

CAPITAL MARKET THEORY AND THE PRICING
OF FINANCIAL SECURITIES

Robert C. Merton
Massachusetts Institute of Technology

Working Paper #1818-86 Revised May 1987

CAPITAL MARKET THEORY AND THE PRICING OF FINANCIAL SECURITIES*

1. Introduction

The core of financial economic theory is the study of individual behavior of households in the intertemporal allocation of their resources in an environment of uncertainty and of the role of economic organizations in facilitating these allocations. The intersection between this specialized branch of microeconomics and macroeconomic monetary theory is most apparent in the theory of capital markets [cf. Fischer and Merton (1984)]. It is therefore appropriate on this occasion to focus on the theories of portfolio selection, capital asset pricing and the roles that financial markets and intermediaries can play in improving allocational efficiency.

The complexity of the interaction of time and uncertainty provide intrinsic excitement to study of the subject, and as we will see, the mathematics of capital market theory contains some of the most interesting applications of probability and optimization theory. As exemplified by option pricing and modern portfolio theory, the research with all its seemingly obstrusive mathematics has nevertheless had a direct and significant influence on practice. This conjoining of intrinsic intellectual interest with extrinsic application is, indeed, a prevailing theme of theoretical research in financial economics.

Forthcoming: B. Friedman, F. Hahn (eds.), Handbook of Monetary Economics, Amsterdam: North-Holland.

The tradition in economic theory is to take the existence of households, their tastes, and endowments as exogenous to the theory. This tradition does not, however, extend to economic organizations and institutions. They are regarded as existing primarily because of the functions they serve instead of functioning primarily because they exist. Economic organizations are endogenous to the theory. To derive the functions of financial instruments, markets and intermediaries, a natural starting point is, therefore, to analyze the investment behavior of individual households.

It is convenient to view the investment decision by households as having two parts: (1) the "consumption-saving" choice where the individual decides how much wealth to allocate to current consumption and how much to save for future consumption; and (2) the "portfolio selection" choice where the investor decides how to allocate savings among the available investment opportunities. In general, the two decisions cannot be made independently. However, many of the important findings in portfolio theory can be more easily derived in an one-period environment where the consumption-savings allocation has little substantive impact on the results. Thus, we begin in Section 2 with the formulation and solution of the basic portfolio selection problem in a static framework, taking as given the individual's consumption decision.

Using the analysis of Section 2, we derive necessary conditions for financial equilibrium that are used to determine restrictions on equilibrium security prices and returns in Sections 3 and 4. In Sections 4 and 5, these restrictions are used to derive spanning or mutual fund theorems that provide a basis for an elementary theory of financial intermediation.

In Section 6, the combined consumption-portfolio selection problem is formulated in a more-realistic and more-complex dynamic setting. As shown in

Section 7, dynamic models in which agents can revise their decisions continuously in time produce significantly sharper results than their discrete-time counterparts and do so without sacrificing the richness of behavior found in an intertemporal decision-making environment.

The continuous-trading model is used in Section 8 to derive a theory of option, corporate-liability, and general contingent-claim pricing. The dynamic portfolio strategies used to derive these prices are also shown to provide a theory of production for the creation of financial instruments by financial intermediaries. The closing section of the paper examines intertemporal general-equilibrium pricing of securities and analyzes the conditions under which allocations in the continuous-trading model are Pareto efficient.

As is evident from this brief overview of content, the paper does not cover a number of important topics in capital market theory. For example, no attempt is made to make explicit how individuals and institutions acquire the information needed to make their decisions, and in particular how they modify their behavior in environments where there are significant differences in the information available to various participants. Thus, we do not cover either the informational efficiency of capital markets or the principal-agent problem and theory of auctions as applied to financial contracting, intermediation and markets.¹

2. One-Period Portfolio Selection

The basic investment choice problem for an individual is to determine the optimal allocation of his or her wealth among the available investment opportunities. The solution to the general problem of choosing the best investment mix is called portfolio selection theory. The study of portfolio selection theory begins with its classic one-period formulation.

There are n different investment opportunities called securities and the random variable one-period return per dollar on security j is denoted by Z_j ($j = 1, \dots, n$) where a "dollar" is the "unit of account." Any linear combination of these securities which has a positive market value is called a portfolio. It is assumed that the investor chooses at the beginning of a period that feasible portfolio allocation which maximizes the expected value of a von Neumann-Morgenstern utility function² for end-of-period wealth. Denote this utility function by $U(W)$, where W is the end-of-period value of the investor's wealth measured in dollars. It is further assumed that U is an increasing strictly-concave function on the range of feasible values for W and that U is twice-continuously differentiable.³ Because the criterion function for choice depends only on the distribution of end-of-period wealth, the only information about the securities that is relevant to the investor's decision is his subjective joint probability distribution for (Z_1, \dots, Z_n) .

In addition, it is assumed that:

Assumption 1: "Frictionless Markets"

There are no transactions costs or taxes, and all securities are perfectly divisible.

Assumption 2: "Price Taker"

The investor believes that his actions cannot affect the probability distribution of returns on the available securities. Hence, if w_j is the fraction of the investor's initial wealth W_0 , allocated to security j , then $\{w_1, \dots, w_n\}$ uniquely determines the probability distribution of his terminal wealth.

A riskless security is defined to be a security or feasible portfolio of securities whose return per dollar over the period is known with certainty.

Assumption 3: "No-Arbitrage Opportunities"

All riskless securities must have the same return per dollar. This common return will be denoted by R .

Assumption 4: "No-Institutional Restrictions"

Short-sales of all securities, with full use of proceeds, is allowed without restriction. If there exists a riskless security, then the borrowing rate equals the lending rate.⁴

Hence, the only restriction on the choice for the $\{w_j\}$ is the budget constraint that $\sum_1^n w_j = 1$.

Given these assumptions, the portfolio selection problem can be formally stated as:

$$\max_{\{w_1, \dots, w_n\}} E\left\{ U\left(\sum_1^n w_j Z_j W_0 \right) \right\}, \quad (2.1)$$

subject to $\sum_1^n w_j = 1$, where E is the expectation operator for the

subjective joint probability distribution. If (w_1^*, \dots, w_n^*) is a solution to (2.1), then it will satisfy the first-order conditions:

$$E\{U'(Z^*W_0)Z_j\} = \lambda, \quad j = 1, 2, \dots, n, \quad (2.2)$$

where the prime denotes derivative; $Z^* \equiv \sum_{j=1}^n w_j^* Z_j$ is the random variable return per dollar on the optimal portfolio; and λ is the Lagrange multiplier for the budget constraint. Together with the concavity assumptions on U , if the $n \times n$ variance-covariance matrix of the returns (Z_1, \dots, Z_n) is nonsingular and an interior solution exists, then the solution is unique.⁵ This non-singularity condition on the returns distribution eliminates "redundant" securities (i.e., securities whose returns can be expressed as exact linear combinations of the returns on other available securities).⁶ It also rules out that any one of the securities is a riskless security.

If a riskless security is added to the menu of available securities (call it the $(n + 1)$ st security), then it is the convention to express (2.1) as the following unconstrained maximization problem:

$$\max_{\{w_1, \dots, w_n\}} E\{ U[(\sum_{j=1}^n w_j (Z_j - R) + R)W_0] \}, \quad (2.3)$$

where the portfolio allocations to the risky securities are unconstrained because the fraction allocated to the riskless security can always be chosen to satisfy the budget constraint (i.e., $w_{n+1}^* = 1 - \sum_{j=1}^n w_j^*$).

The first-order conditions can be written as:

$$E\{U'(Z^*W_0)(Z_j - R)\} = 0, \quad j = 1, 2, \dots, n, \quad (2.4)$$

where Z^* can be rewritten as $\sum_1^n w_j^* (Z_j - R) + R$. Again, if it is assumed that the variance-covariance matrix of the returns on the risky securities is non-singular and an interior solution exists, then the solution is unique.

As formulated, neither (2.1) nor (2.3) reflects the physical constraint that end-of-period wealth cannot be negative. That is, no explicit consideration is given to the treatment of bankruptcy. To rule out bankruptcy, the additional constraint that with probability one, $Z^* \geq 0$, could be imposed on the choices for (w_1^*, \dots, w_n^*) .⁷ If, however, the purpose of this constraint is to reflect institutional restrictions designed to avoid individual bankruptcy, then it is too weak, because the probability assessments on the $\{Z_j\}$ are subjective. An alternative treatment is to forbid borrowing and short-selling in conjunction with limited-liability securities where, by law, $Z_j \geq 0$. These rules can be formalized as restrictions on the allowable set of $\{w_j\}$, such that $w_j \geq 0$, $j = 1, 2, \dots, n + 1$, and (2.1) or (2.3) can be solved using the methods of Kuhn and Tucker (1951) for inequality constraints. In Section 8, we formally analyze portfolio behavior and the pricing of securities when both investors and security lenders recognize the prospect of default. Thus, until that section, it is simply assumed that there exists a bankruptcy law which allows for $U(W)$ to be defined for $W < 0$, and that this law is consistent with the continuity and concavity assumptions on U .

The optimal demand functions for risky securities, $\{w_j^* W_0\}$, and the resulting probability distribution for the optimal portfolio will, of course, depend on the risk preferences of the investor, his initial wealth, and the joint distribution for the securities' returns. It is well known that

the von Neumann-Morgenstern utility function can only be determined up to a positive affine transformation. Hence, the preference orderings of all choices available to the investor are completely specified by the Pratt-Arrow⁸ absolute risk-aversion function, which can be written as:

$$A(W) \equiv \frac{-U''(W)}{U'(W)} \quad , \quad (2.5)$$

and the change in absolute risk aversion with respect to a change in wealth is, therefore, given by:

$$\frac{dA}{dW} = A'(W) = A(W) \left[A(W) + \frac{U'''(W)}{U''(W)} \right] \quad . \quad (2.6)$$

By the assumption that $U(W)$ is increasing and strictly concave, $A(W)$ is positive, and such investors are called risk-averse. An alternative, but related, measure of risk aversion is the relative risk-aversion function defined to be:

$$\underline{R}(W) \equiv - \frac{U''(W)W}{U'(W)} = A(W)W \quad , \quad (2.7)$$

and its change with respect to a change in wealth is given by:

$$\underline{R}'(W) = A'(W)W + A(W) \quad . \quad (2.8)$$

The certainty-equivalent end-of-period wealth, W_c , associated with a given portfolio for end-of-period wealth whose random variable value is denoted by W , is defined to be that value such that:

$$U(W_c) = E\{U(W)\} \quad , \quad (2.9)$$

i.e., W_c is the amount of money such that the investor is indifferent between having this amount of money for certain or the portfolio with random variable outcome W . The term "risk-averse" as applied to investors with

concave utility functions is descriptive in the sense that the certainty equivalent end-of-period wealth is always less than the expected value of the associated portfolio, $E\{W\}$, for all such investors. The proof follows directly by Jensen's Inequality: if U is strictly concave, then:

$$U(W_c) = E\{U(W)\} < U(E\{W\}) \quad ,$$

whenever W has positive dispersion, and because U is a non-decreasing function of W , $W_c < E\{W\}$.

The certainty-equivalent can be used to compare the risk-aversions of two investors. An investor is said to be more risk averse than a second investor if for every portfolio, the certainty-equivalent end-of-period wealth for the first investor is less than or equal to the certainty equivalent end-of-period wealth associated with the same portfolio for the second investor with strict inequality holding for at least one portfolio.

While the certainty equivalent provides a natural definition for comparing risk aversions across investors, Rothschild and Stiglitz⁹ have in a corresponding fashion attempted to define the meaning of "increasing risk" for a security so that the "riskiness" of two securities or portfolios can be compared. In comparing two portfolios with the same expected values, the first portfolio with random variable outcome denoted by W_1 is said to be less risky than the second portfolio with random variable outcome denoted by W_2 if:

$$E\{U(W_1)\} \geq E\{U(W_2)\} \quad (2.10)$$

for all concave U with strict inequality holding for some concave U . They bolster their argument for this definition by showing its equivalence to the following two other definitions:

There exists a random variable Z such that W_2 has the same distribution as $W_1 + Z$ where the conditional expectation of Z given the outcome on W_1 is zero (i.e., W_2 is equal in distribution to W_1 plus some "noise").

(2.11)

If the points of F and G , the distribution functions of W_1 and W_2 are confined to the closed interval $[a, b]$, and $T(y) \equiv \int_a^y [G(x) - F(x)] dx$, then $T(y) \geq 0$ and $T(b) = 0$ (i.e., W_2 has more "weight in its tails" than W_1).

(2.12)

A feasible portfolio with return per dollar Z will be called an efficient portfolio if there exists an increasing, strictly concave function V such that $E\{V'(Z)(Z_j - R)\} = 0$, $j = 1, 2, \dots, n$. Using the Rothschild-Stiglitz definition of "less risky," a feasible portfolio will be an efficient portfolio only if there does not exist another feasible portfolio which is less risky than it is. All portfolios that are not efficient are called inefficient portfolios.

From the definition of an efficient portfolio, it follows that no two portfolios in the efficient set can be ordered with respect to one another. From (2.10), it follows immediately that every efficient portfolio is a possible optimal portfolio, i.e., for each efficient portfolio there exists an increasing, concave U and an initial wealth W_0 such that the efficient portfolio is a solution to (2.1) or (2.3). Furthermore, from (2.10), all risk-averse investors will be indifferent between selecting their optimal portfolios from the set of all feasible portfolios or from the set of

efficient portfolios. Hence, without loss of generality, assume that all optimal portfolios are efficient portfolios.

With these general definitions established, we now turn to the analysis of the optimal demand functions for risky assets and their implications for the distributional characteristics of the underlying securities. A note on notation: the symbol " Z_e " will be used to denote the random variable return per dollar on an efficient portfolio, and a bar over a random variable (e.g., \bar{Z}) will denote the expected value of that random variable.

Theorem 2.1: If Z denotes the random variable return per dollar on any feasible portfolio and if $(Z_e - \bar{Z}_e)$ is riskier than $(Z - \bar{Z})$ in the Rothschild and Stiglitz sense, then $\bar{Z}_e > \bar{Z}$.

Proof: By hypothesis, $E\{U([Z - \bar{Z}]W_0)\} > E\{U([Z_e - \bar{Z}_e]W_0)\}$. If $\bar{Z} \geq \bar{Z}_e$, then trivially, $E\{U(ZW_0)\} > E\{U(Z_e W_0)\}$. But Z is a feasible portfolio and Z_e is an efficient portfolio. Hence, by contradiction, $\bar{Z}_e > \bar{Z}$.

Corollary 2.1a: If there exists a riskless security with return R , then $\bar{Z}_e \geq R$, with equality holding only if Z_e is a riskless security.

Proof: The riskless security is a feasible portfolio with expected return R . If Z_e is riskless, then by Assumption 3, $\bar{Z}_e = R$. If Z_e is not riskless, then $(Z_e - \bar{Z}_e)$ is riskier than $(R - R)$. Therefore, by Theorem 2.1, $\bar{Z}_e > R$.

Theorem 2.2: The optimal portfolio for a non-satiated, risk-averse investor will be the riskless security (i.e., $w_{n+1}^* = 1$, $w_j^* = 0$, $j = 1, 2, \dots, n$) if and only if $\bar{Z}_j = R$ for $j = 1, 2, \dots, n$.

Proof: From (2.4), $\{w_1^*, \dots, w_n^*\}$ will satisfy $E\{U'(Z^*W_0)(Z_j - R)\} = 0$, If $\bar{Z}_j = R$, $j = 1, 2, \dots, n$, then $Z^* = R$ will satisfy these first-order conditions. By the strict concavity of U and the non-singularity of the variance-covariance matrix of returns, this solution is unique. This proves the "if" part. If $Z^* = R$ is an optimal solution, then we can rewrite (2.4) as $U'(RW_0)E(Z_j - R) = 0$. By the non-satiation assumption, $U'(RW_0) > 0$. Therefore, for $Z^* = R$ to be an optimal solution, $\bar{Z}_j = R$, $j = 1, 2, \dots, n$. This proves the "only if" part.

Hence, from Corollary 2.1a and Theorem 2.2, if a risk-averse investor chooses a risky portfolio, then the expected return on that portfolio exceeds the riskless rate, and a risk-averse investor will choose a risky portfolio if, at least, one available security has an expected return different from the riskless rate.

Define the notation $E(Y|X_1, \dots, X_q)$ to mean the conditional expectation of the random variable Y , conditional on knowing the realizations for the random variables (X_1, \dots, X_q) .

Theorem 2.3: If there exists a feasible portfolio with return Z_p such that for security s , $Z_s = Z_p + \epsilon_s$, where $E(\epsilon_s) = E(\epsilon_s | Z_p, Z_j, j = 1, \dots, n, j \neq s) = 0$, then the fraction of every efficient portfolio allocated to security s is the same and equal to zero.

Proof: The proof follows by contradiction. Suppose Z_e is the return on an efficient portfolio with fraction $\delta_s \neq 0$ allocated to security s . Let Z be the return on a portfolio with the same fractional holdings as Z_e except instead of security s , it holds the fraction δ_s in feasible portfolio Z_p . Hence, $Z_e = Z + \delta_s(Z_s - Z_p)$ or $Z_e = Z + \delta_s \epsilon_s$. By hypothesis, $\bar{Z}_e = \bar{Z}$ and by construction $E(\epsilon_s | Z) = 0$. Therefore, for $\delta_s \neq 0$, Z_e is riskier than Z in the Rothschild-Stiglitz sense. But this contradicts the hypothesis that Z_e is an efficient portfolio. Hence, $\delta_s = 0$ for every efficient portfolio.

Corollary 2.3a: Let ψ denote the set of n securities with returns $(Z_1, \dots, Z_{s-1}, Z_s, Z_{s+1}, \dots, Z_n)$ and ψ' denote the same set of securities, except Z_s is replaced with $Z_{s'}$. If $Z_{s'} = Z_s + \epsilon_s$ and $E(\epsilon_s) = E(\epsilon_s | Z_1, \dots, Z_{s-1}, Z_s, Z_{s+1}, \dots, Z_n) = 0$, then all risk-averse investors would prefer to choose their optimal portfolios from ψ rather than ψ' .

The proof is essentially the same as the proof of Theorem 2.3, with Z_s replacing Z_p . Unless the holdings of Z_s in every efficient portfolio are zero, ψ will be strictly preferred to ψ' .

Theorem 2.3 and its corollary demonstrate that all risk-averse investors would prefer any "unnecessary" uncertainty or "noise" to be eliminated. In particular, by this theorem, the existence of lotteries is shown to be inconsistent with strict risk aversion on the part of all investors.¹⁰

While the inconsistency of strict risk aversion with observed behavior such as betting on the numbers can be "explained" by treating lotteries as consumption goods, it is difficult to use this argument to explain other implicit lotteries such as callable, sinking fund bonds where the bonds to be redeemed are selected at random.

As illustrated by the partitioning of the feasible portfolio set into its efficient and inefficient parts and the derived theorems, the Rothschild-Stiglitz definition of increasing risk is quite useful for studying the properties of optimal portfolios. However, it is important to emphasize that these theorems apply only to efficient portfolios and not to individual securities or inefficient portfolios. For example, if $(Z_j - \bar{Z}_j)$ is riskier than $(Z - \bar{Z})$ in the Rothschild-Stiglitz sense and if security j is held in positive amounts in an efficient or optimal portfolio (i.e., $w_j^* > 0$), then it does not follow that \bar{Z}_j must equal or exceed \bar{Z} . In particular, if $w_j^* > 0$, it does not follow that \bar{Z}_j must equal or exceed R . Hence, to know that one security is riskier than a second security using the Rothschild-Stiglitz definition of increasing risk provides no normative restrictions on holdings of either security in an efficient portfolio. And because this definition of riskier imposes no restrictions on the optimal demands, it cannot be used to derive properties of individual securities' return distributions from observing their relative holdings in an efficient portfolio. To derive these properties, a second definition of risk is required. Development of this measure is the topic of Section 3.

3. Risk Measures for Securities and Portfolios in the One-Period Model

In the previous section, it was suggested that the Rothschild-Stiglitz measure is not a natural definition of risk for a security. In this section, a second definition of increasing risk is introduced, and it is argued that this second measure is a more appropriate definition for the risk of a security. Although this second measure will not in general provide the same orderings as the Rothschild-Stiglitz measure, it is further argued that the two measures are not in conflict, and indeed, are complimentary.

If Z_e^K is the random variable return per dollar on an efficient portfolio K, then let $V_K(Z_e^K)$ denote an increasing, strictly concave function such that, for $V_K' \equiv dV_K/dZ_e^K$,

$$E\{(V_K'(Z_j) - R)\} = 0, \quad j = 1, 2, \dots, n,$$

i.e., V_K is a concave utility function such that an investor with initial wealth $W_0 = 1$ and these preferences would select this efficient portfolio as his optimal portfolio. While such a function V_K will always exist, it will not be unique. If $\text{cov}[x_1, x_2]$ is the functional notation for the covariance between the random variables x_1 and x_2 , then define the random variable, Y_K , by:

$$Y_K \equiv \frac{V_K' - E\{V_K'\}}{\text{cov}[V_K', Z_e^K]} \quad (3.1)$$

Y_K is well defined as long as Z_e^K has positive dispersion because $\text{cov}[V_K', Z_e^K] < 0$.¹¹ It is understood that in the following discussion "efficient portfolio" will mean "efficient portfolio with positive dispersion." Let Z_p denote the random variable return per dollar on any feasible portfolio p.

Definition: The measure of risk of portfolio p relative to efficient portfolio K with random variable return Z_e^K , b_p^K , is defined by:

$$b_p^K \equiv \text{cov}[Y_K, Z_p]$$

and portfolio p is said to be riskier than portfolio p' relative to efficient portfolio K if $b_p^K > b_{p'}^K$.

Theorem 3.1: If Z_p is the return on a feasible portfolio p and Z_e^K is the return on efficient portfolio K , then $\bar{Z}_p - R = b_p^K(\bar{Z}_e^K - R)$.

Proof: From the definition of V_K , $E\{V_K'(Z_j - R)\} = 0$,

$j = 1, 2, \dots, n$. Let δ_j be the fraction of portfolio p allocated to security j . Then, $Z_p = \sum_1^n \delta_j (Z_j - R) + R$, and $\sum_j \delta_j E\{V_K'(Z_j - R)\} = E\{V_K'(Z_p - R)\} = 0$. By a similar argument,

$E\{V_K'(Z_e^K - R)\} = 0$. Hence, $\text{cov}[V_K', Z_e^K] = (R - \bar{Z}_e^K)E\{V_K'\}$ and

$\text{cov}[V_K', Z_p] = (R - \bar{Z}_p)E\{V_K'\}$. By Corollary 2.1a, $\bar{Z}_e^K > R$. Therefore,

$\text{cov}[Y_K, Z_p] = (R - \bar{Z}_p)/(R - \bar{Z}_e^K)$.

Hence, the expected excess return on portfolio p , $\bar{Z}_p - R$, is in direct proportion to its risk, and because $\bar{Z}_e^K > R$, the larger is its risk, the larger is its expected return. Thus, Theorem 3.1 provides the first argument why b_p^K is a natural measure of risk for individual securities.

A second argument goes as follows. Consider an investor with utility function U and initial wealth W_0 who solves the portfolio selection problem:

$$\max_w E\{U([wZ_j + (1 - w)Z]W_0)\} \quad ,$$

where Z is the return on a portfolio of securities and Z_j is the return on the security j . The optimal mix, w^* , will satisfy the first-order condition:

$$E\{U'([w^*Z_j + (1 - w^*)Z]W_0)(Z_j - Z)\} = 0 \quad . \quad (3.2)$$

If the original portfolio of securities chosen was this investor's optimal portfolio (i.e., $Z = Z^*$), then the solution to (3.2) is $w^* = 0$. However, an optimal portfolio is an efficient portfolio. Therefore, by Theorem 3.1, $\bar{Z}_j - R = b_j^*(\bar{Z}^* - R)$. Hence, the "risk-return tradeoff" provided in Theorem 3.1 is a condition for personal portfolio equilibrium. Indeed, because security j may be contained in the optimal portfolio, w^*W_0 is similar to an excess demand function. b_j^* measures the contribution of security j to the Rothschild-Stiglitz risk of the optimal portfolio in the sense that the investor is just indifferent to a marginal change in the holdings of security j provided that $\bar{Z}_j - R = b_j^*(\bar{Z}^* - R)$. Moreover, by the Implicit Function Theorem, we have from (3.2) that:

$$\frac{\partial w^*}{\partial \bar{Z}_j} = \frac{w^*W_0 E\{U''(Z - Z_j)\} - E\{U'\}}{E\{U''(Z - Z_j)^2\}} > 0 \quad , \quad \text{at } w^* = 0 \quad . \quad (3.3)$$

Therefore, if \bar{Z}_j lies above the "risk-return" line in the (\bar{Z}, b^*) plane, then the investor would prefer to increase his holdings in security j , and if \bar{Z}_j lies below the line, then he would prefer to reduce his holdings. If the risk of a security increases, then the risk-averse investor must be "compensated" by a corresponding increase in that security's expected return

if his current holdings are to remain unchanged.

A third argument for why b_p^K is a natural measure of risk for individual securities is that the ordering of securities by their systematic risk relative to a given efficient portfolio will be identical to their ordering relative to any other efficient portfolio. That is, given the set of available securities, there is an unambiguous meaning to the statement "security j is riskier than security i ." To show this equivalence along with other properties of the b_p^K measure, we first prove a lemma.

Lemma 3.1: (a) $E[Z_p | V_K'] = E[Z_p | Z_e^K]$ for efficient portfolio K . (b) If $E[Z_p | Z_e^K] = \bar{Z}_p$, then $\text{cov}[Z_p, V_K'] = 0$. (c) $\text{cov}[Z_p, V_K'] = 0$ for efficient portfolio K if and only if $\text{cov}[Z_p, V_L'] = 0$ for every efficient portfolio L .

Proof: (a) V_K' is a continuous, monotonic function of Z_e^K and hence, V_K' and Z_e^K are in one-to-one correspondence.

(b) $\text{cov}[Z_p, V_K'] = E[V_K'(Z_p - \bar{Z}_p)] = E\{V_K' E[Z_p - \bar{Z}_p | Z_e^K]\} = 0$. (c) By definition, $b_p^K = 0$ if and only if $\text{cov}[Z_p, V_K'] = 0$. From Theorem 3.1, if $b_p^K = 0$, then $\bar{Z}_p = R$. From Corollary 2.1a, $\bar{Z}_e^L > R$ for every efficient portfolio L . Thus, from Theorem 3.1, $b_p^L = 0$ if and only if $\bar{Z}_p = R$.

Properties of the b_p^K measure of risk are:

Property 1: If L and K are efficient portfolios, then for any portfolio p , $b_p^K = b_{Lp}^{K,L}$.

From Corollary 2.1a, $\bar{Z}_e^K > R$ and $\bar{Z}_e^L > R$. From Theorem 3.1,

$$b_L^K = (\bar{Z}_e^L - R)/(\bar{Z}_e^K - R), \quad b_p^K = (\bar{Z}_p - R)/(\bar{Z}_e^K - R), \quad \text{and} \quad b_p^L = (\bar{Z}_p - R)/(\bar{Z}_e^L - R).$$

Hence, the b_p^K measure satisfies a type of "chain rule," with respect to different efficient portfolios.

Property 2: If L and K are efficient portfolios, then $b_K^K = 1$ and $b_K^L > 0$.

Property 2 follows from Theorem 3.1 and Corollary 2.1a. Hence, all efficient portfolios have positive systematic risk, relative to any efficient portfolio.

Property 3: $\bar{Z}_p = R$ if and only if $b_p^K = 0$ for every efficient portfolio K.

Property 3 follows from Theorem 3.1 and Properties 1 and 2.

Property 4: Let p and q denote any two feasible portfolios and let K and L denote any two efficient portfolios. $b_p^K \geq b_q^K$ if and only if $b_p^L \geq b_q^L$.

Property 4 follows from Property 3 if $b_p^L = b_q^L = 0$. Suppose $b_p^L \neq 0$.

Then Property 4 follows from Properties 1 and 2 because $(b_q^L/b_p^L) = (b_K^L b_q^K)/(b_K^L b_p^K) = (b_q^K/b_p^K)$. Thus, the b_p^K measure provides the same orderings of risk for any reference efficient portfolio.

Property 5: For each efficient portfolio K and any feasible portfolio p ,

$$Z_p = R + b_p^K (Z_e^K - R) + \epsilon_p \quad \text{where } E(\epsilon_p) = 0 \quad \text{and}$$

$$E[\epsilon_p V_L'(Z_e^L)] = 0 \quad \text{for every efficient portfolio } L.$$

From Theorem 3.1, $E(\epsilon_p) = 0$. If portfolio q is constructed by holding

\$1 in portfolio p , $\$b_p^K$ in the riskless security, and shortselling $\$b_p^K$ of efficient portfolio K , then $Z_q = R + \epsilon_p$. From Property 3,

$Z_q = R$ implies that $b_q^L = 0$ for every efficient portfolio L . But

$$b_q^L = 0 \quad \text{implies} \quad 0 = \text{cov}[Z_q, V_L'] = E[\epsilon_p V_L'] \quad \text{for every}$$

efficient portfolio L .

Property 6: If a feasible portfolio p has portfolio weights

$$(\delta_1, \dots, \delta_n), \quad \text{then} \quad b_p^K = \sum_{j=1}^n \delta_j b_j^K.$$

Property 6 follows directly from the linearity of the covariance operator with respect to either of its arguments. Hence, the systematic risk of a portfolio is the weighted sum of the systematic risks of its component securities.

The Rothschild-Stiglitz measure of risk is clearly different from the

b_j^K measure here. The Rothschild-Stiglitz measure provides only for a partial ordering while the b_j^K measure provides a complete ordering.

Moreover, they can give different rankings. For example, suppose the return on security j is independent of the return on efficient portfolio K , then

$$b_j^K = 0 \quad \text{and} \quad \bar{Z}_j = R. \quad \text{Trivially, } b_R^K = 0 \quad \text{for the riskless security.}$$

Therefore, by the b_j^K measure, security j and the riskless security

have equal risk. However, if security j has positive variance, then by the Rothschild-Stiglitz measure, security j is more risky than the riskless security. Despite this, the two measures are not in conflict and, indeed, are complementary. The Rothschild-Stiglitz definition measures the "total risk" of a security in the sense that it compares the expected utility from holding a security alone with the expected utility from holding another security alone. Hence, it is the appropriate definition for identifying optimal portfolios and determining the efficient portfolio set. However, it is not useful for defining the risk of securities generally because it does not take into account that investors can mix securities together to form portfolios. The b_j^K measure does take this into account because it measures the only part of an individual security's risk which is relevant to an investor: namely, the part that contributes to the total risk of his optimal portfolio. In contrast to the Rothschild-Stiglitz measure of total risk, the b_j^K measures the "systematic risk" of a security (relative to efficient portfolio K). Of course, to determine the b_j^K , the efficient portfolio set must be determined. Because the Rothschild-Stiglitz measure does just that, the two measures are complementary.

Although the expected return of a security provides an equivalent ranking to its b_p^K measure, the b_p^K measure is not vacuous. There exist non-trivial information sets which allow b_p^K to be determined without knowledge of \bar{Z}_p . For example, consider a model in which all investors agree on the joint distribution of the returns on securities. Suppose we know the utility function U for some investor and the probability distribution of his optimal portfolio, Z^*W_0 . From (3.2) we therefore know the distribution of $Y(Z^*)$. For security j , define the random variable $\epsilon_j \equiv Z_j - \bar{Z}_j$. Suppose, furthermore, that we have enough information about the joint distribution of $Y(Z^*)$ and ϵ_j to compute $\text{cov}[Y(Z^*), \epsilon_j] = \text{cov}[Y(Z^*), Z_j] = b_j^*$, but do not know \bar{Z}_j .¹² However, Theorem 3.1 is a necessary condition for equilibrium in the securities market. Hence, we can deduce the equilibrium expected return on security j from $\bar{Z}_j = R + b_j^*(\bar{Z}^* - R)$. Analysis of the necessary information sets required to deduce the equilibrium structure of security returns is an important topic in portfolio theory and one that will be explored further in succeeding sections.

The manifest behavioral characteristic shared by all risk-averse utility maximizers is to diversify (i.e., to spread one's wealth among many investments). The benefits of diversification in reducing risk depends upon the degree of statistical interdependence among returns on the available investments. The greatest benefits in risk reduction come from adding a security to the portfolio whose realized return tends to be higher when the

return on the rest of the portfolio is lower. Next to such "counter-cyclical" investments in terms of benefit are the non-cyclic securities whose returns are orthogonal to the return on the portfolio. Least beneficial are the pro-cyclical investments whose returns tend to be higher when the return on the portfolio is higher and lower when the return on the portfolio is lower. A natural summary statistic for this characteristic of a security's return distribution is its conditional expected-return function, conditional on the realized return of the portfolio. Because the risk of a security is measured by its marginal contribution to the risk of an optimal portfolio, it is perhaps not surprising that there is a direct relation between the risk measure of portfolio p , b_p , and the behavior of the conditional expected-return function, $G_p(Z_e) \equiv E[Z_p | Z_e]$, where Z_e is the realized return on an efficient portfolio.

Theorem 3.2: If Z_p and Z_q denote the returns on portfolios p and q respectively, and if for each possible value of Z_e , $dG_p(Z_e)/dZ_e > dG_q(Z_e)/dZ_e$ with strict inequality holding over some finite probability measure of Z_e , then portfolio p is riskier than portfolio q and $\bar{Z}_p > \bar{Z}_q$.

Proof: From (3.1) and the linearity of the covariance operator, $b_p - b_q = \text{Cov}[Y(Z_e), Z_p - Z_q] = E[Y(Z_e)(Z_p - Z_q)]$ because $E[Y(Z_e)] = 0$. By the property of conditional expectations, $E[Y(Z_e)(Z_p - Z_q)] = E(Y(Z_e)[G_p(Z_e) - G_q(Z_e)]) = \text{Cov}[Y(Z_e), G_p(Z_e) - G_q(Z_e)]$.

Thus, $b_p - b_q = \text{Cov}[Y(Z_e), G_p(Z_e) - G_q(Z_e)]$. From (3.1), $Y(Z_e)$ is a strictly increasing function of Z_e and by hypothesis, $G_p(Z_e) - G_q(Z_e)$, is a nondecreasing function of Z_e for all Z_e and a strictly increasing function of Z_e over some finite probability measure of Z_e . From Theorem 236 in Hardy, Littlewood, and Pölya (1959), it follows that $\text{Cov}[Y(Z_e), G_p(Z_e) - G_q(Z_e)] > 0$, and therefore, $b_p > b_q$. From Theorem 3.1, it follows that $\bar{Z}_p > \bar{Z}_q$.

Theorem 3.3: If Z_p and Z_q denote the returns on portfolio p and q , respectively and if, for each possible value of Z_e , $dG_p(Z_e)/dZ_e - dG_q(Z_e)/dZ_e = a_{pq}$, a constant, then $b_p = b_q + a_{pq}$ and $\bar{Z}_p = \bar{Z}_q + a_{pq}(\bar{Z}_e - R)$.

Proof: By hypothesis, $G_p(Z_e) - G_q(Z_e) = a_{pq}Z_e + h$ where h does not depend on Z_e . As in the proof of Theorem 3.2, $b_p - b_q = \text{Cov}[Y(Z_e), G_p(Z_e) - G_q(Z_e)] = \text{Cov}[Y(Z_e), a_{pq}Z_e + h]$. Thus, $b_p - b_q = a_{pq}$ because $\text{Cov}[Y(Z_e), Z_e] = 1$ and $\text{Cov}[Y(Z_e), h] = 0$. From Theorem 3.1, $\bar{Z}_p = R + b_q(\bar{Z}_e - R) + a_{pq}(\bar{Z}_e - R) = \bar{Z}_q + a_{pq}(\bar{Z}_e - R)$.

Theorem 3.4: If, for all possible values of Z_e ,

(i) $dG_p(Z_e)/dZ_e > 1$, then $\bar{Z}_p > \bar{Z}_e$.

$$(ii) \quad 0 < dG_p(Z_e)/dZ_e < 1, \quad \text{then } R < \bar{Z}_p < \bar{Z}_e$$

$$(iii) \quad dG_p(Z_e)/dZ_e < 0, \quad \bar{Z}_p < R \quad .$$

$$(iv) \quad dG_p(Z_e)/dZ_e = a_p, \text{ a constant, then } \bar{Z}_p = R + a_p(\bar{Z}_e - R) \quad .$$

The proof follows directly from Theorems 3.3 and 3.4 by substituting either Z_e or R for Z_q and noting that $dG_q(Z_e)/dZ_e = 1$ for $Z_q = Z_e$ and $dG_q(Z_e)/dZ_e = 0$ for $Z_q = R$.

As Theorems 3.2-3.4 demonstrate, the conditional expected-return function provides considerable information about a security's risk and equilibrium expected return. It is, moreover, common practice for security analysts to provide conditioned forecasts of individual security returns, conditioned on the realized return of a broad-based stock portfolio such as the Standard & Poor's 500. As is evident from these theorems, the conditional expected-return function does not in general provide sufficient information to determine the exact risk of a security. As follows from Theorems 3.3 and 3.4(iv), the exception is the case where this function is linear in Z_e . Although surely a special case, it is a rather important one as will be shown in Section 4.

4. Spanning, Separation, and Mutual Fund Theorems

Definition: A set of M feasible portfolios with random variable returns (X_1, \dots, X_M) are said to span the space of portfolios contained in the set Ψ if and only if for any portfolio in Ψ with return denoted by Z_p , there exists numbers $(\delta_1, \dots, \delta_M)$, $\sum_1^M \delta_j = 1$, such that $Z_p = \sum_1^M \delta_j X_j$. If N is the number of securities available to generate the portfolios in Ψ and if M^* denotes the smallest number of feasible portfolios that span the space of portfolios contained in Ψ , then $M^* \leq N$.

Fischer (1972) and Merton (1982, pp. 611-614) use comparative statics analysis to show that little can be derived about the structure of optimal portfolio demand functions unless further restrictions are imposed on the class of investors' utility functions or the class of probability distributions for securities' returns. A particularly fruitful set of such restrictions is the one that provides for a non-trivial (i.e., $M^* < N$) spanning of either the feasible or efficient portfolio sets. Indeed, the spanning property leads to a collection of "mutual fund" or "separation" theorems that are fundamental to modern financial theory.

A mutual fund is a financial intermediary that holds as its assets a portfolio of securities and issues as liabilities shares against this collection of assets. Unlike the optimal portfolio of an individual investor, the portfolio of securities held by a mutual fund need not be an efficient portfolio. The connection between mutual funds and the spanning property can be seen in the following theorem:

Theorem 4.1: If there exist M mutual funds whose portfolios span the portfolio set Ψ , then all investors will be indifferent between selecting their optimal portfolios from Ψ or from portfolio combinations of just the M mutual funds.

The proof of the theorem follows directly from the definition of spanning. If Z^* denotes the return on an optimal portfolio selected from Ψ and if X_j denotes the return on the j th mutual fund's portfolio, then there exist portfolio weights $(\delta_1^*, \dots, \delta_M^*)$ such that $Z^* = \sum_{j=1}^M \delta_j^* X_j$. Hence, any investor would be indifferent between the portfolio with return Z^* and the $(\delta_1^*, \dots, \delta_M^*)$ combination of the mutual fund shares.

Although the theorem states "indifference," if there are information-gathering or other transactions costs and if there are economies of scale, then investors would prefer the mutual funds whenever $M < N$. By a similar argument, one would expect that investors would prefer to have the smallest number of funds necessary to span Ψ . Therefore, the smallest number of such funds, M^* , is a particularly important spanning set. Hence, the spanning property can be used to derive an endogenous theory for the existence of financial intermediaries with the functional characteristics of a mutual fund. Moreover, from these functional characteristics a theory for their optimal management can be derived.

For the mutual fund theorems to have serious empirical content, the minimum number of funds required for spanning M^* must be significantly smaller than the number of available securities N . When such spanning obtains, the investor's portfolio selection problem can be separated into two

steps: first, individual securities are mixed together to form the M^* mutual funds; second, the investor allocates his wealth among the M^* funds' shares. If the investor knows that the funds span the space of optimal portfolios, then he need only know the joint probability distribution of (X_1, \dots, X_{M^*}) to determine his optimal portfolio. It is for this reason that the mutual fund theorems are also called "separation" theorems. However, if the M^* funds can be constructed only if the fund managers know the preferences, endowments, and probability beliefs of each investor, then the formal separation property will have little operational significance.

In addition to providing an endogenous theory for mutual funds, the existence of a non-trivial spanning set can be used to deduce equilibrium properties of individual securities' returns and to derive optimal rules for business firms making production and capital budgeting decisions. Moreover, in virtually every model of portfolio selection in which empirical implications beyond those presented in Sections 2 and 3 are derived, some non-trivial form of the spanning property obtains.

While the determination of conditions under which non-trivial spanning will obtain is, in a broad sense, a subset of the traditional economic theory of aggregation, the first rigorous contributions in portfolio theory were made by Arrow (1953,1964), Markowitz (1959), and Tobin (1958). In each of these papers, and most subsequent papers, the spanning property is derived as an implication of the specific model examined, and therefore such derivations provide only sufficient conditions. In two notable exceptions, Cass and Stiglitz (1970) and Ross (1978) "reverse" the process by deriving necessary conditions for non-trivial spanning to obtain. In this section necessary and sufficient conditions for spanning are developed along the lines of Cass and Stiglitz and

Ross, leaving until Section 5 discussion of the specific models of Arrow, Markowitz, and Tobin.

Let Ψ^f denote the set of all feasible portfolios that can be constructed from a riskless security with return R and n risky securities with a given joint probability distribution for their random variable returns (Z_1, \dots, Z_n) . Let Ω denote the $n \times n$ variance-covariance matrix of the returns on the n risky assets.

Theorem 4.2: Necessary conditions for the M feasible portfolios with returns (X_1, \dots, X_M) to span the portfolio set Ψ^f are (i) that the rank of $\Omega \leq M$ and (ii) that there exists numbers $(\delta_1, \dots, \delta_M)$, $\sum_1^M \delta_j = 1$, such that the random variable $\sum_1^M \delta_j X_j$ has zero variance.

Proof: (i) The set of portfolios Ψ^f defines a $(n + 1)$ dimensional vector space. By definition, if (X_1, \dots, X_M) spans Ψ^f , then each risky security's return can be represented as a linear combination of (X_1, \dots, X_M) . Clearly, this is only possible if the rank of $\Omega \leq M$.

(ii) The riskless security is contained in Ψ^f . Therefore, if (X_1, \dots, X_M) spans Ψ^f , then there must exist a portfolio combination of (X_1, \dots, X_M) which is riskless.

Proposition 4.1: If $Z_p = \sum_1^n a_j Z_j + b$ is the return on some security or portfolio and if there are no "arbitrage opportunities" (Assumption 3), then (1) $b = [1 - \sum_1^n a_j]R$ and (2) $Z_p = R + \sum_1^n a_j (Z_j - R)$.

Proof: Let Z^\dagger be the return on a portfolio with fraction δ_j^\dagger allocated to security j , $j = 1, \dots, n$; δ_p allocated to the security with return Z_p ; $(1 - \delta_p - \sum_1^n \delta_j^\dagger)$ allocated to the riskless security with return R . If δ_j^\dagger is chosen such that $\delta_j^\dagger = -\delta_p a_j$, then $Z^\dagger = R + \delta_p (b - R[1 - \sum_1^n a_j])$. Z^\dagger is a riskless security, and therefore, by Assumption 3, $Z^\dagger = R$. But δ_p can be chosen arbitrarily. Therefore, $b = [1 - \sum_1^n a_j]R$. Substituting for b , it follows directly that $Z_p = R + \sum_1^n a_j (Z_j - R)$.

As long as there are no arbitrage opportunities, from Theorem 4.2 and Proposition 4.1, it can be assumed without loss of generality that one of the portfolios in any candidate spanning set is the riskless security. If, by convention $X_M = R$, then in all subsequent analyses the notation (X_1, \dots, X_m, R) will be used to denote an M -portfolio spanning set where $m \equiv M - 1$ is the number of risky portfolios (together with the riskless security) that span Ψ^f .

Theorem 4.3: A necessary and sufficient condition for (X_1, \dots, X_m, R) to span Ψ^f is that there exist numbers (a_{ij}) such that $Z_j = R + \sum_1^m a_{ij} (X_i - R)$, $j = 1, 2, \dots, n$.

Proof: If (X_1, \dots, X_m, R) span Ψ^f , then there exist portfolio weights $(\delta_{1j}, \dots, \delta_{mj})$, $\sum_1^m \delta_{ij} = 1$, such that $Z_j = \sum_1^m \delta_{ij} X_i$. Noting that $X_M = R$ and substituting $\delta_{Mj} = 1 - \sum_1^m \delta_{ij}$, we have that $Z_j = R + \sum_1^m \delta_{ij} (X_i - R)$. This proves necessity. If there exist

numbers (a_{ij}) such that $Z_j = R + \sum_1^m a_{ij}(X_i - R)$, then pick the portfolio weights $\delta_{ij} = a_{ij}$ for $i = 1, \dots, m$, and $\delta_{Mj} = 1 - \sum_1^m \delta_{ij}$, from which it follows that $Z_j = \sum_1^M \delta_{ij} X_i$. But every portfolio in ψ^f can be written as a portfolio combination of (Z_1, \dots, Z_n) and R . Hence, (X_1, \dots, X_m, R) spans ψ^f and this proves sufficiency.

Let Ω_X be the $m \times m$ variance-covariance matrix of the returns on the m portfolios with returns (X_1, \dots, X_m) .

Corollary 4.3a. A necessary and sufficient condition for (X_1, \dots, X_m, R) to be the smallest number of feasible portfolios that span (i.e., $M^* = m + 1$) is that the rank of Ω equals the rank of $\Omega_X = m$.

Proof: If (X_1, \dots, X_m, R) span ψ^f and m is the smallest number of risky portfolios that does, then (X_1, \dots, X_m) must be linearly independent, and therefore $\text{rank } \Omega_X = m$. Hence, (X_1, \dots, X_m) form a basis for the vector space of security returns (Z_1, \dots, Z_n) . Therefore, the rank of Ω must equal Ω_X . This proves necessity. If the rank of $\Omega_X = m$, then (X_1, \dots, X_m) are linearly independent. Moreover, $(X_1, \dots, X_m) \in \psi^f$. Hence, if the rank of $\Omega = m$, then there exist

numbers (a_{ij}) such that $Z_j - \bar{Z}_j = \sum_1^m a_{ij}(X_i - \bar{X}_i)$ for $j = 1, 2, \dots, n$. Therefore, $Z_j = b_j + \sum_1^m a_{ij} X_i$, where $b_j \equiv \bar{Z}_j - \sum_1^m a_{ij} \bar{X}_i$. By the same argument used to prove Proposition 4.1, $b_j = [1 - \sum_1^m a_{ij}]R$. Therefore, $Z_j = R + \sum_1^m a_{ij}(X_i - R)$. By Theorem 4.3, (X_1, \dots, X_m, R) span ψ^f .

It follows from Corollary 4.3a that a necessary and sufficient condition for non-trivial spanning of Ψ^f is that some of the risky securities are redundant securities. Note, however that this condition is sufficient only if securities are priced such that there are no arbitrage opportunities.

In all these derived theorems the only restriction on investors' preferences was that they prefer more to less. In particular, it was not assumed that investors are necessarily risk averse. Although Ψ^f was defined in terms of a known joint probability distribution for (Z_1, \dots, Z_n) , which implies homogeneous beliefs among investors, inspection of the proof of Theorem 4.3 shows that this condition can be weakened. If investors agree on a set of portfolios (X_1, \dots, X_m, R) such that $Z_j = R + \sum_{i=1}^m a_{ij}(X_i - R)$, $j = 1, 2, \dots, n$, and if they agree on the numbers (a_{ij}) , then by Theorem 4.3 (X_1, \dots, X_m, R) span Ψ^f even if investors do not agree on the joint distribution of (X_1, \dots, X_m) . These appear to be the weakest restrictions on preferences and probability beliefs that can produce non-trivial spanning and the corresponding mutual fund theorem. Hence, to derive additional theorems it is now further assumed that all investors are risk averse and that investors have homogeneous probability beliefs.

Define Ψ^e to be the set of all efficient portfolios contained in Ψ^f .

Proposition 4.2: If Z_e is the return on a portfolio contained in Ψ^e , then any portfolio that combines positive amounts of Z_e with the riskless security is also contained in Ψ^e .

Proof: Let $Z = \delta(Z_e - R) + R$ be the return on a portfolio with positive fraction δ allocated to Z_e and fraction $(1 - \delta)$ allocated to the riskless security. Because Z_e is an efficient portfolio, there exists a strictly concave, increasing function V such that $E\{V'(Z_e)(Z_j - R)\} = 0$, $j = 1, 2, \dots, n$. Define $U(W) \equiv V(aW + b)$, where $a \equiv 1/\delta > 0$ and $b \equiv (\delta - 1)R/\delta$. Because $a > 0$, U is a strictly concave and increasing function. Moreover, $U'(Z) = aV'(Z_e)$. Hence, $E\{U'(Z)(Z_j - R)\} = 0$, $j = 1, 2, \dots, n$. Therefore, there exists a utility function such that Z is an optimal portfolio, and thus Z is an efficient portfolio.

It follows immediately from Proposition 4.2, that for every number \bar{Z} such that $\bar{Z} \geq R$, there exists at least one efficient portfolio with expected return equal to \bar{Z} . Moreover, we also have that if (X_1, \dots, X_M) are the returns on M candidate portfolios to span the space of efficient portfolios Ψ^e , then without loss of generality it can be assumed that one of the portfolios is the riskless security.

Theorem 4.4: Let (X_1, \dots, X_m) denote the returns on m feasible portfolios. If for security j , there exist numbers (a_{ij}) such that $Z_j = \bar{Z}_j + \sum_1^m a_{ij}(X_i - \bar{X}_i) + \epsilon_j$ where $E[\epsilon_j V'_K(Z_e^K)] = 0$ for some efficient portfolio K , then $\bar{Z}_j = R + \sum_1^m a_{ij}(\bar{X}_i - R)$.

Proof: Let Z_p be the return on a portfolio with fraction δ allocated to security j ; fraction $\delta_i = -\delta a_{ij}$ allocated to portfolio X_i , $i = 1, \dots, m$; and $1 - \delta - \sum_1^m \delta_i$ allocated to the riskless security. By hypothesis, Z_p can be written as $Z_p = R + \delta[\bar{Z}_j - R - \sum_1^m a_{ij}(\bar{X}_i - R)] + \delta\epsilon_j$ where

$E[\delta \epsilon_j V'_K] = \delta E(\epsilon_j V'_K) = 0$. By construction, $E(\epsilon_j) = 0$, and hence, $\text{cov}[Z_p, V'_K] = 0$. Therefore, the systematic risk of portfolio p , $b_p^K = 0$. From Theorem 3.1, $\bar{Z}_p = R$. But δ can be chosen arbitrarily. Therefore, $\bar{Z}_j = R + \sum_1^m a_{ij}(\bar{X}_i - R)$.

Hence, if the return on a security can be written in this linear form relative to the portfolios (X_1, \dots, X_m) , then its expected excess return is completely determined by the expected excess returns on these portfolios and the weights (a_{ij}) .

Theorem 4.5: If, for every security j , there exist numbers (a_{ij}) such that $Z_j = R + \sum_1^m a_{ij}(X_i - R) + \epsilon_j$, where $E[\epsilon_j | X_1, \dots, X_m] = 0$, then (X_1, \dots, X_m, R) span the set of efficient portfolios Ψ^e .

Proof: Let w_j^K denote the fraction of efficient portfolio K allocated to security j , $j = 1, \dots, n$. By hypothesis, we can write $Z_e^K = R + \sum_1^m \delta_i^K (X_i - R) + \epsilon^K$ where $\delta_i^K \equiv \sum_1^n w_j^K a_{ij}$ and $\epsilon^K \equiv \sum_1^n w_j^K \epsilon_j$ where $E[\epsilon^K | X_1, \dots, X_m] = \sum_1^n w_j^K E[\epsilon_j | X_1, \dots, X_m] = 0$. Construct the portfolio with return Z by allocating fraction δ_i^K to portfolio $X_i, i = 1, \dots, m$ and fraction $1 - \sum_1^m \delta_i^K$ to the riskless security. By construction, $Z_e^K = Z + \epsilon^K$ where $E[\epsilon^K | Z] = E[\epsilon^K | \sum_1^m \delta_i^K X_i] = 0$ because $E[\epsilon^K | X_1, \dots, X_m] = 0$. Hence, for $\epsilon^K \neq 0$, Z_e^K is riskier than Z in the Rothschild-Stiglitz sense, which contradicts that Z_e^K is an efficient portfolio. Thus, $\epsilon^K \equiv 0$ for every efficient portfolio K , and all efficient portfolios can be generated by a portfolio combination of (X_1, \dots, X_m, R) .

Therefore, if we can find a set of portfolios (X_1, \dots, X_m) such that every security's return can be expressed as a linear combination of the returns (X_1, \dots, X_m, R) , plus noise relative to these portfolios, then we have a set of portfolios that span Ψ^e . The following theorem, first proved by Ross (1978), shows that security returns can always be written in a linear form relative to a set of spanning portfolios.

Theorem 4.6: Let w_j^K denote the fraction of efficient portfolio K allocated to security j , $j = 1, \dots, n$. (X_1, \dots, X_m, R) span Ψ^e if and only if there exist numbers (a_{ij}) for every security j such that $Z_j = R + \sum_1^m a_{ij}(X_i - R) + \epsilon_j$, where $E[\epsilon_j | \sum_1^m c_i^K X_i] = 0$, $c_i^K \equiv \sum_1^n w_j^K a_{ij}$, for every efficient portfolio K .

Proof: The "if part" follows directly from the proof of Theorem 4.5. In that proof, we only needed that $E[\epsilon^K | \sum_1^m c_i^K X_i] = 0$ for every efficient portfolio K to show that (X_1, \dots, X_m) span Ψ^e . The proof of the "only if" part is long and requires the proof of four specialized lemmas [see Ross (1978, appendix)]. It is therefore, not presented here.

Corollary 4.6: (X, R) span Ψ^e if and only if there exist a number a_j for each security j , $j = 1, \dots, n$, such that $Z_j = R + a_j(X - R) + \epsilon_j$ where $E(\epsilon_j | X) = 0$.

Proof: The "if part" follows directly from Theorem 4.5. The "only if" part is as follows: By hypothesis, $Z_e^K = \delta^K(X - R) + R$ for every efficient portfolio K . If $X = R$, then from Corollary 2.1a, $\delta^K = 0$ for every efficient portfolio K and R spans Ψ^e . Otherwise, from Theorem 2.2, $\delta^K \neq 0$ for every efficient portfolio. By Theorem 4.6, $E[\epsilon_j | \delta^K X] = 0$, for $j = 1, \dots, n$ and every efficient portfolio K . But, for $\delta^K \neq 0$, $E[\epsilon_j | \delta^K X] = 0$ if and only if $E[\epsilon_j | X] = 0$.

In addition to Ross (1978), there have been a number of studies of the properties of efficient portfolios [cf., Chen and Ingersoll (1983), Dybvig and Ross (1982), and Neilsen (1986)]. However, there is still much to be determined. For example, from Theorem 4.6, a necessary condition for (X_1, \dots, X_m, R) to span Ψ^e is that $E[\epsilon_j | Z_e^K] = 0$ for $j = 1, \dots, n$ and every efficient portfolio K . For $m > 1$, this condition is not sufficient to ensure that (X_1, \dots, X_m, R) span Ψ^e . The condition that $E[\epsilon_j | \sum_1^m \lambda_i X_i] = 0$ for all numbers λ_i implies that $E[\epsilon_j | X_1, \dots, X_m] = 0$. If, however, the $\{\lambda_i\}$ are restricted to the class of optimal portfolio weights $\{\delta_i^K\}$ as in Theorem 4.6 and $m > 1$, it does not follow that $E[\epsilon_j | X_1, \dots, X_m] = 0$. Thus, $E[\epsilon_j | X_1, \dots, X_m] = 0$ is sufficient, but not necessary, for (X_1, \dots, X_m, R) to span Ψ^e . It is not known whether any material cases of spanning are ruled out by imposing this stronger condition. Empirical application of the spanning conditions generally assume that the condition $E[\epsilon_j | X_1, \dots, X_m] = 0$ obtains.

Since ψ^e is contained in ψ^f , any properties proved for portfolios that span ψ^e must be properties of portfolios that span ψ^f . From Theorems 4.3, 4.5, and 4.6, the essential difference is that to span the efficient portfolio set it is not necessary that linear combinations of the spanning portfolios exactly replicate the return on each available security. Hence, it is not necessary that there exist redundant securities for non-trivial spanning of ψ^e to obtain. Of course, both theorems are empty of any empirical content if the size of the smallest spanning set M^* is equal to $(n + 1)$.

As discussed in the introduction to this section, all the important models of portfolio selection exhibit the non-trivial spanning property for the efficient portfolio set. Therefore, for all such models that do not restrict the class of admissible utility functions beyond that of risk aversion, the distribution of individual security returns must be such that

$Z_j = R + \sum_{i=1}^m a_{ij}(X_i - R) + \epsilon_j$, where ϵ_j satisfies the conditions of Theorem 4.6 for $j = 1, \dots, n$. Moreover, given some knowledge of the joint

distribution of a set of portfolios that span ψ^e with $(Z_j - \bar{Z}_j)$,

there exists a method for determining the a_{ij} and \bar{Z}_j .

Proposition 4.3: If, for every security j , $E(\epsilon_j | X_1, \dots, X_m) = 0$ with (X_1, \dots, X_m) linearly independent with finite variances and if the return on security j , Z_j , has a finite variance, then the (a_{ij}) , $i = 1, 2, \dots, m$ in Theorems 4.5 and 4.6 are given by:

$$a_{ij} = \sum_{k=1}^m v_{ik} \text{cov}[X_k, Z_j] \quad ,$$

where v_{ik} is the i -kth element of Ω_X^{-1} .

The proof of Proposition 4.3 follows directly from the condition that $E(\varepsilon_j | X_k) = 0$, which implies that $\text{cov}[\varepsilon_j, X_k] = 0$, $k = 1, \dots, m$. The condition that (X_1, \dots, X_m) be linearly independent is trivial in the sense that knowing the joint distribution of a spanning set one can always choose a linearly independent subset. The only properties of the joint distributions required to compute the a_{ij} are the variances and covariances of the X_1, \dots, X_m and the covariances between Z_j and X_1, \dots, X_m . In particular, knowledge of \bar{Z}_j is not required because $\text{cov}[X_k, Z_j] = \text{cov}[X_k, Z_j - \bar{Z}_j]$. Hence, for $m < n$ (and especially so for $m \ll n$), there exists a non-trivial information set which allows the a_{ij} to be determined without knowledge of \bar{Z}_j . If $\bar{X}_1, \dots, \bar{X}_m$ are known, then \bar{Z}_j can be computed by the formula in Theorem 4.4. By comparison with the example in Section 3, the information set required there to determine \bar{Z}_j was a utility function and the joint distribution of its associated optimal portfolio with $(Z_j - \bar{Z}_j)$. Here, we must know a complete set of portfolios that span Ψ^e . However, here only the second-moment properties of the joint distribution need be known, and no utility function information other than risk aversion is required.

A special case of no little interest is when a single risky portfolio and the riskless security span the space of efficient portfolios and Corollary 4.6 applies. Indeed, the classic model of Markowitz and Tobin, which is discussed in Section 5, exhibits this strong form of separation. Moreover, most macroeconomic models have highly aggregated financial sectors where investors' portfolio choices are limited to simple combinations of two securities: "bonds" and "stocks." The rigorous microeconomic foundation for such aggregation is precisely that

Ψ^e is spanned by a single risky portfolio and the riskless security.

If X denotes the random variable return on a risky portfolio such that (X, R) spans Ψ^e , then the return on any efficient portfolio, Z_e , can be written as if it had been chosen by combining the risky portfolio with return X with the riskless security. Namely, $Z_e = \delta(X - R) + R$, where δ is the fraction allocated to the risky portfolio and $(1 - \delta)$ is the fraction allocated to the riskless security. By Corollary 2.1a, the sign of δ will be same for every efficient portfolio, and therefore all efficient portfolios will be perfectly positively correlated. If $\bar{X} > R$, then by Proposition 4.2, X will be an efficient portfolio and $\delta > 0$ for every efficient portfolio.

Proposition 4.4: If (Z_1, \dots, Z_n) contain no redundant securities,

δ_j denotes the fraction of portfolio X allocated to security j , and w_j^* denotes the fraction of any risk-averse investor's optimal portfolio allocated to security j , $j = 1, \dots, n$, then for every such risk-averse investor,

$$w_j^*/w_k^* = \delta_j/\delta_k, \quad j, k = 1, 2, \dots, n \quad .$$

The proof follows immediately because every optimal portfolio is an efficient portfolio, and the holdings of risky securities in every efficient portfolio are proportional to the holdings in X . Hence, the relative holdings of risky securities will be the same for all risk-averse investors. Whenever

Proposition 4.4 holds and if there exist numbers δ_j^* where $\delta_j^*/\delta_k^* = \delta_j/\delta_k$, $j, k, = 1, \dots, n$ and $\sum_1^n \delta_j^* = 1$, then the portfolio with proportions $(\delta_1^*, \dots, \delta_n^*)$ is called the Optimal Combination of Risky Assets. If such a portfolio exists, then without loss of generality it can always be assumed that $X = \sum_1^n \delta_j^* Z_j$.

Proposition 4.5: If (X, R) spans Ψ^e , then Ψ^e is a convex set.

Proof: Let Z_e^1 and Z_e^2 denote the returns on two distinct efficient portfolios. Because (X, R) spans Ψ^e , $Z_e^1 = \delta_1(X - R) + R$ and $Z_e^2 = \delta_2(X - R) + R$. Because they are distinct, $\delta_1 \neq \delta_2$, and so assume $\delta_1 \neq 0$. Let $Z \equiv \lambda Z_e^1$ denote the return on a portfolio which allocates fraction λ to Z_e^1 and $(1 - \lambda)$ to Z_e^2 , where $0 \leq \lambda \leq 1$. By substitution, the expression for Z can be rewritten as $Z = \delta(Z_e^1 - R) + R$, where $\delta \equiv [\lambda + (\delta_2/\delta_1)(1 - \lambda)]$. Because Z_e^1 and Z_e^2 are efficient portfolios, the sign of δ_1 is the same as the sign of δ_2 . Hence, $\delta \geq 0$. Therefore, by Proposition 4.2, Z is an efficient portfolio. It follows by induction that for any integer k and numbers λ_i such that $0 \leq \lambda_i \leq 1$, $i = 1, \dots, k$ and $\sum_1^k \lambda_i = 1$, $Z^k \equiv \sum_1^k \lambda_i Z_e^i$ is the return on an efficient portfolio. Hence, Ψ^e is a convex set.

Definition: A market portfolio is defined as a portfolio that holds all available securities in proportion to their market values. To avoid the problems of "double counting" caused by financial intermediaries and

inter-investor issues of securities, the equilibrium market value of a security for this purpose is defined to be the equilibrium value of the aggregate demand by individuals for the security. In models where all physical assets are held by business firms and business firms hold no financial assets, an equivalent definition is that the market value of a security equals the equilibrium value of the aggregate amount of that security issued by business firms. If V_j denotes the market value of security j and V_R denotes the value of the riskless security, then:

$$\delta_j^M = \frac{V_j}{\sum_1^n V_j + V_R}, \quad j = 1, 2, \dots, n,$$

where δ_j^M is the fraction of security j held in a market portfolio.

Theorem 4.7: If Ψ^e is a convex set, and if the securities' market is in equilibrium, then a market portfolio is an efficient portfolio.

Proof: Let there be K risk-averse investors in the economy with the initial wealth of investor k denoted by w_0^k . Define $Z^k \equiv R + \sum_1^n w_j^k (Z_j - R)$ to be the return per dollar on investor k 's optimal portfolio, where w_j^k is the fraction allocated to security j . In equilibrium, $\sum_1^n w_j^k w_0^k = V_j$, $j = 1, 2, \dots, n$, and $\sum_1^K w_0^k \equiv W_0 = \sum_1^n V_j + V_R$. Define $\lambda_k \equiv w_0^k / W_0$, $k = 1, \dots, K$. Clearly, $0 \leq \lambda_k \leq 1$ and $\sum_1^K \lambda_k = 1$. By definition of a market portfolio, $\sum_1^K w_j^k \lambda_k = \delta_j^M$, $j = 1, 2, \dots, n$. Multiplying by $(Z_j - R)$ and summing over j , it

follows that $\sum_{1k}^k \lambda_{1k} \sum_{1j}^n w_j^k (Z_j - R) = \sum_{1k}^K \lambda_{1k} (Z^k - R) = \sum_{1j}^n \delta_j^M (Z_j - R) = Z_M - R$,

where Z_M is defined to be the return per dollar on the market portfolio.

Because $\sum_{1k}^k \lambda_{1k} = 1$, $Z_M = \sum_{1k}^K \lambda_{1k} Z^k$. But every optimal

portfolio is an efficient portfolio. Hence, Z_M is a convex combination of

the returns on K efficient portfolios. Therefore, if Ψ^e is convex,

then the market portfolio is contained in Ψ^e .

Because a market portfolio can be constructed without the knowledge of preferences, the distribution of wealth, or the joint probability distribution for the outstanding securities, models in which the market portfolio can be shown to be efficient are more likely to produce testable hypotheses. In addition, the efficiency of the market portfolio provides a rigorous microeconomic justification for the use of a "representative man" in aggregated economic models, i.e., if the market portfolio is efficient, then there exists a concave utility function such that maximization of its expected value with initial wealth equal to national wealth, would lead to the market portfolio as the optimal portfolio. Moreover, it is currently fashionable in the real world to advise "passive" investment strategies that simply mix the market portfolio with the riskless security. Provided that the market portfolio is efficient, by Proposition 4.2 no investor following such strategies could ever be convicted of "inefficiency." Unfortunately, necessary and sufficient conditions for the market portfolio to be efficient have not as yet been derived.

However, even if the market portfolio were not efficient, it does have the following important property:

Proposition 4.6: In all portfolio models with homogeneous beliefs and risk-averse investors, the equilibrium expected return on the market portfolio exceeds the return on the riskless security.

The proof follows directly from the proof of Theorem 4.7 and Corollary 2.1a. Clearly, $\bar{z}_M - R = \sum_1^K \lambda_k (\bar{z}^k - R)$. By Corollary 2.1a, $\bar{z}^k \geq R$ for $k = 1, \dots, K$, with strict inequality holding if z^k is risky. But, $\lambda_k > 0$. Hence, $\bar{z}_M > R$, if any risky securities are held by any investor. Note that using no information other than market prices and quantities of securities outstanding, the market portfolio (and combinations of the market portfolio and the riskless security) is the only risky portfolio where the sign of its equilibrium expected excess return can always be predicted.

Returning to the special case where ψ^e is spanned by a single risky portfolio and the riskless security, it follows immediately from Proposition 4.5 and Theorem 4.7 that the market portfolio is efficient. Because all efficient portfolios are perfectly positively correlated, it follows that the risky spanning portfolio can always be chosen to be the market portfolio (i.e., $X = Z_M$). Therefore, every efficient portfolio (and hence, every optimal portfolio) can be represented as a simple portfolio combination of the market portfolio and the riskless security with a positive fraction allocated to the market portfolio. If all investors want to hold risky securities in the same relative proportions, then the only way in which this is possible is if these relative proportions are identical to those in the market portfolio. Indeed, if there were one best investment strategy, and if this "best" strategy were widely known, then whatever the original statement of the

strategy, it must lead to simply this imperative: "hold the market portfolio."

Because for every security $\delta_j^M > 0$, it follows from Proposition 4.4, that in equilibrium, every investor will hold non-negative quantities of risky securities, and therefore, it is never optimal to short-sell risky securities. Hence, in models where $m = 1$, the introduction of restrictions against short-sales will not affect the equilibrium.

Theorem 4.8: If (Z_M, R) span Ψ^e , then the equilibrium expected return on security j , can be written as:

$$\bar{Z}_j = R + \beta_j (\bar{Z}_M - R) \quad ,$$

where

$$\beta_j \equiv \frac{\text{cov}[Z_j, Z_M]}{\text{var}(Z_M)} \quad .$$

The proof follows directly from Corollary 4.6 and Proposition 4.3. This relation, called the Security Market Line, was first derived by Sharpe (1964) as a necessary condition for equilibrium in the mean-variance model of Markowitz and Tobin when investors have homogenous beliefs. This relation has been central to most empirical studies of securities' returns published during the last two decades. Indeed, the switch in notation from a_{ij} to β_j in this special case reflects the almost universal adoption of the term, "the 'beta' of a security" to mean the covariance of that security's return with the market portfolio divided by the variance of the return on the market portfolio.

In the special case of Theorem 4.8, β_j measures the systematic risk of security j relative to the efficient portfolio Z_M (i.e., $\beta_j = b_j^M$

as defined in Section 3), and therefore beta provides a complete ordering of the risk of individual securities. As is often the case in research, useful concepts are derived in a special model first. The term "systematic risk" was first coined by Sharpe and was measured by beta. The definition in Section 3 is a natural generalization. Moreover, unlike the general risk measure of Section 3, β_j can be computed from a simple covariance between Z_j and Z_M . Securities whose returns are positively correlated with the market are pro-cyclical, and will be priced to have positive equilibrium expected excess returns. Securities whose returns are negatively correlated are counter-cyclical, and will have negative equilibrium expected excess returns.

In general, the sign of b_j^k cannot be determined by the sign of the correlation coefficient between Z_j and Z_e^k . However, as shown in Theorems 3.2-3.4, because $\partial Y(Z_e^k)/\partial Z_e^k > 0$ for each realization of Z_e^k , $b_j^k > 0$ does imply a generalized positive "association" between the return on Z_j and Z_e^k . Similarly, $b_j^k < 0$ implies a negative "association."

Let Ψ_{\min} denote the set of portfolios contained in Ψ^f such that there exists no other portfolio in Ψ^f with the same expected return and a smaller variance. Let $Z(\mu)$ denote the return on a portfolio contained in Ψ_{\min} such that $Z(\mu) = \mu$, and let δ_j^μ denote the fraction of this portfolio allocated to security j , $j = 1, \dots, n$.

Theorem 4.9: If (Z_1, \dots, Z_n) contain no redundant securities, then (a)

for each value μ , δ_j^μ , $j = 1, \dots, n$ are unique; (b) there exists a portfolio contained in Ψ_{\min} with return X such that (X, R) span Ψ_{\min} ;

(c) $\bar{Z}_j - R = a_j(\bar{X} - R)$, where $a_j \equiv \text{cov}(Z_j, X)/\text{var}(X)$, $j = 1, 2, \dots, n$.

Proof: Let σ_{ij} denote the i - j th element of Ω and because (Z_1, \dots, Z_n) contain no redundant securities, Ω is non-singular. Hence, let v_{ij} denote the i - j th element of Ω^{-1} . All portfolios in Ψ_{\min} with expected return μ must have portfolio weights that are solutions to the problem: minimize $\sum_{i=1}^n \sum_{j=1}^n \delta_i \delta_j \sigma_{ij}$

subject to the constraint $\bar{Z}(\mu) = \mu$. Trivially, if $\mu = R$, then $Z(R) = R$ and $\delta_j^R = 0$, $j = 1, 2, \dots, n$. Consider the case when $\mu \neq R$.

The n first-order conditions are:

$$0 = \sum_{j=1}^n \delta_j^\mu \sigma_{ij} - \lambda_\mu (\bar{Z}_i - R), \quad i = 1, 2, \dots, n,$$

where λ_μ is the Lagrange multiplier for the constraint. Multiplying by δ_i^μ and summing, we have that $\lambda_\mu = \text{var}[Z(\mu)]/(\mu - R)$.

By definition of Ψ_{\min} , λ_μ must be the same for all $Z(\mu)$. Because Ω is non-singular, the set of linear equations has the unique solution:

$$\delta_j^\mu = \lambda_\mu \sum_{i=1}^n v_{ij} (\bar{Z}_i - R), \quad i = 1, 2, \dots, n,$$

This proves (a). From this solution, $\delta_j^\mu / \delta_k^\mu, j, k = 1, 2, \dots, n$, are the same for every value of μ . Hence, all portfolios in Ψ_{\min} with $\mu \neq R$ are perfectly correlated. Hence, pick any portfolio in Ψ_{\min} with $\mu \neq R$ and call its return X . Then every $Z(\mu)$ can be written in the

form $Z(\mu) = \delta_\mu(X - R) + R$. Hence, (X, R) span Ψ_{\min} which proves (b), and from Corollary 4.6 and Proposition 4.3, (c) follows directly.

From Theorem 4.9, a_k will be equivalent to b_k^K as a measure of a security's systematic risk provided that the $Z(\mu)$ chosen for X is such that $\mu > R$. Like β_k , the only information required to compute a_k are the joint second moments of Z_k , and X . Which of the two equivalent measures will be more useful obviously depends upon the information set that is available. However, as the following theorem demonstrates, the a_k measure is the natural choice in the case when there exists a spanning set for Ψ^e with $m = 1$.

Theorem 4.10: If (X, R) span Ψ^e and if X has a finite variance, then Ψ^e is contained in Ψ_{\min} .

Proof: Let Z_e be the return on any efficient portfolio. By hypothesis, Z_e can be written as $Z_e = R + a_e(X - R)$. Let Z_p be the return on any portfolio in Ψ^f such that $\bar{Z}_e = \bar{Z}_p$. By Corollary 4.6, Z_p can be written as $Z_p = R + a_p(X - R) + \epsilon_p$, where $E(\epsilon_p) = E(\epsilon_p | X) = 0$. Therefore, $a_p = a_e$ if $\bar{Z}_p = \bar{Z}_e$; $\text{var}(Z_p) = a_p^2 \text{var}(X) + \text{var}(\epsilon_p) > a_p^2 \text{var}(X) = \text{var}(Z_e)$. Hence, Z_e is contained in Ψ_{\min} . Moreover, Ψ^e will be the set of all portfolios in Ψ_{\min} such that $\mu \geq R$.

Thus, whenever there exists a spanning set for Ψ^e with $m = 1$, the means, variances, and covariances of (Z_1, \dots, Z_n) are sufficient statistics to completely determine all efficient portfolios. Such a strong set of conclusions suggests that the class of joint probability distributions for (Z_1, \dots, Z_n) which admit a two-fund separation theorem will be highly specialized. However, as the following theorems demonstrate, the class is not empty.

Theorem 4.11: If (Z_1, \dots, Z_n) have a joint normal probability distribution, then there exists a portfolio with return X such that (X, R) span Ψ^e .

Proof: Using the procedure applied in the proof of Theorem 4.9, construct a risky portfolio contained in Ψ_{\min} , and call its return X . Define the random variables, $\epsilon_k \equiv Z_k - R - a_k(X - R)$, $k = 1, \dots, n$. By part (c) of that theorem, $E(\epsilon_k) = 0$, and by construction, $\text{cov}[\epsilon_k, X] = 0$. Because Z_1, \dots, Z_n are normally distributed, X will be normally distributed. Hence, ϵ_k is normally distributed, and because $\text{cov}[\epsilon_k, X] = 0$, ϵ_k and X are independent. Therefore, $E(\epsilon_k) = E(\epsilon_k | X) = 0$. From Corollary 4.6, it follows that (X, R) span Ψ^e .

It is straightforward to prove that if (Z_1, \dots, Z_n) can have arbitrary means, variances and covariances, and can be mutually independent, then a necessary condition for there to exist a portfolio with return X such that (X, R) span Ψ^e is that (Z_1, \dots, Z_n) be joint normally

distributed. However, it is important to emphasize both the word "arbitrary" and the prospect for independence. For example, consider a joint distribution for (Z_1, \dots, Z_n) such that the joint probability density function, $p(Z_1, \dots, Z_n)$ is a symmetric function. That is, for each set of admissible outcomes for (Z_1, \dots, Z_n) , $p(Z_1, \dots, Z_n)$ remains unchanged when any two arguments of p are interchanged. An obvious special case is when (Z_1, \dots, Z_n) are independently and identically distributed and $p(Z_1, \dots, Z_n) = p(Z_1)p(Z_2) \dots p(Z_n)$.

Theorem 4.12: If $p(Z_1, \dots, Z_n)$ is a symmetric function with respect to all its arguments, then there exists a portfolio with return X such that (X, R) spans Ψ^e .

Prcof: By hypothesis, $p(Z_1, \dots, Z_1, \dots, Z_n) = p(Z_1, \dots, Z_1, \dots, Z_n)$ for each set of given values (Z_1, \dots, Z_n) . Therefore, from the first-order conditions for portfolio selection, (2.4), every risk-averse investor will choose $\delta_1^* = \delta_i^*$. But, this is true for $i = 1, \dots, n$. Hence, all investors will hold all risky securities in the same relative proportions. Therefore, if X is the return on a portfolio with an equal dollar investment in each risky security, then (X, R) will span Ψ^e .

Samuelson (1967) was the first to examine this class of symmetric density functions in a portfolio context. Chamberlain (1983) has shown that the class of spherically-symmetric distributions characterize the distributions that imply mean-variance utility functions for all risk-averse expected utility maximizers. However, for distributions other than Gaussian to obtain, the

security returns cannot be independently distributed.

The Arbitrage Pricing Theory (APT) model developed by Ross (1976a) provides an important class of linear-factor models that generate (at least approximate) spanning without assuming joint normal probability distributions. Suppose the returns on securities are generated by:

$$Z_j = \bar{Z}_j + \sum_1^m a_{ij} Y_i + \epsilon_j, \quad j = 1, \dots, n, \quad (4.1)$$

where $E(\epsilon_j) = E(\epsilon_j | Y_1, \dots, Y_m) = 0$ and without loss of generality, $E(Y_i) = 0$ and $\text{cov}[Y_i, Y_j] = 0$, $i \neq j$. The random variables $\{Y_i\}$ represent common factors that are likely to affect the returns on a significant number of securities. If it is possible to construct a set of m portfolios with returns (X_1, \dots, X_m) such that X_i and Y_i are perfectly correlated, $i = 1, 2, \dots, m$, then the conditions of Theorem 4.5 will be satisfied and (X_1, \dots, X_m, R) will span Ψ^e .

Although in general, it will not be possible to construct such a set, by imposing some mild additional restrictions on $\{\epsilon_j\}$, Ross (1976a) derives an asymptotic spanning theorem as the number of available securities, n , becomes large. While the rigorous derivation is rather tedious, a rough description goes as follows: let Z_p be the return on a portfolio with fraction δ_j allocated to security $j, j = 1, 2, \dots, n$. From (4.1), Z_p can be written as:

$$Z_p = \bar{Z}_p + \sum_1^m a_{ip} Y_i + \epsilon_p, \quad (4.2)$$

where $\bar{Z}_p = R + \sum_1^n \delta_j (\bar{Z}_j - R)$; $a_{ip} \equiv \sum_1^n \delta_j a_{ij}$; $\epsilon_p \equiv \sum_1^n \delta_j \epsilon_j$. Consider the set of portfolios (called well-diversified portfolios) that have the property $\delta_j \equiv \mu_j/n$, where $|\mu_j| \leq M_j < \infty$ and M_j is independent of $n, j = 1, \dots, n$. Virtually by the definition of a common factor, it is reasonable to assume that for every $n \gg m$, a significantly positive fraction of all securities, λ_i , have $a_{ij} \neq 0$, and this will be true for each common factor $i, i = 1, \dots, m$. Similarly, because the $\{\epsilon_j\}$ denote the variations in securities' returns not explained by common factors, it is also reasonable to assume for large n that for each j , ϵ_j is uncorrelated with virtually all other securities' returns. Hence, if the number of common factors, m , is fixed, then for all $n \gg m$, it should be possible to construct a set of well-diversified portfolios $\{X_k\}$ such that for X_k , $a_{ik} = 0$, $i = 1, \dots, m, i \neq k$ and $a_{kk} \equiv 0$. It follows from (4.2), that X_k can be written as:

$$X_k = \bar{X}_k + a_{kk} Y_k + \frac{1}{n} \sum_1^n \mu_j^k \epsilon_j, \quad k = 1, \dots, m.$$

But $|\mu_j^k|$ is bounded, independently of n , and virtually all the ϵ_j are uncorrelated. Therefore, by the Law of Large Numbers, as $n \rightarrow \infty$, $X_k \rightarrow \bar{X}_k + a_{kk} Y_k$. So, as n becomes very large, X_k and Y_k become perfectly correlated, and by Theorem 4.5, asymptotically (X_1, \dots, X_m, R) will span Ψ^e . In particular, if $m = 1$, then asymptotically two-fund separation will obtain independent of any other distributional characteristics of Y_1 or the $\{\epsilon_j\}$.

As can be seen from Theorem 2.3 and its corollary, all efficient portfolios in the APT model are well-diversified portfolios. Unlike in the

mean-variance model, returns on all efficient portfolios need not, however, be perfectly correlated. The model is also attractive because at least in principle, the equilibrium structure of expected returns and risks of securities can be derived without explicit knowledge of investors' preferences or endowments. Indeed, whenever non-trivial spanning of Ψ^e obtains and the set of risky spanning portfolios can be identified, much of the structure of individual securities returns can be empirically estimated. For example, if we know of a set of portfolios $\{X_i\}$ such that $E(\epsilon_j | X_1, \dots, X_m) = 0$, $j = 1, \dots, n$, then by Theorem 4.5, (X_1, \dots, X_m, R) span Ψ^e . By Proposition 4.3, ordinary-least-squares regression of the realized excess returns on security j , $Z_j - R$, on the realized excess returns of the spanning portfolios, $(X_1 - R, \dots, X_m - R)$, will always give unbiased estimates of the a_{ij} . Of course, to apply time-series estimation, it must be assumed that the spanning portfolios (X_1, \dots, X_m) and $\{a_{ij}\}$ are intertemporally stable. For these estimators to be efficient, further restrictions on the $\{\epsilon_j\}$ are required to satisfy the Gauss-Markov Theorem.

Early empirical studies of stock market securities' returns rarely found more than two or three statistically-significant common factors.¹³ Given that there are tens of thousands of different corporate liabilities traded in U.S. securities markets, there appears to be empirical foundation for the assumptions of the APT model. More-recent studies have, however, concluded that the number of common factors may be considerably larger, and some have raised serious questions about the prospect for identifying the factors by using stock-return data alone.¹⁴

Although the analyses derived here have been expressed in terms of

restrictions on the joint distribution of security returns without explicitly mentioning security prices, it is obvious that these derived restrictions impose restrictions on prices through the identity that $Z_j \equiv V_j/V_{j0}$, where V_j is the random variable, end-of-period aggregate value of security j and V_{j0} is its initial value. Hence, given the characteristics of any two of these variables, the characteristics of the third are uniquely determined. For the study of equilibrium pricing, the usual format is to derive equilibrium V_{j0} given the distribution of V_j .

Theorem 4.13: If (X_1, \dots, X_m) denote a set of linearly independent portfolios that satisfy the hypothesis of Theorem 4.5, and all securities have finite variances, then a necessary condition for equilibrium in the securities' market is that:

$$V_{j0} = \frac{V_j - \sum_{k=1}^m \sum_{l=1}^m v_{lk} \text{cov}[X_k, V_j](X_l - R)}{R}, \quad j = 1, \dots, n \quad (4.3)$$

where v_{ik} is the i - k th element of Ω_X^{-1} .

Proof: By linear independence, Ω_X is non-singular. From the identity

$$V_j \equiv Z_j V_{j0} \text{ and Theorem 4.5, } V_j = V_{j0}[R + \sum_{i=1}^m a_{ij}(X_i - R) + \epsilon_j],$$

where $E(\epsilon_j | X_1, \dots, X_m) = E(\epsilon_j) = 0$. Taking expectations

$$\bar{V}_j = V_{j0}[R + \sum_{i=1}^m a_{ij}(\bar{X}_i - R)]. \text{ Noting that } \text{cov}[X_k, V_j] = V_{j0} \text{cov}[X_k, Z_j],$$

we have from Proposition 4.3, that $V_{j0} a_{ij} = \sum_{k=1}^m v_{ik} \text{cov}[X_k, V_j]$.

By substituting for a_{ij} in the \bar{V}_j expression and rearranging terms, the theorem is proved.

Hence, from Theorem 4.13, a sufficient set of information to determine the equilibrium value of security j is the first and second moments for the joint distribution of (X_1, \dots, X_m, V_j) . Moreover, the valuation formula has the following important "linearity" properties:

Corollary 4.13a: If the hypothesized conditions of Theorem 4.13 hold and if the end-of-period value of a security is given by $V = \sum_1^n \lambda_j V_j$, then in equilibrium:

$$V_0 = \sum_1^n \lambda_j V_{j0} \quad .$$

The proof of the corollary follows by substitution for V in formula (4.3). This property of formula (4.3) is called "value-additivity."

Corollary 4.13b: If the hypothesized conditions of Theorem 4.13 hold and if the end-of-period value of a security is given by $V = qV_j + u$, where $E(u) = E(u|X_1, \dots, X_m) = \bar{u}$ and $E(q) = E(q|X_1, \dots, X_m, V_j) = \bar{q}$, then in equilibrium:

$$V_0 = \bar{q}V_{j0} + \bar{u}/R \quad .$$

The proof follows by substitution for V in formula (4.3) and by applying the hypothesized conditional-expectation conditions to show that $\text{Cov}[X_k, V] = \bar{q}\text{Cov}[X_k, V_j]$. Hence, to value two securities whose end-of-period values

differ only by multiplicative or additive "noise," we can simply substitute the expected values of the noise terms.

As discussed in Merton (1982a, pp. 642-651), Theorem 4.13 and its corollaries are central to the theory of optimal investment decisions by business firms. To finance new investments, the firm can use internally available funds, issue common stock or issue other types of financial claims (e.g., debt, preferred stock, and convertible bonds). The selection from the menu of these financial instruments is called the firm's financing decision. Although, the optimal investment and financing decisions by a firm generally require simultaneous determination, under certain conditions, the optimal investment decision can be made independently of the method of financing.

Consider firm j with random variable end-of-period value V^j and q different financial claims. The k th such financial claim is defined by the function $f_k[V^j]$, which describes how the holders of this security will share in the end-of-period value of the firm. The production technology and choice of investment intensity, $V_j(I_j; \theta_j)$ and I_j , are taken as given where θ_j is a random variable. If it is assumed that the end-of-period value of the firm is independent of its choice of financial liabilities,¹⁵ then $V^j = V_j(I_j; \theta_j)$, and $\sum_{k=1}^q f_k \equiv V_j(I_j; \theta_j)$ for every outcome θ_j .

Suppose that if firm j were all-equity-financed, there exists an equilibrium such that the initial value of firm j is given by $V_{j0}(I_j)$.

Theorem 4.14: If firm j is financed by q different claims defined by the functions $f_k(V^j)$, $k = 1, \dots, q$, and if there exists an equilibrium such that the return distributions of the efficient portfolio set remains unchanged

from the equilibrium in which firm j was all-equity-financed, then:

$$\sum_1^q f_{k0} = V_{j0}(I_j) \quad ,$$

where f_{k0} is the equilibrium initial value of financial claim k .

Proof: In the equilibrium in which firm j is all equity-financed, the end-of-period random variable value of firm j is $V_j(I_j; \theta_j)$ and the initial value, $V_{j0}(I_j)$, is given by formula (4.3) where (X_1, \dots, X_m, R) span the efficient set. Consider now that firm j is financed by the q different claims. The random variable end-of-period value of firm j , $\sum_1^q f_k$, is still given by $V_j(I_j; \theta_j)$. By hypothesis, there exists an equilibrium such that the distribution of the efficient portfolio set remains unchanged, and therefore, the distribution of (X_1, \dots, X_m, R) remains unchanged. By inspection of formula (4.3), the initial value of firm j will remain unchanged, and therefore $\sum_1^q f_{k0} = V_{j0}(I_j)$.

Hence, for a given investment policy, the way in which the firm finances its investment will not affect the market value of the firm unless the choice of financial instruments changes the return distributions of the efficient portfolio set. Theorem 4.14 is representative of a class of theorems that describe the impact of financing policy on the market value of a firm when the investment decision is held fixed, and this class is generally referred to as the Modigliani-Miller Hypothesis, after the pioneering work in this direction by Modigliani and Miller.¹⁶

Clearly, a sufficient condition for Theorem 4.14 to obtain is that each of the financial claims issued by the firm are "redundant securities" whose

payoffs can be replicated by combining already-existing securities. This condition is satisfied by the subclass of corporate liabilities that provide for linear sharing rules (i.e., $f_k(V) = a_k V + b_k$ where $\sum_1^q a_k = 1$ and $\sum_1^q b_k = 0$). Unfortunately, as will be shown in Section 8, most common types of financial instruments issued by corporations have nonlinear payoff structures. As Stiglitz (1969,1974) has shown for the Arrow-Debreu and Capital Asset Pricing Models, linearity of the sharing rules is not a necessary condition for Theorem 4.14 to obtain. Nevertheless, the existence of nonlinear payoff structures among such a wide class of securities makes the establishment of conditions under which the hypothesis of Theorem 4.14 is valid no small matter.

Beyond the issue of whether firms can optimally separate their investment and financing decisions, the fact that many securities have nonlinear sharing rules raises serious questions about the robustness of spanning models. As already discussed, the APT model, for example, has attracted much interest because it makes no explicit assumptions about preferences and places seemingly few restrictions on the joint probability distribution of security returns. In the APT model, (X_1, \dots, X_m, R) span the set of optimal portfolios and there exist m numbers (a_{1k}, \dots, a_{mk}) for each security $k, k = 1, \dots, n$, such that $Z_k = \sum_1^m a_{ik} (X_i - R) + R + \epsilon_k$ where $E(\epsilon_k) = E(\epsilon_k | X_1, \dots, X_m) = 0$.

Suppose that security k satisfies this condition and security q has a payoff structure that is given by $Z_q = f(Z_k)$, where f is a nonlinear function. If security q is to satisfy this condition, then there must exist numbers (a_{1q}, \dots, a_{mq}) so that for all possible values of (X_1, \dots, X_m) , $E[f(\sum_1^m a_{ik} (X_i - r) + R + \epsilon_k) | X_1, \dots, X_m] = \sum_1^m a_{iq} (X_i - R) + R$. However, unless

$\epsilon_k \equiv 0$ and $\epsilon_q \equiv 0$, such a set of numbers cannot be found for a general nonlinear function f .

Since the APT model only has practical relevance if for most securities, $\text{Var}(\epsilon_k) > 0$, it appears that the reconciliation of nontrivial spanning models with the wide-spread existence of securities with nonlinear payoff structures requires further restrictions on either the probability distributions of securities returns or investor preferences. How restrictive these conditions are cannot be answered in the abstract. First, the introduction of general-equilibrium pricing conditions on securities will impose some restrictions on the joint distribution of returns. Second, the discussed benefits to individuals from having a set of spanning mutual funds may induce the creation of financial intermediaries or additional financial securities, that together with pre-existing securities will satisfy the conditions of Theorem 4.6. The intertemporal models of Sections 7-9 will explore these possibilities in detail.

An alternative approach to the development of non-trivial spanning theorems is to derive a class of utility functions for investors such that even with arbitrary joint probability distributions for the available securities, investors within the class can generate their optimal portfolios from the spanning portfolios. Let Ψ^u denote the set of optimal portfolios selected from Ψ^f by investors with strictly concave von Neumann-Morgenstern utility functions $\{U_i\}$. Cass and Stiglitz (1970) have proved the following theorem.

Theorem 4.15: There exists a portfolio with return X such that (X, R) span Ψ^u if and only if $A_i(W) = 1/(a_i + bW) > 0$, where A_i is

the absolute risk aversion function for investor i in ψ^u .¹⁷

The family of utility functions whose absolute risk-aversion functions can be written as $1/(a + bW) > 0$ is called the "HARA" (Hyperbolic Absolute Risk Aversion) family.¹⁸ By appropriate choices for a and b , various members of the family will exhibit increasing, decreasing, or constant absolute and relative risk aversion. Hence, if each investor's utility function could be approximated by some member of the HARA family, then it might appear that this alternative approach would be fruitful. However, it should be emphasized that the b in the statement of Theorem 4.15 does not have a subscript i , and therefore, for separation to obtain, all investors in ψ^u must have virtually the same utility function.¹⁹ Moreover, they must agree on the joint probability distribution for (Z_1, \dots, Z_n) . Hence, the only significant way in which investors can differ is in their endowments of initial wealth.

Cass and Stiglitz (1970) also examine the possibilities for more-general non-trivial spanning (i.e., $1 \leq m < n$) by restricting the class of utility functions and conclude, "...it is the requirement that there be any mutual funds, and not the limitation on the number of mutual funds, which is the restrictive feature of the property of separability." (p. 144) Hence, the Cass and Stiglitz analysis is essentially a negative report on this approach to developing spanning theorems.

In closing this section, two further points should be made. First, although virtually all the spanning theorems require the generally implausible assumption that all investors agree upon the joint probability distribution

for securities, it is not so unreasonable when applied to the theory of financial intermediation and mutual fund management. In a world where the economic concepts of "division of labor" and "comparative advantage" have content, then it is quite reasonable to expect that an efficient allocation of resources would lead to some individuals (the "fund managers") gathering data and actively estimating the joint probability distributions and the rest either buying this information directly or delegating their investment decisions by "agreeing to agree" with the fund managers' estimates. If the distribution of returns is such that non-trivial spanning of Ψ^e does not obtain, then there are no gains to financial intermediation over the direct sale of the distribution estimates. However, if non-trivial spanning does obtain and the number of risky spanning portfolios, m , is small, then a significant reduction in redundant information processing and transactions can be produced by the introduction of mutual funds. If a significant coalition of individuals can agree upon a common source for the estimates and if they know that, based on this source, a group of mutual funds offered spans Ψ^e , then they need only be provided with the joint distribution for these mutual funds to form their optimal portfolios. On the supply side, if the characteristics of a set of spanning portfolios can be identified, then the mutual fund managers will know how to structure the portfolios of the funds they offer.

The second point concerns the riskless security. It has been assumed throughout that there exists a riskless security. Although some of the specifications will change slightly, virtually all the derived theorems can be shown to be valid in the absence of a riskless security.²⁰ However, the existence of a riskless security vastly simplifies many of the proofs.

5. Two Special Models of One-Period Portfolio Selection

The two most-cited models in the literature of portfolio selection are the Time-State Preference Model of Arrow (1953,1964) and Debreu (1959) and the Mean-Variance Model of Markowitz (1959) and Tobin (1958). Because these models have been central to the development of the microeconomic theory of investment, there are already many review and survey articles devoted just to each of these models.²¹ Hence, only a cursory description of each model is presented here, with specific emphasis on how each model fits within the framework of the analyses presented in the other sections. Moreover, while, under appropriate conditions, both models can be interpreted as multiperiod, intertemporal portfolio-selection models, such an interpretation is delayed until later sections.

The structure of the Arrow-Debreu model is described as follows. Consider an economy where all possible configurations for the economy at the end of the period can be described in terms of M possible states of nature. The states are mutually exclusive and exhaustive. It is assumed that there are N risk-averse individuals with initial wealth W_0^k and a von Neumann-Morgenstern utility function $U^k(W)$ for investor $k, k = 1, \dots, N$. Each individual acts on the basis of subjective probabilities for the states of nature denoted by $\Pi_k(\theta)$, $\theta = 1, \dots, M$. While these subjective probabilities can differ across investors, it is assumed for each investor that $0 < \Pi_k(\theta) < 1$, $\theta = 1, \dots, M$. As was assumed in Section 2, there are n risky securities with returns per dollar Z_j and initial market value, V_{j0} , $j = 1, \dots, n$, and the "perfect market" assumptions of that section, Assumptions 1-4, are assumed here as well. Moreover, if state θ obtains, then the return on security j will be $Z_j(\theta)$, and all

investors agree on the functions $\{Z_j(\theta)\}$. Because the set of states is exhaustive, $[Z_j(1), \dots, Z_j(M)]$ describe all the possible outcomes for the returns on security j . In addition, there are available M "pure" securities with the properties that, $i = 1, \dots, M$, one unit (share) of pure security i will be worth \$1 at the end of the period if state i obtains and will be worthless if state i does not obtain. If P_i denotes the price per share of pure security i and if X_i denotes its return per dollar, then for $i = 1, \dots, M$, X_i as a function of the states of nature can be written as $X_i(\theta) = 1/P_i$ if $\theta = i$ and $X_i(\theta) = 0$ if $\theta \neq i$. All investors agree on the functions $\{X_i(\theta)\}$, $i, \theta = 1, \dots, M$.

Let $Z = Z(N_1, \dots, N_M)$ denote the return per dollar on a portfolio of pure securities that holds N_j shares of pure security j , $j = 1, \dots, M$. If $V_0(N_1, \dots, N_M) \equiv \sum_1^M N_j P_j$ denotes the initial value of this portfolio, then the return per dollar on the portfolio, as a function of the states of nature, can be written as $Z(\theta) = N_\theta / V_0$, $\theta = 1, \dots, M$.

Proposition 5.1: There exists a riskless security, and its return per dollar R equals $1/(\sum_1^M P_j)$.

Proof: Consider the pure-security portfolio that holds one share of each pure security ($N_j = 1$, $j = 1, \dots, M$). The return per dollar Z is the same in every state of nature and equals $1/V_0(1, \dots, 1)$. Hence, there exists a riskless security and by Assumption 3, its return R is given by $1/(\sum_1^M P_j)$.

Proposition 5.2: For each security j with return Z_j , there exists a portfolio of pure securities, whose return per dollar exactly replicates Z_j .

Proof: Let $Z^j \equiv Z(Z_j(1), \dots, Z_j(M))$ denote the return on a portfolio of pure securities with $N_\theta = Z_j(\theta)$, $\theta = 1, \dots, M$. It follows that $V_0(Z_j(1), \dots, Z_j(M)) = \sum_{i=1}^M P_i Z_j(i)$ and $Z^j(\theta) = Z_j(\theta)/V_0$, $\theta = 1, \dots, M$. Consider a three-security portfolio with return Z_p where fraction V_0 is invested in Z^j ; fraction -1 is invested in Z_j ; and fraction $1 - V_0 - (-1) = (2 - V_0)$ is invested in the riskless security. The return per dollar on this portfolio as a function of the states of nature can be written as:

$$Z_p(\theta) = (2 - V_0)R + V_0 Z^j(\theta) - Z_j(\theta) = (2 - V_0)R \quad ,$$

which is the same for all states. Hence, Z_p is a riskless security, and by Assumption 3, $Z_p(\theta) = R$. Therefore, $V_0 = 1$, and $Z^j(\theta) = Z_j(\theta)$, $\theta = 1, \dots, M$.

Proposition 5.3: The set of pure securities with returns (X_1, \dots, X_M) span the set of all feasible portfolios that can be constructed from the M pure securities and the n other securities.

The proof follows immediately from Propositions 5.1 and 5.2. Hence, whenever a complete set of pure securities exists or can be constructed from the available securities, then every feasible portfolio can be replicated by a portfolio of pure securities. Models in which such a set of pure securities exists are called complete-markets models in the sense that any additional securities or markets would be redundant. Necessary and sufficient conditions for such a set to be constructed from the available n risky securities alone

and therefore, for markets to be complete, are that: $n \geq M$: a riskless asset can be created and Assumption 3 holds; and the rank of the variance-covariance matrix of returns, Ω , equals $M - 1$.

The connection between the pure securities of the Arrow-Debreu model and the mutual fund theorems of Section 4 is obvious. To put this model in comparable form, we can choose the alternative spanning set (X_1, \dots, X_m, R) where $m = M - 1$. From Theorem 4.3, the returns on the risky securities can be written as:

$$Z_j = R + \sum_{i=1}^M a_{ij}(X_i - R), \quad j = 1, \dots, m, \quad (5.1)$$

where the numbers (a_{ij}) are given by Proposition 4.3.

Note that nowhere in the derivation were the subjective probability assessments of the individual investors required. Hence, individual investors need not agree on the joint distribution for (X_1, \dots, X_m) . However, by Theorem 4.3, investors cannot have arbitrary beliefs in the sense that they must agree on the (a_{ij}) in (5.1).

Proposition 5.4: If $V_j(\theta)$ denotes the end-of-period value of security j , if state θ obtains, then a necessary condition for equilibrium in the securities' market is that:

$$V_{j0} = \sum_{k=1}^M P_k V_j(k), \quad j = 1, \dots, n.$$

The proof follows immediately from the proof of Proposition 5.2. It was shown there that $V_0 = \sum_{k=1}^M P_k Z_j(k) = 1$. Multiplying both sides by V_{j0} and noting the identity $V_j(k) \equiv V_{j0} Z_j(k)$, it follows that $V_{j0} = \sum_{k=1}^M P_k V_j(k)$.

However, by Theorem 4.3 and Proposition 5.3, it follows that the $\{V_{j0}\}$ can also be written as:

$$V_{j0} = \frac{\bar{V}_j - \sum_{l=1}^m \sum_{k=1}^m v_{lk} \text{cov}[X_k, V_j] (\bar{X}_l - R)}{R}, \quad j = 1, \dots, n, \quad (5.2)$$

where v_{ik} is the i - k th element of Ω_X^{-1} . Hence, from (5.2) and Proposition 5.4, it follows that the a_{ij} in (5.1) can be written as:

$$a_{ij} = [Z_j(i) - R] / [1/P_i - R], \quad i = 1, \dots, m; \quad j = 1, \dots, n. \quad (5.3)$$

From (5.3), given the prices of the securities $\{P_i\}$ and $\{V_{i0}\}$, the $\{a_{ij}\}$ will be agreed upon by all investors if and only if they agree upon the $\{V_j(i)\}$ functions.

While it is commonly believed that the Arrow-Debreu model is completely general with respect to assumptions about investors' beliefs, the assumption that all investors agree on the $\{V_j(i)\}$ functions can impose non-trivial restrictions on these beliefs. In particular, when there is production, it will in general be inappropriate to define the states, tautologically, by the end-of-period values of the securities, and therefore, investors will at least have to agree on the technologies specified for each firm.²² However, as discussed in Section 4, it is unlikely that a model without some degree of homogeneity in beliefs (other than agreement on currently-observed variables) can produce testable restrictions. Among models that do produce such testable restrictions, the assumptions about investors' beliefs in the Arrow-Debreu model are among the most general.

Finally, for the purposes of portfolio theory, the Arrow-Debreu model is a special case of the spanning models of Section 4, which serves to illustrate

the generality of the linear structure of those models.

The most elementary type of portfolio selection model in which all securities are not perfect substitutes is one where the attributes of every optimal portfolio can be characterized by two numbers: its "risk" and its "return." The mean-variance portfolio selection model of Markowitz (1959) and Tobin (1958) is such a model. In this model, each investor chooses his optimal portfolio so as to maximize a utility function of the form $H[E(W), \text{var}(W)]$, subject to his budget constraint, where W is his random variable end-of-period wealth. The investor is said to be "risk averse in a mean-variance sense" if $H_1 > 0$; $H_2 < 0$; $H_{11} < 0$; $H_{22} < 0$; and $H_{11}H_{22} - H_{12}^2 > 0$, where subscripts denote partial derivatives.

In an analogous fashion to the general definition of an efficient portfolio in Section 2, a feasible portfolio will be called a mean-variance efficient portfolio if there exists a risk-averse mean-variance utility function such that this feasible portfolio would be preferred to all other feasible portfolios. Let Ψ_{mv}^e denote the set of mean-variance efficient portfolios. As defined in Section 4, Ψ_{\min} is the set of feasible portfolios such that there exists no other portfolio with the same expected return and a smaller variance. For a given initial wealth W_0 , every risk-averse investor would prefer the portfolio with the smallest variance among those portfolios with the same expected return. Hence, Ψ_{mv}^e is contained in Ψ_{\min} .

Proposition 5.5: If (Z_1, \dots, Z_n) are the returns on the available risky securities, then there exists a portfolio contained in Ψ_{mv}^e with return X such that (X, R) span Ψ_{mv}^e and $\bar{Z}_j - R = a_j(\bar{X} - R)$ where $a_j \equiv \text{cov}(Z_j, X) / \text{var}(X)$, $j = 1, 2, \dots, m$.

The proof follows immediately from Theorem 4.9.²³ Hence, all the properties derived in the special case of two-fund spanning ($m = 1$) in Section 4, apply to the mean-variance model. Indeed, because all such investors would prefer a higher expected return for the same variance of return, Ψ_{mv}^e is the set of all portfolios contained in Ψ_{min} , such that their expected returns are equal to or exceed R . Hence, as with the complete-markets model, the mean-variance model is also a special case of the spanning models developed in Section 4.

If investors have homogeneous beliefs, then the equilibrium version of the mean-variance model is called the Capital Asset Pricing Model.²⁴ It follows from Proposition 4.5, and Theorem 4.7 that, in equilibrium, the market portfolio can be chosen as the risky spanning portfolio. From Theorem 4.8, the equilibrium structure of expected returns must satisfy the Security Market Line.

Because of the mean-variance model's attractive simplicity and its strong empirical implications, a number of authors²⁵ have studied the conditions under which such a criterion function is consistent with the expected utility maxim. Like the studies of general spanning properties cited in Section 4, these studies examined the question in two parts. (i) What is the class of probability distributions such that the expected value of an arbitrary concave utility function can be written solely as a function of mean and variance?

(ii) What is the class of strictly-concave von Neumann-Morgenstern utility functions whose expected value can be written solely as a function of mean and variance for arbitrary distributions? Since the class of distributions in (i) was shown in Section 4 to be equivalent to the class of finite-variance distributions that admit two-fund spanning of the efficient set, the analysis will not be repeated here. To answer (ii), it is straightforward to show that a necessary condition is that U have the form, $W - bW^2$, with $b > 0$. This member of the HARA family is called the quadratic, and will satisfy the von Neumann axioms only if $W \leq 1/2b$, for all possible outcomes for W . Even if U is defined to be $\max[W - bW^2, 1/4b]$, so that U satisfies the axioms for all W , its expected value for general distributions can be written as a function of just $E(W)$ and $\text{var}(W)$ only if the maximum possible outcome for W is less than $1/2b$.

Although both the Arrow-Debreu and Markowitz-Tobin models were shown to be special cases of the spanning models in Section 4, they deserve special attention because they are unquestionably the genesis of these general models.

6. Intertemporal Consumption and Portfolio Selection Theory

As in the preceding analyses here, the majority of papers on investment theory under uncertainty have assumed that individuals act so as to maximize the expected utility of end-of-period wealth and that intra-period revisions are not feasible. Therefore, all events which take place after next period are irrelevant to their decisions. Of course, investors do care about events beyond "next period," and they can review and change their allocations periodically. Hence, the one-period, static analyses will only be valid under those conditions such that an intertemporally-maximizing individual acts, each period, as if he were a one-period, expected utility-of-wealth maximizer. In this section, the lifetime consumption-portfolio selection problem is solved, and conditions are derived under which the one-period static portfolio problem will be an appropriate "surrogate" for the dynamic, multi-period portfolio problem.

As in the early contributions by Hakansson (1970), Samuelson (1969), and Merton (1969), the problem of choosing optimal portfolio and consumption rules for an individual who lives T years is formulated as follows. The individual investor chooses his consumption and portfolio allocation for each period so as to maximize:²⁶

$$E_0 \left\{ \sum_0^{T-1} U[C(t),t] + B[W(T),T] \right\} \quad , \quad (6.1)$$

where $C(t)$ is consumption chosen at age t ; $W(t)$ is wealth at age t ; E_t is the conditional expectation operator conditional on knowing all relevant information available as of time t ; the utility function (during life) U is assumed to be strictly concave in C ; and the "bequest" function B is also assumed to be concave in W .

It is assumed that there are n risky securities with random variable returns between time t and $t + 1$ denoted by $Z_1(t + 1), \dots, Z_n(t + 1)$, and there is a riskless security whose return between t and $t + 1$, $R(t)$, will be known with certainty as of time t .²⁷ When the investor "arrives" at date t , he will know the value of his portfolio, $W(t)$. He chooses how much to consume, $C(t)$, and then reallocates the balance of his wealth, $W(t) - C(t)$, among the available securities. Hence, the accumulation equation between t and $t + 1$ can be written as:²⁸

$$W(t+1) = [\sum_1^n w_j [Z_j(t+1) - R(t)] + R(t)] [W(t) - C(t)], \quad (6.2)$$

where $w_j(t)$ is the fraction of his portfolio allocated to security j at date t , $j = 1, \dots, n$. Because the fraction allocated to the riskless security can always be chosen to equal $1 - \sum_1^n w_j(t)$, the choices for $w_1(t), \dots, w_n(t)$ are unconstrained.

It is assumed that there exist m state variables, $\{S_k(t)\}$, such that the stochastic processes for $\{Z_1(t + 1), \dots, Z_n(t + 1), R(t + 1), S_1(t + 1), \dots, S_m(t + 1)\}$ are Markov with respect to $S_1(t), \dots, S_m(t)$, and $S(t)$ denotes the m -vector of state-variable values at time t .²⁹

The method of stochastic dynamic programming is used to derive the optimal consumption and portfolio rules. Define the function $J[W(t), S(t), t]$ by:

$$J[W(t), S(t), t] \equiv \max E_t \left\{ \sum_t^{T-1} U[C(\tau), \tau] + B[W(T), T] \right\} \quad (6.3)$$

J , therefore, is the (utility) value of the balance of the investor's optimal consumption-investment program from date t forward, and, in this context, is called the "derived" utility of wealth function. By the Principle of

Optimality, (6.3) can be rewritten as:

$$J[W(t), S(t), t] = \max\{U[C(t), t] + E_t (J[W(t+1), S(t+1), t+1])\} \quad , \quad (6.4)$$

where "max" is over the current decision variables $[C(t), w_1(t), \dots, w_n(t)]$.

Substituting for $W(t + 1)$ in (6.4) from (6.2) and differentiating with respect to each of the decision variables, we can write the $n + 1$ first-order conditions for a regular interior maximum as:³⁰

$$0 = U_C[C^*(t), t] - E_t \left\{ J_W[W(t+1), S(t+1), t+1] \left(\sum_{j=1}^n w_j^* (Z_j - R) + R \right) \right\} \quad (6.5)$$

and

$$0 = E_t \{ J_W[W(t+1), S(t+1), t+1] (Z_j - R) \}, \quad j = 1, 2, \dots, n \quad , \quad (6.6)$$

where $U_C \equiv \partial U / \partial C$; $J_W \equiv \partial J / \partial W$; and (C^*, w_j^*) are the optimum values for the decision variables. Henceforth, except where needed for clarity, the time indices will be dropped. Using (6.6), (6.5) can be written as:

$$0 = U_C[C^*, t] - RE_t \{ J_W \} \quad . \quad (6.7)$$

To solve for the complete optimal program, one first solves (6.6) and (6.7) for C^* and w^* as functions of $W(t)$ and $S(t)$ when $t = T - 1$. This can be done because $J[W(T), S(T), T] = B[W(T), T]$, a given function. Substituting the solutions for $C^*(T - 1)$ and $w^*(T - 1)$ in the right-hand side of (6.4), (6.4) becomes an equation and therefore, one has $J[W(T - 1), T - 1]$. Using (6.6), (6.7), and (6.4), one can proceed to solve for the optimal rules in earlier periods in the usual "backwards" recursive fashion of dynamic programming. Having done so, one will have a complete schedule of optimal consumption and portfolio rules for each date expressed as functions of the (then) known state variables $W(t)$, $S(t)$, and t . Moreover, as

Samuelson (1969) has shown, the optimal consumption rules will satisfy the "envelope condition" expressed as:

$$J_W[W(t), S(t), t] = U_C[C^*(t), t] \quad , \quad (6.8)$$

i.e., at the optimum, the marginal utility of wealth (future consumption) will just equal the marginal utility of (current) consumption. Moreover, from (6.8), it is straightforward to show that $J_{WW} < 0$ because $U_{CC} < 0$. Hence, J is a strictly concave function of wealth.

A comparison of the first-order conditions for the static portfolio-selection problem, (2.4) in Section 2, with the corresponding conditions (6.6) for the dynamic problem will show that they are formally quite similar. Of course, they do differ in that, for the former case, the utility function of wealth is taken to be exogenous while, in the latter, it is derived. However, the more fundamental difference in terms of derived portfolio-selection behavior is that J is not only a function of W , but also a function of S . The analogous condition in the static case would be that the end-of-period utility function of wealth is also state dependent.

To see that this difference is not trivial, consider the Rothschild-Stiglitz definition of "riskier" that was used in the one-period analysis to partition the feasible portfolio set into its efficient and inefficient parts. Let W_1 and W_2 be the random variable, end-of-period values of two portfolios with identical expected values. If W_2 is equal in distribution to $W_1 + Z$, where $E(Z|W_1) = 0$, then from (2.10) and (2.11), W_2 is riskier than W_1 and every risk-averse maximizer of the expected utility of end-of-period wealth would prefer W_1 to W_2 . However, consider an intertemporal maximizer with a strictly-concave, derived

utility function J . It will not, in general, be true that $E_t\{J[W_1, S(t+1), t+1]\} > E_t\{J[W_2, S(t+1), t+1]\}$. Therefore, although the intertemporal maximizer selects his portfolio for only one period at a time, the optimal portfolio selected may be one that would never be chosen by any risk-averse, one-period maximizer. Hence, the portfolio-selection behavior of an intertemporal maximizer will, in general, be operationally distinguishable from the behavior of a static maximizer.

To adapt the Rothschild-Stiglitz definition to the intertemporal case, a stronger condition is required: namely if W_2 is equal in distribution to $W_1 + Z$, where $E[Z|W_1, S(t+1)] = 0$, then every risk-averse intertemporal maximizer would prefer to hold W_1 rather than W_2 in the period t to $t+1$. The proof follows immediately from the concavity of J and Jensen's Inequality. Namely, $E_t\{J[W_2, S(t+1), t+1]\} = E_t\{E(J[W_2, S(t+1), t+1]|W_1, S(t+1))\}$. By Jensen's Inequality, $E(J[W_2, S(t+1), t+1]|W_1, S(t+1)) < J[E(W_2|W_1, S(t+1)), S(t+1), t+1] = J[W_1, S(t+1), t+1]$, and therefore, $E_t\{J[W_2, S(t+1), t+1]\} < E_t\{J[W_1, S(t+1), t+1]\}$. Hence, "noise" as denoted by Z must not only be noise relative to W_1 , but noise relative to the state variables $S_1(t+1), \dots, S_m(t+1)$. All the analyses of the preceding sections can be formally adapted to the intertemporal framework by simply requiring that the "noise" terms there, ϵ , have the additional property that $E_t(\epsilon|S(t+1)) = E_t(\epsilon) = 0$. Hence, in the absence of further restrictions on the distributions, the resulting efficient portfolio set for intertemporal maximizers will be larger than in the static case.

However, under certain conditions,³¹ the portfolio selection behavior of intertemporal maximizers will be "as if" they were one-period maximizers. For

example, if $E_t[Z_j(t+1)] \equiv \bar{Z}_j(t+1) = E_t[Z_j(t+1)|S(t+1)]$, $j = 1, 2, \dots, n$, then the additional requirement that $E_t(\epsilon|S(t+1)) = 0$ will automatically be satisfied for any feasible portfolio, and the original Rothschild-Stiglitz "static" definition will be valid. Indeed, in the cited papers by Hakansson, Samuelson, and Merton, it is assumed that the security returns $\{Z_1(t), \dots, Z_n(t)\}$ are serially independent and identically distributed in time which clearly satisfies this condition.

Define the investment opportunity set at time t to be the joint distribution for $\{Z_1(t+1), \dots, Z_n(t+1)\}$ and the return on the riskless security, $R(t)$. The Hakansson et al, papers assume that the investment opportunity set is constant through time. The condition $\bar{Z}_j(t+1) = E_t[Z_j(t+1)|S(t+1)]$, $j = 1, \dots, n$, will also be satisfied if changes in the investment opportunity set are either completely random or time dependent in a non-stochastic fashion. Moreover, with the possible exception of a few perverse cases, these are the only conditions on the investment opportunity set under which $\bar{Z}_j(t+1) = E_t[Z_j(t+1)|S(t+1)]$, $j = 1, \dots, n$. Hence, for arbitrary concave utility functions, the one-period analysis will be a valid surrogate for the intertemporal analysis only if changes in the investment opportunity set satisfy these conditions.

Of course, by inspection of (6.6), if J were of the form $V[W(t), t] + H[S(t), t]$ so that $J_W = V_W$ is only a function of wealth and time, then for arbitrary investment opportunity sets such an intertemporal investor will act "as if" he is a one-period maximizer. Unfortunately, the only concave utility function that will produce such a J function and satisfy the

additivity specification in (6.1) is $U[C,t] = a(t)\log[C]$ and $B[W,T] = b(T)\log[W]$, where either $a = 0$ and $b > 0$ or $a > 0$ and $b \geq 0$.

While some have argued that this utility function is of special normative significance,³² any model whose results depend singularly upon all individuals having the same utility function and where, in addition, the utility function must have a specific form, can only be viewed as an example, and not the basis for a general theory.

Hence, in general, the one-period static analysis will not be rich enough to describe the investor behavior in an intertemporal framework. Indeed, without additional assumptions, the only derived restrictions on optimal demand functions and equilibrium security returns are the ones that rule out arbitrage. Hence, to deduce additional properties, further assumptions about the dynamics of the investment opportunity set are needed.

7. Consumption and Portfolio Selection Theory in Continuous-Time Models

There are three time intervals or horizons involved in the consumption-portfolio problem.³³ First, there is the trading horizon, which is the minimum length of time between which successive transactions by economic agents can be made in the market. In a sequence-of-markets analysis, it is the length of time between successive market openings, and is therefore part of the specification of the structure of markets in the economy. While this structure will depend upon the tradeoff between the costs of operating the market and its benefits, this time scale is not determined by the individual investor, and is the same for all investors in the economy. Second, there is the decision horizon, which is the length of time between which the investor makes successive decisions, and it is the minimum time between which he would take any action. For example, an investor with a fixed decision interval of one month, who makes a consumption decision and portfolio allocation today, will under no conditions make any new decisions or take any action prior to one month from now. This time scale is determined by the costs to the individual of processing information and making decisions, and is chosen by the individual. Third, there is the planning horizon, which is the maximum length of time for which the investor gives any weight in his utility function. Typically, this time scale would correspond to the balance of his lifetime and is denoted by T in the formulation (6.1).

The static approach to portfolio selection implicitly assumes that the individual's decision and planning horizons are the same: "one period." While the intertemporal approach distinguishes between the two, when individual demands are aggregated to determine market equilibrium relations, it is implicitly assumed in both approaches that the decision interval is the

same for all investors, and therefore corresponds to the trading interval.

If h denotes the length of time in the trading interval, then every solution derived has, as an implicit argument, h . Clearly, if h changes, then the derived behavior of investors would change, as indeed would any deduced equilibrium relations.³⁴ I might mention, somewhat parenthetically, that empirical researchers almost uniformly neglect to recognize that h is part of a model's specification. For example, in Theorem 4.6, the returns on securities were shown to have a linear relation to the returns on a set of spanning portfolios. However, because the n -period return on a security is the product (and not the sum) of the one-period returns, this linear relation can only obtain for the single time interval, h . If we define a fourth time interval, the observation horizon, to be the length of time between successive observations of the data by the researcher, then the usual empirical practice is to implicitly assume that the decision and trading intervals are equal to the observation interval. This is done whether the observation interval is daily, weekly, monthly, or annually!

If the frictionless-markets assumption (Assumption 1) is extended to include no costs of information processing or operating the markets, then it follows that all investors would prefer to have h as small as physically possible. Indeed, the aforementioned general assumption that all investors have the same decision interval will, in general, only be valid if all such costs are zero. This said, it is natural to examine the limiting case when h tends to zero and trading takes place continuously in time.

Consider an economy where the trading interval, h , is sufficiently small that the state description of the economy can change only "locally" during the interval $(t, t + h)$. Formally, the Markov stochastic processes for

the state variables, $S(t)$ are assumed to satisfy the property that one-step transitions are permitted only to the nearest neighboring states. The analogous condition in the limiting case of continuous time is that the sample paths for $S(t)$ are continuous functions of time, i.e., for every realization of $S(t+h)$ except possibly on a set of measure zero, $\lim_{h \rightarrow 0} [S_k(t+h) - S_k(t)] = 0$, $k = 1, \dots, m$. If, however, in the continuous limit, the uncertainty of "end-of-period" returns is to be preserved, then an additional requirement is that $\lim_{h \rightarrow 0} [S_k(t+h) - S_k(t)]/h$ exists almost nowhere, i.e., even though the sample paths are continuous, the increments to the states are not, and therefore, in particular, "end-of-period" rates of return will not be "predictable" even in the continuous time limit. The class of stochastic processes that satisfy these conditions are called diffusion processes.³⁵

Although such processes are almost nowhere differentiable in the usual sense, under some mild regularity conditions, there is a generalized theory of stochastic differential equations which allows their instantaneous dynamics to be expressed as the solution to the system of equations:³⁶

$$dS_i(t) = G_i(S,t)dt + H_i(S,t)dq_i(t), \quad i = 1, \dots, m, \quad (7.1)$$

where $G_i(S,t)$ is the instantaneous expected change in $S_i(t)$ per unit time at time t ; H_i^2 is the instantaneous variance of the change in $S_i(t)$, where it is understood that these statistics are conditional on $S(t) = S$. The $dq_i(t)$ are Weiner processes with the instantaneous correlation coefficient per unit of time between $dq_i(t)$ and $dq_j(t)$ given by the function $\eta_{ij}(S,t)$, $i, j = 1, \dots, m$.³⁷ Moreover, specifying the functions $\{G_i, H_i, \eta_{ij}\}$, $i, j = 1, \dots, m$ is sufficient to completely determine the transition probabilities for $S(t)$ between any two dates.³⁸

Under the assumption that the returns on securities can be described by diffusion processes, Merton (1969,1971) has solved the continuous-time analog to the discrete-time formulation in (6.1), namely:

$$\max E_0 \left\{ \int_0^T U[C(t),t]dt + B[W(T),T] \right\} \quad (7.1)$$

Adapting the notation in the 1971 paper,³⁹ the rate of return dynamics on security j can be written as:

$$dP_j/P_j = \alpha_j(S,t)dt + \sigma_j(S,t)dZ_j, \quad j = 1, \dots, n \quad (7.3)$$

where α_j is the instantaneous conditional expected rate of return per unit time; σ_j^2 is its instantaneous conditional variance per unit time; and dZ_j are Weiner processes, with the instantaneous correlation coefficient per unit time between $dZ_j(t)$ and $dZ_k(t)$ given by the function $\rho_{jk}(S,t), j,k = 1, \dots, n$. In addition to the n risky securities, there is a riskless security whose instantaneous rate of return per unit time is the interest rate $r(t)$.⁴⁰ To complete the model's dynamics description, define the functions $\mu_{ij}(S,t)$ to be the instantaneous correlation coefficients per unit time between $dq_i(t)$ and $dZ_j(t)$, $i = 1, \dots, m; j = 1, \dots, n$.⁴¹

If J is defined by:

$$J[W(t),S(t),t] \equiv \max E_t \left\{ \int_t^T U[C(\tau),\tau]d\tau + B[W(T),T] \right\} \quad (7.4)$$

then the continuous-time analog to (6.4) can be written as:⁴²

$$0 = \max \left\{ U[C,t] + J_t + J_W \left[\left(\sum_1^n w_j (\alpha_j - r) + r \right) W - C \right] + \sum_1^m J_i G_i \right. \quad (7.5)$$

$$\left. + \frac{1}{2} J_{WW} \sum_1^n \sum_1^n w_i w_j \sigma_{ij} W^2 + \frac{1}{2} \sum_1^m \sum_1^m J_{ij} H_i H_j \eta_{ij} + \sum_1^m \sum_1^n J_{iW} w_j \sigma_{ij} H_i \mu_{ij} W \right\} .$$

where the subscripts t , W , and i on J denote partial derivatives with respect to the arguments, t , W , and S_i , ($i = 1, \dots, m$) of J , respectively, and $\sigma_{ij} \equiv \sigma_i \sigma_j \rho_{ij}$ is the instantaneous covariance of the returns of security i with security j , $i, j = 1, \dots, n$. As was the case in (6.4), the "max" in (7.5) is over the current decision variables $[C(t), w_1(t), \dots, w_n(t)]$. If C^* and w^* are the optimum rules, then the $(n + 1)$ first-order conditions for (7.5) can be written as:

$$0 = U_C[C^*, t] - J_W[W, S, t] \quad (7.6)$$

and

$$0 = J_W(\alpha_j - r) + J_{WW} \sum_{i=1}^n w_i^* \sigma_{ij}^W + \sum_{i=1}^m J_{iW} \sigma_j^H \mu_{ij}, \quad j = 1, \dots, n \quad (7.7)$$

Eq. (7.6) is identical to the "envelope condition," (6.8), in the discrete-time analysis. However, unlike (6.6) in the discrete-time analysis, (7.7) is a system of equations which is linear in the optimal demands for risky securities. Hence, if none of the risky securities is redundant, then (7.7) can be solved explicitly for the optimal demand functions using standard matrix inversion, i.e.:

$$w_j^*(t)W(t) = K \sum_{k=1}^n v_{kj}(\alpha_k - r) + \sum_{i=1}^m B_i \zeta_{ij}, \quad j = 1, \dots, n, \quad (7.8)$$

where v_{kj} is the k - j th element of the inverse of the instantaneous variance-covariance matrix of returns $[\sigma_{ij}]$;

$$\zeta_{ij} \equiv \sum_{k=1}^n v_{kj} \sigma_k^H \mu_{ik}; \quad K \equiv -J_W/J_{WW}; \quad \text{and} \quad B_i \equiv -J_{iW}/J_{WW}, \quad i = 1, \dots, m.$$

As an immediate consequence of (7.8), we have the following mutual fund theorem:

Theorem 7.1: If the returns dynamics are described by (7.1) and (7.3), then there exist $(m + 2)$ mutual funds constructed from linear combinations of the available securities such that, independent of preferences, wealth distribution, or planning horizon, individuals will be indifferent between choosing from linear combinations of just these $(m + 2)$ funds or linear combinations of all n risky securities and the riskless security.

Proof: Let mutual fund #1 be the riskless security; let mutual fund #2 hold fraction, $\delta_j \equiv \sum_1^n v_{kj}(\alpha_k - r)$, in security j , $j = 1, \dots, n$, and the balance $(1 - \sum_1^n \delta_j)$ in the riskless security; let mutual fund $\#(2 + i)$ hold fraction $(\delta_j^i \equiv \zeta_{ij})$ in security j , $j = 1, \dots, n$ and the balance $(1 - \sum_1^n \delta_j^i)$ in the riskless security for $i = 1, \dots, m$. Consider a portfolio of these mutual funds which allocates $d_2(t) = K$ dollars to fund #2; $d_{2+i}(t) = B_i$ dollars to fund $\#(2 + i)$, $i = 1, \dots, m$; and $d_1(t) = W(t) - \sum_2^{2+m} d_i(t)$ dollars to fund #1. By inspection of (7.8), this portfolio of funds exactly replicates the optimal portfolio holdings chosen from among the original n risky securities and the riskless security. However, the fractional holdings of these securities by the $(m + 2)$ funds do not depend upon the preferences, wealth, or planning horizon of the individuals investing in the funds. Hence, every investor can replicate his optimal portfolio by investing in the $(m + 2)$ funds.

Of course, as with the mutual fund theorems of Section 4, Theorem 7.1 is vacuous if $m \geq n + 1$. However, for $m \ll n$, the $(m + 2)$ portfolios provide for a non-trivial spanning of the efficient portfolio set, and it is straightforward to show that the instantaneous returns on individual securities will satisfy the same linear specification relative to these spanning portfolios as was derived in Theorem 4.6 for the one-period analysis.

It was shown in the discrete-time analysis of Section 6 that if $E_t[Z_j(t + 1)|S(t)] = \bar{Z}_j(t + 1)$, $j = 1, \dots, n$, then the intertemporal maximizer's demand behavior is "as if" he were a static maximizer of the expected utility of end-of-period wealth. The corresponding condition in the continuous-time case is that the instantaneous rates of return on all available securities are uncorrelated with the unanticipated changes in all state variables $S(t)$ (i.e., $\mu_{ij} = 0$, $i = 1, \dots, m$, and $j = 1, \dots, n$). Under this condition, the optimal demand functions in (7.8) can be rewritten as:

$$w_j^*(t)W(t) = K \sum_1^n v_{jk} (\alpha_k - r), \quad j = 1, \dots, n \quad . \quad (7.9)$$

A special case of this condition occurs when the investment opportunity set is nonstochastic (i.e., either $H_i = 0$, $i = 1, \dots, m$ or $(\alpha_i, \sigma_{ij}, r)$ are, at most, deterministic functions of time, $i, j = 1, \dots, n$). Optimal demands will also satisfy (7.9) if preferences are such that the marginal utility of wealth of the derived-utility function does not depend on $S(t)$ (i.e., $B_i = 0$, $i = 1, \dots, m$). By inspection of (7.6), this condition will obtain if the optimal consumption function C^* does not depend on $S(t)$. In direct correspondence to the discrete-time finding in Section 6, the only time-additive and independent utility function to satisfy this condition is $U[C, t] = a(t)\log[C(t)]$.

By inspection of (7.9), the relative holdings of risky securities, $w_j^*(t)/w_i^*(t)$, are the same for all investors, and thus, under these conditions, the efficient portfolio set will be spanned by just two funds: a single risky fund and a riskless fund. Moreover, by the procedure used to prove Theorem 4.9 and Theorem 4.10 in the static analysis, the efficient portfolio set here can be shown to be generated by the set of portfolios with minimum (instantaneous) variance for a given expected rate of return. Hence, under these conditions, the continuous-time intertemporal maximizer will act "as if" he were a static, Markowitz-Tobin mean-variance maximizer. Although the demand functions are formally identical to those derived from the mean-variance model, the analysis here assumes neither quadratic preferences nor spherically-symmetric or normally-distributed security returns. Indeed, if for example, the investment opportunity set $\{\alpha_j, r, \sigma_{ij}, i, j = 1, 2, \dots, n\}$ is constant through time, then from (7.3), the return on each risky security will be lognormally distributed, which implies that all securities have limited liability.⁴³

In the general case described in Theorem 7.1, the qualitative behavioral differences between an intertemporal maximizer and a static maximizer can be clarified further by analyzing the characteristics of the derived spanning portfolios.

As already shown, fund #1 and fund #2 are the "usual" portfolios that would be mixed to provide an optimal portfolio for a static maximizer. Hence, the intertemporal behavioral differences are characterized by funds #(2 + i), $i = 1, \dots, m$. At the level of demand functions, the "differential demand" for risky security j , ΔD_j^* , is defined to be the difference between the demand for that security by an intertemporal maximizer at time t and the

demand for that security by a static maximizer of the expected utility of "end-of-period" wealth where the absolute risk aversion and current wealth of the two maximizers are the same. Noting that $K \equiv -J_W/J_{WW}$ is the reciprocal of the absolute risk aversion of the derived utility of wealth function, from (7.8) we have that:

$$\Delta D_j^* = \sum_1^m B_i \zeta_{ij} \quad , \quad j = 1, \dots, n \quad . \quad (7.10)$$

Lemma 7.1: Define:

$$dY_i \equiv dS_i - \left(\sum_1^n \delta_j^+ \left(\frac{dP_j}{P_j} - rdt \right) + rdt \right) \quad .$$

The set of portfolio weights $\{\delta_j^+\}$ that minimize the (instantaneous) variance of dY_i are given by $\delta_j^+ = \zeta_{ij}$, $j = 1, \dots, n$ and $i = 1, \dots, m$.

Proof: The instantaneous variance of dY_i is equal to $[H_i^2 - 2 \sum_1^n \delta_j^+ H_i \sigma_j \mu_{ij} + \sum_1^n \sum_1^n \delta_j^+ \delta_k^+ \sigma_{jk}]$. Hence, the minimizing set of $\{\delta_j^+\}$ will satisfy $0 = -H_i \sigma_j \mu_{ij} + \sum_1^n \delta_k^+ \sigma_{jk}$, $j = 1, \dots, n$. By matrix inversion, $\delta_j^+ = \zeta_{ij}$.

The instantaneous rate of return on fund $\#(2+i)$ is exactly $[rdt + \sum_1^n \zeta_{ij} (dP_j/P_j - rdt)]$. Hence, fund $\#(2+i)$ can be described as that feasible portfolio whose rate of return most closely replicates the stochastic part of the instantaneous change in state variable $S_i(t)$, and this is true for $i = 1, \dots, m$.

Consider the special case where there exist securities that are instantaneously perfectly correlated with changes in each of the state variables. Without loss of generality, assume that the first m securities are the securities such that dP_i/P_i is perfectly positively correlated with dS_i , $i = 1, \dots, m$. In this case,⁴⁴ the demand function (7.8) can be rewritten in the form:

$$\begin{aligned}
 w_i^*(t)W(t) &= K \sum_1^n v_{ik}(\alpha_k - r) + B_i H_i / \sigma_i, \quad i = 1, \dots, m, \\
 &= K \sum_1^n v_{ik}(\alpha_k - r), \quad i = m + 1, \dots, n.
 \end{aligned}
 \tag{7.11}$$

Hence, the relative holdings of securities $m + 1$ through n will be the same for all investors, and the differential demand functions can be rewritten as:

$$\begin{aligned}
 \Delta D_i^* &= B_i H_i / \sigma_i, \quad i = 1, \dots, m, \\
 &= 0, \quad i = m + 1, \dots, n.
 \end{aligned}
 \tag{7.12}$$

The composition of fund $\#(2 + i)$ reduces to a simple combination of security i and the riskless security.

The behavior implied by the demand functions in (7.8) can be more easily interpreted if they are rewritten in terms of the direct-utility and optimal-consumption functions. The optimal-consumption function has the form $C^*(t) = C^*(W, S, t)$, and from (7.6), it follows immediately that:

$$K = -U_C[C^*,t]/(U_{CC}[C^*,t]\partial C^*/\partial W) \quad , \quad (7.13)$$

$$B_i = -(\partial C^*/\partial S_i)/(\partial C^*/\partial W) \quad , \quad i = 1, \dots, m \quad . \quad (7.14)$$

Because $\partial C^*/\partial W > 0$, it follows that the sign of B_i equals the sign of $(-\partial C^*/\partial S_i)$. An unanticipated change in a state variable is said to be unfavorable if, ceteris paribus, such a change would reduce current optimal consumption, e.g., an unanticipated increase in S_i would be unfavorable if $\partial C^*/\partial S_i < 0$. Inspection of (7.12), for example, shows that for such an individual, the differential demand for security i (which is perfectly positively correlated with changes in S_i) will be positive. If there is an unanticipated increase in S_i , then, ceteris paribus, there will be an unanticipated increase in his wealth. Because $\partial C^*/\partial W > 0$, this increase in wealth will tend to offset the negative impact on C^* caused by the increase in S_i , and therefore the unanticipated variation in C^* will be reduced. In effect, by holding more of this security, the investor expects to be "compensated" by larger wealth in the event that S_i changes in the unfavorable direction. Of course, if $\partial C^*/\partial S_i > 0$, then the investor takes a differentially short position. However, in all cases investors will allocate their wealth to the funds $\#(2 + i)$, $i = 1, \dots, m$, so as to "hedge" against unfavorable changes in the state variables $S(t)$.⁴⁵

Analysis of the usual static model does not produce such hedging behavior because the utility function is posited to depend only on end-of-period wealth and therefore, implicitly assumes that $\partial C^*/\partial S_i = 0$, $i = 1, \dots, m$. Thus, in addition to their manifest function of providing an "efficient" risk-return tradeoff for end-of-period wealth, securities in the intertemporal model have

a latent function of allowing consumers to hedge against other uncertainties.⁴⁶ The effect on equilibrium security prices from these "hedging demands" is examined in Section 9.

As a consequence of the richer role played by securities in the intertemporal model, the number of securities required to span the set of optimal portfolios will, in general, be larger than in the corresponding one-period model. It is, therefore, somewhat surprising that non-trivial spanning can obtain in the continuous-trading model. In the one-period analysis of Section 4, it was shown that for general preferences, a necessary and sufficient condition for a set of portfolios to span the efficient portfolio set is that the returns on every security can be written as a linear function of the returns on the spanning portfolios plus noise. As discussed in Section 4, in the absence of complete markets in the Arrow-Debreu sense, the wide-spread existence of corporate liability and other securities with nonlinear payoff structures appears to virtually rule out non-trivial spanning unless further restrictions are imposed on either preferences or the probability distributions of security returns. The hypothesized conditions of Theorem 7.1 require only that investors be risk averse with smooth preferences. Thus, it follows that the key to the spanning result is the combination of continuous trading and diffusion processes for the dynamics of security returns. As will be shown in Sections 8 and 9, diffusion processes are "closed" under nonlinear transformations. That is, the dynamics of a reasonably well-behaved function of diffusion-driven random variables will also be described by a diffusion process. Thus, unlike in the static and discrete-time dynamic models, the creation of securities whose payoff structures are nonlinear functions of existing security prices will not, in general, cause the size of the portfolio spanning set to increase.

8. Options, Contingent Claims Analysis and the Modigliani-Miller Theorem

Futures contracts, options, loan guarantees, mortgage-backed securities and virtually all corporate liabilities are among the many types of securities with the feature that their payoffs are contractually-linked to the prices of other securities at some future date. Contingent claims analysis (CCA) is a technique for determining the price of such "derivative" securities. As indicated in Sections 4 and 7, the fact that these contractual arrangements often involve nonlinear sharing rules has important implications for both corporate finance and the structure of equilibrium asset prices. For this reason and because such securities represent a significant and growing fraction of the outstanding stock of financial instruments, CCA is a mainstream topic in financial economic theory.

Although closely connected with the continuous-time portfolio models analyzed in the previous section, the origins of CCA are definitely rooted in the pioneering work of Black and Scholes (1973) on the theory of option pricing. Thus, we begin the study of contingent-claim pricing with an analysis of option securities.

An "European-type call (put) option" is a security that gives its owner the right to buy (sell) a specified quantity of a financial or real asset at a specified price, the "exercise price," on a specified date, the "expiration date." An American-type option allows its owner to exercise the option on or before the expiration date. If the owner chooses not to exercise the option on or before the expiration date, then it expires and becomes worthless.

If $V(t)$ denotes the price of the underlying asset at time t and E denotes the exercise price, then from the contract terms, the price of the call option at the expiration date T is given by $\max[0, V(T) - E]$ and the price of the put option is $\max[0, E - V(T)]$. If there is positive probability that $V(T) > E$, and positive probability that $V(T) < E$, then these options provide examples of securities with contractually-derived nonlinear sharing rules with respect to the underlying asset.

Although academic study of option pricing can be traced back to at least the turn of the century, the "watershed" in this research is the Black-Scholes (1973) model, which uses arbitrage arguments to derive option prices. It was, of course, well known before 1973, that if a portfolio of securities can be constructed to exactly match the payoffs to some security, then that security is redundant, and to rule out arbitrage, its price is uniquely determined by the prices of securities in the replicating portfolio. It was also recognized that because the price of an option at its expiration date is perfectly functionally related to the price of its underlying asset, the risk of an option position could be reduced by taking an offsetting position in the underlying asset. However, because portfolios involve linear combinations of securities and because the option has a nonlinear payoff structure, there is no static (i.e., "buy-and-hold") portfolio strategy in the underlying asset that can exactly replicate the payoff to the option. Thus, it would seem that an option cannot be priced by arbitrage conditions alone. Black and Scholes had the fundamental insight that a dynamic portfolio strategy in the underlying asset and the riskless security can be used to hedge the risk of an option position. With the idea in mind that the precision of the hedge can be improved by increasing the frequency of portfolio revisions, they focused on

the limiting case of continuous revisions. By assuming that the price dynamics of the underlying asset are described by a geometric Brownian motion and that the interest rate is constant, Black and Scholes derive a trading strategy that perfectly hedges the option position. They are, thus, able to determine the option price from the equilibrium condition that the return on a perfectly-hedged portfolio must equal the interest rate.

Under the assumption that the dynamics for the underlying asset price are described by a diffusion process with a continuous sample path, Merton (1973a,1977) uses the mathematics of Itô stochastic integrals to prove that with continuous trading, the Black-Scholes dynamic portfolio strategy will exactly replicate the payoff to an option held until exercise or expiration. Therefore, under these conditions, the Black-Scholes option price is a necessary condition to rule out arbitrage. Using a simplified version of the arbitrage proof in Merton (1977), we derive the Black-Scholes price for an European call option.

Following the notation in Section 7, we assume that the dynamics of the underlying asset price are described by a diffusion process given by:

$$dV = \alpha V dt + \sigma V dZ \quad (8.1)$$

where σ is, at most, a function of V and t . No cash payments or other distributions will be made to the owners of this asset prior to the expiration date of the option.

Let $F(V,t)$ be the solution to the partial differential equation:

$$1/2 \sigma^2 V^2 F_{11} + rVF_1 - rF + F_2 = 0 \quad (8.2)$$

subject to the boundary conditions:

- (a) $F(0,t) = 0$
- (b) F/V bounded (8.3)
- (c) $F(V,T) = \max[0, V - E]$

where subscripts denote partial derivatives with respect to the arguments of F . A solution to (8.2)-(8.3) exists and is unique.⁴⁷

Consider a continuous-time portfolio strategy where the investor allocates the fraction $w(t)$ to the underlying asset and $[1 - w(t)]$ to the riskless security. If $w(t)$ is right-continuous function and $P(t)$ denotes the value of the portfolio at time t , then, from Section 7, the dynamics for P can be written as:

$$dP = [w(\alpha - r) + r]Pdt + w\sigma PdZ \quad . \quad (8.4)$$

Suppose the investor selects the particular portfolio strategy, $w(t) = F_1(V,t)V(t)/P(t)$. Note that the strategy rule $w(t)$ for each t depends on the partial derivative of the known function F , the current price of the underlying asset and the current value of the portfolio. By substitution into (8.4), we have that:

$$dP = [F_1V(\alpha - r) + rP]dt + F_1V\sigma dZ \quad . \quad (8.5)$$

Since F is twice continuously differentiable, Itô's Lemma⁴⁸ can be applied to express the stochastic process for $F(V(t),t)$ as:

$$dF = [1/2 \sigma^2 V^2 F_{11} + \alpha VF_1 + F_2]dt + F_1V\sigma dZ \quad (8.6)$$

But F satisfies (8.2) and therefore, (8.6) can be rewritten as:

$$dF = [F_1V(\alpha - r) + rF]dt + F_1V\sigma dZ \quad . \quad (8.7)$$

From (8.5) and (8.7), $dP - dF = [P - F]rdt$, an ordinary differential equation with the well-known solution:

$$P(t) - F(V(t),t) = [P(0) - F(V(0),0)]e^{rt} \quad . \quad (8.8)$$

If the initial investment in the portfolio is chosen so that $P(0) = F(V(0),0)$,

then from (8.8), $P(t) = F(V(t),t)$ for $0 \leq t \leq T$. From (8.3), we have that $P(t) = 0$, if $V(t) = 0$ and $P(T) = \max[0, V(T) - E]$. Thus, we have constructed a feasible portfolio strategy in the underlying asset and the riskless security that exactly replicates the payoff structure to an European call option with exercise price E and expiration date T . By the standard no-arbitrage condition, two securities with identical payoff structures must have the same price. Thus, the equilibrium call option price at time t is given by $F(V(t),t)$, the Black-Scholes price.

The derivation did not assume that equilibrium option price depends only on the price of the underlying asset and the riskless interest rate. Thus, if the option price is to depend on other prices or stochastic variables, then, by inspection of (8.2), it must be because either σ^2 or r is a function of these prices. Similarly, the findings that the option price is a twice continuously-differentiable function of the underlying asset price and that its dynamics follow a diffusion process are derived results and not assumptions.

Because Black and Scholes derived (8.2)-(8.3) for the case where σ^2 is a constant (i.e., geometric Brownian motion), they were able to obtain a closed-form solution, given by:

$$F(V,t) = V\Phi(x_1) - Ee^{-r(T-t)}\Phi(x_2) \quad , \quad (8.9)$$

where $x_1 \equiv [\log(V/E) + (r + \sigma^2/2)(T - t)]/\sigma \sqrt{T - t}$; $x_2 \equiv$

$x_1 - \sigma \sqrt{T - t}$; and $\Phi(\cdot)$ is the cumulative Gaussian density function.

From (8.9), it follows that the portfolio-construction rule is given by $w(t) = F_1 V/F = \Phi(x_1)V/F(V,t)$.

By inspection of (8.2) or (8.9), a striking feature of the Black-Scholes analysis is that the determination of the option price and the replicating

portfolio strategy does not require knowledge of either the expected return on the underlying asset α or investor risk preferences and endowments. Indeed, the only variable or parameter required that is not directly observable is the variance-rate function, σ^2 . This feature, together with the relatively-robust nature of an arbitrage derivation, gives the Black-Scholes model an important practical significance, and it has been widely adopted in the practicing financial community.

In the derivation of the equilibrium call-option price, the only place that the explicit features of the call option enter is in the specification of the boundary conditions (8.3). Hence, by appropriately adjusting the boundary conditions, the same methodology can be used to derive the equilibrium prices of other derivative securities with payoff structures contingent on the price of the underlying asset. For example, to derive the price of an European put option, one need only change (8.3) so that F satisfies $F(0,t) = E \exp[-r(T - t)]$; $F(V,t)$ bounded; and $F(V,T) = \max[0, E - V]$.

Although options are rather specialized financial instruments, the Black-Scholes option pricing methodology can be applied to a much broader class of securities.

A prototypical example analyzed in Black and Scholes (1973) and Merton (1973a;1974) is the pricing of debt and equity of a corporation. Consider the case of a firm financed by equity and a single homogeneous zero-coupon debt issue. The contractual obligation of the firm is to pay B dollars to the debtholders on the maturity date T and in the event that the firm does not pay (i.e., defaults), then ownership of the firm is transferred to the debtholders. The firm is prohibited from making payments or transferring assets to the equityholders prior to the debt being retired. Let $V(t)$

denote the market value of the firm at time t . If at the maturity date of the debt, $V(T) \geq B$, then the debtholders will receive their promised payment B and the equityholders will have the "residual" value, $V(T) - B$. If, however, $V(T) < B$, then there are inadequate assets within the firm to pay the debtholders their promised amount. By the limited-liability provision of corporate equity, the equityholders cannot be assessed to make up the shortfall, and it is clearly not in their interests to do so voluntarily. Hence, if $V(T) < B$, then default occurs and the value of the debt is $V(T)$ and the equity is worthless. Thus, the contractually-derived payoff function for the debt at time T , f_1 , can be written as:

$$f_1(V,T) = \min[V(T),B] \quad , \quad (8.10)$$

and the corresponding payoff function for equity, f_2 , can be written as:

$$f_2(V,T) = \max[0, V(T) - B] \quad . \quad (8.11)$$

Provided that default is possible but not certain, we have from (8.10) and (8.11), that the sharing rule between debtholders and equityholders is a nonlinear function of the value of the firm. Moreover, the payoff structure to equity is isomorphic to an European call option where the underlying asset is the firm; the exercise price is the promised debt payment; and the expiration date is the maturity date. Because $\min[V(T),B] = V(T) - \max[0, V(T) - B]$, the debtholders' position is functionally equivalent to buying the firm outright from the equityholders at the time of issue and simultaneously, giving them an option to buy back the firm at time T for B . Hence, provided that the conditions of continuous-trading opportunities and a diffusion-process representation for the dynamics of the firm's value are satisfied, the Black-Scholes option pricing theory can be applied directly to the pricing of levered equity and corporate debt with default risk.

The same methodology can be applied quite generally to the pricing of derivative securities by adjusting the boundary conditions in (8.3) to match the contractually-derived payoff structure. Cox, Ingersoll and Ross (1985b) use this technique to price default-free bonds in their widely-used model of the term structure of interest rates. The survey articles by Smith (1976) and Mason and Merton (1985) and the excellent Cox-Rubinstein book (1985) provide applications of CCA in a broad range of areas, including the pricing of general corporate liabilities; project evaluation and financing; pension fund and deposit insurance; and employee-compensation contracts such as guaranteed wage floors and tenure. Although, in most applications, (8.2)-(8.3) will not yield closed-form solutions, powerful computational methods have been developed to provide high-speed numerical solutions for both the security price and its first derivative.

As shown in Section 4, the linear generating process for security returns which is required for non-trivial spanning in Theorem 4.6 is generally not satisfied by securities with nonlinear sharing rules. However, if the underlying asset-price dynamics are diffusions, we have shown that the dynamics of equilibrium derivative-security prices will also follow diffusion processes. The existence of such securities is, therefore, consistent with the hypothesized conditions of Theorem 7.1. Hence, the creation of securities with nonlinear sharing rules will not adversely affect the spanning results derived for the continuous-time portfolio selection model of Section 7. Using a replication argument similar to the one presented here, Merton (1977) proves that Theorem 4.14, the Modigliani-Miller Theorem, will obtain under the conditions of continuous trading and a diffusion representation for the dynamics of the market value of the firm.

In the introduction to the basic portfolio-selection problem in Section 2, it was pointed out that the standard model allows borrowing and short selling, but does not explicitly take account of personal bankruptcy (i.e., that an investor's wealth cannot be negative). However, unlike in the static and discrete-time dynamic portfolio models, the continuous-time model of Section 7 can be easily adapted to include the effects of bankruptcy on portfolio choice and on the return distributions of both investor and creditor portfolios. Moreover, such explicit recognition of bankruptcy will have no material impact on derived investor behavior. Just as we were able to determine the price and return characteristics of corporate debt with default possibilities, so we can evaluate the price and return characteristics of loans of cash or securities to an investor whose portfolio is the sole collateral for these loans. As shown, the introduction of nonlinear sharing rules (in this case, between investors and their creditors) does not by itself violate the diffusion assumption for security and portfolio returns. Therefore, the nonlinearities induced by taking account of personal bankruptcy do not affect the conclusions of Theorem 7.1.

In Theorem 7.1, as in the mutual fund theorems of Section 4, investors are shown to be indifferent between selecting their portfolios either from the $m + 2$ portfolios that span the efficient set or from all $n + 1$ available securities. Similarly, investors are indifferent as to whether or not derivative securities are available because they can use portfolio prescription (8.2)-(8.3) to replicate the payoffs to these securities. It would, thus, seem that the rich menu of financial intermediaries and financial instruments observed in the real world has no important economic function in the environment posited in Section 7. Such indifference is indeed the case

if, as assumed, all investors can gather information and transact without cost. The introduction of transactions costs where financial intermediaries and market makers have significantly lower costs than individual investors and general business firms, does however provide an adequate raison d'etre for financial intermediation and markets for derivative securities.

From this perspective, the portfolio rules used to establish the mutual funds in Theorem 7.1 and the replication formula provided by (8.2)-(8.3) can be viewed as blueprints for the production technologies in the financial-intermediation industry. For example, the Black-Scholes trading strategy and price function provide the technology and the production cost for an intermediary to create an option on a traded asset.

If the input (in this case, traded-securities) markets are competitive and there is free entry into the financial-services industry, then the equilibrium prices of financial products will equal the production costs to the lowest-cost producers. Furthermore, if there is a sufficient number of potential producers who can (to a reasonable approximation) trade continuously without significant marginal costs or restrictions, then equilibrium prices for derivative securities will be well-approximated by the solution to (8.2) with the appropriate boundary conditions.⁴⁹ Thus, with well-functioning capital markets and financial intermediation, the creation of mutual funds and derivative-security markets can provide important economic benefits to individual investors and corporate issuers, even though these securities are priced as if they were redundant.

Although clearly a powerful tool for determining the price and return characteristics of many important classes of securities, CCA cannot be used directly to determine the prices of all types of securities. Moreover,

because it simply assumes that underlying asset price dynamics can be described by a Markov diffusion process, CCA is partial equilibrium in nature. Therefore, we turn now to the issues of general-equilibrium pricing and the efficiency of allocations for the continuous-trading model of Section 7.

9. Intertemporal Capital Asset Pricing

Using the continuous-time model of portfolio selection described in Section 7, Merton (1973b) and Breeden (1979) aggregate individual investor demand functions and impose market-clearing conditions to derive an intertemporal model of equilibrium asset prices. By assuming constant-returns-to-scale production technologies with stochastic outputs and technical progress described by diffusion processes, Cox, Ingersoll and Ross (1985a) develop a general-equilibrium version of the model, which explicitly integrates the real and financial sectors of the economy. Huang (1985;1987) further strengthens the foundation for these models by showing that if information in an economy with continuous-trading opportunities evolves according to diffusion processes, then intertemporal-equilibrium security prices will also evolve according to diffusion processes.

In Proposition 5.5, it was shown that if X denotes the return on any mean-variance efficient portfolio (with positive dispersion), then the expected returns on each of the n risky assets used to construct this portfolio will satisfy $\bar{Z}_j - R = a_j(\bar{X} - R)$ where $a_j \equiv \text{cov}(Z_j, X)/\text{Var}(X)$, $j = 1, \dots, n$. Because the returns on all mean-variance efficient portfolios are perfectly-positively correlated, this relation will apply with respect to any such portfolio. Moreover, because Proposition 5.5 is purely a mathematical result, it follows immediately for the model of Section 7 that at time t :

$$\alpha_j(t) - r(t) = a_j^*(t)[\alpha^*(t) - r(t)], \quad j = 1, \dots, n \quad (9.1)$$

where α^* is the expected return on an (instantaneously) mean-variance efficient portfolio and a_j^* equals the instantaneous covariance of the

return on security j with this portfolio divided by the instantaneous variance of the portfolio's return. Thus, knowledge of the expected return and variance of a mean-variance efficient portfolio together with the covariance of that portfolio's return with the return on asset j is sufficient information to determine the risk and expected return on asset j . It is, however, generally difficult to identify an ex-ante mean-variance efficient portfolio by statistical estimation alone and hence, the practical application of (9.1) is limited.

As shown in Sections 4 and 5, the Sharpe-Lintner-Mossin Capital Asset Pricing Model provides an example where identification without estimation is possible. Because all investors in that model hold the same relative proportions of risky assets, market-clearing conditions for equilibrium imply that the market portfolio is mean-variance efficient. The mathematical identity (9.1) is, thus, transformed into the Security Market Line (Theorem 4.8) which has economic content. The market portfolio can, in principle, be identified without knowledge of the joint distribution of security returns.

From (7.8), the relative holdings of risky assets will not in general be the same for all investors in the continuous-trading model. Thus, the market portfolio need not be mean-variance efficient as a condition for equilibrium and therefore, the Security Market Line will not obtain in general. However, Merton (1973b,p. 881) and Breeden (1979,p. 273) show that equilibrium expected returns will satisfy:

$$\alpha_j(t) - r(t) = \sum_1^{m+1} B_{ij}(t)[\alpha^i(t) - r(t)], \quad j = 1, \dots, n \quad (9.2)$$

where α^1 is the expected return on the market portfolio; α^i is the expected return on a portfolio with the maximum feasible correlation of its

return with the change in state variable S_{i-1} , $i = 2, \dots, m+1$; and the $\{B_{ij}\}$ correspond to the theoretical multiple-regression coefficients from regressing the (instantaneous) returns of security j on the returns of these $m+1$ portfolios. (9.2) is a natural generalization of the Security Market Line and is therefore aptly called the Security Market Hyperplane.

Let dX^i/X^i , $i = 1, \dots, m+1$, denote the instantaneous rate of return on the i th portfolio, whose expected return is represented on the right-hand side of (9.2). It follows immediately from the definition of the $\{B_{ij}\}$ that the return dynamics on asset j can be written as:

$$dP_j/P_j = r(t)dt + \sum_1^{m+1} B_{ij}(t)(dX^i/X^i - r(t)dt) + d\epsilon_j \quad (9.3)$$

where $d\epsilon_j$ is a diffusion process such that $E_t(d\epsilon_j) = E_t(d\epsilon_j | dX^1/X^1, \dots, dX^{m+1}/X^{m+1}) = 0$. (9.3) is the continuous-trading dynamic analog to the result derived for static models of spanning in Theorem 4.6. If the $\{B_{ij}(t)\}$ are sufficiently slowly-varying functions of time relative to the intervals over which successive returns are observed, then from (9.3), these risk-measure coefficients can, in principle, be estimated using time-series regressions of individual security returns on the spanning portfolios' returns.

From Lemma 7.1, we have that $dX^{i+1}/X^{i+1} = dS_i - dY_i$, $i = 1, \dots, m$, where dY_i is uncorrelated with all speculative securities' returns and therefore, is uncorrelated with dX^k/X^k , and $d\epsilon_j$, $k = 1, \dots, m+1$, and $j = 1, \dots, n$. It follows from (9.3) that:

$$dP_j/P_j = A_j(t)dt + B_{1j}(t)[dX^1/X^1 - r(t)dt] + \sum_2^{m+1} B_{ij}(t)dS_{i-1} + d\epsilon'_j \quad (9.4)$$

where $A_j(t)$ is a locally-nonstochastic drift term and $d\epsilon'_j \equiv d\epsilon_j - \sum_2^{m+1} B_{ij}(t)[dY_{i-1} - E_t(dY_{i-1})]$. Although $E_t(d\epsilon'_j) = 0$ and $E_t(d\epsilon'_j | dX^1/X^1, dS_1, \dots, dS_m) = 0$, it is not the case that $E_t(d\epsilon'_j | dX^1/X^1, dS_1, \dots, dS_m) = 0$ (unless $dY_i \equiv 0$, $i = 1, \dots, m$). Hence, unlike (9.3), (9.4) is not a properly-specified regression equation, because dS_i and dY_k are not uncorrelated for every $i, k = 1, \dots, m$. Thus, it is not in general valid to regress speculative-price returns on the change in nonspeculative-price state variables to obtain estimates of the $\{B_{ij}(t)\}$.

By Itô's Lemma, the unanticipated change in investor q 's optimal consumption rate can be written as:

$$dC_q - E_t(dC_q) = (\partial C_q / \partial W_q) \sum_1^n w_j^* W_q \sigma_j dZ_j + \sum_1^m (\partial C_q / \partial S_i) H_i dq_i, \quad (9.5)$$

where $\{w_j^*\}$ is his optimal holding of security j as given in (7.8). Let dX^*/X^* denote the return on the mean-variance efficient portfolio, which allocates fraction $\sum_1^n v_{ij}(\alpha_i - r)$ to security j , $j = 1, \dots, n$ and the balance to the riskless security. By substitution for $w_j^* W_q$ from (7.8) and rearranging terms, we can rewrite (9.5) as:

$$dC_q - E_t(dC_q) = V_q (dX^*/X^* - \alpha^* dt) - \sum_1^m (\partial C_q / \partial S_i) [dY_i - E_t(dY_i)] \quad , \quad (9.6)$$

where $V_q \equiv [\partial C_q / \partial W_q] K_q W_q$ and dY_i is as defined in Lemma 7.1. From Lemma 7.1, dP_j/P_j and dY_i are uncorrelated for $j = 1, \dots, n$ and $i = 1, \dots, m$. It follows, therefore, that $\text{cov}[dC_q, dP_j/P_j] = V_q \text{cov}[dX^*/X^*, dP_j/P_j]$. If $\bar{C} \equiv \sum_q C_q$ denotes aggregate consumption,

then by the linearity of the covariance operator, we have that:

$$\text{cov}[d\bar{C}, dP_j/P_j] = V \text{cov}[dX^*/X^*, dP_j/P_j] \quad , \quad j = 1, \dots, n \quad , \quad (9.7)$$

where $V \equiv \sum_q V_q$. But, from (9.1), $(\alpha_j - r) = \text{cov}[dX^*/X^*, dP_j/P_j](\alpha^* - r) / \text{var}[dX^*/X^*]$. Hence, from (9.1) and (9.7), we have that:

$$\alpha_k - r = \beta_{kC}(\alpha_j - r) / B_{jC} \quad , \quad k, j = 1, \dots, n \quad , \quad (9.8)$$

where $\beta_{kC} \equiv \text{cov}[d\bar{C}, dP_k/P_k] / \text{Var}[d\bar{C}]$. Thus, from (9.8), a security's risk can be measured by a single composite statistic: namely, the covariance between its return and the change in aggregate consumption. Breeden (1979, pp. 274-276) was the first to derive this relation which combines the generality of (9.2)-(9.4) with the simplicity of the classic Security Market Line.⁵⁰

If in the model of Section 7, the menu of available securities is sufficiently rich that investors can perfectly hedge against unanticipated changes in each of the state variables S_1, \dots, S_m , then from Lemma 7.1, $\text{var}(dY_i) = 0$ for $i = 1, \dots, m$. From (9.6), unanticipated changes in each investor's optimal consumption rate are instantaneously perfectly correlated with the returns on a mean-variance efficient portfolio and therefore, are instantaneously perfectly correlated with unanticipated changes in aggregate consumption. This special case analyzed in (7.11) and (7.12) takes on added significance because Breeden (1979) among others has shown that inter-temporal-equilibrium allocations will be Pareto efficient if such perfect-hedging opportunities are available.

This efficiency finding for general preferences and endowments is perhaps surprising, because it is well known that a competitive equilibrium does not in general produce Pareto-optimal allocations without complete Arrow-Debreu

markets. Because the dynamics of the model in Section 7 are described by diffusion processes, there are a continuum of possible states over any finite interval of time. Therefore, complete markets in this model would seem to require an uncountable number of pure Arrow-Debreu securities. However, as we know from the work of Radner (1972), an Arrow-Debreu equilibrium allocation can be achieved without a full set of pure time-state contingent securities if the menu of available securities is sufficient for agents to use dynamic-trading strategies to replicate the payoff structures of the pure securities. Along the lines of the contingent-claims analysis of Section 8, we now show that if perfect hedging of the state variables is feasible, then the Radner conditions are satisfied by the continuous-trading model of Section 7.

By hypothesis, it is possible to construct portfolios whose returns are instantaneously perfectly correlated with changes in each of the state variables, $[dS_1(t), \dots, dS_m(t)]$, as described by (7.1). For notational simplicity and without loss of generality, assume that the first m available risky securities are these portfolios so that $dZ_i = dq_i$, $i = 1, \dots, m$.

With subscripts denoting partial derivatives of Π , with respect to (S_1, \dots, S_m, t) , let $\Pi(S, t)$ satisfy the linear partial differential equation:

$$0 = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m H_i H_j \eta_{ij} \Pi_{ij} + \sum_{j=1}^m [G_j - H_j(\alpha_j - r)/\sigma_j] \Pi_j + \Pi_{m+1} - r\Pi \quad (9.9)$$

subject to the boundary conditions: $0 \leq \Pi(S, t) \leq 1$ for all S and $t < T$;
 $\Pi(S, T) = 1$ if $\bar{S}_k - \epsilon \leq S_k(T) \leq \bar{S}_k + \epsilon$ for each $k = 1, 2, \dots, m$ and
 $\Pi(S, T) = 0$ otherwise.⁵¹

Consider the continuous-trading portfolio strategy that allocates the fraction $\delta_j(t) = \Pi_j H_j / [\sigma_j V(t)]$ to security j , $j = 1, \dots, m$, and the balance to the riskless asset, where $V(t)$ denotes the value of the portfolio at time t . From (7.3), the dynamics of the portfolio value can be written as:

$$\begin{aligned} dV &= V \left\{ \left[\sum_{j=1}^m \delta_j (\alpha_j - r) + r \right] dt + \sum_{j=1}^m \delta_j \sigma_j dZ_j \right\} \\ &= \left[\sum_{j=1}^m \Pi_j H_j (\alpha_j - r) / \sigma_j + rV \right] dt + \sum_{j=1}^m \Pi_j H_j dq_j, \end{aligned} \quad (9.10)$$

because $dZ_j = dq_j$, $j = 1, \dots, m$.

Because Π is twice-continuously differentiable, we can use Itô's Lemma to write the dynamics of $\Pi(S(t), t)$ as:

$$d\Pi = \left(\frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m H_i H_j \rho_{ij} \Pi_{ij} + \sum_{j=1}^m [G_j \Pi_j] + \Pi_{m+1} \right) dt + \sum_{j=1}^m \Pi_j H_j dq_j. \quad (9.11)$$

But, $\Pi(S, t)$ satisfies (9.9) and hence, (9.11) can be rewritten as:

$$d\Pi = \left[\sum_{j=1}^m \Pi_j H_j (\alpha_j - r) / \sigma_j + r\Pi \right] dt + \sum_{j=1}^m \Pi_j H_j dq_j, \quad (9.12)$$

From (9.10) and (9.12), $d\Pi - dV = r(\Pi - V)dt$. Therefore, if the initial investment in the portfolio is chosen so that $V(0) = \Pi(S(0), 0)$, then $V(t) = \Pi(S(t), t)$ for $0 \leq t \leq T$.

Thus, by using $(m + 1)$ available securities, a dynamic portfolio strategy has been constructed with a payoff structure that matches the boundary conditions of (9.9). By taking the appropriate limit as $\epsilon \rightarrow 0$, the solution of (9.9) provides the portfolio prescription to exactly replicate the payoff to a pure Arrow-Debreu security, which pays \$1 at time T if

$S_k(T) = \bar{S}_k$, $k = 1, \dots, m$ and pays 0, otherwise.⁵² By changing the parameters \bar{S} and T , one can generate the portfolio rules to replicate all of the uncountable number of Arrow-Debreu securities using just $m + 1$ securities.⁵³

As in the similar analysis of derivative-security pricing in Section 8, we have here that $\Pi(S, t)$ will also be the equilibrium price for the corresponding Arrow-Debreu security. Note, however, that unlike the analysis in Section 8, the solution to (9.9) requires knowledge of the expected returns $(\alpha_1, \dots, \alpha_m)$. The reason is that the state variables of the system are not speculative prices. If they were, then to avoid arbitrage, $G_j, H_j, \alpha_j, \sigma_j$ and r would have to satisfy the condition $[G_j - rS_j]/H_j = (\alpha_j - r)/\sigma_j$, $j = 1, \dots, m$. In that case, the coefficient of Π_j in (9.9) could be rewritten as rS_j , $j = 1, \dots, m$, and the solution of (9.9) would not require explicit knowledge of either G_j or α_j .

As we saw in Section 8, options are a fundamental security in the theory of contingent-claims pricing. As demonstrated in a general context by Ross (1976b), options can also be used in an important way to complete markets and thereby, to improve allocational efficiency.

The close connection between pure-state securities and options in the continuous-time model is also exemplified by the work of Breeden and Litzenberger (1978). They analyze the pricing of state-contingent claims in the scalar case where a single variable is sufficient to describe the state of the economy. A "butterfly-spread" option strategy holds a long position in two call options, one with exercise price $E - \Delta$ and the other with exercise price $E + \Delta$, and a short position in two call options with exercise price E , where the expiration dates of the options are the same. If options are available on a security whose price $V(t)$ is in one-to-one correspondence with the state variable, then the payoff to a state-contingent claim which pays \$1 at time T if $V(T) = E$ and \$0 otherwise, can be approximated by $1/\Delta$ units of a butterfly spread. This approximation becomes exact in the limit as $\Delta \rightarrow dE$, the infinitesimal differential. Breeden and Litzenberger, thus, show that the pure-state security price is given by $(\partial^2 F / \partial E^2) dE$, where F is the call-option pricing function derived in Section 8. Under the specialized conditions for which the Black-Scholes formula (8.9) applies, the solution for the pure-state security price has a closed-form given by $\exp[-r(T - t)] \Phi'(x_2) dE / (\sigma E \sqrt{T - t})$.

In the intertemporal version of the Arrow-Debreu complete-markets model, there is a security for every possible state of the economy, but markets need only be open "once" because agents will have no need for further trade. In the model of this section, there are many fewer securities, but agents trade continuously. Nevertheless, both models have many of the same properties. It appears, therefore, that a good substitute for having a large number of markets and securities is to have the existing markets open more frequently for trade.

In addressing this point as well as the robustness of the continuous-time model, Duffie and Huang (1985) derive necessary and sufficient conditions for continuous-trading portfolio strategies with a finite number of securities to effectively complete markets in a Radner economy. As discussed in Section 5, the Arrow-Debreu model permits some degree of heterogeneity in beliefs among agents. Just so, Duffie and Huang show that the spanning results derived here for continuous trading are robust with respect to heterogeneous probability assessments among agents provided that their subjective probability measures are uniformly absolutely continuous. In later work (forthcoming), they derive conditions where these results obtain in the more-general framework of differential information among agents.

Although continuous trading is, of course, only a theoretical proposition, the continuous-trading solutions will be an asymptotically-valid approximation to the discrete-time solutions as the trading interval becomes small.⁵⁴ An in-depth discussion of the mathematical and economic assumptions required for the valid application of the continuous-time analysis is beyond the scope of this paper.⁵⁵ However, actual securities markets are open virtually all the time, and hence, the required assumptions are rather reasonable when applied in that context.

In summary, we have seen that all the interesting models of portfolio selection and capital market theory share in common, the property of nontrivial spanning. If, however, a model is to be broadly applicable, then it should also satisfy the further conditions that: (i) the number of securities required for spanning be considerably smaller than both the number of agents and the number of possible states for the economy; and (ii) the creation of securities with nonlinear sharing rules by an individual investor

or firm should not, in general, alter the size of the spanning set. As we have seen, the continuous-trading model with vector diffusions for the underlying state variables meets these criteria. Motivated in part by the work of Harrison and Kreps (1979), Duffie and Huang (1985) use martingale representation theorems to show that with continuous trading, these conditions will also obtain for a class of non-Markov, path-dependent processes, some of which do not have continuous sample paths.⁵⁶ It remains, however, an open and important research question as to whether in the absence of continuous trading, these criteria can be satisfied in interesting models with general preferences and endowments.

FOOTNOTES

*This paper is a revised and expanded version of Merton (1982a).

1. On the informational efficiency of the stock market, see Fama (1965,1970a), Samuelson (1965), Hirshleifer (1973), Grossman (1976), Grossman and Stiglitz (1976), Black (1986), and Merton (1987a,1987b). On financial markets and incomplete information generally, see the excellent survey paper by Bhattacharya (forthcoming). On financial markets and auction theory, see Hansen (1985), Parsons and Raviv (1985), and Rock (1986). On the role of behavioral theory in finance, see Hogarth and Reder (1986).
2. von Neumann and Morgenstern (1947). For an axiomatic description, see Herstein and Milnor (1953). Although the original axioms require that U be bounded, the continuity axiom can be extended to allow for unbounded functions. See Samuelson (1977) for a discussion of this and the St. Petersburg Paradox.
3. The strict concavity assumption implies that investors are everywhere risk averse. Although strictly convex or linear utility functions on the entire range imply behavior that is grossly at variance with observed behavior, the strict concavity assumption also rules out Friedman-Savage type utility functions whose behavioral implications are reasonable. The strict concavity also implies $U'(W) > 0$, which rules out individual satiation.
4. Borrowings and short-sales are demand loans collateralized by the investor's total portfolio. The "borrowing rate" is the rate on riskless-in-terms-of-default loans. Although virtually every individual loan involves some chance of default, the empirical "spread" in the rate on actual margin loans to investors suggests that this assumption is not a "bad approximation" for portfolio selection analysis. However, an explicit analysis of risky loan evaluation is provided in Section 8.
5. The existence of an interior solution is assumed throughout the analyses in the paper. For a complete discussion of necessary and sufficient conditions for the existence of an interior solution, see Leland (1972) and Bertsekas (1974).
6. For a trivial example, shares of IBM with odd serial numbers are distinguishable from ones with even serial numbers and are, therefore, technically different securities. However, because their returns are identical, they are perfect substitutes from the point of view of investors. In portfolio theory, securities are operationally defined by their return distributions, and therefore, two securities with identical returns are indistinguishable.

7. If U is such that $U'(0) = \infty$, and by extension, $U'(W) = \infty, W < 0$, then from (2.2) or (2.4), it is easy to show that the probability of $Z^* \leq 0$ is a set of measure zero. Mason (1981) has studied the effects of various bankruptcy rules on portfolio behavior.
8. The behavior associated with the utility function $V(W) \equiv aU(W) + b$, $a > 0$, is identical to that associated with $U(W)$. Note: $A(W)$ is invariant to any positive affine transformation of $U(W)$. See Pratt (1964).
9. Rothschild and Stiglitz (1970,1971). There is an extensive literature, not discussed here, that uses this type of risk measure to determine when one portfolio "stochastically dominates" another. Cf. Hadar and Russell (1970,1971), Hanoch and Levy (1969), and Bawa (1975).
10. I believe that Christian von Weizsäcker proved a similar theorem in unpublished notes some years ago. However, I do not have a reference.
11. For a proof, see Theorem 236 in Hardy, Littlewood, and Pölya (1959).
12. A sufficient amount of information would be the joint distribution of Z^* and ϵ_j . What is necessary will depend on the functional form of U' . However, in no case will knowledge of \bar{Z}_j be a necessary condition.
13. Cf. King (1966), Livingston (1977), Farrar (1962), Feeney and Hester (1967), and Farrell (1974). Unlike standard "factor analysis," the number of common factors here does not depend upon the fraction of total variation in an individual security's return that can be "explained." Rather, what is important is the number of factors necessary to "explain" the covariation between pairs of individual securities.
14. There is considerable controversy on this issue. See Chamberlain and Rothschild (1983), Dhrymes, Friend, and Gultekin (1984,1985), Roll and Ross (1980), Rothschild (1986), and Trzcinka (1986).
15. This assumption formally rules out financial securities that alter the tax liabilities of the firm (e.g., interest deductions) or ones that can induce "outside" costs (e.g., bankruptcy costs). However, by redefining $V_j(I_j; \theta_j)$ as the pre-tax-and-bankruptcy value of the firm and letting one of the f_k represent the government's tax claim and another the lawyers' bankruptcy-cost claim, the analysis in the text will be valid for these extended securities as well.
16. Miller and Modigliani (1958). See also Stiglitz (1969,1974), Fama (1978), Miller (1977). The "MM" concept has also been applied in other parts of monetary economics as in Wallace (1981).
17. For this family of utility functions, the probability distribution for securities cannot be completely arbitrary without violating the von Neumann-Morgenstern axioms. For example, it is required that for every

realization of W , $W > -a/b$ for $b > 0$ and $W < -a/b$ for $b < 0$. The latter condition is especially restrictive.

18. A number of authors have studied the properties of this family. See Merton (1971, p. 389) for references.
19. As discussed in footnote 17, the range of values for a_i cannot be arbitrary for a given b . Moreover, the sign of b uniquely determines the sign of $A'(W)$.
20. Cf. Ross (1978) for spanning proofs in the absence of a riskless security. Black (1972) and Merton (1972) derive the two-fund theorem for the mean-variance model with no riskless security.
21. For the Arrow-Debreu model, see Hirshleifer (1965, 1966, 1970), Myers (1968), and Radner (1970). For the mean-variance model, see Jensen (1972), Jensen, ed. (1972), and Sharpe (1970).
22. If the states are defined in terms of end-of-period values of the firm in addition to "environmental" factors, then the firms' production decisions will, in general, alter the state-space description which violates the assumptions of the model. Moreover, I see no obvious reason why individuals are any more likely to agree upon the $\{V_j(i)\}$ function than upon the probability distributions for the environmental factors. If sufficient information is available to partition the states into fine-enough categories to produce agreement on the $\{V_j(i)\}$ functions, then, given this information, it is difficult to imagine how rational individuals would have heterogeneous beliefs about the probability distributions for these states. As with the standard certainty model, agreement on the technologies is necessary for Pareto optimality in this model. However, as Peter Diamond has pointed out to me, it is not sufficient. Sufficiency demands the stronger requirement that everyone be "right" in their assessment of the technologies. See Varian (1985) and Black (1986, footnote 5) on whether differences of opinion among investors can be supported in this model.
23. In particular, the optimal portfolio demand functions are of the form derived in the proof of Theorem 4.9. For a complete analytic derivation, see Merton (1972).
24. Sharpe (1964), Lintner (1965), and Mossin (1966) are generally credited with independent derivations of the model. Black (1972) extended the model to include the case of no riskless security.
25. Cf. Borch (1969), Feldstein (1969), Tobin (1969), and Samuelson (1967).
26. The additive independence of the utility function and the single-consumption good assumptions are made for analytic simplicity and because the focus of the paper is on capital market theory and not the theory of consumer choice. Fama (1970b) in discrete time and Huang and Kreps (1985) in continuous time, analyze the problem for non-additive utilities. Although T is treated as known in the text, the analysis is essentially

the same for an uncertain lifetime with T a random variable. Cf. Richard (1975) and Merton (1971). The analysis is also little affected by making the direct-utility function "state dependent" (i.e., having U depend on other variables in addition to consumption and time).

27. This definition of a riskless security is purely technical and without normative significance. For example, investing solely in the riskless security will not allow for a certain consumption stream because $R(t)$ will vary stochastically over time. On the other hand, a T -period, riskless-in-terms-of-default coupon bond, which allows for a certain consumption stream is not a riskless security, because its one-period return is uncertain. For further discussion, see Merton (1973b).
28. It is assumed that all income comes from investment in securities. The analysis would be the same with wage income provided that investors can sell shares against future income. However, because institutionally this cannot be done, the "non-marketability" of wage income will cause systematic effects on the portfolio and consumption decisions.
29. Many non-Markov stochastic processes can be transformed to fit the Markov format by expanding the number of state variables. Cf. Cox and Miller (1968, pp. 16-18). To avoid including "surplus" state variables, it is assumed that $\{S(t)\}$ represent the minimum number of variables necessary to make $\{Z_j(t+1)\}$ Markov.
30. Cf. Dreyfus (1965) for the dynamic programming technique. Sufficient conditions for existence are described in Bertsekas (1974). Uniqueness of the solutions is guaranteed by: (1) strict concavity of U and B ; (2) no redundant securities; and (3) no arbitrage opportunities. See Cox, Ingersoll and Ross (1985a) for corresponding conditions in the continuous-time version of the model.
31. See Fama (1970b) for a general discussion of these conditions.
32. See Latane (1959), Markowitz (1976), and Rubinstein (1976) for arguments in favor of this view, and Samuelson (1971), Goldman (1974), and Merton and Samuelson (1974) for arguments in opposition to this view.
33. These introductory paragraphs are adapted from Merton (1975a, pp. 662-663).
34. If investor behavior were invariant to h , then investors would choose the same portfolio if they were "frozen" into their investments for ten years as they would if they could revise their portfolios everyday.
35. See Feller (1966), Itô and McKean (1964) and Cox and Miller (1968).
36. (7.1) is a short-hand expression for the stochastic integral:

$$S_i(t) = S_i(0) + \int_0^t G_i(S, \tau) d\tau + \int_0^t H_i(S, \tau) dq_i \quad ,$$

where $S_j(t)$ is the solution to (7.1) with probability one. For a general discussion and proofs, see Itô and McKean (1964), McKean (1969), McShane (1974) and Harrison (1985).

37. $\int_0^t dq_j = q_j(t) - q_j(0)$ is normally distributed with a zero mean and variance equal to t .
38. See Feller (1966, pp. 320-321); Cox and Miller (1968, p. 215). The transition probabilities will satisfy the Kolmogorov or Fokker-Planck partial differential equations.
39. Merton (1971, p. 377), dP_j/P_j in continuous time corresponds to $Z_j(t+1) - 1$ in the discrete time analysis.
40. $r(t)$ corresponds to $R(t) - 1$ in the discrete-time analysis, and is the "force-of-interest," continuous rate. While the rate earned between t and $(t + dt)$, $r(t)$, is known with certainty as of time t , $r(t)$ can vary stochastically over time.
41. Unlike in the Arrow-Debreu model, for example, it is not assumed here that the returns are necessarily completely described by the changes in the state variables, dS_i , $i = 1, \dots, m$, i.e., the dZ_j need not be instantaneously perfectly correlated with some linear combination of dq_1, \dots, dq_m . Rather, it is only assumed that $(dP_1/P_1, \dots, dP_n/P_n, dS_1, \dots, dS_m)$ is Markov in $S(t)$.
42. See Merton (1972, p. 381) and Kushner (1967, Ch. IV, Theorem 7).
43. See Merton (1971, p. 384-388). It is also shown there that the returns will be lognormal on the risky fund which, together with the riskless security, spans the efficient portfolio set. Joint lognormal distributions are not spherically-symmetric distributions.
44. As will be shown in Section 9, this case is similar in spirit to the Arrow-Debreu complete-markets model.
45. This behavior obtains even when the return on fund $\#(2 + i)$ is not instantaneously perfectly correlated with dS_i .
46. For further discussion of this analysis, descriptions of specific sources of uncertainty, and extensions to discrete-time examples, see Merton (1973b, 1975a, 1975b). Breeden (1979) shows that similar behavior obtains in the case of multiple consumption goods with uncertain relative prices. However, C^* is a vector and J_W is the "shadow" price of the "composite" consumption bundle. Hence, the corresponding derived "hedging" behavior is to minimize the unanticipated variations in J_W .
47. (8.2) is a classic linear partial-differential equation of the parabolic type. If σ^2 is a continuous function, then there exists a unique solution that satisfies boundary conditions (8.3). The usual method for solving this equation is Fourier transforms.

48. Itô's Lemma is for stochastic differentiation, the analog to the Fundamental Theorem of the calculus for deterministic differentiation. For a statement of the Lemma and applications in economics, see Merton (1971,1973a). For its rigorous proof, see McKean (1969,p. 44).
49. Although small, the transactions costs faced by even large securities-trading firms are, of course, not zero. As is evident from the work of Kandel and Ross (1983), Constantinides (1984), and Leland (1985), the analysis of optimal portfolio selection and derivative-security pricing with transactions costs is technically complex. Development of a satisfactory theory of equilibrium security prices in the presence of such costs promises to be even more complicated, because it requires a simultaneous determination of prices and the least-cost form of market structure and financial intermediation.
50. Although equilibrium condition (9.2) will apply in the cases of state-dependent direct utility, $U(C,S,t)$, and utilities which depend on the path of past consumption, (9.8) will no longer obtain under these conditions.
51. Under mild regularity conditions on the functions H , η , g , α , σ and r , a solution exists and is unique.
52. Of course, with a continuum of states, the price of any one Arrow-Debreu security, like the probability of a state, is infinitesimal. The solution to (9.9) is analogous to a probability density and therefore, the actual Arrow-Debreu price is $\pi(S,t)d\bar{S}_k$. The limiting boundary condition for (9.9) is a vector, generalized Dirac delta function.
53. The derivation can be generalized to the case in Section 7, where dZ_{m+1}, \dots, dZ_n are not perfectly correlated with the state variables by adding the mean-variance efficient portfolio to the $m+1$ portfolios used here. Cox, Ingersoll, and Ross (1985a) present a more-general version of partial differential equation (9.9), which describes general-equilibrium pricing for all assets and securities in the economy. See Duffie (1986) for discussion of existence of equilibrium in general models.
54. See Samuelson (1970) and Merton and Samuelson (1974). Merton (1975a,p. 663) discusses special cases in which the limiting discrete-time solutions do not approach the continuous-time solutions.
55. Merton (1982b) discusses in detail, the economic assumptions required for the continuous-time methodology. Moreover, most of the mathematical tools for manipulation of these models are derived using only elementary probability theory and the calculus.
56. If the underlying dynamics of the system includes Poisson-driven processes with discontinuous sample paths, then the resulting equilibrium prices will satisfy a mixed partial difference-differential equation. In the case of non-Markov path-dependent processes, the valuation conditions cannot be represented as a partial differential equation.

REFERENCES

- Arrow, K.J. (1953), "Le Rôle des Valeurs Boursières pour la repartition la Meilleure des Risques," Econometrie, Colloques Internationaux du Centre National de la Recherche Scientifique, Vol. XI, Paris, 41-47.
- Arrow, K.J. (1964), "The Role of Securities in the Optimal Allocation of Risk Bearing," Review of Economic Studies, 31:91-96.
- Bawa, V.S. (1975), "Optimal Rules for Ordering Uncertain Prospects," Journal of Financial Economics, 2:95-121.
- Bertsekas, D.P. (1974), "Necessary and Sufficient Conditions for Existence of An Optimal Portfolio," Journal of Economic Theory, 8:235-247.
- Bhattacharya, S. (forthcoming), "Financial Markets and Incomplete Information: A Review of Some Recent Developments," in S. Bhattacharya and G.M. Constantinides (eds.), A Course in Financial Theory, Totowa, NJ: Littlefield and Adams.
- Black, F. (1972), "Capital Market Equilibrium With Restricted Borrowing," Journal of Business, 45:444-455.
- Black, F. (1986), "Noise," Journal of Finance, 41:529-543.
- Black, F. and M. Scholes (1973), "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, 81:637-654.
- Borch, K. (1969), "A Note on Uncertainty and Indifference Curves," Review of Economic Studies, 36:1-4.
- Breeden, D.T. (1979), "An Intertemporal Asset Pricing Model With Stochastic Consumption and Investment Opportunities," Journal of Financial Economics, 7:265-296.
- Breeden, D.T. and R. Litzenberger (1978), "Prices of State-Contingent Claims Implicit in Option Prices," Journal of Business, 51:621-651.
- Cass, D. and J.E. Stiglitz (1970), "The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds," Journal of Economic Theory, 2:122-160.
- Chamberlain, G. (1983), "A Characterization of the Distributions That Imply Mean-Variance Utility Functions," Journal of Economic Theory, 29:185-201.
- Chamberlain, G. and M. Rothschild (1983), "Arbitrage and Mean-Variance Analysis on Large Asset Markets," Econometrica, 51:1281-1301.

- Chen, N. and J.E. Ingersoll (1983), "Exact Pricing in Linear Factor Models With Finitely-Many Assets: A Note," Journal of Finance, 38:985-988.
- Constantinides, G. (1984), "Capital Market Equilibrium With Transactions Costs," Working Paper CRSP, University of Chicago (October).
- Cox, D.A. and H.D. Miller (1968), The Theory of Stochastic Processes, New York: John Wiley & Sons, Inc.
- Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985a), "An Intertemporal General Equilibrium Model of Asset Prices," Econometrica, 53:363-384.
- Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985b), "A Theory of the Term Structure of Interest Rates," Econometrica, 53:385-408.
- Cox, J.C. and M. Rubinstein (1985), Options Markets, Englewood Cliffs, New Jersey: Prentice Hall.
- Debreu, G. (1959), Theory of Value, New York: John Wiley & Sons, Inc.
- Dhrymes, P., I. Friend, and N. Gultekin (1984), "A Critical Examination of the Empirical Evidence on the Arbitrage Pricing Theory," Journal of Finance, 39:323-347.
- Dhrymes, P., I. Friend, and N. Gultekin (1985), "New Tests of the APT and Their Implications," Journal of Finance, 40:659-674.
- Dreyfus, S.E. (1965), Dynamic Programming and the Calculus of Variations, New York: Academic Press.
- Duffie, D. (1986), "Stochastic Equilibria: Existence, Spanning Number, and the 'No Expected Financial Gain from Trade' Hypothesis," Econometrica 54:1161-1184.
- Duffie, D. and C. Huang (1985), "Implementing Arrow-Debreu Equilibria by Continuous Trading of a few Long-Lived Securities," Econometrica, 53:1337-1356.
- Duffie, D. and C. Huang (forthcoming), "Multiperiod Securities Markets With Differential Information: Martingales and Resolution Times," Journal of Mathematical Economics.
- Dybvig, P. and S.A. Ross (1982), "Portfolio Efficient Sets," Econometrica, 50:1525-1546.
- Fama, E. (1965), "The Behavior of Stock Market Prices," Journal of Business, 38:34-105.
- Fama, E. (1970a), "Efficient Capital Markets: A Review of Theory and Empirical Work," Journal of Finance, 25:383-417.

- Fama, E. (1970b), "Multiperiod Consumption-Investment Decisions," American Economic Review, 60:163-174.
- Fama, E. (1978), "The Effects of a Firm's Investment and Financing Decisions on the Welfare of Its Securityholders," American Economic Review, 68.
- Farrar, D.E. (1962), The Investment Decision Under Uncertainty, Englewood Cliffs, New Jersey: Prentice Hall.
- Farrell, J.L. (1974), "Analyzing Covariation of Returns to Determine Homogeneous Stock Groupings," Journal of Business, 47:186-207.
- Feeney, G.J., and D. Hester (1967), "Stock Market Indices: A Principal Components Analysis," in: D. Hester and J. Tobin, (eds.), Risk Aversion and Portfolio Choice, New York: John Wiley & Sons, Inc.
- Feldstein, M.S. (1969), "Mean-Variance Analysis in the Theory of Liquidity Preference and Portfolio Selection," Review of Economic Studies, 36:5-12.
- Feller, W. (1966), An Introduction to Probability Theory and Its Applications, Vol. 2, New York: John Wiley & Sons, Inc.
- Fischer, S. (1972), "Assets, Contingent Commodities, and the Slutsky Equations," Econometrica, 40:371-385.
- Fischer, S. and R.C. Merton (1984), "Macroeconomics and Finance: The Role of the Stock Market," in K. Brunner and A.H. Melzer (eds.), Essays on Macroeconomic Implications of Financial and Labor Markets and Political Processes, Vol. 21, Amsterdam: North-Holland Publishing Company.
- Friend, I. and J. Bicksler (eds.) (1975), Studies in Risk and Return, Vols. I&II, Cambridge, MA: Ballinger.
- Goldman, M.B. (1974), "A Negative Report on the 'Near-Optimality' of the Max-Expected-Log Policy as Applied to Bounded Utilities for Long-Lived Programs," Journal of Financial Economics, 1:97-103.
- Grossman, S. (1976), "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," Journal of Finance, 31:573-585.
- Grossman, S. and J.E. Stiglitz (1976), "Information and Competitive Price Systems," American Economic Review, 66:246-253.
- Hadar, J. and W.R. Russell (1969), "Rules for Ordering Uncertain Prospects," American Economic Review, 59:25-34.
- Hadar, J. and W.R. Russell (1971), "Stochastic Dominance and Diversification," Journal of Economic Theory, 3:288-305.
- Hakansson, N. (1970), "Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions," Econometrica, 38:587-607.

- Hanoch, G. and H. Levy (1969), "The Efficiency Analysis of Choices Involving Risk," Review of Economic Studies, 36:335-346.
- Hansen, L. (1985), "Auctions With Contingent Payments," American Economic Review, 75:862-865.
- Hardy, G.H., J.E. Littlewood, and G. Pölya (1959), Inequalities, Cambridge: Cambridge University Press.
- Harrison, J.M. (1985), Brownian Motion and Stochastic Flow Systems, New York: John Wiley & Sons.
- Harrison, J.M. and D. Kreps (1979), "Martingales and Arbitrage in Multiperiod Securities Markets," Journal of Economic Theory, 20:381-408.
- Herstein, I. and J. Milnor (1953), "An Axiomatic Approach to Measurable Utility," Econometrica, 21:291-297.
- Hirshleifer, J. (1965), "Investment Decision Under Uncertainty: Choice-Theoretic Approaches," Quarterly Journal of Economics, 79:509-536.
- Hirshleifer, J. (1966), "Investment Decision Under Uncertainty: Applications of the State-Preference Approach," Quarterly Journal of Economics, 80:252-277.
- Hirshleifer, J. (1970), Investment, Interest and Capital, Englewood Cliffs, New Jersey: Prentice-Hall.
- Hirshleifer, J. (1973), "Where are We in The Theory of Information?" American Economic Review, 63:"31-39.
- Hogarth, R.M. and M.W. Reder, eds. (1986), "The Behavioral Foundations of Economic Theory," Journal of Business No. 4, Part 2, 59:S181-S505.
- Huang, C. (1985), "Information Structure and Equilibrium Asset Prices," Journal of Economic Theory, 53:33-71.
- Huang, C. (1987), "An Intertemporal General Equilibrium Asset Pricing Model: The Case of Diffusion Information," Econometrica, 55:117-142.
- Huang, C. and K. Kreps (1985), "Intertemporal Preferences With a Continuous Time Dimension: An Exploratory Study," Mimeo, Massachusetts Institute of Technology (December).
- Itô, K. and H.P. McKean, Jr. (1964), Diffusion Processes and Their Sample Paths, New York: Academic Press.
- Jensen, M.C. (ed.) (1972), Studies in the Theory of Capital Markets, New York: Praeger Publishers.

- Jensen, M.C. (1972), "Capital Markets: Theory and Evidence," Bell Journal of Economics and Management Science, 3:357-398.
- Kandel, S. and S. Ross (1983), "Some Intertemporal Models of Portfolio Selection With Transactions Costs," Working Paper 107, CRSP, University of Chicago (September).
- King, B.R. (1966), "Market and Industry Factors in Stock Price Behavior," Journal of Business, 39, Supplement:139-190.
- Kuhn H.W. and A.W. Tucker (1951), "Nonlinear Programming," in: J. Neyman (ed.), Proceedings of the Second Berkeley Symposium of Mathematical Statistics and Probability, Berkeley: University of California Press.
- Kushner, H.J. (1967), Stochastic Stability and Control, New York: Academic Press.
- Latane, H. (1959), "Criteria for Choice Among Risky Ventures," Journal of Political Economy, 67:144-155.
- Leland, H. (1972), "On the Existence of Optimal Policies Under Uncertainty," Journal of Economic Theory, 4:35-44.
- Leland, H. (1985), "Option Pricing and Replication With Transactions Costs," Journal of Finance, 40:1283-1301.
- Lintner, J. (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," Review of Economics and Statistics, 47:13-37.
- Livingston, M. (1977), "Industry Movements of Common Stocks," Journal of Finance, 32:861-874.
- Markowitz, H. (1959), Portfolio Selection: Efficient Diversification of Investment, New York: John Wiley & Sons, Inc.
- Markowitz, H. (1976), "Investment for the Long Run: New Evidence for an Old Rule," Journal of Finance, 31:1273-1286.
- Mason, S. (1981), "Consumption and Investment Incentives Associated With Welfare Programs," Working Paper 79-34, Harvard Business School, Cambridge, Massachusetts (December).
- Mason, S. and R.C. Merton (1985), "The Role of Contingent Claims Analysis in Corporate Finance," in E. Altman and M. Subrahmanyam (eds.), Recent Advances in Corporate Finance, Homewood, Illinois: Richard D. Irwin.
- McKean, Jr., H.P. (1969) Stochastic Integrals, New York: Academic Press.

- McShane, E.J. (1974), Stochastic Calculus and Stochastic Models, New York: Academic Press.
- Merton, R.C. (1969), "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case," Review of Economics and Statistics, 51:247-257.
- Merton, R.C. (1971), "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," Journal of Economic Theory, 3:373-413.
- Merton, R.C. (1972), "An Analytic Derivation of the Efficient Portfolio Frontier," Journal of Financial and Quantitative Analysis, 7:1851-1872.
- Merton, R.C. (1973a), "Theory of Rational Option Pricing," Bell Journal of Economics and Management Science, 4:141-183.
- Merton, R.C. (1973b), "An Intertemporal Capital Asset Pricing Model," Econometrica, 41:867-887.
- Merton, R.C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," Journal of Finance, 29:449-470.
- Merton, R.C. (1975a), "Theory of Finance From the Perspective of Continuous Time," Journal of Financial and Quantitative Analysis, 10:659-674.
- Merton, R.C. (1975b), "A Reexamination of the Capital Asset Pricing Model," in: Friend and Bicksler (eds.) (1975).
- Merton, R.C. (1977), "On the Pricing of Contingent Claims and the Modigliani-Miller Theorem," Journal of Financial Economics, 5:241-249.
- Merton, R.C. (1982a), "On the Microeconomic Theory of Investment Under Uncertainty," in: K. Arrow and M. Intriligator (eds.), Handbook of Mathematical Economics, Volume II, Amsterdam: North-Holland.
- Merton, R.C. (1982b), "On the Mathematics and Economic Assumptions of Continuous-Time Financial Models," in: W.F. Sharpe (ed.), Financial Economics: Essays in Honor of Paul Cootner, Englewood Cliffs: Prentice Hall, Inc.
- Merton, R.C. (1987a), "On the Current State of the Stock Market Rationality Hypothesis," in: S. Fischer and R. Dornbusch (eds.), Macroeconomics and Finance: Essays in Honor of Franco Modigliani, Cambridge, Massachusetts: MIT Press.
- Merton, R.C. (1987b), "A Simple Model of Capital Market Equilibrium With Incomplete Information," Journal of Finance, 42:_____.
- Merton, R.C. and P.A. Samuelson (1974), "Fallacy of the Log-Normal Approximation to Optimal Portfolio Decision Making Over Many Periods," Journal of Financial Economics, 1:67-94.

- Miller, M.H. (1977), "Debt and Taxes," Journal of Finance, 32:261-276.
- Modigliani, F. and M.H. Miller (1958), "The Cost of Capital, Corporation Finance, and the Theory of Investment," American Economic Review, 48:261-297.
- Mossin, J. (1966), "Equilibrium in a Capital Asset Market," Econometrica, 35:768-783.
- Myers, S.C. (1968), "A Time-State-Preference Model of Security Valuation," Journal of Financial and Quantitative Analysis, 3:1-33.
- Nielsen, L.T. (1986), "Mutual Fund Separation: Factor Structure and Robustness," Working Paper 86/87-2-3, Graduate School of Business, University of Texas at Austin (September).
- Parsons, J. and A. Raviv (1985), "Underpricing of Seasoned Issues," Journal of Financial Economics, 14:377-397.
- Pratt, J.W. (1964), "Risk Aversion in the Small and in the Large," Econometrica, 32:122-136.
- Radner, R. (1970), "Problems in the Theory of Markets Under Uncertainty," American Economic Review, 60:454-460.
- Radner, R. (1972), "Existence of Plans, Prices, and Price Expectations in a Sequence of Markets," Econometrica, 40:289-303.
- Richard, S. (1975), "Optimal Consumption, Portfolio, and Life Insurance Rules for An Uncertain Lived Individual in a Continuous-Time Model," Journal of Financial Economics, 2:187-204.
- Rock, K. (1986), "Why New Issues Are Underpriced," Journal of Financial Economics, 15:187-212.
- Roll, R. and S.A. Ross (1980), "An Empirical Investigation of the Arbitrage Pricing Theory," Journal of Finance, 35:1073-1103.
- Ross, S.A. (1976a), "Arbitrage Theory of Capital Asset Pricing," Journal of Economic Theory, 13:341-360.
- Ross, S.A. (1976b), "Options and Efficiency," Quarterly Journal of Economics, 90:75-89.
- Ross, S.A. (1978), "Mutual Fund Separation in Financial Theory: The Separating Distributions," Journal of Economic Theory, 17:254-286.
- Rothschild, M. (1986), "Asset Pricing Theories," in W.P. Heller, R.M. Starr and D.A. Starrett (eds.), Uncertainty, Information and Communication: Essays in Honor of Kenneth J. Arrow, Vol. III, Cambridge, England: Cambridge University Press.

- Rothschild, M. and J.E. Stiglitz (1970), "Increasing Risk I: A Definition," Journal of Economic Theory, 2:225-243.
- Rothschild, M. and J.E. Stiglitz (1971), "Increasing Risk II: Its Economic Consequences," Journal of Economic Theory, 3:66-84.
- Rubinstein, M. (1976), "The Strong Case for the Generalized Logarithmic Utility Model as the Premier Model of Financial Markets," Journal of Finance, 31:551-572.
- Samuelson, P.A. (1965), "Proof That Properly Anticipated Prices Fluctuate Randomly," Industrial Management Review, 6:41-49, Reprinted in Samuelson (1972).
- Samuelson, P.A. (1967), "General Proof That Diversification Pays," Journal of Financial and Quantitative Analysis, 2:1-13, Reprinted in Samuelson (1972).
- Samuelson, P.A. (1969), "Lifetime Portfolio Selection by Dynamic Stochastic Programming," Review of Economics and Statistics, 51:239-246, Reprinted in Samuelson (1972).
- Samuelson, P.A. (1970), "The Fundamental Approximation Theory of Portfolio Analysis in Terms of Means, Variances, and Higher Moments," Review of Economic Studies, 37:537-542, Reprinted in Samuelson (1972).
- Samuelson, P.A. (1971), "The 'Fallacy' of Maximizing the Geometric Mean in Long Sequences of Investing or Gambling," Proceedings of the National Academy of Sciences 68:2493-2496, Reprinted in Samuelson (1972).
- Samuelson, P.A. (1972), in: R.C. Merton (ed.), The Collected Scientific Papers of Paul A. Samuelson, Vol. III, Cambridge: MIT Press.
- Samuelson, P.A. (1977), "St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described," Journal of Economic Literature, 15:24-55.
- Sharpe, W. (1964), "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," Journal of Finance, 19:425-442.
- Sharpe, W. (1970), Portfolio Theory and Capital Markets, New York: McGraw-Hill.
- Smith, Jr., C.W. (1976), "Option Pricing: A Review," Journal of Financial Economics, 3:3-52.
- Stiglitz, J.E. (1969), "A Re-Examination of the Modigliani-Miller Theorem," American Economic Review, 59:78-93.
- Stiglitz, J.E. (1974), "On the Irrelevance of Corporate Financial Policy," American Economic Review, 64:851-886.

Tobin, J. (1958), "Liquidity Preference as Behavior Towards Risk," Review of Economic Studies, 25:68-85.

Tobin, J. (1969), "Comment on Borch and Feldstein," Review of Economic Studies, 36:13-14.

Trzcinka, C. (1986), "On the Number of Factors in the Arbitrage Pricing Model," Journal of Finance, 41:347-368.

Varian, H. (1985), "Divergence of Opinion in Complete Markets: A Note," Journal of Finance, 40:309-318.

von Neumann, J. and O. Morgenstern (1947), Theory of Games and Economic Behavior, 2nd edn., Princeton: Princeton University Press.

Wallace, N. (1981), "A Modigliani-Miller Theorem for Open-Market Operations," American Economic Review, 71:267-274.