revision might lead to differential pricing of factor risk between Japan and other economies, the lack of barriers after the revision should lead to consistent pricing of factor risk [$\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ in (14) where i denotes the $i^{th}$ country].

Weekly common stock returns on 110 stocks traded in Japan and 110 stocks traded in the United States over the period 1977-1984 are used for the tests. The capital control period is 1977-1980 and the integrated period is 1981-1984. Gultekin, Gultekin, and Penati (1989) use both prespecified factors and factor analysis to estimate the assets' factor sensitivities, B. They find that they are able to reject the hypothesis of equal prices of risk across countries in the 1977-1980 period but are not able to reject the hypothesis in the 1981-1984 period. They

interpret the results as indicating capital market segregation before the revision in the FEFTCL and integration afterward. There is also some evidence that the risk premia are estimated less precisely in the 1981-1984 period, which might mean that the failure to reject in that period is due to the test having low power.

Another implication of the international versions of the APT is that the risk premia on forward positions in currencies should be explained by the currencies' sensitivities to the pervasive factors. There exists a substantial literature indicating time-varying returns on forward currency positions [e.g., Bilson (1981), Fama (1984), Korajczyk (1985), and Hodrick (1987)] which has been interpreted by some as a market inefficiency and by others as evidence of time-varying risk premia in the forward currency market. Korajczyk and Viallet (1992) test whether the observed premia can be explained by an international version of the APT. They form factor-mimicking portfolios from data on monthly common stock returns for 23,587 firms from Australia, France, Japan, the United States, and the United Kingdom over the period from January 1974 to December 1988. The number of firms, with return data available in a given month, ranges from 8,010 to 11,659. The asymptotic principal components technique is used to estimate the returns on factor-mimicking portfolios, $F_t$. The test asset returns, $R_t$, are the excess returns on forward positions in eight foreign currencies (the exchange rates are all relative to the U.S. dollar and are from Canada, France, Germany, Italy, Japan, the Netherlands, Switzerland, and the United Kingdom). They estimate time-series regressions such as (25) in which the instrument, $Z_{t-1}$, is the differential between the forward and spot exchange rates at the end of the previous month. If this implementation of the APT is successful in pricing currency returns, then $\alpha$ and $\delta$ in (25) should be zero.

Korajczyk and Viallet (1992) find that the factor model explains a large part of the risk premia in currency returns. However, they are able to reject the joint hypothesis that $\alpha = 0$ and $\delta = 0$ for the forward currency positions. Thus, the model does not provide a complete characterization of forward currency risk premia.

Heston, Rouwenhorst, and Wessels (1992) test for capital market integration between the United States and twelve European markets. They use monthly common stock returns on 4,490 stocks in the United States and 1,863 stocks on European markets, over the period 1978 through 1990, to estimate excess returns on factor-mimicking portfolios, $F_t$. The asymptotic principal components procedure is applied to the entire cross-sectional sample to estimate international factors and is applied to each country's assets to estimate domestic factor-mimicking portfolios.

Capital market integration is tested through time-series regressions of the form (22). The factors, $F_t$, are the excess returns on the international factor-mimicking portfolios. There are several sets of test assets. The first set of test asset excess returns, $R_t$, is composed of the equal-weighted market portfolios for each of the thirteen countries. The second set of test asset returns is composed of the value-weighted market portfolios for each of the thirteen countries. Then there are thirteen sets of test asset returns, one for each country, which are the first five domestic factor-mimicking portfolios. The null hypothesis, that $\alpha = 0$ in (22), finds mixed support. The null hypothesis is generally not rejected using the equal-weighted market portfolios or the domestic factor-mimicking portfolios, but is rejected using the value-weighted market portfolios as test assets.

Heston, Rouwenhorst, and Wessels (1992) also test whether forward currency returns are explained by the international factor-mimicking portfolios by estimating (22) and testing whether $\alpha = 0$ for the forward returns. This is similar to the tests of Korajczyk and Viallet (1992) except that Korajczyk and Viallet also include lagged instruments in the tests [as in (25)]. The results reject the hypothesis that $\alpha = 0$ for the forward currency returns.

Bansal, Hsieh, and Viswanathan (1993) apply the non-linear APT of Bansal and Viswanathan (1993) to the pricing of international equity indices (from Germany, Japan, the United Kingdom, and the United States), short-term bonds (U.S. Treasury bills and Eurodollar deposits), and forward currency contracts. They use weekly returns data from January 1975 through December 1990. They find that a non-linear single-factor model (with a world equity

index as the factor) is not rejected, while linear single-factor models are rejected.

## VI.    Applications

Asset pricing models have uses in a variety of applications in investments and corporate finance.  The APT has been used as an alternative to other asset pricing models for many applied problems, a few of which we discuss here.

### VI.1    Portfolio Performance Evaluation

A standard application of asset pricing models is the evaluation of the performance of professionally managed portfolios.  If the APT is the appropriate model of the risk/return tradeoff for securities, then all individual assets and portfolios formed on the basis of public information should have values of $\alpha$ in (22) equal to zero.  This corresponds to the case where all expected returns above the riskless rate are due to factor risk premia.  On the other hand, if a portfolio manager has superior ability in choosing assets, then one would expect that the manager's portfolio would earn higher rates of return than is warranted by its level of risk.  That is, superior ability should lead to values of $\alpha$ greater than zero.  Conversely, large transactions costs caused by excessive turnover should lead to negative values for $\alpha$.  Thus, $\alpha$ is one metric of risk-adjusted portfolio performance.  This measure has been used extensively in the context of the CAPM and has come to be known as Jensen's measure of portfolio performance [see Jensen (1968, 1969)].  Given the excess returns on factor-mimicking portfolios, $F_t$, $\alpha$ in (22) is simply the multi-factor, APT analog of Jensen's measure.

Lehmann and Modest (1987) provide an extensive comparison of APT-based and CAPM-based portfolio performance measures.  The equal-weighted and value-weighted NYSE portfolios are used as proxies for the market portfolio.  A variety of alternative implementations of the APT are used by Lehmann and Modest (1987).  For each estimation method, they estimate a version of the APT that assumes the existence of a riskless asset (the riskless rate version) and

a version that does not make this assumption (the zero-beta version). The matrix of factor sensitivities, B, is estimated by four alternative methods: (i) maximum likelihood factor analysis; (ii) restricted maximum likelihood factor analysis [where the restriction is that $E(r_t)$ is given by (13)]; (iii) principal components; and (iv) instrumental variables factor analysis [see Madansky (1964)]. Given the estimate, B, factor-mimicking portfolios are formed using the minimum idiosyncratic risk procedure described above [see (24)].

The sample used to estimate $F_t$ is essentially the same as the sample in Lehmann and Modest (1988). The returns used for $R_t$ are the monthly returns on 130 mutual funds over the period from January 1968 to December 1982. Lehmann and Modest (1987) find that the rankings of mutual funds and the average size of Jensen's measure is sensitive to whether the APT or CAPM benchmarks are used and to the type of factor estimation procedure used. The measured performance using the APT benchmarks was not sensitive to the number of factors beyond five factors. The CAPM-based performance measures were more highly related to simple average returns without risk adjustment than to the APT-based measures. The average Jensen measure, across funds, was consistently negative.

Connor and Korajczyk (1991) evaluate the performance of the same set of mutual funds used in Lehmann and Modest (1987) using a hybrid approach to constructing the factor-mimicking portfolios. The asymptotic principal components procedure is used to estimate excess returns on factor-mimicking portfolios. Then, linear combinations of these portfolios are formed so that they are maximally correlated with a set of macroeconomic factors, similar to those chosen by Chen, Roll, and Ross (1986). This combines the advantages of statistical estimation of the factors with the advantage of interpretability of the macroeconomic factors. As in Lehmann and Modest (1987), Connor and Korajczyk (1991) find that the average APT-based estimates of Jensen's measure for various portfolio classes (e.g., income, growth, maximum capital gain, etc.) are consistently negative as well as being different from the CAPM-based measures using the value-weighted NYSE/AMEX portfolio. Lehmann and Modest (1987) and

Connor and Korajczyk (1991) also address some issues related to the effects of market timing activities on the part of portfolio managers on Jensen's measure. We will not address those issues here [see also Admati, Bhattacharya, Pfleiderer, and Ross (1986)]. Rubio (1992) applies similar methods to a sample of Spanish mutual funds. He also finds negative fund performance, on average.

The negative average performance of mutual funds might be related to the size anomaly. Mutual funds tend to hold high capitalization stocks which have underperformed low capitalization stocks, on average.

Sharpe (1988, 1992) suggests a multifactor model of returns for portfolio evaluation where the factors are defined to be various asset classes. He adds the constraint that the factor benchmarks, against which the portfolios are compared, do not have short positions in assets. Other empirical studies of mutual fund performance using the APT include Chang and Lewellen (1985), Berry, Burmeister, and McElroy (1988a), and Frohlich (1991).

## VI.2    Cost of Capital Estimation

Another major use of asset pricing models is the estimation of costs of capital for use in capital budgeting problems. As in the portfolio performance evaluation literature, the CAPM has traditionally been the workhorse of risk adjustment in corporate finance texts. However, the APT is becoming a more common alternative to the CAPM [e.g., see Copeland and Weston (1988), Copeland, Koller, and Murrin (1990), Brealey and Myers (1991), and Ross, Westerfield, and Jaffe (1993)]. To the extent that one believes that the APT provides a better description of the risk/return tradeoff demanded by the capital market, the argument can be made for the use of the APT instead of the CAPM for cost of capital estimation.

The empirical literature on testing the APT, discussed in Section IV, and the extensive empirical literature on the CAPM, provide the most extensive set of information on the performance of the models. However, many studies investigate only one of the models, so that

making cross-model comparisons is sometimes difficult.

On a more pragmatic level, it is certainly of some interest to determine if costs of capital implied by the CAPM and APT are very different. Copeland, Koller, and Murrin (1990, exhibit 6.7) and Brealey and Myers (1991, Table 8-2) provide some comparisons for various industries, while Roll and Ross (1983) and Bower, Bower, and Logue (1984) provide estimates for utilities. While the CAPM and APT estimated costs of capital can be quite close to each other for some industries, they can be quite different for others. Thus, the choice of the appropriate model can be a substantive issue.

## VI.3   Event Studies

Single index models are used extensively in studies of market reaction to firm-specific or industry-specific events. This method was originally developed by Fama, Fisher, Jensen and Roll (1969).[26] The notion is that firm-specific news should be reflected in the idiosyncratic component of returns, $\varepsilon$, in (1). If we wish to study the market's reaction to a firm-specific (or at least nonpervasive) announcement,[27] then $\varepsilon_t$ provides a less noisy estimate of the reaction than $r_t$. If including multiple factors reduces the variability of $\varepsilon_t$ attributable to news other than the event in question, then using multiple factors might increase the accuracy of the estimated effect and the power of any related hypothesis tests. Merely adding factors, however, does not guarantee more precise estimates of $\varepsilon_t$, since the variance of $\hat{\varepsilon}_t$ is determined by the population variance of $\varepsilon_t$ *and* the sampling error of B. Adding factors would decrease the population variance but could increase or decrease the sampling variance. Thus, the use of multifactor models in event studies does not necessarily lead to unambiguous improvement. Brown and Weinstein (1985) and Chen, Copeland, and Mayers (1987) compare single and multiple factor approaches to estimating the valuation effects of news.

Brown and Weinstein (1985) simulate abnormal returns in a manner similar to that of Brown and Warner (1980, 1985) and tabulate the size and power of single and multiple factor

models for detecting these abnormal returns. They find that there is not an appreciable difference between single and multiple factor results. The multiple factor models seem to perform marginally better in their simulations.

Chen, Copeland, and Mayers (1987) apply single factor and multiple factor models to portfolios formed on the basis of assets' ranking of forecasted performance by *Value Line* and on the basis of firm size. They find that neither procedure has a particular bias. In terms of the variance of the estimate $\hat{\varepsilon}_t$, they find that single factor models tend to perform better when the test portfolio return, $r_t$, is poorly diversified, while multiple factor models tend to perform better when the test portfolio is diversified. This is due to the fact that diversification of the portfolio leads to lower estimation error in B, which in turn leads to a smaller variance in $\hat{\varepsilon}_t$.

The applications of multifactor models to event studies are somewhat peripheral to the question of whether the APT, the CAPM, or some other model is a better model for assets' expected returns. This is due to the fact that the event study applications rarely impose the restrictions implied by the various pricing models. This strand of the literature is more in the spirit of the early studies on the factor structure of asset returns, which were primarily interested in a parsimonious description of the primary variables influencing returns.

## VII.    Conclusion

The APT is based on a simple and intuitive concept. Ross's basic insight was that a linear factor model of asset returns, in an economy with a large number of available assets, implies that idiosyncratic risk is diversifiable and that the equilibrium prices of securities will be approximately linear in their factor exposures. This idea has spawned a literature which has pushed the scientific frontiers in several directions. It has led to new work in mathematical economics on infinite-dimensional vector spaces as models of many-asset portfolio returns, and the properties of continuous pricing operators on these vector spaces. It has led to econometric insights about what constitutes a factor model, and how to efficiently estimate factor models with

large cross-sectional data sets. It has underpinned an enormous body of empirical research on asset pricing relationships, and on related topics such as performance measurement and cost of capital estimation.

Lack of arbitrage opportunities implies that assets can be priced by a single random variable, variously referred to in the literature as the pricing kernel, stochastic discount factor, intertemporal marginal rate of substitution, or state price density [see Ross (1978), Dybvig and Ross (1989), Ferson (1993)]. One might wonder, then, what the advantage would be to using a multiple factor model. Particular asset pricing models differ in their specification of the stochastic discount factor. If there is an advantage to using multifactor models, it must be that the multifactor models provide a closer approximation to the stochastic discount factor than alternative approaches. To date, the empirical literature has tended to emphasize tests of the restrictions of a single model rather than emphasize comparisons across models. When comparisons across models have been made, the APT has tended to do well against the competing models. More of these cross-model comparisons are needed to assess relative performance across models. Many studies have rejected the strict restrictions of various asset pricing models, including the APT. The persistence and size of these asset pricing anomalies may not be total explicable within the paradigm of frictionless markets [MacKinlay (1993)]. The existence of frictions in asset markets has potential for explaining some of the observed failures of existing models [e.g., see Luttmer (1993)].

As Fama (1991) stresses, one cannot expect any particular asset pricing model to completely describe reality; an asset pricing model is a success if it improves our understanding of security market returns. By this standard, the APT is a success. The APT does have weaknesses and gaps. Current statistical methods are not amenable to testing an approximate pricing relation. As a result, our tests of the exact multifactor pricing relation are joint tests of the APT and additional assumptions necessary to obtain exact pricing. The empirical work on identifying the factor structure in security returns has had mixed success, and the econometric

techniques in this area are insufficiently developed, particularly with respect to incorporating conditioning information. The APT would be a better model if we could relate the factors more closely to identifiable sources of economic risk. Understanding the relationship between return factors and economic risks requires more work in asset pricing theory, macroeconomics, and econometrics. The APT will continue to evolve and may eventually be changed beyond recognition. Yet whatever changes occur, Ross's creative insight will endure as a fundamental building block in asset pricing theory.

# References

Abeysekera, Sarath P. and Arvind Mahajan, 1987, A test of the APT in pricing UK stocks, *Journal of Business Finance & Accounting* **14**, 377-391.

Admati, Anat R., Sudipto Bhattacharya, Paul Pfleiderer, and Stephen A. Ross, 1986, On timing and selectivity, *Journal of Finance* **41**, 715-730.

Admati, Anat R. and Paul Pfleiderer, 1985, Interpreting the factor risk premia in the arbitrage pricing theory, *Journal of Economic Theory* **35**, 191-195.

Andersen, Torben G., 1989, Estimation of systematic risk in the presence of non-trading: Comments and extensions, Working paper, Yale University.

Anderson, T. W., 1984, *An introduction to multivariate statistical analysis* 2nd ed. (Wiley, New York).

Bansal, Ravi, and S. Viswanathan, 1993, No arbitrage and arbitrage pricing: A new approach, Forthcoming *Journal of Finance*.

Bansal, Ravi, David A. Hsieh, and S. Viswanathan, 1993, A new approach to international arbitrage pricing, Forthcoming *Journal of Finance*.

Berges-Lobera, Angel, undated, An empirical study on international asset pricing models and capital market integration, Working paper, Universidad Autonoma de Madrid.

Berry, Mike, Ed Burmeister, and Marjorie McElroy, 1988a, A practical perspective on evaluating mutual fund risk, *Investment Management Review* **2**, 78-86.

Berry, Michael A., Edwin Burmeister, and Marjorie B. McElroy, 1988b, Sorting out risks using known APT factors, *Financial Analysts Journal* **44**, 29-42.

Bilson, John F. O., 1981, The 'speculative efficiency' hypothesis, *Journal of Business* **54**, 435-452.

Binder, John J., 1985, On the use of the multivariate regression model in event studies, *Journal of Accounting Research* **23**, 370-383.

Black, Fischer, Michael C. Jensen, and Myron Scholes, 1972, The capital asset pricing model: Some empirical tests, in: Michael C. Jensen (ed.), *Studies in the theory of capital markets*, (Praeger, New York).

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* **31**, 307-327.

Bossaerts, Peter and Richard C. Green, 1989, A general equilibrium model of changing risk premia: Theory and tests, *Review of Financial Studies* **2**, 467-493.

Bower, Dorothy H., Richard S. Bower, and Dennis E. Logue, 1984, Arbitrage pricing theory and utility stock returns, *Journal of Finance* **39**, 1041-1054.

Brealey, Richard A. and Stewart C. Myers, 1991, *Principles of corporate finance*, 4th ed. (McGraw-Hill, New York).

Breeden, Douglas T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, *Journal of Financial Economics* **7**, 265-296.

Brennan, M. J., 1971, Capital asset pricing and the structure of security returns, Working paper, University of British Columbia.

Brock, William A., 1982, Asset prices in a production economy, in: John J. McCall (ed.), *The economics of information and uncertainty*, (University of Chicago Press, Chicago).

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

Brown, Stephen J. and Toshiyuki Otsuki, 1990, Macroeconomic factors and the japanese equity markets: The CAPMD project, in: Edwin J. Elton and Martin J. Gruber, *Japanese capital markets*, (Harper & Row, New York).

Brown, Stephen J. and Jerold B. Warner, 1980, Measuring security price performance, *Journal of Financial Economics* **8**, 205-258.

Brown, Stephen J. and Jerold B. Warner, 1985, Using daily stock returns: The case of event studies, *Journal of Financial Economics* **14**, 3-31.

Brown, Stephen J. and Mark I. Weinstein, 1983, A new approach to asset pricing models: The bilinear paradigm, *Journal of Finance* **38**, 711-743.

Brown, Stephen J. and Mark I. Weinstein, 1985, Derived factors in event studies, *Journal of Financial Economics* **14**, 491-495.

Burmeister, Edwin and Marjorie B. McElroy, 1988, Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory, *Journal of Finance* **43**, 721-733.

Burmeister, Edwin and Kent D. Wall, 1986, The arbitrage pricing theory and macroeconomic factor measures, *The Financial Review* **21**, 1-20.

Chamberlain, Gary, 1983, Funds, factors and diversification in arbitrage pricing models, *Econometrica* **51**, 1305-1323.

Chamberlain, Gary, 1988, Asset pricing in multiperiod securities markets, *Econometrica* **51**, 1283-1300.

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

Chan, K. C., Nai-fu Chen, and David Hsieh, 1985, An exploratory investigation of the firm size effect, *Journal of Financial Economics* **14**, 451-471.

Chan, Kam-Fai, and Michael Beenstock, 1984, Testing the arbitrage pricing theory in the U.K. 1961-1982, in: *Proceedings of the 11th annual meeting of the European Finance Association*, (Manchester).

Chang, Eric C. and Wilbur G. Lewellen, 1985, An arbitrage pricing approach to evaluating

mutual fund perfomance, *Journal of Financial Research* **8**, 15-30.

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

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

Chen, Nai-fu and Jonathan E. Ingersoll, Jr., 1983, Exact pricing in linear factor models with finitely many assets:  A note, *Journal of Finance* **38**, 985-988.

Chen, Nai-fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* **59**, 383-403.

Chen, Zhiwu, and Peter J. Knez, 1992, A pricing operator-based testing foundation for the arbitrage pricing theory, Working paper, University of Wisconsin.

Cho, D. Chinhyung, 1984, On testing the arbitrage pricing theory:  Inter-battery factor analysis, *Journal of Finance* **39**, 1485-1502.

Cho, D. Chinhyung, Edwin J. Elton, and Martin J. Gruber, 1984, On the robustness of the Roll and Ross arbitrage pricing theory, *Journal of Financial Quantitative Analysis* **19**, 1-10.

Cho, D. Chinhyung, Cheol S. Eun, and Lemma W. Senbet, 1986, International arbitrage pricing theory:  An empirical investigation, *Journal of Finance* **41**, 313-329.

Clyman, Dana R., Michael E. Edelson, and Randall S. Hiller, 1991, International arbitrage pricing, risk premia and exchange rate drift, Working paper 92-019, (Harvard Business School, Boston, MA).

Cohen, Kalman J., Gabriel A. Hawawini, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb, 1983, Friction in the trading process and the estimation of systematic risk, *Journal of Financial Economics* **12**, 263-278.

Cohen, Kalman J. and Jerry A. Pogue, 1967, An empirical evaluation of alternative portfolio-selection models, *Journal of Business* **40**, 166-193.

Connor, Gregory, 1982, Asset pricing theory in factor economies, Ph.D. dissertation, (Yale University, New Haven, CT).

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

Connor, Gregory, 1989, Notes on the arbitrage pricing theory, in:  Sudipto Bhattacharya and George M. Constantinides (eds.), *Theory of valuation:  Frontiers of modern financial theory*, Vol. 1 (Rowman & Littlefield, Totowa, NJ).

Connor, Gregory and Robert A. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory:  A new framework for analysis, *Journal of Financial Economics* **15**, 373-394.

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

Connor, Gregory and Robert A. Korajczyk, 1988b, Estimating pervasive economic factors with missing observations, Working paper 34, (Department of Finance, Northwestern University, Evanston, IL).

Connor, Gregory and Robert A. Korajczyk, 1989, An intertemporal equilibrium beta pricing model, *Review of Financial Studies* **2**, 373-392.

Connor, Gregory and Robert A. Korajczyk, 1991, The attributes, behavior, and performance of U.S. mutual funds, *Review of Quantitative Finance and Accounting* **1**, 5-26.

Connor, Gregory and Robert A. Korajczyk, 1993, A test for the number of factors in an approximate factor model, Forthcoming, *Journal of Finance* **48**.

Connor, Gregory and Robert Uhlaner, 1988, New cross-sectional regression tests of beta pricing models, Working paper, (School of Business Administration, University of California, Berkeley, CA).

Connor, Gregory and Robert Uhlaner, 1989, A synthesis of two approaches to factor estimation, Working paper, (School of Business Administration, University of California, Berkeley, CA).

Constantinides, George M., 1989, Theory of valuation:  Overview and recent developments, in: Sudipto Bhattacharya and George M. Constantinides (eds.), *Theory of valuation: Frontiers of modern financial theory*, Vol. 1 (Rowman & Littlefield, Totowa, NJ).

Conway, Delores A. and Marc R. Reinganum, 1988, Stable factors in security returns: Identification through cross validation, *Journal of Business & Economic Statistics* **6**, 1-15.

Copeland, Thomas E. and J. Fred Weston, 1988, *Financial theory and corporate policy*, 3rd ed. (Addison-Wesley, Reading).

Copeland, Tom, Tim Koller, and Jack Murrin, 1990, *Valuation:  Measuring and managing the value of companies*, (Wiley, New York).

Cragg, John G. and Stephen G. Donald, 1992, Testing and determining arbitrage pricing structure from regressions on macro variables, Working paper, (University of British Columbia, Vancouver, BC).

Cragg, John G. and Burton G. Malkiel, 1982, *Expectations and the structure of share prices*, (University of Chicago Press, Chicago).

Davidson, Russell, and James G. Mackinnon, 1981, Several tests for model specification in the presence of alternative hypotheses, *Econometrica* **49**, 781-793.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977, Maximum likelihood from incomplete data via the *EM* algorithm, *Journal of the Royal Statistical Society, Series B* **39**, 1-22.

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

Diacogiannis, George P., 1986, Arbitrage pricing model:  A critical examination of its empirical applicability for the London stock exchange, *Journal of Business Finance & Accounting* **13**, 489-504.

Dumontier, Pascal, 1986, Le modèle d'evaluation par arbitrage des actifs financiers:  Une etude sur le marché financier parisien, *Finance* **7**, 7-21.

Dybvig, Philip H., 1983, An explicit bound on deviations from APT pricing in a finite economy, *Journal of Financial Economics* **12**, 483-496.

Dybvig, Philip H. and Stephen A. Ross, 1985, Yes, the APT is testable, *Journal of Finance* **40**, 1173-1188.

Dybvig, Philip H. and Stephen A. Ross, 1989, Arbitrage, in:  John Eatwell, Murray Milgate, and Peter Newman (eds.), *The new palgrave:  Finance*, (Norton, New York).

Elton, Edwin J. and  Martin J. Gruber, 1973, Estimating the dependence structure of share prices - implications for portfolio selection, *Journal of Finance* **28**, 1203-1232.

Engle, Robert F., Victor K. Ng, and Michael Rothschild, 1990, Asset pricing with a factor-ARCH covariance structure:  Empirical estimates for treasury bills, *Journal of Econometrics* **45**, 213-237.

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

Fama, Eugene F., 1984, Forward and spot exchange rates, *Journal of Monetary Economics* **14**, 319-338.

Fama, Eugene F., 1991, Efficient capital markets:  II, *Journal Finance* **46**, 1575-1617.

Fama, Eugene F., Lawrence Fisher, Michael C. Jensen, and Richard Roll, 1969, The adjustment of stock prices to new information, *International Economic Review* **10**, 1-21.

Fama, Eugene F. and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* **47**, 427-465.

Fama, Eugene F. and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**, 3-56.

Fama, Eugene F. and Michael R. Gibbons, 1984, A comparison of inflation forecasts, *Journal of Monetary Economics* **13**, 327-348.

Fama, Eugene F. and James D. MacBeth, 1973, Risk, return, and equilibrium:  Empirical tests, *Journal of Political Economy* **71**, 607-636.

Farrar, Donald E., 1962, *The investment decision under uncertainy*, (Prentice-Hall, Englewood Cliffs).

Ferson, Wayne E., 1992, Asset pricing models, Working paper 351, Center for Research in Security Prices, University of Chicago, Forthcoming, in: Douglas Greenwald (ed.), *The encyclopedia of economics*, (McGraw-Hill, New York).

Ferson, Wayne E., 1993, Theory and empirical testing of asset pricing models, in: Robert Jarrow, Vojislav Maksimovic, and William Ziemba (eds.), *Finance handbook*, (North-Holland, Amsterdam).

Ferson, Wayne E. and Campbell R. Harvey, 1991a, The variation of economic risk premiums, *Journal of Political Economy* **99**, 385-415.

Ferson, Wayne E. and Campbell R. Harvey, 1991b, Sources of predictability in portfolio returns, *Financial Analysts Journal* **47**, 49-56.

Ferson, Wayne E. and Robert A. Korajczyk, 1993, Do arbitrage pricing models explain the predictability of stock returns?, Working paper 115, (Department of Finance, Northwestern University, Evanston, IL).

Frohlich, C. J., 1991, A performance measure for mutual funds using the Connor-Korajczyk methodology: An empirical study, *Review of Quantitative Finance and Accounting* **1**, 427-434.

Gallant, A. Ronald, 1987, *Nonlinear Statistical Models*, (Wiley, New York).

Gehr, Adam Jr., 1978, Some tests of the arbitrage pricing theory, *Journal of the Midwest Finance Association* **7**, 91-106.

Geweke, John and Guofu Zhou, 1993, Measuring the pricing error of the arbitrage pricing theory, Working paper, (University of Minnesota, Minneapolis, MN).

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* **10**, 3-27.

Gibbons, Michael R., 1986, Empirical examination of the return generating process of the arbitrage pricing theory, Research paper 881, Graduate School of Business, Stanford University.

Gilles, Christian, and Stephen F. LeRoy, 1990, The arbitrage pricing theory: A geometric interpretation, Working paper, (Carleton University, Ottawa, Ontario).

Gilles, Christian, and Stephen F. LeRoy, 1991, On the arbitrage pricing theory, *Economic Theory* **1**, 213-229.

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

Grinblatt, Mark and Sheridan Titman, 1985, Approximate factor structures: Interpretations and implications for empirical tests, *Journal of Finance* **40**, 1367-1373.

Grinblatt, Mark and Sheridan Titman, 1987, The relation between mean-variance efficiency and arbitrage pricing, *Journal of Business* **60**, 97-112.

Gultekin, Mustafa N., N. Bulent Gultekin, and Alessandro Penati, 1989, Capital controls and international capital market segmentation:  The evidence from the Japanese and American stock markets, *Journal of Finance* **44**, 849-869.

Hamao, Yasushi, 1988, An empirical examination of the arbitrage pricing theory:  Using japanese data, *Japan and the World Economy* **1**, 45-61.

Hansen, Lars Peter and Ravi Jagannathan, 1991a, Implications of security market data for models of dynamic economies, *Journal of Political Economy* **99**, 225-262.

Hansen, Lars Peter and Ravi Jagannathan, 1991b, Assessing specification errors in stochastic discount factor models, Working paper (University of Chicago, Chicago, IL).

Harrington, Diana R., 1987, *Modern portfolio theory, the capital asset pricing model & arbitrage pricing theory:  A user's guide*, 2nd ed. (Prentice-Hall, Englewood Cliffs).

Harrison, J. Michael and David M. Kreps, 1979, Martingales and arbitrage in multiperiod securities markets, *Journal Economic Theory* **20**, 381-408.

Heston, Steve, 1991, Testing approximate linear asset pricing models, Working paper, (Yale University, New Haven, CT).

Heston, Steven L., K. Geert Rouwenhorst, and Roberto E. Wessels, 1992, The structure of international stock returns and the integration of capital markets, Working paper, (Yale University, New Haven, CT).

Hodrick, Robert J., 1987, *The empirical evidence on the efficiency of forward and futures foreign exchange markets*, Harwood Academic Publishers.

Hollifield, Burton, 1993, Linear asset pricing with time-varying betas and risk premia, Working paper, (University of British Columbia, Vancouver, BC).

Huberman, Gur, 1982, A simple approach to arbitrage pricing, *Journal of Economic Theory* **28**, 183-191.

Huberman, Gur, 1989, Arbitrage pricing theory, in:  John Eatwell, Murray Milgate, and Peter Newman (eds.), *The new palgrave:  Finance*, (Norton, New York).

Huberman, Gur, and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* **42**, 873-888.

Huberman, Gur, Shmuel Kandel, and G. Andrew Karolyi, 1987, Size and industry related covariations of stock returns, Working paper 202, (CRSP, University of Chicago, Chicago, IL).

Huberman, Gur, Shmuel Kandel, and Robert F. Stambaugh, 1987, Mimicking portfolios and exact asset pricing, *Journal of Finance* **42**, 1-9.

Hughes, Patricia J., 1984, A test of the arbitrage pricing theory using Canadian security returns, *Canadian Journal of Administrative Science* **1**, 195-214.

Ingersoll, Jonathan E., Jr., 1984, Some results in the theory of arbitrage pricing, *Journal of Finance* **39**, 1021-1039.

Ingersoll, Jonathan E., Jr., 1987, *Theory of financial decision making*, (Rowman & Littlefield, Totowa, NJ).

Jagannathan, Ravi and S. Viswanathan, 1988, Linear factor pricing, term structure of interest rates and the small firm anomaly, Working paper 57, (Department of Finance, Northwestern University, Evanston, IL).

Jarrow, Robert A., 1988, Preferences, continuity, and the arbitrage pricing theory, *Review of Financial Studies* **1**, 159-172.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* **23**, 389-419.

Jensen, Michael C., 1969, Risk, the pricing of capital assets, and the evaluation of investment portfolios, *Journal of Business* **42**, 167-247.

Jobson, J. D., 1982, A multivariate linear regression test of the arbitrage pricing theory, *Journal of Finance* **37**, 1037-1042.

Kale, Jivendra K., Nils H. Hakansson, and Gerald W. Platt, 1991, Industry vs. other factors in risk prediction, Finance working paper 201, (University of California, Berkeley, CA).

Kandel, Shmuel and Robert F. Stambaugh, 1989, A mean-variance framework for tests of asset pricing models, *Review of Financial Studies* **2**, 125-156.

Keim, Donald B., 1983, Size-related anomalies and stock return seasonality:  Further empirical evidence, *Journal of Financial Economics* **12**, 13-32.

Ketterer, Juan, 1987, Asset pricing with differential information, Working paper, (Department of Finance, Northwestern University, Evanston, IL).

King, Benjamin F., 1966, Market and industry factors in stock price behavior, *Journal of Business* **39**, 139-190.

Korajczyk, Robert A., 1985, The pricing of forward contracts for foreign exchange, *Journal of Political Economy* **93**, 346-368.

Korajczyk, Robert A. and Claude J. Viallet, 1989, An empirical investigation of international asset pricing, *Review of Financial Studies* **2**, 553-585.

Korajczyk, Robert A. and Claude J. Viallet, 1992, Equity risk premia and the pricing of foreign exchange risk, *Journal of International Economics* **33**, 199-219.

Kreps, David M., 1981, Arbitrage and equilibrium in economies with infinitely many commodities, *Journal of Mathematical Economics* **8**, 15-35.

Kritzman, Mark, 1993, What practitioners need to know about factor methods, *Financial Analysts Journal* **49**, 12-15.

Latham, Mark, 1989, The arbitrage pricing theory and supershares, *Journal of Finance* **44**, 263-281.

Lehmann, Bruce N., 1990, Residual risk revisited, *Journal of Econometrics* **45**, 71-97.

Lehmann, Bruce N., 1992, Notes on dynamic factor pricing models, *Review of Quantitative Finance and Accounting* **2**, 69-87.

Lehmann, Bruce N. and David M. Modest, 1985, The empirical foundations of the arbitrage pricing theory II:  The optimal construction of basis portfolio, Working paper 292, (Department of Economics, Columbia University, New York, NY).

Lehmann, Bruce N. and David M. Modest, 1987, Mutual fund performance evaluation:  A comparison and benchmarks and benchmark comparisons, *Journal of Finance* **42**, 233-265.

Lehmann, Bruce N. and David M. Modest, 1988, The empirical foundations of the arbitrage pricing theory, *Journal of Financial Economics* **21**, 213-254.

Levine, Ross, 1989, An international arbitrage pricing model with PPP deviations, *Economic Inquiry* **27**, 87-599.

Litterman, Robert and José Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* **1**, 54-61.

Long, John B., 1974, Stock prices, inflation and the term structure of interest rates, *Journal of Financial Economics* **1**, 131-170.

Lucas, Robert E., Jr., 1978, Asset prices in an exchange economy, *Econometrica* **46**, 1429-1445.

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

Luttmer, Erzo G. J. 1993, Asset pricing in economies with frictions, Working paper 151, (Department of Finance, Northwestern University, Evanston, IL).

MacKinlay, A. Craig, 1993, Multifactor models do not explain deviations from the CAPM, Working paper, (University of Pennsylvania, Philadelphia, PA).

Madansky, Albert, 1964, Instrumental variables in factor analysis, *Psychometrica* **29**, 105-113.

McCloskey, Donald N., 1985, The loss function has been mislaid:  The rhetoric of significance tests, *American Economic Review* **75**, 201-205.

McCulloch, Robert and Peter E. Rossi, 1990, Posterior, predictive, and utility-based approaches to testing the arbitrage pricing theory, *Journal of Financial Economics* **28**, 7-38.

McCulloch, Robert and Peter E. Rossi, 1991, A Bayesian approach to testing the arbitrage pricing theory, *Journal of Econometrics* **49**, 141-168.

McElroy, Marjorie B. and Edwin Burmeister, 1988, Arbitrage pricing theory as a restricted nonlinear multivariate regression model, *Journal of Business & Economic Statistics* **6**, 29-42.

Mei, Jianping, 1993, A semi-autoregression approach to the arbitrage pricing theory, *Journal of Finance* **48**, 599-620.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* **41**, 867-887.

Merton, Robert C., 1977, A reexamination of the capital asset pricing model, in: I. Friend and J. Bicksler (eds.), *Risk and return in finance*, (Ballinger, Cambridge, MA).

Miller, Merton H. and Myron Scholes, 1972, Rates of return in relation to risk: A re-examination of some recent findings, in: Michael C. Jensen (ed.), *Studies in the theory of capital markets*, (Praeger, New York).

Milne, Frank, 1988, Arbitrage and diversification in a general equilibrium asset economy, *Econometrica* **56**, 815-840.

Morrison, Donald F., 1976, *Multivariate statistical methods*, 2nd ed. (McGraw-Hill, New York).

Ng, Victor, Robert Engle, and Michael Rothschild, 1992, A multi-dynamic-factor model for stock returns, *Journal of Econometrics* **52**, 245-266.

Ohlson, James A. and Mark B. Garman, 1980, A dynamic equilibrium for the Ross arbitrage model, *Journal of Finance* **35**, 675-684.

Raveh, Adi, 1985, A note on factor analysis and arbitrage pricing theory, *Journal of Banking and Finance* **9**, 317-321.

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

Reinganum, Marc R., 1981, The arbitrage pricing theory: Some simple tests, *Journal of Finance* **36**, 313-322.

Reisman, Haim, 1988, A general approach to the arbitrage pricing theory (APT), *Econometrica* **56**, 473-476.

Reisman, Haim, 1992a, Intertemporal arbitrage pricing theory, *Review of Financial Studies* **5**, 105-122.

Reisman, Haim, 1992b, Reference variables, factor structure, and the approximate multibeta representation, *Journal of Finance* **47**, 1303-1314.

Reisman, Haim, 1992c, The APT with proxies, Working paper, (Technion, Haifa).

Roll, Richard and Stephen A. Ross, 1980, An empirical investigation of the arbitrage pricing theory, *Journal of Finance* **35**, 1073-1103.

Roll, Richard and Stephen A. Ross, 1983, Regulation, the capital asset pricing model, and the arbitrage pricing theory, *Public Utilities Fortnightly* **111**, 22-28.

Roll, Richard and Stephen A. Ross, 1984a, The arbitrage pricing theory approach to strategic portfolio planning, *Financial Analysts Journal*, 14-26.

Roll, Richard and Stephen A. Ross, 1984b, A critical reexamination of the empirical evidence on the arbitrage pricing theory:  A reply, *Journal of Finance* **39**, 347-350.

Rosenberg, Barr, 1974, Extra-market components of covariance in security returns, *Journal of Financial and Quantitative Analysis* **9**, 263-274.

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

Ross, Stephen A., 1977, Return, risk and arbitrage, in:  I. Friend and J. Bicksler (ed.), *Risk and return in finance*, (Ballinger, Cambridge, MA).

Ross, Stephen A., 1978a, The current status of the capital asset pricing model (CAPM), *Journal of Finance* **33**, 885-901.

Ross, Stephen A., 1978b, A simple approach to the valuation of risky streams, *Journal of Business* **51**, 453-475.

Ross, Stephen A., 1982, On the general validity of the mean-variance approach in large markets, in:  William F. Sharpe and Cathryn M. Cootner (eds.), *Financial economics:  Essays in honor of Paul Cootner*, (Prentice-Hall, Englewood Cliffs).

Ross, Stephen A. 1990, Arbitrage and the APT:  Some new results, Working paper, School of Management, (Yale University, New Haven, CT).

Ross, Stephen A., and Michael M. Walsh, 1983, A simple approach to the pricing of risky assets with uncertain exchange rates, *Research in International Business and Finance* **3**, 39-54.

Ross, Stephen A., Randolph W. Westerfield, and Jeffrey F. Jaffe, 1993, *Corporate finance*, 3rd ed. (Irwin, Homewood).

Rothschild, Michael, 1986, Asset pricing theories, in:  W. P. Heller, R. M. Starr, and D. A. Starrett (eds.), *Uncertainty, information, and communication:  Essays in honor of Kennth J. Arrow*, Vol. III (Cambridge University Press, Cambridge).

Rubio, Gonzalo, 1992, Further evidence on performance evaluation:  Portfolio holdings, recommendations, and turnover costs, Finance working paper 222, (University of California, Berkeley, CA).

Scholes, Myron and Joseph Williams, 1977, Estimating betas from nonsynchronous data, *Journal of Financial Economics* **5**, 309-327.

Sentana, Enrique, 1992, Identification of multivariate conditionally heteroskedastic factor models, Discussion paper 139, (London School of Economics, London).

Shanken, Jay, 1982, The arbitrage pricing theory:  Is it testable? *Journal of Finance* **37**, 1129-1140.

Shanken, Jay, 1985, Multi-beta CAPM or equilibrium APT:  A reply, *Journal of Finance* **40**, 1189-1196.

Shanken, Jay, 1987a, Multivariate proxies and asset pricing relations:  Living with Roll critique, *Journal of Financial Economics* **18**, 91-110.

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

Shanken, Jay, 1990, Intertemporal asset pricing:  An empirical investigation, *Journal of Econometrics* **45**, 99-120.

Shanken, Jay, 1992a, On the estimation of beta-pricing models, *Review of Financial Studies* **5**, 1-33.

Shanken, Jay, 1992b, The current state of the arbitrage pricing theory, *Journal of Finance* **47**, 1569-1574.

Shanken, Jay, and Mark I. Weinstein, 1990, Macroeconomic variables and asset pricing: Estimation and tests, Working paper, (University of Rochester, Rochester, NY).

Sharpe, William F., 1963, A simplified model for portfolio analysis, *Management Science* **9**, 277-293.

Sharpe, William F., 1977, The capital asset pricing model:  A multi-beta interpretation, in:  H. Levy and M. Sarnat (eds.), *Financial decision making under uncertainty*, (Academic Press, New York).

Sharpe, William F., 1982, Some factors in New York Stock Exchange security returns, 1931-1979, *Journal of Portfolio Management* **8**, 5-19.

Sharpe, William F., 1984, Factor models, CAPMs, and the APT, *Journal of Portfolio Management* **11**, 21-25.

Sharpe, William F., 1988, Determining a fund's effective asset mix, *Investment Management Review* **2**, 59-69.

Sharpe, William F., 1992, Asset allocation:  Management style and performance measurement, *Journal of Portfolio Management* **18**, 7-19.

Solnik, Bruno, 1983, International arbitrage pricing theory, *Journal of Finance* **38**, 449-457.

Stambaugh, Robert F., 1983, Arbitrage pricing with information, *Journal of Financial Economics* **12**, 357-69.

Stroyny, Alvin L., 1992, Still more on *EM* factor analysis, Working paper, (University of Wisconsin, Milwaukee, WI).

Suzuki, Yoshio, ed., 1987, *The Japanese financial system*, (Clarendon Press, Oxford).

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

Warga, Arthur, 1989, Experimental design in tests of linear factor models, *Journal of Business and Economic Statistics* **7**, 191-198.

Wei, K. C. John, 1988, An asset pricing theory unifying the CAPM and APT, *Journal of Finance* **43**, 881-892.

Wei, K. C. John, Cheng-few Lee, and Andrew H. Chen, 1991, Multivariate regression tests of the arbitrage pricing theory:  The instrumental variables approach, *Review of Quantitative Finance and Accounting* **1**, 191-208.

White, Halbert, and Ian Domowitz, 1984, Nonlinear regression with dependent observations, *Econometrica* **52**, 143-161.

Winkelmann, Michael, 1984, Testing APT for the German stock market, in:  *Proceedings of the 11th annual meeting of the European Finance Association*, (Manchester).

Young, S. David, Michael A. Berry, David W. Harvey, and John R. Page, 1987, Systematic risk and accounting information under the arbitrage pricing theory, *Financial Analysts Journal* **43**, 73-76.

## Endnotes

\*     The size of the literature related to Arbitrage Pricing Theory precludes us from summarizing all relevant contributions and we apologize in advance to those whose work has not been discussed here. We have received helpful comments from many colleagues. We owe particular thanks to Torben Andersen, Denis Gromb, Ravi Jagannathan, Jack Treynor, and Mark Weinstein. We also thank Mary Korajczyk for editorial assistance.

1.     The "first k" eigenvectors are the k eigenvectors associated with the k largest eigenvalues. That is, we order the eigenvalues by descending size, and then use the induced ordering on the eigenvectors.

2.     The maximum eigenvalue of this matrix is equal to the maximum eigenvalue of the within-industry covariance matrices. This eigenvalue is less than or equal to h times the maximum idiosyncratic variance of an asset in the industry.

3.     It is easy to show that $B'B = (n/k)I_k$ where $I_k$ is the $k \times k$ identity matrix. The k eigenvalues of this matrix all equal $n/k$, which goes to infinity with n.

4.     See Bollerslev (1986) for a detailed analysis of GARCH models.

5.     If $\iota^n$ and B are linearly dependent, then the $k+1 \times n$ matrix $[\iota^n, B]$ has rank k. In this case, there is a rotation of the factors under which every asset in the economy has unit betas against (at least) one factor. Thus, there is no way to construct a zero-beta portfolio with unit cost (since any asset combination with unit cost also has a beta of unity with respect to the above factor). This situation creates an ambiguity in the definition of $\lambda_0$ since there is no well-defined risk-free return. If a risk-free asset exists separately from the factor model (this assumption is often made), then the ambiguity disappears.

6.     If the asset returns are independent and identically distributed with finite mean and variance, then the return to this portfolio is the expected return of the assets.

7.     This assumption does not appear explicitly in Chen and Ingersoll (1983) because they make an exogenous assumption about equilibrium portfolios.

8.     The first-order condition for the mean-variance efficiency of $\omega$ is $\Sigma\omega = E[r]\gamma_1 + \iota^n\gamma_2$, where $\gamma_1$ and $\gamma_2$ are proportional to Lagrange multipliers for the constrained optimization problem. Rearranging this first-order condition gives (11). See Grinblatt and Titman (1987) for more details.

9.     A special case of this is when $\lambda_t$ is assumed to be constant through time, although the theory does not require this.

10.     Equations (19) and (20) assume that the instruments $Z_{t-1}$ are predetermined relative to $r_t$ and $F_t$. Not all studies use instruments that are strictly predetermined.

11.     In some cases there are multiple passes in which the $\bar{F}_t$ from a cross-sectional regression is used to re-estimate betas in additional time-series regressions. These new betas are then used to re-estimate $\bar{F}_t$ via cross-sectional regressions [see Connor and Uhlaner (1989)].

12.     Solutions to the first-order equations with negative $V_{ii}$ (negative idiosyncratic variances) are called Heywood cases [see Anderson (1984) for proposals for dealing with them].

13.    The tests of $\lambda_0^i = \lambda_0$ and $\lambda^i = \lambda$ in (14) can be viewed as tests of the law of one price (i.e., that the price of risk is the same across subgroups), *conditional* on an estimated factor model. Chen and Knez (1992) propose a test of consistent pricing across subsets of assets that does not require a first-stage estimation of a factor model.

14.    An alternative approach would be to estimate only the restricted factor sensitivity matrix, B, and regress returns on $\iota^{60}$, B, and Z, as in (19) and (20). The top $30 \times (k + 1)$ submatrix of Z is a matrix of zeros and the bottom $30 \times (k + 1)$ submatrix of Z is equal to $[\iota^{30} : B^2]$ where $B^2$ is defined as the last 30 rows of B. Then a test of $\delta = 0$ is a test of consistent pricing across subgroups.

15.    In some specifications, the annual percentage change in industrial production is also included, but is not found to be statistically significant.

16.    The unexpected inflation and change in expected inflation variables require a model of expected inflation. Chen, Roll, and Ross (1986) use the approach to measuring expected inflation developed by Fama and Gibbons (1984).

17.    This is a slightly weaker test than testing whether the mean residuals are zero.

18.    The statistic is adjusted by the EIV correction from Shanken (1992a).

19.    We will assume that such portfolios exist.

20.    Stroyny (1992) suggests a modification to the *EM* algorithm that substantially improves its rate of convergence.

21.    Note that the restriction that V be diagonal is not required by the APT. In approximate factor models, V may be non-diagonal, but this correlation across assets needs to be weak.

22.    The fact that the $R^2$ value for the typical regressions in (22) and (25) is around 98% for the APT models and 75% for the CAPM models gives some indication of the greater precision of the estimated $\alpha$ vector in the former case.

23.    The constant is chosen to make the sample mean of the factor, from 1926 to 1981, equal to zero.

24.    See McElroy and Burmeister (1988) or Gallant (1987, ch. 5) for a discussion of the estimation methods.

25.    Financial market integration does not imply or require that the countries be engaged in producing the same goods. Therefore, financially integrated countries might still have assets that are subject to country-specific productivity shocks. A country-specific, but priced, factor could occur if the country in question is not small relative to the world economy.

26.    Brown and Warner (1980, 1985) analyze the properties of various alternative approaches to estimating the asset-specific reaction to news.

27.    Examples of primarily firm-specific news are announcements related to stock splits, corporate earnings, dividend declarations, and equity issues.