Reprinted from JOURNAL OF ECONOMIC THEORY
All Rights Reserved by Academic Press, New York and London

Vol. 28, No. 1, October 1982
Printed in Belgium

A Simple Approach to Arbitrage Pricing Theory

GUR HUBERMAN*

Graduate School of Business, University of Chicago, Chicago, Illinois 60637

Received August 18, 1980; revised March 16, 1981

1. INTRODUCTION

The arbitrage theory of capital asset pricing was developed by Ross [9, 10, 11] as an alternative to the mean-variance capital asset pricing model (CAPM), whose main conclusion is that the market portfolio is mean-variance efficient. Its formal statement entails the following notation. A given asset i has mean return E_i and the market portfolio has mean return E_m and variance σ_m^2. The covariance between the return on asset i and the return on the market portfolio is σ_{im}, and the riskless interest rate is r. The CAPM asserts that

$$E_i = r + \lambda b_i, \tag{1.1}$$

where

$$\lambda = E_m - r,$$

and

$$b_i = \sigma_{im}/\sigma_m^2 \tag{1.2}$$

is the "beta coefficient" of asset i.

Normality of the returns of the capital assets or quadratic preferences of their holders are the assumptions which lead to (1.1)–(1.2). Theoretically and empirically it is difficult to justify the assumptions of the CAPM. Moreover, the CAPM has been under strong criticism because of its dubious empirical content (cf. [7]). The market portfolio is practically not obser-

* This is a modified version of the third chapter in my Ph.D. dissertation which was written at Yale University. I am grateful to Gregory Connor, who inspired some of the ideas in the paper. Comments from my advisor, Steve Ross, as well as from Jon Ingersoll, Uriel Rothblum and Michael Rothschild were helpful in my attempts to bring this work to a lucid form. This research was supported by NSF Grants ENG-78-25182 and SOC-77-22301.

0022-0531/82/050183–09$02.00/0

vable, and a statement on the market portfolio (such as the CAPM) is difficult to test empirically. Yet the linear relation (1.1) is appealing in its simplicity and in its ready interpretations. The arbitrage pricing theory [10, 11] is an alternative theory to mean-variance theories, an alternative which implies an approximately linear relation like (1.1). In [10] Ross elaborated on the economic interpretation of the arbitrage pricing theory and its relation to other models, whereas in [11] he provided a rigorous treatment of the theory. Recent interest in the APT is evident from papers elaborating on the theory (e.g., Chamberlin and Rothschild [1], Connor [4] and Kwon [5, 6]) as well as on its empirical aspects (e.g., Chen [2, 3] and Roll and Ross [8]).

The main advantage of Ross' arbitrage pricing theory is that its empirical testability does not hinge upon knowledge of the market's portfolio. Unfortunately, Ross' analysis is difficult to follow. He does not provide an explicit definition of arbitrage and his proof—unlike the intuitively appealing introductory remarks in [11]—involves assumptions on agents' preferences as well as "no arbitrage" assumptions.

Here arbitrage is defined and the intuition is formalized to obtain a simple proof that no arbitrage implies Ross' linear-like relation among mean returns and covariances. The main lines of the proof are illustrated in the following paragraphs.

Consider an economy with n risky assets whose returns are denoted by $\tilde{x}_i$ ($i = 1,..., n$) and they are generated by a factor model

$$\tilde{x}_i = E_i + \beta_i \tilde{\delta} + \tilde{\varepsilon}_i \qquad (i = 1,..., n), \tag{1.3}$$

where the expectations $E\tilde{\delta} = E\tilde{\varepsilon}_i = 0$ ($i = 1,..., n$), the $\tilde{\varepsilon}_i$ are uncorrelated and their variances are bounded. Relying on results from linear algebra, express the vector E (whose ith component is E_i) as a linear combination of the vector e (whose ith component is 1), the vector β (whose ith component is β_i) and a third vector c which is *orthogonal* both to e and to β.[1] In other words, one can always find a vector c such that

$$E = \rho e + \gamma\beta + c, \tag{1.4}$$

where ρ and γ are scalars,

$$ec \equiv \sum_{i=1}^{n} c_i = 0, \tag{1.5}$$

and

$$\beta c \equiv \sum_{i=1}^{n} \beta_i c_i = 0. \tag{1.6}$$

[1] Gregory Connor used this idea in an earlier work of his [4].

Next, consider a portfolio which is proportional to c, namely αc (α is a scalar). Note that it costs nothing to acquire such a portfolio because its components (the dollar amount put into each asset) sum to zero by (1.5). We shall call such a portfolio an arbitrage portfolio. Also, by (1.6) this is a zero-beta portfolio. The return on this portfolio is

$$\alpha \tilde{x} c = \alpha \sum_{i=1}^{n} \tilde{x}_i c_i = \alpha \sum_{i=1}^{n} c_i^2 + \alpha \sum_{i=1}^{n} c_i \tilde{\varepsilon}_i, \tag{1.7}$$

by virtue of the decomposition (1.4) and the orthogonality relations (1.5) and (1.6). It is important to notice that the expected return on the portfolio αc is proportional to α (and $\sum_{i=1}^{n} c_i^2$), whereas an upper bound on the variance of its return is proportional to α^2 (and $\sum_{i=1}^{n} c_i^2$).

Suppose now that the number of assets n increases to infinity. Think of arbitrage in this environment as the oppurtunity to create a sequence of arbitrage portfolios whose expected returns increase to infinity while the variances of their returns decrease to zero. If the sum $\sum_{i=1}^{n} c_i^2$ increased to infinity as n did, then one could find such arbitrage opportunities as follows. Set $\alpha = 1/(\sum_{i=1}^{n} c_i^2)^{3/4}$ and use the portfolio αc. The reason why such a choice of α will create the arbitrage is that the expected return on the portfolio is proportional to α (and with $\alpha = 1/(\sum_{i=1}^{n} c_i^2)^{3/4}$ it equals $\alpha \sum_{i=1}^{n} c_i^2 = (\sum_{i=1}^{n} c_i^2)^{1/4}$), while its variance is proportional to α^2 (and with $\alpha = 1/(\sum_{i=1}^{n} c_i^2)^{3/4}$ it equals $\alpha^2 \sum_{i=1}^{n} c_i^2 = 1/(\sum_{i=1}^{n} c_i^2)^{1/2}$).

Therefore, if there are no arbitrage opportunities (as described above) the sum $\sum_{i=1}^{n} c_i^2$ cannot increase to infinity as n does. In particular, when the number of assets n is large, most of the c_i's are small and approximately zero. Going back to the original decomposition (1.4) we conclude that $E_i \approx \rho + \gamma\beta_i$ for most of the assets.

When motivating his proof, Ross [11, p. 342] emphasized the role of "well-diversified" arbitrage portolios. He indicated that the law of large numbers was the driving force behind the diminishing contribution of the idiosyncratic risks $\tilde{\varepsilon}_i$ to the overall risks of the arbitrage portfolios. The portfolios presented above, αc, need not be well diversified, but they satisfy the orthogonality conditions (1.5) and (1.6). It is the judicious choice of the scalar α that enables us to apply an idea, which is in the spirit of the proof of the law of large numbers.

Section 2 of this paper pesents the formal model, a precise statement of the result and a rigorous proof. In the closing section an attempt is made to interpret the linear-like pricing relation and to justify the no-arbitrage assumption in an equilibrated economy of von Neumann–Morgenstern expected utility maximizers.

2. Arbitrage Pricing

The arbitrage pricing theory considers a sequence of economies with increasing sets of risky assets. In the nth economy there are n risky assets whose returns are generated by a k-factor model (k is a fixed number). Loosely speaking, arbitrage is the possibility to have arbitrarily large returns as the number of available assets grows. We will show that in the absence of arbitrage a relation like (1.1) holds, namely (2.9).

Formally, in the nth economy, we consider an array of returns on risky assets $\{\tilde{x}_i^n : i = 1,\ldots, n\}$. These returns are generated by a k-factor linear model of the form

$$\tilde{x}_i^n = E_i^n + \beta_{i1}^n \tilde{\delta}_1^n + \beta_{i2}^n \tilde{\delta}_2^n + \cdots + \beta_{ik}^n \tilde{\delta}_k^n + \tilde{\varepsilon}_i^n \qquad (i = 1, 2,\ldots, n), \quad (2.1)$$

where

$$E\tilde{\delta}_j^n = 0 \qquad (j = 1,\ldots, k), \qquad E\tilde{\varepsilon}_i^n = 0 \qquad (i = 1,\ldots, n), \qquad (2.2)$$

$$E\tilde{\varepsilon}_i^n \tilde{\varepsilon}_j^n = 0 \qquad \text{if} \quad i \neq j, \qquad (2.3)$$

$$\text{and } \operatorname{Var} \tilde{\varepsilon}_i^n \leqslant \bar{\sigma}^2 \qquad (i = 1,\ldots, n), \qquad (2.4)$$

where $\bar{\sigma}^2$ is a fixed (positive) number. Using standard matrix notation we can rewrite (2.1) as

$$\tilde{x}^n = E^n + \beta^n \tilde{\delta}^n + \tilde{\varepsilon}^n, \qquad (2.5)$$

where β^n is the $n \times k$ matrix whose elements are β_{ij}^n ($i = 1,\ldots, n$; $j = 1,\ldots, k$).

A portfolio $c^n \in R^n$ in the nth economy is an *arbitrage portfolio* if $c^n e^n = 0$, where $e^n = (1, 1,\ldots, 1) \in R^n$. The return on a portfolio c is

$$\tilde{z}(c) = c\tilde{x}^n = cE^n + c\beta^n \tilde{\delta}^n + c\tilde{\varepsilon}^n. \qquad (2.6)$$

Arbitrage is the existence of a subsequence n' of arbitrage portfolios whose returns $\tilde{z}(c^{n'})$ satisfy

$$\lim_{n' \to \infty} E\tilde{z}(c^{n'}) = +\infty, \qquad (2.7)$$

and

$$\lim_{n' \to \infty} \operatorname{Var} \tilde{z}(c^{n'}) = 0. \qquad (2.8)$$

In Section 3 we relate (2.7)–(2.8) to standard probabilistic convergence concepts, and discuss how von Neumann–Morgenstern expected utility maximizers view (2.7)–(2.8).

In Theorem 1 we show that the absence of arbitrage implies an approximation to a linear relation like (1.1).

THEOREM 1. *Suppose the returns on the risky investments satisfy* (2.1)–(2.4) *and there is no arbitrage. Then for* $n = 1, 2, \ldots,$ *there exists* $\rho^n, \gamma_1^n, \ldots, \gamma_k^n$, *and an* A *such that*

$$\sum_{i=1}^{n} \left(E_i^n - \rho^n - \sum_{j=1}^{k} \beta_{ij}^n \gamma_j^n \right)^2 \leqslant A, \qquad \textit{for} \quad n = 1, 2, \ldots \tag{2.9}$$

Proof. Using the orthogonal projection of E^n into the linear subspace spanned by e^n and the columns of β^n, one obtains the representation

$$E^n = \rho^n e + \beta^n \gamma^n + c^n, \tag{2.10}$$

where

$$\gamma^n \in R^k,$$
$$e^n c^n = 0, \tag{2.11}$$

and

$$\beta^n c^n = 0. \tag{2.12}$$

Note that $\|c^n\|^2 \equiv \sum_{i=1}^n (c_i^n)^2 = \sum_{i=1}^n (E_i^n - \rho^n - \sum_{j=1}^k \gamma_j^n \beta_{ij}^n)^2$, and assume that the result is false. Consequently, there is an increasing subsequence (n') with

$$\lim_{n' \to \infty} \|c^{n'}\| = +\infty \tag{2.13}$$

Let p be fixed between -1 and $-1/2$, and consider the portfolio $d^{n'} = \alpha_{n'} c^{n'}$, where

$$\alpha_{n'} = \|c^{n'}\|^{2p}. \tag{2.14}$$

By (2.11), $d^{n'}$ is an arbitrage portfolio for each n'. Use (2.10)–(2.12) to see that its return

$$\tilde{z}(d^{n'}) = \alpha_{n'} \|c^{n'}\|^2 + \alpha_{n'} c^{n'} \tilde{\varepsilon}^{n'}. \tag{2.15}$$

Note that

$$E\tilde{z}(d^{n'}) = \alpha_{n'} \|c^{n'}\|^2 = \|c^{n'}\|^{2+2p}, \tag{2.16}$$

so (by (2.13)–(2.14)),

$$\lim_{n' \to \infty} E\tilde{z}(d^{n'}) = +\infty. \tag{2.17}$$

Now look at excess returns of the risky assets (excess relative to the riskless rate), i.e., at

$$\tilde{y}_i^n = \tilde{x}_i^n - r_0^n, \qquad i = 1, 2, \ldots, n.$$

Note that any arbitrage portfolio $(c_0, c_1, \ldots, c_n) \in R^{n+1}$ of $x_0^n, \tilde{x}_1^n, \ldots, \tilde{x}_n^n$ (which of course satisfies $\sum_{i=0}^n c_i = 0$) is equivalent to a vector $(c_1, \ldots, c_n) \in R^n$ indicating a wealth allocation among the risky assets. Using this idea one can go through the same analysis as in Theorem 1 with the excess returns vector $\tilde{y}^n$, the decomposition (2.10) replaced by

$$E^n - r_0^n e = \beta^n \gamma^n + c^n, \tag{2.10$'$}$$

and (2.11) deleted.

Consequently, one has

Corollary. *Suppose the returns on the risky investments satisfy* (2.1)–(2.4), *there is a risk free asset satisfying* (2.19) *and there is no arbitrage. Then there exist* $\gamma_1^n, \gamma_2^n, \ldots, \gamma_k^n$ *such that*

$$\sum_{i=1}^{n} \left(E_i^n - r_0^n - \sum_{j=1}^{k} \beta_{ij}^n \gamma_j^n \right)^2 \leqslant A \qquad \textit{for} \quad n = 1, 2, \ldots. \tag{2.20}$$

Remark. Analogously, a similar result holds for the stationary model.

3. Discussion

The interpretation of (2.9) or (2.9′) is straightforward: for most of the assets in a large economy, the mean return on an asset is approximately linearly related to the covariances of the asset's returns with economy-wide common factors. As the number of assets becomes large, the linear approximation improves and most of the assets' mean returns are almost exact linear functions of the appropriate covariances.

Next, consider the probabilistic implications of arbitrage returns satisfying (2.7)–(2.8). Given a sequence of random returns $\tilde{z}(c^{n'})$ which satisfy (2.7)–(2.8), we can apply Chebychev's inequality to see that along this sequence $\lim_{n' \to \infty} \tilde{z}(c^{n'}) = +\infty$ in probability (i.e., for all $M > 0$, $\lim_{n' \to \infty} \Pr\{\tilde{z}(c^{n'}) \geqslant M\} = 1$). Furthermore, along a subsequence $\hat{n}$, a stronger convergence holds: $\lim_{\hat{n} \to \infty} \tilde{z}(c^{\hat{n}}) = +\infty$ almost surely (i.e., for all $M > 0$, $\Pr\{\liminf_{\hat{n} \to \infty} \tilde{z}(c^{\hat{n}}) \geqslant M\} = 1$).

Are arbitrage portfolio which satisfy (2.7)–(2.8) desirable for an expected utility maximizer? In other words, do (2.7)–(2.8) suffice to assert that $\lim_{n \to \infty} EU(\tilde{z}(c^n)) = U(+\infty)$ for any monotone concave utility function U? The negative answer is illustrated by the following examples.

On the other hand (using (2.3), (2.4))

$$\operatorname{Var} \tilde{z}(d^{n'}) \leqslant \bar{\sigma}^2 \alpha_{n'}^2 \|c^{n'}\|^2 = \bar{\sigma}^2 \|c^{n'}\|^{2+4p}, \tag{2.18}$$

so (by (2.13)),

$$\lim_{n' \to \infty} \operatorname{Var} \tilde{z}(d^{n'}) = 0,$$

thus completing the proof. ∎

Next, consider a stationary model, in which $E_i^n = E_i$ and $\beta_{ij}^n = \beta_{ij}$ for all i, j and n. In other words, (2.5) is replaced by

$$\tilde{x}^n = E + \beta \tilde{\delta}^n + \tilde{\varepsilon}^n. \tag{2.5'}$$

The stationary model is the one considered originally by Ross [11]. The nonstationary model is more general than the stationary model but its result (2.9) is not as elegantly presentable as the result in the stationary case (2.9′).

THEOREM 2. *Suppose the returns on the risky investments satisfy* (2.5′), (2.2)–(2.4), *and there is no arbitrage. Then there exist* $\rho, \gamma_1, \ldots, \gamma_k$ *such that*

$$\sum_{i=1}^{\infty} \left(E_i - \rho - \sum_{j=1}^{k} \beta_{ij} \gamma_j \right)^2 < \infty. \tag{2.9'}$$

Proof. Consider the $n \times (k+1)$ matrix B^n whose (i, j) entry is 1 if $j = 1$ and β_{ij-1} if $1 < j \leqslant k+1$. Let $r(n)$ be the rank of B^n. Since $1 \leqslant r(n) \leqslant r(n+1) \leqslant k+1$ for all n, and $r(n)$ is an integer, there is an $\bar{n}$ such that $r(n) = r(\bar{n})$ for all $n \geqslant \bar{n}$. Let $n \geqslant \bar{n}$ be fixed. By permuting the columns of B^n we may assume that its first $r(\bar{n})$ columns can be expressed as linear combinations of the first $r(\bar{n})$ columns. Define the set H^n by

$$H^n = \left\{ (\rho, \gamma_1, \ldots, \gamma_k) : \sum_{i=1}^{n} \left(E_i - \rho - \sum_{j=1}^{k} \beta_{ij} \gamma_j \right)^2 \leqslant A, \gamma_j = 0, \right.$$
$$\left. \text{for} \quad r(\bar{n}) < j \leqslant k \right\},$$

where A is the A whose existence was asserted in Theorem 1. Note that H^n is nonempty (by Theorem 1), compact for $n \geqslant \bar{n}$ and $H^n \subset H^{n+1}$. Therefore, $\bigcap_{n=1}^{\infty} H^n$ is nonempty. Since every $k+1$ tuple $(\rho, \gamma_1, \ldots, \gamma_k) \in \bigcap_{n=1}^{\infty} H^n$ satisfies (2.9′), the proof is complete. ∎

Finally, we turn attention to the case where a risk free asset exists, i.e., where there is an additional asset in the nth economy, whose return, say, x_0^n, satisfies

$$x_0^n = r_0^n. \tag{2.19}$$

The first example considers a utility function which is $-\infty$ for nonpositive wealth levels, whereas the second example is for an exponential utility which takes finite values for finite wealth levels.

1. The returns $\tilde{z}(c^n)$ are 0, n and $2n$ with probabilities $1/n^3$, $1-2/n^3$, and $1/n^3$, respectively. The utility function $U(x)=-1/x$ for $x>0$ and $U(x)=-\infty$ for $x \leqslant 0$. Then $EU(\tilde{z}(c^n))=-\infty$ although $\tilde{z}(c^n)$ satisfies (2.7)–(2.8).

2. The returns $\tilde{z}(c^n)$ are $-n$, n, and $3n$ with probabilities $1/n^3$, $1-2/n^3$, and $1/n^3$, respectively. The utility function is $U(x)=-\exp(-x)$. Then $EU(\tilde{z}(c^n)) \leqslant -n^3 \exp(n)$, so $\lim_{n\to\infty} EU(\tilde{z}(c^n))=-\infty$, although (2.7)–(2.8) are met.

General conditions which assert that (2.7)–(2.8) imply $\lim_{n'\to\infty} EU(\tilde{z}(c^{n'}))=U(+\infty)$ are not known. As shown in [11, Appendix 2], utility functions which are bounded below or uniformly integrable utility functions will possess this property.

We conclude that one needs to make assumptions on agents' preferences in order to relate existence of equilibria to absence of arbitrage. This task is beyond the scope of this paper. However, it is straightforward to see that if the economies satisfy the assumptions made by Ross (see [11], especially the first paragraph on p. 349 and Appendix 2), then no arbitrage can exist. In fact, a result of the type "no arbitrage implies a certain behavior of returns," should involve no consideration of the preference structure of the agents involved. Our analysis is in this spirit, because it involves no assumptions on utilities. Other than the simple proof, this may be another contribution of this work.

References

1. G. Chamberlain and M. Rothchild, Arbitrage and mean-variance analysis of large asset markets, mimeo, University of Wisconsin–Madison, 1980.
2. Nai-Fu Chen, Empirical evidence of the arbitrage pricing theory, mimeo, UCLA, 1980.
3. Nai-Fu Chen, The arbitrage pricing theory: Estimation and applications, mimeo, UCLA, 1980.
4. G. Connor, "Asset Prices in a Well-Diversified Economy," Technical Report No. 47, Yale School of Organization and Management, July 1980.
5. Y. Kwon, Counterexaples to Ross' arbitrage asset pricing model, mimeo, University of Kansas, 1980.
6. Y. Kwon, "On the negligibility of diversifiable risk components at the capital market equilibrium," mimeo, University of Kansas, 1980.
7. R. Roll, A critique of the asset pricing theory's tests: Part I: On past and potential testability of the theory, *Financial Econ.* 4 (1977), 129–179.
8. R. Roll and S. Ross, An empirical investigation of the arbitrage pricing theory, *J. Finance* 35 (1980), 1073–1103.

9. S. Ross, "The General Validity of the Mean-Variance Approach in Large Markets," Discussion Paper No. 12–72, Rodney L. White Center for Financial Research, University of Pennsylvania.
10. S. Ross, Return, risk and arbitrage, *in* "Risk and Return in Finance" (I. Friend and J. Bicksler, Eds.), Ballinger, Cambridge, Mass., 1977.
11. S. Ross, The arbitrage theory of capital asset pricing, *J. Econ. Theory* **13**, No. 3 (1976), 341–360

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium

discussion | **Notes on the Arbitrage Pricing Theory**

Gregory Connor

University of California, Berkeley

The arbitrage pricing theory (APT) constitutes one of the most important models of security market pricing. The chief aim of Huberman's justly famous paper, reprinted in this volume, is to clarify and simplify Ross's model. He does an admirable job of this. His presentation is so clear that a detailed critical review would be superfluous. I will attempt, instead, a general overview of the APT. A novice to the field might use this broad discussion as a way to solidify his or her understanding of the original papers. An economist familiar with the APT may benefit from seeing another economist's conceptual framework laid out simply.

Section I describes a few results from "pure" arbitrage pricing theory. "Pure" arbitrage pricing theory is only distantly related to *the* arbitrage pricing theory. Part of the intent of Section I is to clarify this relationship. The results of this section also have independent interest; Key Result 2 in Section I must be one of the most elegant theorems in financial economics.

Section II develops the APT using Ross's original "approximate arbitrage" argument. Ross conjectured that several assumptions of his model could be weakened, and acknowledged that some of its features could be refined. Section III describes one of the most important refinements: the generalization of the factor model assumption that Ross uses. Section IV develops the competitive equilibrium version of the APT. Section V summarizes the paper and suggests some fruitful directions for future research. Each section (except V) ends with a selective list of references to guide further reading.

I. PURE ARBITRAGE PRICING THEORY

An arbitrage opportunity is the existence of a collection of assets that can be combined into a costless portfolio with some chance of a positive payoff and no chance of a negative payoff. There are two strong arguments for assuming that arbitrage opportunities will not appear in security market price relationships. First, *any* investor who observes such an opportunity can make limitless profits (unless prices adjust). Second, *all* investors can improve on their current portfolios by reshuffling their holdings to take advantage of the arbitrage opportunity. Chaos follows unless prices adjust.

The absence of arbitrage opportunities is a minimal condition for well-functioning capital markets. The standard tool for price analysis in free markets is the concept of competitive equilibrium. The absence of arbitrage is a more general condition than the presence of competitive equilibrium, as codified below.

KEY RESULT 1. *If the economy is in competitive equilibrium (and there exists at least one nonsatiated investor) then there do not exist any arbitrage opportunities.*

See Harrison and Kreps (1979, p. 385) for a proof. Arbitrage (more correctly "nonarbitrage" or "the absence of arbitrage") places linear restrictions on the relationships among asset prices. (In security market analysis we usually describe price in terms of expected return, which is essentially the reciprocal of price.) A complete characterization of these restrictions is provided in Key Result 2.

For simplicity, assume that the economy lasts one period and there is a finite set of l possible outcomes, or states of the world, denoted by $\theta_1, \theta_2, \ldots, \theta_l$. Assume that there are N assets in the economy; asset i is represented by an l-vector R_i giving the gross return of the asset for each possible state. The set of possible returns for all assets can be represented by an $N \times l$ matrix R. The assumption of no arbitrage creates linear restrictions on R. Let 1^N denote an N-vector of 1's.

KEY RESULT 2. *R allows no arbitrage opportunities if and only if there exists an l-vector p_θ that consists entirely of positive numbers such that $1^N = Rp_\theta$.*

See Ross (1978, p. 474) for the proof of this simple version and Kreps (1981, Theorem 3) for a more complex version allowing for continuous time and a continuum of states. The elements of p_θ can be viewed as the "fundamental prices" of the l states of nature; a security that pays one dollar only in state θ_j costs P_{θ_j}. The theorem can be interpreted to say that there must exist some set of positive state prices. Note, however, that these state prices need not be uniquely defined. In the continuous-time, continuous-state-space version, the vector p_θ is replaced by an infinite-dimensional linear operator; the theorem guarantees the existence of a positive linear operator that relates the price of a security to its payoff density over states.

Key result 1 shows that competitive equilibrium implies the nonexistence of arbitrage opportunities. The next theorem proves, in a restricted sense, that a converse relationship also holds. For any returns that do not permit arbitrage, there exists some economic model (specified by investor preferences and endowments) in which the returns are consistent with competitive equilibrium. Without loss of generality, one can restrict the analysis to representative investor economies—economies in which all investors are identical. One must also assume that the representative investor is nonsatiated in all states.

KEY RESULT 3. *R does not permit arbitrage if and only if there exists a representative investor economy in which R is consistent with competitive equilibrium.*

See Harrison and Kreps (1979, Theorem 1) for a proof. This theorem describes in what sense a competitive equilibrium pricing theory can be reduced to a nonarbitrage pricing theory. Any competitive equilibrium model that does not depend on preferences or endowments can be derived by a nonarbitrage argument. The Black-Scholes option pricing model, originally proven with a competitive equilibrium argument, is an example of such a model. Most other security market pricing models rely on preferences or endowments in some fashion and so cannot be reduced to nonarbitrage models.

Arbitrage theory has produced powerful results using fairly weak assumptions. Research has branched in three directions. First, arbitrage theory has played a key role in developing the mathematical foundations of asset pricing theory. See Kreps (1981), Duffie (1985), and Duffie and Huang (1985) for a taste of this field of research. Second, arbitrage theory has been applied to the pricing of derivative assets. A derivative asset is an asset whose payoff is a function of one or more other observable assets (the primary assets). If all assets are traded continuously and price changes are "smooth," then the relationship between the price of the derivative asset and that of the primary assets is determined by arbitrage. Applications range from the pricing of put and call options (Black and Scholes 1973, Merton 1973) to the valuation of oil drilling rights (Paddock, Siegel, and Smith 1985) to choice of investment in higher education (Dothan and Williams 1981). See Smith (1976) for a general review of this area.

The third direction for research is to place more structure on the relationship among primary asset payoffs so that arbitrage restricts their price relationship. This is the intent of Ross's APT and the topic of the rest of this chapter.

II. APPROXIMATE ARBITRAGE AND THE APT

Many successful theories begin with a simple model that captures some empirical regularity. Ross's APT is based on a readily observable feature of securities markets: asset returns have strong patterns of positive covariation. First, Ross captures this observation with a very simple model called a *noiseless factor model*. All asset returns are assumed to be exact linear combinations of a constant term and a set of K random variates called factors:

$$r = \bar{r} + Bf, \quad r_0 = \bar{r}_0, \tag{1}$$

where r_0 is the riskless return, B is an $N \times k$ matrix of factor betas, $\bar{r}$ is an N-vector of expected returns, and f is a K-vector of mean-zero random variables called factors. Since the returns are linearly related, nonarbitrage guarantees that the expected returns also have a linear relationship.

KEY RESULT 4. *If returns obey (1), then in the absence of arbitrage there exists a K-vector γ such that $\bar{r} = 1^N r_0 + B\gamma$.*

The proof is in Ross (1977, p. 197); briefly, it goes as follows. Any portfolio with zero risk and zero cost must have zero expected return to prevent arbitrage. Letting α denote an arbitrary portfolio, we can state this nonarbitrage condition as follows:

(2) $$\text{If } \alpha' B = 0 \text{ and } \alpha' 1^N = 0, \text{ then } \alpha' \bar{r} = 0.$$

The duality theorem of linear algebra can now be applied.[1] Statement (2) says that the dual space of B and 1^N belongs to the dual space of $\bar{r}$. This implies (by the duality theorem) that $\bar{r}$ is a linear combination of B and 1^N,

$$\bar{r} = 1^N \gamma_0 + B\gamma,$$

for some γ_0 and γ. It is easy to show (by nonarbitrage) that $\gamma_0 = r_0$.

Next, Ross generalizes the model to the case of a standard factor model (sometimes called a *strict factor model* to differentiate it from the *approximate factor model* discussed in Section III). He assumes the assets have uncorrelated, mean-zero random terms added to the noiseless factor model of (1):

$$r = \bar{r} + Bf + \varepsilon.$$

He also assumes that there are many assets (N is large). The weak law of large numbers guarantees that if we take a large convex combination of uncorrelated random variates and each of the linear coefficients is small, then the randomness approximately disappears from the sum. Note that the random return in a portfolio is a convex combination of the random returns to assets. Since there are many assets and the terms in ε are uncorrelated, these terms obey the weak law of large numbers. By choosing each individual element of α to be small, an investor can approximately eliminate the ε variates from his or her portfolio return.

Because of the ε terms, pure arbitrage pricing theory does not give an exact pricing restriction. A portfolio with $\alpha' 1^N = 0$, $\alpha' B = 0$, and $\alpha' \bar{r} > 0$ is no longer an arbitrage portfolio because the investor must incur idiosyncratic risk. However, by relying on the diversification of ε in portfolios, we can get an approximate form of (2) that holds if N is large. If $\alpha_i \approx 0$ for every i, then $\alpha' \varepsilon \approx 0$, where the "$\approx$" will be used loosely to mean "approximately equal." This gives an approximate version of (2):

(3) $$\text{If } \alpha' 1^N = 0, \alpha' B = 0, \text{ and } \alpha' \varepsilon \approx 0, \text{ then } \alpha' \bar{r} \approx 0.$$

Invoking the duality theorem together with (3) gives

(4) $$\bar{r} \approx 1^N \gamma_0 + B\gamma.$$

Equation (4) is the central conclusion of the APT. Unfortunately, the informal intuition given above cannot be formalized directly—the duality theorem is an exact theorem and does not allow for the approximate equality $\alpha' \varepsilon \approx 0$ in (3). Ross uses a quadratic programming problem to mimic (3). This provides an approximate form of the duality theorem. Most recent theoretical papers abandon duality theory entirely and rely on projection theory instead. A clear and simple proof using projection theory appears in Huberman's paper in this volume.

There are two distinct uses of the symbol "$\approx$" in the intuitive description of the APT given above. I will follow Huberman in defining each of these precisely. A sequence of approximate arbitrage opportunities (for shorthand, an approximate arbitrage opportunity) is a sequence of portfolios with expected payoff going to infinity and the probability of negative payoff going to zero. This clarifies the meaning of "$\approx$" in (3).

The symbol "$\approx$" in (4) will be defined as follows. Let d^N denote the N-vector of pricing errors in the equality version of (4): $\bar{r} = 1^N\gamma_0 + B\gamma + d^N$. The sequence obeys the *Ross pricing bound* if the sum of squared pricing errors, $d^{N\prime}d^N$, does not go to infinity with N. That is, there exists some finite number C such that $d^{N\prime}d^N < C$ for all N.

KEY RESULT 5. *If there do not exist any appropriate arbitrage opportunities, then the Ross pricing bound holds.*

See Ross (1976, Theorem 1) or Huberman (1982, Theorem 1) for a proof. Next, we describe in what sense this gives approximate pricing for large N. The proof is straightforward and is left to the reader.

KEY RESULT 6. *Given the Ross pricing bound, the mean square pricing error is less than C/N; for any $\delta > 0$ the proportion of assets with squared pricing errors less than δ is greater than $1 - C/\delta N$; and the squared pricing error of any individual asset is less than C.*

This is not a standard approximation: even for large N it only guarantees accurate pricing on average or for most assets. The error bound for an individual asset ($C < \infty$) does not change with N. For now, let us view this as a useful, although unconventional, form of approximation and return for a second look when we discuss competitive equilibrium.

III. APPROXIMATE FACTOR MODELS

Ross's original derivation assumes that the idiosyncratic risks have zero correlation. This is one of the features of a standard factor model. This obviously allows a diversification of idiosyncratic risk, but, as Ross notes, a weaker condition could also suffice. Chamberlain and Rothschild (1983) and Ingersoll (1984) (working independently) developed a very appealing alternative, called an approximate factor model.

First, recall that in a strict factor model random returns can be written in the form $r - \bar{r} = Bf + \varepsilon$, where ε_i is uncorrelated with ε_j and f_k for every $i, j, k, i \neq j$. Without loss of generality we can set $E[ff'] = I$ (the $K \times K$ identity matrix). The assumption of an exact factor model is identical to assuming the following form for the return covariance matrix:

$$E[(r-\bar{r})(r-\bar{r})'] = \Sigma = BB' + D,$$

where $D = E[\varepsilon\varepsilon']$ is diagonal.

We wish to relax some of the restrictiveness of the strict factor model. We maintain the assumption that f_k, ε_i are uncorrelated, but we drop the assumption that $\varepsilon_i, \varepsilon_j$ are uncorrelated. This gives the more general form

$$\Sigma = BB' + V, \tag{5}$$

where $V = E[\varepsilon\varepsilon']$ need not be diagonal.

We need to place sufficient structure on (5) to derive the APT. Recall that the APT is an asymptotic theory dealing with approximate relationships for large N. Let $V^N = E[\varepsilon^N \varepsilon^{N'}]$ be the idiosyncratic covariance matrix indexed by the number of assets. A basic concept of the APT is that the randomness captured by ε_i, $i = 1, 2, \ldots, N$, should disappear from portfolios with holdings spread evenly over a large collection of assets. Define a *sequence of diversified portfolios* as a sequence of N-vectors α^N such that $\alpha^{N'} 1^N = 1$ for all N and $\text{limit}_{N\to\infty} \alpha^{N'}\alpha^N = 0$. Call a sequence of random variables ε^N *diversifiable* if $\text{limit}_{N\to\infty} E[(\alpha^{N'}\varepsilon^N)^2] = 0$ for all sequences of diversified portfolios α^N.

KEY RESULT 7. *The sequence of random variates ε^N is diversifiable if there exists a $C < \infty$ such that the maximum eigenvalue of V^N is less than C for all N.*

See Chamberlain (1983, Lemma 2) for a proof. There is a useful restriction on $B^N B^{N'}$ that is symmetrical to the restriction on V^N. The factors in Ross's model are intended to represent economywide shocks to asset returns. As such, each factor should have a broad-based influence affecting many assets in the economy. This means that each of the columns of B^N should have many nonzero components. Through a technical argument (see Connor 1982, appendix) this gives rise to a restriction called the *pervasiveness condition*. The pervasiveness condition requires that the minimum eigenvalue of $B^N B^{N'}$ approaches infinity as N goes to infinity.

Asset returns follow an *approximate factor model* if the sequence of covariance matrices can be written in the form (5) and the minimum eigenvalue of $B^N B^{N'}$ approaches infinity with N while the maximum eigenvalue of V^N is bounded for all N.

KEY RESULT 8. *If returns obey an approximate factor model, then the nonexistence of approximate arbitrage implies Ross's pricing bound.*

See Chamberlain and Rothschild (1983, Theorem 3) or Ingersoll (1984, Corollary) for a proof. Besides generalizing Ross's assumptions, the approximate factor model has implications for the econometrics of the APT. Let H^N denote the $N \times K$ matrix consisting of the K eigenvectors of Σ^N associated with the K largest eigenvalues. Chamberlain and Rothschild (1983, Corollary 3) show that, in an important sense, H^N is approximately equal to B^N for large N. Hence, the econometrician can estimate B^N by estimating the first K eigenvectors of Σ^N.

IV. THE COMPETITIVE EQUILIBRIUM VERSION OF THE APT

In the simple case of a noiseless factor model, the APT only requires the nonexistence of arbitrage in its derivation. For strict or approximate factor models, many theorists now believe that assuming competitive equilibrium improves the model. A competitive equilibrium approach avoids two weaknesses of the approximate arbitrage proof. First, approximate arbitrage can only give a pricing bound rather than a pricing approximation in the conventional sense. Second, the Ross-Huberman model of approximate arbitrage, although clever, is very different from the classic economic model of competitive market pricing. This separates the approximate arbitrage version of the APT from a wide range of applications and extensions that rely on classic price theory.

The extra assumptions needed for a competitive equilibrium proof of the APT are very natural to Ross's framework. Rather than weakening the APT by restricting its applicability, these extra assumptions actually strengthen the theory by clarifying the economic principles behind it.

First, for the competitive equilibrium proof, the market portfolio must be well diversified. This merely requires that no single asset in the economy accounts for a significant proportion of market wealth. If aggregate wealth is spread evenly across assets, then, by definition, so are the market portfolio weights.

Second, the market factors must be pervasive. As discussed in Section III, this restriction is justified by the economic rationale behind Ross's model. In the competitive equilibrium model this assumption guarantees that investors can efficiently trade factor risk and idiosyncratic risk by exchanging available securities. It allows investors to diversify away idiosyncratic risk without restricting their choice of factor risk exposure.

The competitive equilibrium proof relies on the most basic notions of economics. The first tool we need is the "invisible hand" of Adam Smith (1776), which guarantees that all mutually beneficial trades will be consummated in a competitive market. Given that investors are risk-averse, idiosyncratic risk is undesirable to them. With the market portfolio diversification and pervasiveness assumptions, it is possible to eliminate this risk from all portfolios via market trading. Rational investors will take advantage of these trading opportunities, and, in competitive equilibrium, all investors' portfolios will be free of idiosyncratic risk.

Another basic principle of economics is the notion that prices accurately reflect the marginal preferences of agents (see, for example, Marshall 1890). The randomness of security returns has two parts in an approximate factor model: a linear combination of factor variates and an additive component of idiosyncratic risk. We deduced above that each investor will be protected from idiosyncratic risk; hence, this risk will not affect his marginal preferences and will not be reflected in expected returns. The remaining risk (factor risk) is in linear combinations across assets. It follows, by the same logic that Ross uses for the noiseless factor model case, that the risk premia for each type of factor risk will be proportional to the linear coefficients.

KEY RESULT 9. *Given the pervasiveness and market portfolio diversification assumptions, then in competitive equilibrium the APT holds with a conventional approximation. In particular, it holds approximately for every asset for large N and exactly for every asset for infinite N.*

See Connor (1982, 1984) for proofs of the large-N and infinite-N cases, respectively, along the lines outlined above. Dybvig (1983) and Grinblatt and Titman (1983) develop alternative versions of the large-N model that simplify the proof and provide explicit bounds on the rate of convergence of the approximation.

The advantages of the competitive equilibrium proof of the APT are obvious: it provides a stronger conceptual basis for the theory by embedding it in a competitive market setting, and it leads to a more conventional approximation of asset prices. One disadvantage is that the model depends on an economywide pricing equilibrium. The approximate arbitrage model only depends on the hedging behavior of any single investor trading any large collection of assets. That some assets (for example, real estate, human capital, corporate bonds) might not be included in the model does not affect the pricing restriction. The equilibrium model analyzes economywide market equilibrium and *all* the trading interactions between investors. It does require that all assets are included in the model; in particular, the model assumes that the market portfolio (the value-weighted portfolio of *all* assets traded in the economy) is well diversified. In this respect, the equilibrium version is similar to the CAPM (capital asset pricing model), which also requires a joint hypothesis on the market portfolio. (For the CAPM the return on the market index used for empirical analysis must be a good proxy for the true market portfolio return.) One must trade off the necessity for a joint hypothesis about economywide aggregates against the fact that the equilibrium version gives a conventional approximation rather than a pricing bound.

Some of the papers in this area are Connor (1982, 1984), Dybvig (1983), Dybvig and Ross (1985), Grinblatt and Titman (1983), Shanken (1982, 1985), and Sharpe (1977).

V. CONCLUSION

Stripping the APT model to its barest essentials, it has two central ideas. The first is contained in Key Result 3—the absence of arbitrage implies linear pricing under (noiseless) factor linearity of returns. This takes the pure APT, which provides an intuitively appealing but empirically sterile result, and gives it usable empirical content.

A noiseless factor model involves an overly restrictive assumption on asset returns. The second central idea of the APT is that asset-specific risks will be diversified out of large portfolios. The weak law of large numbers states that all randomness disappears from a many-term, convex combination of random variables, given that the linear weights are spread "evenly" across terms and there are appropriate limits on the interdependencies between the random variables. A portfolio return is a convex combination of asset returns. If a portfolio consists of weights that are spread "evenly" across many assets, and asset-specific risks have limited independence, then these risks will disappear from the portfolio return. Risks that disappear from portfolio returns ought not affect the market prices of assets. This is the logical tool that allows the analyst to generalize the arbitrage-based pricing result of the noiseless factor model to a strict or approximate factor model.

One feature that is *not* essential to the APT is the approximate arbitrage approach, which Ross and Huberman use to prove the theory. Arbitrage (including approximate arbitrage) deals with the portfolio strategies of individual investors. It cannot be used to analyze the influence of investor interaction on prices. To more fully understand the diversification phenomenon, the analyst must examine marketwide equilibrium. Diversification is then viewed as a consequence of the invisible hand acting on investors. Investors eliminate asset-specific risks from their portfolios because it is mutually beneficial for them to do so. In this form the APT has its basis in the most fundamental principles of economics.

The greatest weakness of the APT is the large amount of ambiguity in its empirical predictions, particularly when compared to the CAPM. The CAPM is explicitly a one-beta model. The APT only guarantees a K-beta form, with K determined (one hopes) empirically. The CAPM specifies the market portfolio return as its "factor." We do not have a perfect proxy for the market portfolio return, but at least we know what we are searching for. The APT gives very little guidance on the identity of the factors beyond the restriction that they should obey the pervasiveness condition. Even with a perfect statistical procedure for extracting factors, a level of ambiguity would remain: any rotation of the factors provides an equally valid set of factors.

The APT assumes a factor model of returns a priori. A stronger model would generate the returns process endogenously, as a consequence of the economic forces in the economy acting on asset prices. In such a model the factors would be "labeled" as money shocks, production shocks, fiscal shocks, and so on, and so identified theoretically rather than empirically. This would greatly diminish the empirical ambiguity in the APT. The paper of Chen, Roll, and Ross (1986), although it does not provide a formal model, constitutes an important first step in this direction.

The APT is winning increasing favor as the best available model of stock market prices. Yet there is a dearth of applications of the model to related areas, such as capital structure, tax effects, dividend policy, and capital budgeting. For these tangential areas the CAPM remains the dominant paradigm. This situation may change with time.

This chapter has been devoted entirely to the theoretical side of the model. One of the most exciting directions for future research is toward tying together the theoretical and empirical approaches to the APT. This tendency is evident, for example, in Chamberlain and Rothschild (1983), Ingersoll (1984), and Stambaugh (1983), where theoretical models lead directly to new econometric insights.

The APT, despite its weaknesses, has made a permanent contribution to our understanding of security market pricing. The model will continue to change and grow due to new theoretical developments and empirical findings. The key insights, however, are here to stay.

NOTE

[1] Let A be an $N \times M$ matrix and b an N-vector. The duality theorem says that if $c'b = 0$ for all N-vectors c such $c'A = 0$, then there exists an M-vector γ such that $b = A\gamma$.

REFERENCES

Black, Fischer, and Scholes, Myron. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81: 637–54.

Chamberlain, Gary. 1983. "Funds, Factors and Diversification in the Arbitrage Pricing Theory." *Econometrica* 51: 1303–24.

Chamberlain, Gary, and Rothschild, Michael. 1983. "Arbitrage and Mean Variance Analysis on Large Asset Markets." *Econometrica* 51: 1281–1304.

Chen, Nai-fu, and Ingersoll, Jonathan, Jr. 1983. "Exact Pricing in Linear Factor Models with Finitely Many Assets: A Note." *Journal of Finance* 38: 985–88.

Chen, N.; Roll, Richard; and Ross, Stephen A. 1986. "Economic Forces and the Stock Market." *Journal of Business* 59: 383–403.

Connor, Gregory. 1982. "A Factor Pricing Theory for Capital Assets." Working paper, Kellogg Graduate School of Management, Northwestern University.

———. 1984. "A Unified Beta Pricing Theory." *Journal of Economic Theory* 34: 13–31.

Dothan, Uri, and Williams, Joseph T. 1981. "Education as an Option." *Journal of Business* 54: 117–39.

Duffie, Darrell. 1985. "Price Operators: Extensions, Potentials, and the Markov Valuation of Securities." Research paper No. 813, Graduate School of Business, Stanford University.

Duffie, Darrell, and Huang, Chi-fu. 1985. "Implementing Arrow-Debreu Equilibria by Continuous Trading of Few Long-Lived Securities." *Econometrica* 53: 1337–56. Reprinted Chapter 4, this volume.

Dybvig, Philip, H. 1985. "An Explicit Bound on Individual Assets' Deviations from APT Pricing in a Finite Economy." *Journal of Financial Economics* 12: 483–96.

Dybvig, Philip, H., and Ross, Stephen A. 1985. "Yes, the APT Is Testable." *Journal of Finance* 40: 1173–88.

Grinblatt, Mark, and Titman, Sheridan. 1983. "Factor Pricing in a Finite Economy." *Journal of Financial Economics* 12: 497–507.

Harrison, J., and Kreps, D. 1979. "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory* 20: 381–408.

Huberman, G. 1982. "A Simple Approach to Arbitrage Pricing Theory." *Journal of Economic Theory* 28: 183–91.

Ingersoll, J., Jr. 1984. "Some Results in the Theory of Arbitrage Pricing." *Journal of Finance* 39: 1021–39.

Kreps, David, M. 1981. "Arbitrage and Equilibrium in Economies with Infinitely Many Commodities." *Journal of Mathematical Economics* 8: 15–35.

Marshall, Alfred. 1920. *Principles of Economics*, 8th ed. Reprinted 1946, London: Macmillan & Co.

Merton, Robert, C. 1973. "Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science* 4: 141–83. Reprinted Chapter 8, this volume.

Paddock, James; Siegel, Daniel; and Smith, James. 1985. "Options Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases." Northwestern University, manuscript.

Ross, S. A. 1976. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory* 13: 341–60.

———. 1977. "Return, Risk, and Arbitrage." In Irwin Friend and James L. Bicksler, eds., *Risk and Return in Finance, I*. Cambridge, Mass.: Ballinger.

———. 1978. "A Simple Approach to the Valuation of Risky Streams." *Journal of Business* 51: 453–75.

Shanken, J. 1982. "The Arbitrage Pricing Theory: Is It Testable." *Journal of Finance* 37: 1129–40.

———. 1985. "Multi-Beta CAPM or Equilibrium APT? A Reply to Dybvig and Ross." *Journal of Finance* 40: 1189–96.

Sharpe, W. 1977. "The Capital Asset Pricing Model: A 'Multi-Beta' Interpretation." In H. Levy and M. Sarnat, eds., *Financial Decision Making under Uncertainty*. New York: Academic Press.

Smith, Adam. 1776. *The Wealth of Nations*. Reprinted 1936. New York: Random House, 10–19.

Smith, Clifford. 1976. "Options Pricing: A Review." *Journal of Financial Economics* 3: 3–51.

Stambaugh, Robert F. 1983. "Arbitrage Pricing with Information." *Journal of Financial Economics* 12: 357–70.