Philip H. Dybvig
Yale University

Jonathan E. Ingersoll, Jr.
University of Chicago

Mean-Variance Theory in Complete Markets*

The complete-markets model of Arrow and Debreu and the mean-variance capital asset pricing model (CAPM) are two paradigms of risky asset markets. This paper provides an integration of these two theories. We prove that the standard mean-variance separation theorem obtains in a complete market only if all investors have quadratic utility. In addition, the familiar CAPM pricing relation can hold for all assets in a complete market only if arbitrage opportunities exist. On the positive side, we show that the validity of this relation for the primary assets in a complete market is unaffected by the distribution of returns on the created financial assets.

I. Introduction

Two paradigms in the pricing of risky assets are the "complete markets" model of Arrow and Debreu and mean-variance pricing as embodied in the capital asset pricing model (CAPM). The former is primarily of theoretical interest. In fact, prior to the recent work of Banz and Miller (1978) and Breeden and Litzenberger (1978) there was little thought or hope of practical applications. On the other hand, the hegemony of the CAPM is due mostly to its apparent ease of applicability and, to a lesser extent, its empirical justification.

In a competitive market with no taxes, transactions costs, or institutional constraints on portfolio holdings, the allocation of the available assets will be Pareto optimal. The end-of-period distribution of wealth (or the consumption goods), however, will typically be Pareto optimal only when there is available an Arrow-Debreu security for each contingency. Since, by definition, a Pareto superior reallocation is weakly preferred by all investors, a complete set of Arrow-Debreu markets should naturally arise provided they are not too costly to establish.

* We would like to thank George Constantinides, Doug Diamond, and Merton Miller for helpful comments on earlier drafts of this paper.

(*Journal of Business*, 1982, vol. 55, no. 2)

0021-9398/82/5502-0003$01.50

As a general rule, mean-variance models have been derived in such perfect economies; therefore the link between these theories is of great interest. An integration could yield a model with both the applicability of the CAPM and the theoretical attractions of Pareto optimal markets. Unfortunately, this synthesis is hampered by the necessity to meet the distributional assumptions of both models. There must be as many linearly independent assets as resultant states, and the returns on all must be related by the separating distribution requirement of Ross (1978*a*).

In Section II we show that the requirements of mean-variance pricing of all Arrow-Debreu securities typically permit arbitrage opportunities to arise. We also demonstrate that the standard portfolio separation result of mean-variance analysis obtains only for the unsatisfactory case of universal quadratic utility. In Sections III and IV we investigate the applicability of mean-variance pricing to a subset of assets. In particular, we address the robustness of the CAPM in the presence of financial assets. In Section III we point out that the introduction of a financial asset into an incomplete-markets economy at a CAPM equilibrium will destroy that equilibrium. In Section IV, however, we demonstrate conditions under which the CAPM equilibrium will hold for a subset of assets whose returns conform to its assumptions. We argue that this result is a broadly applicable justification for ignoring financial assets when applying the CAPM despite the damaging findings in Section III.

II. Mean-Variance Pricing in Complete Markets—Arbitrage and Quadratic Utility

The first step in the integration of mean-variance pricing into an Arrow-Debreu market setting is the formation of a common framework for analysis. In the Arrow-Debreu framework, preferences are expressed as a utility function defined over (some subset of) the commodity space. Furthermore, a given commodity can be defined very narrowly. In principle, the partition of states may be made sufficiently fine so that any and all risks are insurable through the purchase of pure state contingent claims, Arrow-Debreu securities, given subsequent spot markets.

Under the mean-variance analysis of the CAPM, there is effectively only a single good, so utility is a function solely of end-of-period wealth. Consequently the only important characterization of a "state" is the pattern of resultant asset values. To integrate the two theories we adopt the latter convention and use this broader partition of states. For our purposes, then, a terminal state is equivalent to a given vector of payoffs (or returns) on the available assets. Two states are distinct if and only if they are distinguishable through realized asset returns alone.

If there are as many assets with linearly independent returns as there are (broadly defined) states, markets are complete at least for the purposes of unambiguously valuing the existing assets and any potential derivative assets. As we have assumed that each investor's utility depends only on his terminal wealth, this relatively weakly complete market is sufficient to guarantee a Pareto optimal allocation of end-of-period wealth across investors.

Throughout most of this paper we will use the following set of maintained assumptions:

(A1) Perfect markets: The markets for all assets are perfect with no taxes or transactions costs. Unlimited borrowing and short selling are permitted with full use of the proceeds. Each asset is infinitely divisible.
(A2) Competition: All investors act as price takers in all markets.
(A3) Homogeneous expectations: All investors have identical probability beliefs.
(A4) State-independent utility: Investors are risk averse and maximize the expectation of a von Neuman–Morgenstern utility function which depends solely on wealth.

When adopted the assumption of complete markets will mean the following:

(A5) Complete markets: Each competitive investor can obtain any pattern of returns through the purchase of marketed assets (subject only to his own budget constraint). If the number of outcome states is finite, markets are complete if the number of marketed assets with linearly independent returns is equal to the number of states.

These assumptions are standard and require no discussion, with one exception. The presumption that utility depends only on end-of-period wealth is not a trivial one. In a single-period context this assumption is standard and in essence reduces the Arrow-Debreu requirement that every risk be insurable to the more modest restriction that only the risks caused by uncertain changes in asset prices must be insurable. In a multiperiod framework, however, the assumption that the direct utility function depends only on the quantity of wealth consumed is insufficient for our purposes. The indirect utility function must depend only on wealth and time. In addition to the former a nonstochastic investment opportunity set or logarithmic utility is also required for the latter.[1]

Under assumptions A1 through A4 it is well known that the CAPM will obtain if either each investor's utility function is quadratic over the

1. In the presence of a stochastic investment opportunity set, investors will generally modify their demand for various assets to take into consideration potential hedging opportunities (see Merton 1973; Long 1974; Constantinides 1980, 1982).

relevant range of outcomes or all asset returns are drawn from one of the class of "separating distributions" defined by Ross (1978*a*) (which includes multivariate normals). These are the only two known sufficient conditions for mean-variance pricing. Indeed each assumption alone is also necessary in the absence of any other assumptions about distributions or tastes. Nevertheless, a validation of the CAPM might yet be possible under weaker but simultaneous restrictions on both returns distributions *and* utility functions. In a similar regard, Samuelson (1970) and Samuelson and Merton (1975) set forth conditions under which mean-variance analysis is an approximation to expected utility maximization.

To maintain as much generality as possible, we shall simply assume that each marketed asset in the economy has an expected return and covariance with the market which is consistent with the CAPM. That is,

$$E(\tilde{X}_i) - R = \lambda \operatorname{cov}(\tilde{X}_i, \tilde{X}_m), \tag{1}$$

where $\tilde{X}_i$ is the random return (one plus the rate of return) on asset i, $\tilde{X}_m$ is the random return on the market portfolio, $R > 0$ is the return on the riskless asset[2] (one plus the interest rate), and $\lambda \equiv [E(\tilde{X}_m) - R]/\operatorname{var}(\tilde{X}_m) > 0$ is the market price of risk.

Denote the random end-of-period value per share of asset i by $\tilde{V}_i$ and its current price per share by P_i. Then $\tilde{X}_i \equiv \tilde{V}_i/P_i$ and, as P_i is observable and therefore not random, (1) may be rewritten as

$$E(\tilde{V}_i)/P_i - R = \lambda \operatorname{cov}(\tilde{V}_i, \tilde{X}_m)/P_i. \tag{2}$$

Solving for the current price, we get

$$P_i = [E(\tilde{V}_i) - \lambda \operatorname{cov}(\tilde{V}_i, \tilde{X}_m)]/R = E\,\{\tilde{V}_i\,[1 - \lambda(\tilde{X}_m - \bar{X}_m)]\}/R, \tag{3}$$

where $\bar{X}_m \equiv E(\tilde{X}_m)$ is the expected return on the market.

We have only assumed that (1) holds for marketed assets. However, it is easy to show that (1) and (3) must hold for all payoffs. If a random payoff $\tilde{Z}$ can be obtained through buying and selling only marketed assets, that is, there is a portfolio of shares α_i such that[3]

$$\tilde{Z} \equiv \Sigma\, \alpha_i \tilde{V}_i, \tag{4}$$

2. We use the Sharpe-Lintner version of the CAPM here rather than the Black version in which the expected return on a portfolio uncorrelated with the market replaces the riskless return. In a complete market a riskless asset must exist by definition.

3. We are assuming implicitly that there are a finite number of assets and, a fortiori, a finite number of states. More generally, the portfolio vector would be replaced by a measure over all assets. To allow for short sales the measure would not need to be positive. Note that, as in defining demand functions, many net trades are priced even though they fail to satisfy aggregate resource constraints.

then to prevent arbitrage the price $P(Z)$ of the payoff Z must be $\Sigma\, \alpha_i P_i$. This price can be calculated by

$$\begin{aligned} P(Z) &= \Sigma\, \alpha_i P_i \\ &= \Sigma\, \alpha_i E\{\tilde{V}_i\, [1 - \lambda(\tilde{X}_m - \overline{X}_m)]\}/R \\ &= E\{(\Sigma\, \alpha_i \tilde{V}_i)[1 - \lambda(\tilde{X}_m - \overline{X}_m)]\}/R \\ &= E\{\tilde{Z}[1 - \lambda(\tilde{X}_m - \overline{X}_m)]\}/R. \end{aligned} \tag{5}$$

The result above proves that the mean-variance valuation in (3) is valid for all payoffs, not only those which come directly from marketed assets.

Equation (5) defines the unique linear pricing rule consistent with mean-variance theory and complete markets (see Ross 1978*b*). Unfortunately, it is not a positive pricing operator. That is, it can assign a negative value to a payoff which is strictly nonnegative. This actually occurs for any payoff which is positive in some states in which $\tilde{X}_m > \overline{X}_m + 1/\lambda$ and is zero in all other states. Specifically, consider the payoff

$$\tilde{Z}^* \equiv \begin{cases} [\lambda(\tilde{X}_m - \overline{X}_m) - 1]^{-1} & \text{if } \tilde{X}_m > \overline{X}_m + 1/\lambda \\ 0 & \text{if } \tilde{X}_m \leq \overline{X}_m + 1/\lambda. \end{cases} \tag{6}$$

Then from (5),

$$\begin{aligned} P(Z^*) &= E\{\tilde{Z}^*[1 - \lambda(\tilde{X}_m - \overline{X}_m)]\}/R \\ &= E(-1 \,|\, \tilde{X}_m > \overline{X}_m + 1/\lambda)/R \\ &= -\text{prob}\ (\tilde{X}_m > \overline{X}_m + 1/\lambda)/R. \end{aligned} \tag{7}$$

Thus, provided only that prob $(\tilde{X}_m > \overline{X}_m + 1/\lambda) > 0$ (i.e., the return on the market can achieve sufficiently high values), we can create an arbitrage opportunity. One such opportunity is simply the purchase, at a negative price, of $\tilde{Z}^*$. This consists of a cash inflow currently and no cash outflows, but possibly cash inflows, next period. Another example of an arbitrage possibility would be the purchase of a call option on the market with an exercise price in excess of $\overline{X}_m + 1/\lambda$. This option also must have a negative price for the same reason.

Formally we have just proved the following theorem.

Theorem 1: Suppose that (i) mean-variance pricing holds for all assets, that is, $E(\tilde{X}_i) - R = \lambda \operatorname{cov}(\tilde{X}_i, \tilde{X}_m)$ with $R, \lambda > 0$; (ii) markets are complete so that any payoff across states can be purchased as some portfolio of marketed securities; and (iii) the market portfolio generates sufficiently large returns in some state(s), that is, prob $(\tilde{X}_m > \overline{X}_m + 1/\lambda) > 0$. Then there exists an arbitrage possibility.

Proof: The proof follows from the constructive example in equation (7). Q.E.D.

For the purposes of this theorem we do not require the assumption that investors' utilities depend only on wealth (assumption A4). Thus theorem 1 holds in a multiperiod, multigood economy as well.

Because arbitrage possibilities preclude the existence of an equilibrium among nonsatiated investors, we have the following corollary to theorem 1.

Corollary: If in equilibrium any investor's terminal wealth is below his level of satiation in all states and conditions ii and iii of theorem 1 hold, mean-variance pricing cannot obtain for all assets.

Proof: If conditions ii and iii hold and mean-variance pricing is valid, by theorem 1 arbitrage opportunities exist. Any nonsatiated investor would employ one of these opportunities to increase his expected utility which shows that previously we could not have been in equilibrium. Q.E.D.

Condition iii of theorem 1 and its corollary will hold whenever the market return density has an unbounded tail in the positive direction. This is a relatively mild restriction, and the multivariate normal model is only the most familiar case for which it is true.

We have shown that mean-variance pricing of all assets in a complete market with unbounded outcomes on the market portfolio is impossible unless each investor has achieved or surpassed his point of satiation. Of course, it is also possible to construct mean-variance models in which the market's return is bounded below the limit given in theorem 1. Theorem 2 below is useful in establishing conditions under which such a bounded outcome equilibrium could be established. Specifically, it is proved that, in a complete market, the standard portfolio separation result of mean-variance pricing holds only among investors with quadratic utility.

Theorem 2: Suppose that as in theorem 1, (i) mean-variance pricing holds and (ii) markets are complete. Further assume that (iii) some von Neuman–Morgenstern expected utility maximizer possessed of a utility function $U(W)$ with continuous first derivative chooses a portfolio in which the risky assets are held in their market proportions (i.e., $\tilde{W} \equiv \alpha\tilde{X}_m + (W_0 - \alpha)R$, for $\alpha \neq 0$). Then $U'(W)$ is linear on the set of values taken on by the agent's wealth, namely, on $\{\alpha Y + (W_0 - \alpha)R \mid Y \in \text{supp}\,(\tilde{X}_m)\}$.[4] Furthermore, if (iv) the return on the market $\tilde{X}_m$ takes on all real values in some open interval, then $U(W)$ is quadratic in the corresponding range for W.

Proof: By condition ii the market is complete. Therefore the first-order condition for any investor's optimal portfolio is that the expected marginal utility in each state is proportional to that state's Arrow-

4. Precisely supp $(\tilde{X}_m)$ is the closure of the set of values taken on by $\tilde{X}_m$ with positive probability or positive density, which is the set of values Y for which the distribution function of X_m is nonconstant in every neighborhood of Y.

Debreu price. Condition i and equation (5) require that each state price be linear in the market return realized in that state[5]

$$\begin{aligned}\pi(s)U'[W(s)] &= \gamma p(s) \\ &= \gamma\pi(s)\{1 - \lambda[X_m(s) - \bar{X}_m]\}/R,\end{aligned} \tag{8}$$

where $\pi(s)$ is the probability or density for the occurrence of state s and $p(s)$ is the state price. By condition iii the portfolio payoff in state s is

$$W(s) = \alpha X_m(s) + (W_0 - \alpha)R. \tag{9}$$

Solving (9) for $X_m(s)$ and combining with (8) gives

$$U'[W(s)] = \frac{\gamma}{R}\left(1 + \lambda\bar{X}_m - \lambda\,\frac{W_0 - \alpha}{\alpha}\,R\right) - \frac{\gamma\lambda}{\alpha R}\,W(s). \tag{10}$$

From (10) marginal utility is linear on the support of $\tilde{W}$. If, in addition, condition iv holds, the support of $\tilde{W}$ is some open interval and $U'(\cdot)$ is linear everywhere in that interval. Equivalently, utility is quadratic in that interval. Q.E.D.

Loosely interpreted, theorem 2 says that for standard two-fund separation to obtain in a complete mean-variance market, all investors must have quadratic utility or utility which is indistinguishable from quadratic on the relevant range of outcomes. Thus it provides a sense in which quadratic utility is necessary as well as sufficient for portfolio separation under mean-variance pricing.

Ross (1978*a*) has characterized all returns distributions under which universal two-fund portfolio separation obtains. Each of these distributions also results in mean-variance pricing provided these expectations exist. From theorem 2 mean-variance pricing and separation can obtain in a complete market only for "quadratic-like" utility. One separating distribution under which any utility function is quadratic-like is for an outcome space with only two states.[6] With just two states, any marginal utility function can be represented by the line through the only two obtainable values. Unfortunately, this case, which is important didactically, is of little practical consequence. Outside of this two-state case, it is only by accident that a nonquadratic utility function would be quadratic-like. With this exception, then, the bounded-outcome separating distributions, which were not eliminated by theorem 1, now can be generally precluded as a potential justification of mean-variance analysis in complete markets.

5. From eq. (5) the state price for a given state is $p(s) = E(Z_s(\sigma)\{1 - \lambda[X_m(\sigma) - \bar{X}_m]\})/R$, where $Z_s(\sigma) = 1$ for $\sigma = s$ and $Z_s(\sigma) = 0$ for $\sigma \neq s$. Thus $p(s) = \pi(s)\{1 - \lambda[X_m(s) - \bar{X}_m]\}/R$.

6. A two-state outcome space is spanned by the riskless asset and a single risky asset. Thus since all risky assets are perfectly correlated the economy trivially satisfies the separating-distribution requirement.

Together theorems 1 and 2 impose severe restrictions on potential mean-variance, complete-market economies. Apart from the two-state case and other degenerate cases, the only known[7] single-period model which allows mean-variance analysis in complete markets is that of universal quadratic utility. For this model theorem 2 is obviously satisfied, and theorem 1 does not require the returns to be bounded as quadratic utility investors can be satiated (or prefer less to more), in which case they will not capitalize on arbitrage opportunities. Of course, quadratic utility is not very attractive for exactly this reason as well as for its increasing absolute risk aversion.

In multiperiod models there is an alternative, the continuous-time lognormal diffusion model.[8] By continuity, the market return cannot take on "large" values over an instant. Therefore, the arbitrage in theorem 1 is no problem. Furthermore, over an instant the derived utility of wealth function, which is fully characterized by its first two derivatives, is locally "quadratic" so theorem 2 is satisfied as well. In fact, we know that with "continuous trading" and unrestricted short sales, these models conform to mean-variance pricing, permit no arbitrage,[9] and exhibit effectively complete markets. By the arguments of Ross (1976) it suffices to note that only options are required to complete the market. In this model all options are redundant assets and can be priced by the no-arbitrage technique originated by Black and Scholes (1973). Furthermore, their instantaneous expected returns are governed by the mean-variance relation.

The continuous-time lognormal model is sometimes used to justify mean-variance analysis as an approximate pricing scheme for short time intervals. We know that the instantaneous expected rates of return and covariances of assets can be approximated by the corresponding measures over short time intervals.[10] This would appear to be sufficient justification. However, in this model, the market achieves a return greater than any given level with positive probability over any time interval of positive length. Therefore by theorem 1 and no arbitrage, mean-variance pricing must be a very bad approximation for some assets. Hence one interpretation of theorem 1 is that when using discrete-time realizations of a lognormal diffusion, the validity of the approximation to mean-variance pricing depends not only on the brev-

7. It is possible that complete-market models may be found in which mean-variance pricing obtains but separation does not. Theorem 2 would not be applicable to them.

8. See, e.g., Merton 1973. For universal logarithmic utility the diffusion need only be locally lognormal, i.e., the means and covariances of returns need not be nonstochastic over time.

9. There is no arbitrage given certain technical restrictions on agents' choices of trading strategies (see Harrison and Kreps 1979).

10. The appropriateness of different "corresponding" measures is an issue on which we are currently at work. This question has direct bearing on the interpretation of many previous empirical tests of the CAPM.

ity of the time interval but also on the particular assets under consideration.

III. Financial Contracting and the Collapse of Mean-Variance Pricing

In the previous section we demonstrated that mean-variance analysis could not generally be used for all assets in a complete market except in the cases of diffusion or quadratic utility. In this section we begin our examination of the applicability of mean-variance analysis to a subset of assets in a complete market. In particular we consider the mean-variance pricing of the primary assets.

Primary assets are those which are in positive net supply to the investing public, that is, they are part of the market portfolio. Common stock is the prototype of a primary asset. Assets which are created, in zero net supply, by the investing public and are not in the market portfolio we term financial or derivative assets. Futures contracts, options, and the securities of financial intermediaries are examples. Note that this definition leaves some flexibility in the classification of assets as primary or financial. Our results are valid with any classification such that the market portfolio can be duplicated by holding only primary assets.

The distinction between primary and financial assets provides a natural point to begin our analysis for the following reason. Unless utility is quadratic some restrictions on all asset returns distributions must be imposed to derive mean-variance pricing. However, no restrictions can be placed on the possible distributions for the financial assets unless the freedom to enter into any arbitrary financial contract, which two or more investors care to create, is artificially abrogated. This, of course, is counter to the spirit of a complete market economy. On the other hand, models in which the returns on the primary assets are all of a given class (e.g., normal) are possible.[11]

Before proving our results we show how the introduction of financial assets into a "mean-variance" economy can upset the previous equilibrium. Theorem 1 and its corollary demonstrated implicitly that added financial assets will not always obey mean-variance pricing since if enough assets were added to complete the market, mean-variance pricing would not hold in equilibrium. We now show that there is no assurance that the primary assets will still be priced according to the CAPM after the addition of financial assets.

We examine below a very simple economy in which the returns on

11. This is obviously true in a single-period model when the end-of-period values of the primary assets are exogenous. Since current prices are observable, the returns on the primary assets are also "exogenous," given their initial equilibrium prices. It can also obtain in a single period or multiperiod model if each primary asset is a physical investment opportunity with stochastic constant returns to scale.

the primary assets are jointly normal. It is well known that if no other assets were traded, the CAPM of Sharpe and Lintner would obtain. However, the introduction of even one financial asset, with a nonnormal return, may obviate this result.

Example: Consider an economy with a single good for consumption and investment. The primary assets are physical investment opportunities with stochastic constant returns to scale. Three such opportunities are available: riskless storage, $R = 1$, and two physical investments whose returns are independent normals, $\tilde{X}_1 \sim N(2,4)$ and $\tilde{X}_2 \sim N(3,8)$. Assuming that each investor is sufficiently risk averse so that he invests in the riskless asset, the lack of a borrowing opportunity will be immaterial.

The relative weights in the optimal risky-asset-only (tangent) portfolio will be proportional to $(\bar{X}_i - R)/\sigma_i^2$. In this case the dollar holding will be equal or $\alpha_1 = \alpha_2$. The optimal portfolio has an expected return of 2.5 and a variance of 3. The betas of the two assets are $2/3$ and $4/3$, respectively, and the Sharpe-Lintner CAPM relation is verified.[12]

Define "one share" of asset 1 or 2 to be an initial investment of \$1.00. As before $\tilde{Z}_i$ represents the end-of-period value. As a financial asset consider a call option with an exercise price of \$2.00 on one share of asset 1. Call this option asset 3; then

$$\tilde{Z}_3 = \max(\tilde{Z}_1 - 2, 0). \tag{11}$$

As $\bar{Z}_1 = 2$, the nonzero portion of $\tilde{Z}_3$ is the right half of a normal distribution. Its statistical description is therefore

$$\bar{Z}_3 = \sigma_1/\sqrt{2\pi} = \sqrt{2/\pi} \approx 0.798, \tag{12a}$$

$$\begin{aligned}\operatorname{cov}(\tilde{Z}_3, \tilde{Z}_1) = E[\tilde{Z}_3(\tilde{Z}_1 - \bar{Z}_1)] = E[&\max(\tilde{Z}_1 - 2, 0)\\ &(\tilde{Z}_1 - 2)] = \sigma_1^2/2 = 2,\end{aligned} \tag{12b}$$

$$\operatorname{cov}(Z_3, Z_2) = 0, \tag{12c}$$

$$\begin{aligned}\operatorname{var}(\tilde{Z}_3) = E\{[\max(\tilde{Z}_1 - \bar{Z}_1, 0)]^2\} - \bar{Z}_3^2 &= \sigma_1^2/2 - \sigma_1^2/2\pi\\ &= 2(1 - 1/\pi) \approx 1.363.\end{aligned} \tag{12d}$$

Properties (12a), (12b), and (12d) are standard results which follow from the convenience that $\tilde{Z}_3$ is $\tilde{Z}_1$ truncated at its mean; (12c) follows from the independence of $\tilde{Z}_1$ and $\tilde{Z}_2$.

Case 1: Homogeneous quadratic utility. Suppose that all investors have the utility function $U(W) = W - bW^2/2$ and assume that $bW_0 \geqslant 2/9$ so that there is positive demand for the riskless asset.

12. In this example the riskless asset will not be in zero net supply. The true market portfolio will not be at the tangency but at a lower risk position on the capital market line. Nevertheless, the form of the CAPM relation will be unchanged since the true market's excess expected return and the betas with respect to it will be proportionately smaller than 2.5 and greater than $2/3$ and $4/3$.

Since all investors are identical and the option is in zero net supply, everyone must be holding exactly zero of it. Consequently, its existence does not affect equilibrium. All investors hold the equally weighted portfolio. We can, however, compute the (shadow) price and return on the option at which there will be no demand. All investors have quadratic utility so the CAPM holds even though the option's returns are not normal. Using (5) and (12) the option's price can be computed as

$$P_3 = [E(\tilde{Z}_3) - \lambda \operatorname{cov}(\tilde{Z}_3, X_m)]/R = \bar{Z}_3 - {}^1/_2 = \sqrt{2/\pi} - {}^1/_2 \approx 0.298. \tag{13}$$

This price implies an expected return of $\bar{X}_3 = \bar{Z}_3/P_3 \approx 2.679$ and a beta of 0.843, which, of course, are consistent with the returns form of the CAPM relation (1).

Case 2: Heterogeneous utilities. Suppose that some investors have quadratic utility and others have power utility $U(W) = W^\gamma/\gamma$ with $\gamma < 0$. Assume for the moment that financial assets are prohibited. For the power utility investors $U(0) = -\infty$; therefore they will consider only portfolios with limited liability. The only possible limited liability portfolio consists entirely of the riskless asset. Consequently they cannot affect the prices of the risky assets. The quadratic utility investors will behave as in case 1. Again the Sharpe-Lintner CAPM will obtain.

Now consider the introduction of the call option. From case 1, the quadratic utility investors will purchase (sell) options if the price is less (greater) than 0.298. As options have potentially unlimited upside gains, power utility investors will clearly never sell them. But options do have limited liability so they might purchase them. In fact it is easy to see that they will hold a positive amount.

A power utility investor will not hold any of assets 1 or 2 to maintain personal limited liability. Thus his portfolio is characterized simply by the dollar holdings of the option α_3. The optimal portfolio is the solution to

$$0 = \frac{dEU(\tilde{W})}{d\alpha_3} = E\{U'[(W_0 - \alpha_3)R + \alpha_3\tilde{X}_3](\tilde{X}_3 - R)\}. \tag{14}$$

Assume for the moment that $\alpha_3 = 0$. But if power utility investors hold no options, neither can the quadratic utility investors. From case 1 the latter occurs only if $\bar{X}_3 \approx 2.678$. Thus at $\alpha_3 = 0$, $dEU(\tilde{W})/d\alpha_3 \approx 1.678\, U'(W_0R) > 0$. By continuity and concavity of the utility function $\alpha_3^* > 0$.

From the preceding paragraph, the power utility investors will, in equilibrium, be long in options and the riskless asset. The quadratic utility investors will be short in options. Their exact holdings of the various assets will depend on risk aversion and the relative wealths of the two classes. All of these factors will determine the call option's price P_3.

Calculating the exact equilibrium price given these factors is a tedious exercise so we will avoid that step and finish the example in terms of the option price. At a given option price, a quadratic utility investor faces the opportunity set characterized by mean returns $\overline{\mathbf{X}} = (2, 3, \sqrt{2/\pi}/P_3)$, $R = 1$, and the variance-covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 0 & 2/P_3 \\ 0 & 8 & 0 \\ 2/P_3 & & \dfrac{2(\pi - 1)}{\pi P_3^2} \end{pmatrix}. \tag{15}$$

His optimal risky-asset-only portfolio weights (assuming he is sufficiently risk averse to hold the riskless asset) are given by

$$\mathbf{w} \equiv (\mathbf{1}'\alpha)^{-1}\,\alpha = \frac{\Sigma^{-1}\,(\overline{\mathbf{X}} - R\mathbf{1})}{\mathbf{1}'\Sigma^{-1}\,(\overline{\mathbf{X}} - R\mathbf{1})} \tag{16}$$

with expected return $\bar{X} = \mathbf{w}'\overline{\mathbf{X}}$ and variance $\sigma^2 = \mathbf{w}'\Sigma\mathbf{w}$.

For example, if $P_3 = 1/3$ the optimal portfolio for quadratic utility is $\mathbf{w}' \approx (.579, .484, -.063)$ with an expected return and variance of 2.459 and 3.802. The (risky-asset) market portfolio is therefore in proportion .579 to .484 to 0 or $\mathbf{w}_m' \approx (.545, .455, 0)$, with an expected return and variance of $\bar{X}_m \approx 2.455$ and $\sigma_m^2 \approx 2.844$. The assets' betas are

$$\boldsymbol{\beta} = \Sigma\,\mathbf{w}_m/\sigma_m^2 \approx (0.767, 1.280, 1.150)', \tag{17}$$

so the predicted (by the CAPM) expected returns would be (2.116, 2.862, 2.673), whereas the true expected returns are (2, 3, 2.393).

This example confirms our earlier statement that the introduction of financial assets into an economy, in which previously all asset returns could be characterized by some separating distribution, will cause the mean-variance equilibrium to break down. Not only will the expected returns on the financial assets differ from those predicted by the CAPM, but also it will no longer predict returns on the primary assets correctly. This example also demonstrates that, by implicitly assuming that financial assets are not traded, existing distributional based mean-variance models require investors to hold portfolios which in general would be Pareto dominated by others including financial assets.

The failure of mean-variance pricing upon the introduction of financial assets is potentially an extremely damaging one. Joint normality of stock returns may be an "objective" property, at least as a working approximation. However, returns on some financial assets, particularly options, are necessarily just as "objectively" far from normal. The active markets for options cast severe doubt on the joint-normality justification for the CAPM.

IV. Mean-Variance Pricing and Complete Markets: An Integration

The previous section documented the failure of mean-variance pricing caused by the introduction of financial assets. In this section we show that in a complete market (or under some other regimes) this failure does not occur. The CAPM equilibrium correctly prices the primary assets.

To keep this paper self-contained we restate several previous definitions and results.

Definition: A portfolio will be termed *efficient* if it is the portfolio which maximizes the expected value of some increasing, concave von Neuman–Morgenstern utility function. Any portfolio which is not efficient will be termed *inefficient*.

This definition of efficiency should not be confused with mean-variance efficiency which is usually associated with mean-variance pricing. The latter is a narrower concept. For some distributions of returns, such as multivariate normals, both types of efficiency are the same. For arbitrary distributions (for which the variance is defined), however, mean-variance efficient portfolios are those chosen by investors with quadratic utility. An efficient portfolio need only be optimal for some investor whose utility function is in the broad concave and monotone class. In general, the optimal portfolio for an investor with utility that is other than quadratic will not be mean-variance efficient. To avoid possible confusion we shall refrain from using the term "mean-variance efficient." Instead we shall refer to the class of minimum-variance portfolios. It consists of each portfolio which has the smallest variance for a given mean return and includes all mean-variance efficient portfolios.

Lemma 1 (Ross): Assume that the return on every risky asset can be characterized as belonging to a *separating distribution* with finite means and variances, that is,

$$\begin{aligned} \tilde{X}_i &= R + b_i\tilde{\delta} + \tilde{\epsilon}_i \quad & i = 1, \ldots, n; \\ E(\tilde{\epsilon}_i \mid \delta) &= 0 & i = 1, \ldots, n; \text{ all } \delta; \end{aligned} \tag{18}$$

and that there exists an n vector $\boldsymbol{\omega}$ such that $\boldsymbol{\omega}'\mathbf{1} = 1$ and $\boldsymbol{\omega}'\tilde{\boldsymbol{\epsilon}} \equiv 0$. Define the set of portfolios E to be all those with investment weights $\mathbf{w}' = \gamma(0, \boldsymbol{\omega}') + (1 - \gamma)(1, 0, \ldots, 0)'$ for $\gamma \geq 0$. Then a portfolio is efficient if and only if it is in the set E. Furthermore, every portfolio not in E is (second degree) stochastically dominated by some portfolio in E. Also, the expected return on each asset is given by

$$\bar{X}_i = R + \frac{\operatorname{cov}(\tilde{X}_i, \tilde{X}_\omega)}{\operatorname{var}(\tilde{X}_\omega)} (\bar{X}_\omega - R), \tag{19}$$

where $\tilde{X}_\omega \equiv \boldsymbol{\omega}'\tilde{\mathbf{X}}$ is the return on the portfolio free of any idiosyncratic risk.

Proof: Ross (1978*a*, theorem 2 and corollary 2) proves that these conditions are sufficient for strong two-fund separation. By definition this guarantees the required dominance.[13] Equation (19) follows directly from (18) and $\tilde{X}_\omega = R + b_\omega \tilde{\delta}$. Q.E.D.

Lemma 2 (Dybvig and Ross): If markets are complete, the set of efficient portfolios is convex.

Proof: See Dybvig and Ross (1982, lemma 3). Q.E.D.

Definition: A random vector $\tilde{\mathbf{Y}}$ is said to have a spherical or *radial distribution* if it is symmetric about the origin and, in the positive orthant, its cumulative distribution function depends only on the length of the vector. If the density function $f(\mathbf{Y})$ is defined, in any orthant

$$f(\mathbf{Y}) = g(\|\mathbf{Y}\|) \tag{20}$$

for all radially distributed random vectors. ($\|\cdot\|$ denotes the Euclidean norm of "length" of a vector.)

Any element in a radially distributed vector has zero mean and is uncorrelated with the remaining elements. However, the elements are not statistically independent except in the important special case of multivariate spherical normal random variables.

Lemma 3 (Chamberlain): If the vector of risky-asset returns can be expressed as a translated, nonsingular transformation of a radially distributed vector, $\tilde{\mathbf{X}} = \overline{\mathbf{X}} + \mathbf{T}\tilde{\mathbf{Y}}$, then the distribution of any portfolio formed with these assets and the riskless asset is completely specified by its mean and variance.[14] Portfolio selection is thus a simple matter of Markowitz mean-variance analysis, and all efficient portfolios belong to the minimum-variance class. These distributions are special cases in the class of separating distributions defined above.

Proof: See Chamberlain (1982). The intuition behind his proof is that all linear combinations of $\tilde{\mathbf{Y}}$ have identical distributions apart from a factor of scaling. This property follows immediately from the spherical symmetry of $\tilde{\mathbf{Y}}$ when we consider each linear combination as a rotation and rescaling of the axes. The deviation of the return on any portfolio of risky assets about its mean return is a linear combination of the radially distributed variables

$$\mathbf{w}'\tilde{\mathbf{X}} - \mathbf{w}'\overline{\mathbf{X}} = \mathbf{w}'\mathbf{T}\tilde{\mathbf{Y}}. \tag{21}$$

Thus the distribution of the return on this portfolio is completely characterized by its mean and the scaling factor $\|\mathbf{T}'\mathbf{w}\|$. Since the

13. Ross's proof provides an even stronger result. In the absence of a riskless asset, R can be replaced by a second common random variable. Ross also shows that these conditions are essentially necessary for strong two-fund separation as well.

14. If the mean or variance is undefined, the portfolio return is completely specified by its median return and some measure of dispersion.

variance-covariance matrix of $\tilde{\mathbf{Y}}$ is $E(\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}') = \sigma^2\mathbf{I}$, the variance of any portfolio's return is

$$\begin{aligned} \text{var}\,(\mathbf{w}'\tilde{\mathbf{X}}) &= E(\mathbf{w}'\mathbf{T}\mathbf{Y}\mathbf{Y}'\mathbf{T}'\mathbf{w}) \\ &= \sigma^2\mathbf{w}'\mathbf{T}\mathbf{T}'\mathbf{w}, \end{aligned} \tag{22}$$

which uniquely determines the scaling factor. Q.E.D.

We can now describe conditions under which the Sharpe-Lintner CAPM is a valid characterization of the expected returns on the primary assets.

Theorem 3: Assume that (i) markets are complete, and (ii) the returns on the primary assets can be characterized as belonging to a separating distribution as defined in lemma 1. Then the equilibrium expected return on each primary asset is

$$\bar{X}_i = R + \frac{\text{cov}\,(\tilde{X}_i, \tilde{X}_m)}{\text{var}\,(\tilde{X}_m)}\,(\bar{X}_m - R). \tag{23}$$

No assumptions are made about the return distributions on the financial assets. For convenience we assume that the riskless asset is a financial asset in zero net supply.[15]

Proof: As Fama (1976), Roll (1977), and others have argued, (23) is equivalent to the assertion that the market portfolio, $\boldsymbol{\mu}$, is in the minimum variance class. We know from our assumption of a separating distribution that the zero residual risk portfolio $\boldsymbol{\omega}$ identified in lemma 1 is in this class. Thus we shall have our proof if we can show that $\boldsymbol{\omega}$ is the market portfolio.

Assume that this is not the case, that is, $\boldsymbol{\mu} \neq \boldsymbol{\omega}$. As the financial assets are all in zero net supply, the market portfolio consists solely of primary assets and, thus, fits within the separating distribution. By assumption, $\boldsymbol{\mu} \neq \boldsymbol{\omega}$. Clearly $\boldsymbol{\mu}$ cannot be any other element of the set E as they all involve the riskless asset. Therefore, by lemma 1 $\boldsymbol{\mu}$ is stochastically dominated by some portfolio(s) in E. But markets are complete; hence, by lemma 2 the set of efficient portfolios is convex. In equilibrium the market portfolio is a convex, linear combination of individuals' optimal (efficient) portfolios; consequently, the market portfolio must be efficient. A portfolio cannot be efficient and stochastically dominated simultaneously. Therefore, the assumption $\boldsymbol{\mu} \neq \boldsymbol{\omega}$

15. This assumption assures that the market portfolio will contain only risky assets. If the riskless asset is in positive net supply, the same proof applies; however, we are then required to interpret the portfolio $\boldsymbol{\omega}$ as the market portfolio margined to reduce net lending to zero.

must be incorrect. The market portfolio is ω, which is minimum variance, and (23) follows. Q.E.D.

Corollary: The results of theorem 3 obtain even if markets are nominally incomplete, as long as investors are unconstrained in the types of financial assets they may create and trade.

Proof: If the introduction of new financial assets is unrestricted, the resulting equilibrium must be one in which no agent desires to create any additional ones. That is, the shadow price of each "uncreated" financial asset must be the same for all agents, or some agents could profit through its introduction. Therefore, even if markets are incomplete, we could create all the Arrow-Debreu state contingent claims (each trading at its shadow price) without affecting the equilibrium, since all agents would have a zero demand for each. Markets are now complete and the results of theorem 3 obtain. Q.E.D.

Within the context of the no-cost models usually examined, this corollary provides a very general justification for ignoring financial assets and proceeding with a mean-variance analysis of the primary securities. This result can be applied broadly because it depends on the Pareto optimality of an unrestricted market place and two-fund separation of the primary assets.

Two cautions should be invoked in the interpretation of theorem 3 and its corollary. First, neither is a portfolio separation theorem. That is, they do not imply that each investor (or, in fact, any actual investor) constructs his optimal portfolio by combining the market with borrowing or lending. In general, separation would obtain in the economy described only if all investors had similar or identical tastes, and it would then follow from the results of Cass and Stiglitz (1970) rather than Ross (1978*a*).

Second, the mean-variance relation among the primary assets in the complete-market economy need not be the same relation which would hold if trading in financial contracts were prohibited and markets were therefore incomplete. In fact, the pricing equation in the corresponding incomplete market might not even be mean-variance. Of course, the opposite can be true as well. A separating distribution may characterize the returns on the primary assets before but not after the financial assets are created.

To apply theorem 3 or its corollary, the distribution of the primary assets' returns *in the presence of the financial assets* must be from the separating-distributions class. We now examine conditions under which this is true.

The existence of a separating distribution characterization for the returns on the primary assets in the absence of any financial assets is not sufficient to guarantee a similar characterization after the markets are completed. For example, suppose that each primary asset represents a fixed investment with a given distribution for its next period payoff Z_i which is unaffected by the introduction of financial assets. To achieve a separating distribution of returns we must have

$$\tilde{Z}_i \equiv P_i\tilde{X}_i = A_i + B_i\tilde{\delta} + \tilde{e}_i,$$
$$E(\tilde{e}_i \mid \delta) = 0, \tag{24}$$
$$\boldsymbol{\alpha}'\tilde{\mathbf{e}} \equiv 0,$$

for some nonzero $\boldsymbol{\alpha}$. Even assuming (24), only prices which satisfy $P_i = RA_i$ will permit a separating distribution as in (18) for arbitrary A_i.[16] Therefore, if a separating distribution obtains in the absence of financial assets, one will obtain after their introduction only if all prices, including that of the riskless asset, adjust in constant proportion. While this would be true if investors were identical (and the original market were therefore effectively complete so prices did not change), it will not be so in general.

For those separating distributions which conform to Chamberlain's (1980) radial distribution requirement, the CAPM will hold both before and after trading in financial assets is permitted. This result is proved below.

Theorem 4: Assume that (i) each primary asset is either a fixed investment with a given distribution for its end-of-period value or discretionary investment with stochastic constant returns to scale; and (ii) in the absence of financial assets the returns on the primary assets are distributed as $\tilde{\mathbf{X}} = \overline{\mathbf{X}} + \mathbf{T}\tilde{\mathbf{Y}}$, where $\mathbf{T}$ is any nonsingular matrix and $\tilde{\mathbf{Y}}$ is a radially distributed vector of random variables with finite variances. Then the CAPM equilibrium

$$\overline{X}_i = R + \frac{\operatorname{cov}(\tilde{X}_i, \tilde{X}_m)}{\operatorname{var}(\tilde{X}_m)}(\overline{X}_m - R) \tag{25}$$

obtains for the primary assets both in the original and completed markets for any distribution of the financial assets' returns. The market portfolio may differ in the two cases as may the covariance structure. If there is no riskless asset in the original market, then, in that equilibrium, R is replaced by the expected return on an asset uncorrelated with the market.

Proof: That (25) obtains in the original market follows immediately from lemma 3. To prove that it obtains in the completed market we need only show that a radial distribution characterization of the primary asset returns is still possible. Since these distributions, described in lemma 3, are special cases of Ross's separating distributions, (25) will then follow from theorem 3.

We denote the returns, prices, and end-of-period values in the com-

16. If $P_i = RA_i$, then (24) and (18) are equivalent for $B_i = P_ib_i$, $\tilde{e}_i = P_i\tilde{\epsilon}_i$, and $\omega_i = \alpha_iP_i/\Sigma\alpha_iP_i$.

pleted market by the small letters $\tilde{x}_i$, p_i, $\tilde{z}_i$. In both economies we partition these vectors into fixed and discretionary components, $\tilde{\mathbf{X}}' = (\tilde{\mathbf{X}}_1', \tilde{\mathbf{X}}_2')$, etc. For the discretionary, constant-returns-to-scale investments $\tilde{\mathbf{x}}_2 = \tilde{\mathbf{X}}_2$. For the fixed investments $\tilde{\mathbf{z}}_1 = \tilde{\mathbf{Z}}_1$, and

$$\mathbf{a}_1\tilde{\mathbf{x}}_1 \equiv \tilde{\mathbf{z}}_1 = \tilde{\mathbf{Z}}_1 \equiv \mathbf{A}_1\tilde{\mathbf{X}}_1, \tag{26}$$

where $\mathbf{a}_1$ and $\mathbf{A}_1$ are diagonal matrices containing the prices $\mathbf{p}_1$ and $\mathbf{P}_1$, respectively. Therefore

$$\begin{aligned} \tilde{\mathbf{x}} &\equiv \begin{pmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^{-1}\mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{pmatrix} \equiv \mathbf{B}\tilde{\mathbf{X}} \\ \tilde{\mathbf{x}} &= \mathbf{B}(\overline{\mathbf{X}} + \mathbf{T}\tilde{\mathbf{Y}}) = \overline{\mathbf{x}} + \hat{\mathbf{T}}\tilde{\mathbf{Y}}. \end{aligned} \tag{27}$$

Since $\mathbf{B}$ is nonsingular so is $\hat{\mathbf{T}} = \mathbf{BT}$. Q.E.D.

V. Conclusion

We have found that there are compelling theoretical arguments why mean-variance pricing cannot be valid for all financial assets in an economy at its unconstrained Pareto-optimal allocation of risk. Nonetheless, we have also found that mean-variance pricing *is* valid for the primary assets in the same economy provided that they satisfy the distributional requirements for two-fund separation.

Our results offer several challenges. If mean-variance pricing is valid only for the primary assets, how are we to price financial assets? Option pricing theory gives some partial answers, but past results have been shown to be valid only under strong distributional assumptions on asset returns.

Another challenge is motivated by the special role our analysis attributes to the continuous-time lognormal diffusion model as a motivation for mean-variance analysis. Empirical tests of mean-variance theory should explicitly address the implications of the continuous-time model.

References

Banz, R. W., and Miller, M. H. 1978. Prices for state-contingent claims: some estimates and applications. *Journal of Business* 51 (October): 653–72.

Black, F., and Scholes, M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (May–June): 637–59.

Breeden, D. T., and Litzenberger, R. H. 1978. Prices of state-contingent claims implicit in option prices. *Journal of Business* 51 (October): 621–51.

Cass, D., and Stiglitz, J. E. 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 (June): 122–60.

Chamberlain, G. 1982. A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory*, in press.

Constantinides, G. M. 1980. Admissible uncertainty in the intertemporal asset pricing model. *Journal of Financial Economics* 8 (March): 71–86.

Constantinides, G. M. 1982. Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. In this issue.

Dybvig, P. H., and Ross, S. A. 1982. Portfolio efficient sets. *Econometrica*, in press.

Fama, E. F. 1976. *Foundations of Finance*. New York: Basic.

Harrison, J. M., and Kreps, D. M. 1979. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory* 20 (June): 381–408.

Long, J. B., Jr. 1974. Stock prices, inflation and the term structure of interest rates. *Journal of Financial Economics* 1 (July): 131–70.

Merton, R. C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41 (September): 867–87.

Roll, R. 1977. A critique of the asset pricing theory's tests; part 1: on past and potential testability of the theory. *Journal of Financial Economics* 4 (March): 129–76.

Ross, S. A. 1976. Options and efficiency. *Quarterly Journal of Economics* 90 (February): 75–89.

Ross, S. A. 1978*a*. Mutual fund separation in financial theory—the separating distributions. *Journal of Economic Theory* 17 (April): 254–86.

Ross, S. A. 1978*b*. A simple approach to the valuation of risky streams. *Journal of Business* 51 (July): 453–75.

Samuelson, P. A. 1970. The fundamental approximation theory of portfolio analysis in terms of means, variances, and higher moments. *Review of Economic Studies* 37 (October): 537–42.

Samuelson, P. A., and Merton, R. C. 1975. Generalized mean-variance tradeoffs for best perturbation corrections to approximate portfolio decisions. *Journal of Finance* 30 (March): 27–40.