TI 2000-108/2
Tinbergen Institute Discussion Paper

Inherent Efficiency, Security Markets, and the Pricing of Investment Strategies

Liang Zou

INHERENT EFFICIENCY, SECURITY MARKETS, AND THE PRICING OF INVESTMENT STRATEGIES

By Liang ZOU[1]

INHERENT EFFICIENCY, SECURITY MARKETS, AND THE PRICING OF INVESTMENT STRATEGIES

Abstract: This paper applies the dichotomous theory of choice by Zou (2000a) to the analysis of investment strategies and security markets. Issues concerning individual optimality, (approximate) arbitrage, capital market equilibrium, and Pareto efficiency are studied under various market conditions. Among the main results are (i) a unique dichotomous pricing model, unifying and generalizing the existing models, that can be used for pricing *any* financial securities under both complete and incomplete markets, (ii) conditions for individual optimality that hold for general utilities (including expected utility as a special case), (iii) the existence and uniqueness of capital market equilibrium, and (iv) implications of capital market equilibrium, including a separation theorem, *inherent efficiency* of the market portfolio, Pareto efficiency, and several testable hypotheses that predict securities' equilibrium *up-market potentials* and *down-market potentials*, respectively.

Key words: Perception of reward and risk, Reward-risk utility, Inherent efficiency, Quasi-complete market, Dichotomous pricing model, Approximate arbitrage, Up-market and Down-market potentials.

JEL Classification numbers: D46, D81,G10, G11, G12.

Correspondence Address: Dr. Liang ZOU, Associate Professor of Finance, Finance Group, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. Phone: (31)-20-5254160; Fax: (31)-20-5255285. Email: liang@fee.uva.nl.

1 INTRODUCTION

In a preceding paper (Zou, 2000a), a dichotomous theory of choice under risk is presented.[2] The theory assumes that choice under risk involves a two-stage decision: first assessing the reward and risk of each feasible investment strategy, and then choosing the best strategy – the one that maximizes the investor's reward-risk utility. Individuals may differ in their preferences over, as well as in their perception of, reward and risk for the feasible strategies. An investment strategy also has its own *inherent reward* and *inherent risk* that are independent of any particular investor's preference or behavior. Axiomatizing this notion leads to a unique measure of inherent reward and risk for any investment strategy, and to the definition of *inherent efficiency*, a criterion of non-dominance in terms of the inherent reward-to-risk ratios of the relevant strategies. The present paper applies this theory to the analysis of investment strategies and the securities market. Among the main results to be shown are:

1. A general analytical framework in which (i) preference is represented by the individual's perception of reward and risk and a reward-risk utility that includes expected utilities as a special case, (ii) contingencies are represented by an n-dimentional vector of states of which the probability distribution is allowed to be discrete, continuous, or both, (iii) leverage and short-selling are allowed, but limited liability is explicitly taken into account to prevent negative wealth, and (iv) new concepts are introduced, such as "co-domain of investment strategies", "quasi-completeness of capital markets", and "flat options".

2. Necessary conditions for individual optimality and optimal investment strategies, including (i) maximization of the (perceived) reward-to-risk ratio, and (ii) sharing perfect co-domain with an inherently efficient strategy. These results hold for the general reward-risk utilities

among which the expected utility is a special case.

3. A simple model for pricing general investment strategies – broadly defined to be any combination of primitive and derivative securities. The model is consistent with no-arbitrage, market equilibrium conditions, and individual optimality. It holds under both complete and incomplete markets, and it unifies the existing (two-date) models (various pricing kernels) into a coherent framework.

4. The existence and uniqueness of capital market equilibrium on the assumption that individuals are *inherently rational* – i.e., their utilities depend only on the inherent reward and inherent risk of investments. And the implication of equilibrium for Pareto efficiency, defined in terms of the general preferences of investors.

It has long been recognized that reward and risk are two basic parameters of investment choice problems (e.g., Knight, 1921; and Hicks, 1939). In contrast to expected utility theory (EUT), the reward-risk approach is appealing for its intuitive simplicity as well as for the ease of applicability of its results. The mean-variance theory is undoubtedly the most successful theory in the reward-risk genre. It has generated important insights in the general issues of investments and capital markets, among which, for instance, are the separation theorem, the role of co-relations among assets in the selection of optimal portfolios (Markowitz, 1952, 1959; and Tobin, 1958, 1965), and the capital asset pricing model (CAPM) (Sharpe, 1963, 1964; and Lintner, 1965a, 1965b).[3]

The limitations of the mean-variance theory, however, are also well known (e.g., Fishburn, 1977 and the references therein). In addition to concerns of using the return variance or standard deviation as a measure of risk, there are also challenges to the validity of some of its implications. For instance, Dybvig and Ingersoll (1982) show that CAPM necessarily involves risk-free arbitrage if it is to hold for all assets in a complete capital market; moreover, the security market line of CAPM can also be problematic if used to measure investment performance (Dybvig and Ross, 1985a, 1985b).

Indeed, if asked what the risk of an investment really is, any investor would say that it is the risk of losing one's money. This common notion of risk is reflected in the mean-downside risk literature, where investment risk is defined as the risk of under-performing a fixed and deterministic benchmark return. Among others, e.g., Domar and Musgrave (1944) measure risk as probability-weighted losses and study the effect of progressive income taxation on risk taking; Porter (1974), Bawa (1975, 1976, 1978), and Fishburn (1977) show that the lower-partial moments as a measure of risk are consistent with stochastic dominance rules; and Bawa and Lindenberg (1977), and Harlow and Rao (1989) examine the capital asset pricing relationship in the lower-partial moment framework. This mean-downside risk literature, unfortunately, has not gained much popularity to date. Suffering the same lack of normative foundation as the mean-variance theory, the assessment of reward and downside risk remains a matter of opinion. Critics of the downside-risk models also could point to their added complexity in both theoretical and empirical analysis.[4]

EUT, on the other hand (e.g., von Neumann and Morgenstern, 1947; Marschak, 1950; Samuelson, 1952; and Herstein and Milnor, 1953; and Savage, 1954), is considered to be the most rigorous approach to decision making under risk.[5] Apparently, expected utility models are more general than reward-risk models since they use information on whole distributions whereas the latter use only a few statistical parameters of the distributions. Much attention has been thus devoted to justifications for reward-risk models in the framework of EUT. Whatever reward-risk model one selects, however, it can be justified in the expected utility framework only for a small class of utility functions (e.g., quadratic utility function for the mean-variance), or for a small class of the return distributions (e.g., normal distribution for the mean-variance). Sacrificing generality for simplicity is unavoidable if one confines himself to the paradigm of expected utility. On the other hand, EUT has some cause for concern regarding its empirical validity (e.g., Tversky, 1967a, 1967b, 1969, 1975; De Bondt and Thaler, 1995). This concern has prompted a large body of literature on non-expected utility analysis (e.g., the survey by Starmer, 2000).

The new dichotomous theory (Zou, 2000a) is a non-expected utility theory. Comprising both expected utility models and reward-risk models, however, it overcomes the above problems these models confront. The general results obtained in this paper illustrate how the new theory integrates expected utility and reward-risk utility into a single consistent framework.

Section 2 considers the market structure, individual preferences and beliefs, and two basic theorems of individual optimality. Section 3 defines inherent efficiency and derives several of its implications. Section 4 introduces the *quasi-complete* market as a bridge between incomplete and complete markets. It presents general pricing models for investment strategies that are derived under a weak assumption of no-arbitrage or no-approximate arbitrage, which is equivalent to inherent efficiency of the capital market. Section 5 studies the equilibrium of the capital market, with an existence theorem showing also uniqueness of equilibrium for inherently rational investors. Pareto efficiency is defined for general preferences, and is shown to be implied by market equilibrium. Section 6 concludes with some suggestions for future research. More technical proofs are relegated to the Appendix.

2 THE MODEL

2.1 The Market and Information Structures

The model considered is best seen as a truncation of an economy spanned by two dates ($t = 0, 1$).[6] The focus is the capital market in this economy where there is a set J of investors and a set K of securities traded at time 0. In general, K includes both *primitive* securities and *derivative* securities. Let $I \subseteq K$ denote the finite set of primitive securities in this economy defined as those having a positive aggregate supply (e.g., by the firms) and directly representing ownership of some real assets, such as stocks, bonds, real estate, etc. The rest of the securities in K are derivative securities that are created by the traders such as stock options, index options, forward contracts, etc., whose payoff depend on the realized payoffs of the underlying securities.[7] The number of derivative securities in K can be infinite, but since every long position of a derivative is matched by a short position of the same derivative, the aggregate supply of any derivative security is zero. "Cash" is considered as a special primitive security that has also a zero aggregate supply.

Let $m_i \geq 0$ denote the aggregate supply, i.e., the number of shares, of each primitive security $i \in I$. The price of each share of the primitive security $i \in I$ is normalized to be 1 at time 0, so that its gross return (1 plus the return) over the period, $r_i \in [0, M] \subset \mathbb{R}_+$, is also the security's share price at time 1. Restricting prices of the primitives to be nonnegative and uniformly bounded from above helps avoid the paradoxical concept of "negative wealth", while allowing strategies to involve *adequate* leverage and short-selling.[8] It implies that investors enjoy limited liability in this market. Let m denote the market portfolio defined as the total value of securities in I; the total time-0 market value is thus $m = \sum_{i \in I} m_i$. The gross return at time 1 of the market is $r_m = \sum_{i \in I}(m_i/m)r_i$, where m_i/m is the value weight of primitive security i in m. The capital market is assumed to be competitive, and to be perfect in that there is no tax, no transaction costs or friction (e.g., Stoll, 2000), no agency problems, etc. Investors can borrow and lend at a gross risk-free interest rate $r_0 \geq 1$ *provided there is no risk of default.*

For consistency, indeed, the default risk of any feasible strategy should be properly taken into account. Otherwise, the negligent creditor who lends out money or shares will create an incentive for someone to "cheat" or, in a more friendly expression, to enjoy "free lunch" (for a formal definition of free lunch see, e.g., Kreps, 1981). The current model will explicitly take such default risks into account. This is done by viewing any strategy that involves default risk as a properly hedged portfolio so that the minimum gross return of the strategy at time 1 is nonnegative. The hedging position in such a portfolio can be either voluntarily or compulsorily taken, the costs of which are the same. Put differently, in an efficient market any potential free lunch will be correctly "charged".[9] For instance, debt with default risk will be subject to a default risk premium (over the risk-free rate), which is equivalent to the (time-1) hedging cost of the debtor to protect himself from losing more than his wealth. The following example clarifies this thought.

Example 1 *(Leverage and short selling).*

Suppose there are two possible states, 1 and 2. Let A denote a risky security that costs 1 dollar at time 0, pays 2 dollars in state 1 and 0.5 dollars in state 2 at time 1. Let B denote a risk-free asset that pays 1 dollar in both states. Let $aA + (1 - a)B$ denote a strategy (portfolio) with a proportion of a invested in A and $(1 - a)$ in B. For $a > 1$, the strategy involves leverage; for $a < 0$, the strategy involves short-selling (of security A). The linear-combination rule says that the portfolio's return is equal to the value-weighted average of its components' returns, i.e., $r_{aA+(1-a)B} = ar_A + (1-a)r_B$. However, this is only true when the portfolio does not involve default risk. See the following table.

Stretegy	1	2	3	4	5	6	7
Payoff	r_A	r_B	$r_{0.5A+0.5B}$	r_{2A-B}	r_{2B-A}	$r_{3A-2\widetilde{B}}$	$r_{3B-2\widetilde{A}}$
State 1	2	1	1.5	3	0	$6-2r_{\widetilde{B}}=3$	0 (default)
State 2	0.5	1	0.75	0	1.5	0 (default)	$3-2r_{\widetilde{A}}=1.5$

It is easily seen that Strategies 1-5 do not have default risk so that $r_{aA+(1-a)B} = ar_A+(1-a)r_B$ $(-1 \leq a \leq 2)$. Strategies 6 and 7 do involve default risk, under which the investor enjoys limited liability in State 2 and State 1 respectively. Since Strategy 4 has also a zero gross return in State 2, to avoid free lunch the payoff of Strategy 6 must be the same as Strategy 4 in State 1, yielding $r_{\widetilde{B}} = 1.5 > r_B = 1$. That is, the investor must borrow at a 50% credit-risk premium above the risk-free rate. Similarly, Strategy 7 must generate the same return as Strategy 5, yielding $r_{\widetilde{A}} = 0.75 > r_A = 0.5$ in state 2. That is, Strategy 7 involves a contingent payment of 0.25 per share to the counter party in the winning state to cover the risk of default in the losing state. $\square$

Now I move on to the information structure of the model. The common information at time 1 is a realized vector of returns on the primitive securities $r = (r_1, r_2, ..., r_n) \in \mathbb{M} = [0, M)^n \subset \mathbb{R}^n_+$, where $\mathbb{M}$ is assumed to include all the relevant contingencies at time 1. Uncertainty about these contingencies are assumed to be represented at time 0 by a probability space $\{\mathbb{M}, \mathcal{B}, P\}$, where $\mathcal{B}$ denotes the σ-field of Borel sets of $\mathbb{M}$, and P denotes the probability distribution of r on $(\mathbb{M}, \mathcal{B})$. For most of the analysis in this paper, P can be generally interpreted as an investor's subjective belief. Only at a later stage, an assumption of common belief of the conditional probability distribution (in a sense to be clear later) will be made for the analysis of equilibrium. In order to simultaneously discuss discrete and continuous distributions, it is understood that for any Borel set $B \subseteq \mathbb{M}$, $P(B)$ will denote the probability that $r \in B$. More specifically, for any integrable function x on $\mathbb{R}^n_+$, $\int_B x(r)dF(r)$ is the Lebesgue-Stieltjes integral with F being the n-dimensional distribution function on the set $\mathbb{M}$ that corresponds to P.

Let Ω denote the set of all feasible investment strategies (sometimes simply called the market). An investment strategy $x \in \Omega$ is broadly defined as a combination of any primitive and derivative securities available in the market place that turns a 1-dollar investment at time 0 into x dollars (possibly random) at time 1.[10] The basic strategies are the buy-and-hold strategies of any asset, portfolio, derivative contract, etc. More sophisticated strategies may involve portfolios of derivatives and compound derivatives whose payoffs are general functions of $r \in \mathbb{M}$.[11] Cash settlement is allowed, so that profits or losses on a derivative contract do not have to involve delivery of the underlying assets. A mild condition to be imposed is that the gross return payoff of any strategy can be represented by some integrable function $x : r \in \mathbb{M} \to x(r) \in [0, \overline{M}) \subset \mathbb{R}_+$ where $\overline{M}$ is any arbitrarily large number serving as the uniform upper bound of the returns of all feasible strategies.[12] Since standard derivatives are typically piecewise linear functions of prices of the underlying securities, this condition is easily met in practice. The condition that $x(r) \geq 0$ for all $r \in \mathbb{M}$ will be called the *nonnegative-wealth constraint*. It is also natural to assume that Ω is convex in that if $x_1 \in \Omega$ and $x_2 \in \Omega$ then $ax_1 + (1-a)x_2 \in \Omega$ and $ax_1 \oplus (1-a)x_2 \in \Omega$ for all $a \in [0,1]$. Here $ax_1 + (1-a)x_2$ means a portfolio with a proportion a of the capital invested in x_1 and the rest in x_2; whereas $ax_1 \oplus (1-a)x_2$ means a strategy that assigns a probability a to playing x_1 and $1-a$ to playing x_2. Put differently, the former is a payoff mixture and the latter is a strategy mixture. Two polar cases of the feasible investment set, for example, are

$$\Omega_0 = \{x : r \in \mathbb{M} \to \sum_{i \in I} \theta_i r_i \in [0, \overline{M}); \sum_{i \in I} \theta_i = 1\}$$

and

$$\Omega_\infty = \{x : r \in \mathbb{M} \to x(r) \in [0, \overline{M}); r_0 \int_{\mathbb{M}} x(r)\psi(r)dF(r) \leq 1, \int_{\mathbb{M}} \psi(r)dF(r) = \frac{1}{r_0}\}.$$

The set Ω_0 represents a canonical portfolio choice set where strategies are restricted to be the linear combinations of the primitives. Note that there is no restriction on the sign of θ_i and the range of the return on a strategy may exceed the upper bound of the primitives M, provided the lower bound 0 is respected. The set Ω_∞ represents the largest set of strategies that can be constructed contingent on every realized return vector of the primitives, where $F(\cdot)$ and $\psi(\cdot)r_0$ are the distribution function and the state prices implied by the market prices of all securities (thus not necessarily the same as an investor's subjective assessments). A state price is for a contingent claim that pays 1 dollar on the realization of r and zero otherwise (e.g., Arrow, 1964 and Debreu, 1959). When r is continuous, $\psi(r)r_0$ is interpreted as a state-price density function with each r regarded as a realized *primitive state* (an n-dimensional vector). For ease of exposition $\psi(r)r_0$ will be simply called the *state-price density* that refers to both cases.

In general, the market may not be complete in that not every primitive state $r \in \mathbb{M}$ is separately contractible. In some sense even the set Ω_∞ may be considered incomplete since it requires the payoff of any derivative security to depend only on the payoff(s) of other traded securities, and not on possible states of nature other than the primitives' payoff states at time 1. For instance, contracting on weather conditions may not be possible unless there exist primitive securities whose payoffs are linked with weather conditions in a predictable manner. However, since it has become conventional to call the capital market complete as long as investors *can* trade every claim whose payoff at time 1 depends on the payoffs of any traded securities (see, e.g., Dybvig, 1988 for more discussion), Ω_∞ will be called a complete market.

There is another type of market, though, that is incomplete but there is a "would-be" state-price density (in the eyes of each investor) consistent with the prices of non-tradable as well as tradable securities. Such markets will be called *quasi-complete* and will be introduced when the state-price density is studied.

2.2 Investor Preferences

For any investor $j \in J$, let $w_j > 0$ denote his initial wealth.[13] With any strategy $x \in \Omega$, the investor's time-1 wealth will be $w_j x(r)$ when $r \in \mathbb{M}$ is realized. The investor's choice problem is assumed to be represented by a utility function of the investment's reward and risk *perceived* by the investor (Zou 2000a).

Assumption 1 (perception of reward and risk) *The investor's perception of investment reward and risk are represented, respectively, by a reward function $U_j : (x, w_j) \in \Omega \times R_{++} \to R_+$, where $w_j > 0$ is the investor's initial capital and $x \in \Omega$ is a chosen strategy, such that*

$$U_j(x, w_j) = \int_{\mathbb{M}\cap\{r:x(r)\geq r_0\}} [g_j(w_j x(r)) - g_j(w_j r_0)]dF_j(r), \tag{1}$$

and a risk function $D_j : (x, w_j) \in \Omega \times R_{++} \to R_+$ such that

$$D_j(x, w_j) = \int_{\mathbb{M}\cap\{r:x(r)\leq r_0\}} [\ell_j(w_j r_0) - \ell_j(w_j x(r))]dF_j(r), \tag{2}$$

where $g_j : R_+ \to R_+$ (the perceived gain function) and $\ell_j : R_+ \to R_+$ (the perceived loss function) are positive monotonic functions satisfying $g_j(w_j r_0) = \ell_j(w_j r_0)$, $g'_j(w_j r_0) = \ell'_j(w_j r_0)$, and $0 < g'_j(w^+) \leq \ell'_j(w^-) < \infty$ for all $w^+ \in [w_j r_0, \infty)$ and $w^- \in [0, w_j r_0]$.

The condition in the last line of this assumption, $0 < g'_j(w^+) \leq \ell'_j(w^-) < \infty$, will not be invoked until Theorems 8 and 9. It is worth remarking, though, that any von Neumman-Morgenstern utility function $V_{NM}(\cdot)$ can be generally written as:

$$V_{NM}(wx(r)) = \max[g(wx(r), g(wr_0)] - \max[\ell(wr_0), \ell(wx(r)]$$

and if V_{NM} is weakly concave than this condition holds.

For brevity, the subscript j will be dropped in most of subsequent text (until Section 5).

Assumption 2 (reward-risk utility) *The investor's preference over any perceived reward and risk of investment strategies* $(U, D) \in R^2_+$ *is represented by a reward-risk utility* $V : (U, D) \in R^2_+ \to V(U, D) \in R$*, where* V *is strictly concave on* R^2_+ *and satisfies* $\frac{\partial V}{\partial U} > 0$ *and* $\frac{\partial V}{\partial U} + \frac{\partial V}{\partial D} < 0$ *for all* $(U, D) \in R^2_{++}$.

In this way, every investor's preference over the investment strategies $x \in \Omega$, given his current investment capital (wealth) $w > 0$, is indirectly represented by a utility function $V(U(x, w), D(x, w))$ with U and D defined in (1)–(2). This representation of investor's choice problem is derived from a set of weaker axioms, in particular a weaker independence axiom, than those adopted in the development of expected utility theory (Zou, 2000a). As a result, a broader class of investor preferences are represented of which the expected utility representation is a special linear case: $V(U, D) = U - D$. The general class of preferences postulated in Assumptions 1 and 2 will be denoted by $\Psi = \{V : (U, D) \to R;\ (U, D) \text{ satisfying (1)–(2)}\}$.

A measure of investors' degree of risk aversion can be defined as, given any level of (U, D),

$$\rho(U, D) = -\left(\frac{\partial V(U, D)}{\partial D}\right) / \left(\frac{\partial V(U, D)}{\partial U}\right) \tag{3}$$

This measure is positive by Assumption 2 and the strict concavity of V ensures that its indifference curves on the (U, D) plane are strictly convex. By conventional terminology, an investor is said to be locally risk neutral, risk averse, or risk loving if and only if his degree of risk aversion $\rho = 1$, $\rho > 1$, or $\rho < 1$ at a given point (U, D); and is said to be globally risk neutral, risk averse, or risk loving if and only if his $\rho = 1$, $\rho > 1$, or $\rho < 1$ for all (U, D) on R^2_+. An investor is considered locally more risk averse than another investor if the former has a higher ρ than the latter at (U, D). If this holds for all (U, D) on R^2_+ then the former investor is globally more risk averse than the latter. Assumption 2 implies that the investors are globally risk averse ($\rho > 1$). It is important to note, however, that this measure of risk aversion is different from the Arrow-Pratt measures (Arrow, 1971, Pratt, 1964), of which an analogy can be made for the investor's perception of reward and risk (Zou, 2000a).

In general, the investor's investment problem is given by the following program $\mathbb{P}$:

$$\max_{x\in\Omega} V(U(x,w), D(x,w))$$

with U and D defined as in (1) and (2).

2.3 Individual Optimality

I first derive some necessary conditions for a solution to $\mathbb{P}$. To start with, let us call $Z(x,w) = U(x,w)/D(x,w)$ the *reward-to-risk ratio* for $D(x,w) \neq 0$ (as assessed by any relevant investor). By definition, for the risk-free asset $D(r_0,w) = U(r_0,w) = 0$ and by Assumption 1 it can be defined that $Z(r_0,w) = 1$.

Theorem 1 *Let any $V \in \Psi$ be given. If investment strategy $x^* \in \Omega$ is a solution to $\mathbb{P}$, then for all $x \in \Omega$,*

$$Z(x,w) = \frac{U(x,w)}{D(x,w)} \leq \frac{U(x^*,w)}{D(x^*,w)} = Z(x^*,w) \tag{4}$$

Proof: Let any investor $V \in \Psi$ be given. Let x^* denote the optimal solution to $\mathbb{P}$ and consider an alternative strategy as follows. Select any other strategy $x \in \Omega$ and let a lottery decide which strategy to choose, with probability θ of choosing x and $1-\theta$ of choosing x^*. It is easy to verify that the reward and risk measures are linear in such mixtures, that is,

$$\begin{aligned} U(\theta x \oplus (1-\theta)x^*, w) &= \theta U(x,w) + (1-\theta)U(x^*,w) \\ D(\theta x \oplus (1-\theta)x^*, w) &= \theta D(x,w) + (1-\theta)D(x^*,w) \end{aligned}$$

where the reader may recall that $\theta x \oplus (1-\theta)x^*$ denotes the strategy with probability θ of choosing x and $1-\theta$ of choosing x^*. Since x^* is optimal, the first-order condition with respect to θ evaluated at $\theta = 0$ must satisfy

$$\frac{\partial V}{\partial \theta} = \frac{\partial V}{\partial U}[U(x,w) - U(x^*,w)] + \frac{\partial V}{\partial D}[D(x,w) - D(x^*,w)] \leq 0$$

or, equivalently,

$$\frac{\partial V}{\partial U}[U(x,w) - U(x^*,w)] \leq -\frac{\partial V}{\partial D}[D(x,w) - D(x^*,w)] \tag{5}$$

for all $x \in \Omega$, where the inequality stems from the constraint that the probability $\theta \geq 0$.

Choosing $x = r_0$ in (5) yields

$$\frac{\partial V}{\partial U}U(x^*,w) \geq -\frac{\partial V}{\partial D}D(x^*,w). \tag{6}$$

Dividing (5) by (6) and re-arranging terms yields (4). □

By Theorem 1, it is necessary for all types of investors to maximize the reward-to-risk ratio $Z(x,w)$. To ease writing, for any $x \in \Omega$ let $\mathbb{G}(x) = \mathbb{M} \cap \{r : x(r) > r_0\}$, $\mathbb{L}(x) = \mathbb{M} \cap \{r : x(r) < r_0\}$, and $\mathbb{N}(x) = \mathbb{M} \cap \{r : x(r) = r_0\}$. They are called the *gain-domain*, *loss-domain*, and *neutral-domain* of investment strategy x, respectively.

Theorem 2 *Let any $V \in \Psi$ be given. If strategy $x^* \in \Omega$ is a solution to $\mathbb{P}$, then for all $x \in \Omega$,*

$$\Gamma_- \leq \int_{\mathbb{G}(x^*)} g'(wx^*(r))(x(r) - r_0)dF(r) - Z(x^*,w)\int_{\mathbb{L}(x^*)} \ell'(wx^*(r))(r_0 - x(r))dF(r) \leq \Gamma_+ \tag{7}$$

where

$$\begin{aligned} \Gamma_+ &= Z(x^*,w)\int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)(r_0 - x(r))dF(r) - \int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} g'(wr_0)(x(r) - r_0)dF(r) \\ \Gamma_- &= Z(x^*,w)\int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)(r_0 - x(r))dF(r) - \int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} g'(wr_0)(x(r) - r_0)dF(r) \end{aligned}$$

Proof: See Appendix.

The relationships in (7) can be interpreted as an investor's subjective believe about all strategy prices that are available. If $P(\mathbb{N}(x^*)) = 0$, then $\Gamma_+ = \Gamma_- = 0$ and (7) gives an exact pricing relationship for all $x \in \Omega$ with respect to x^*. Since Ω may contain (portfolios of) derivative contracts, however, it is fairly possible that $P(\mathbb{N}(x^*)) > 0$ for some investors. This may be the case even where returns on the primitive-securities are continuously distributed. For one thing, a strategy with 100% in cash that yields r_0 combined with a self-financed reversal (a long call and a short put with the same underlying asset or portfolio, both out of the money and with equal premiums) will have a positive probability of breaking even (yielding r_0).

If the second inequality in (7) holds strictly for some asset $x \in \Omega$, there will be no demand for this asset by the investor; instead, the investor may consider short-selling some amount of this asset. It might be the case, however, that at the same time the first inequality in (7) also holds strictly so that short-selling of x is not interesting either. In such a case, the asset x is truly redundant in the eyes of the investor – even if its price moves up or down within a certain range.[14]

The pricing relationships in (7) is quite general. They are derived without much restriction on preferences and beliefs. Thus, (7) can be interpreted in terms of an investor's subjective assessment of the expected security returns (e.g., the distribution function F can be subjective) and subjective perception of reward and risk (e.g., the functional forms of g and ℓ may differ across the investors). Unfortunately, however, these pricing relationships are difficult to test empirically and the inequality relations, if hold strictly, provide only ranges of security prices and not a precise pricing relationship. This motivates the consideration in the next section of a special type of investors who perceive investment reward and risk objectively.

3 INHERENT EFFICIENCY

An investor is called *inherently rational* if he perceives investment rewards and risks *inherently*, i.e., without personal bias, and his reward-risk utility depends only on the inherent reward and risk of his investment choice and his current wealth.[15]

Definition 1 (inherent reward) *The inherent reward of a strategy $x \in \Omega$, $u : x \in \Omega \to u(x) \in R_+$, is defined as*

$$u(x) = E\max(x(r) - r_0, 0) = \int_{\mathbb{G}(x)} [x(r) - r_0] dF(r) \tag{8}$$

Definition 2 (inherent risk) *The inherent risk of a strategy $x \in \Omega$, $v : x \in \Omega \to v(x) \in R_+$, is defined as*

$$v(x) = E\max(r_0 - x(r), 0) = \int_{\mathbb{L}(x)} [r_0 - x(r)] dF(r) \tag{9}$$

For any investment strategy, its inherent reward and risk are a pair of *unique* measures common to all investors (Zou, 2000a). Such measures can be justified under two axioms additional to those in the derivation of individually perceived reward and risk. One is called the *allocation-independence axiom*, which says that if strategy x_1 is "more-rewarding" (or "more-risky") than strategy x_2 then $ax_1 + (1-a)r_0$ is "more-rewarding" (or "more-risky") than $ax_2 + (1-a)r_0$ for all $a \in (0, 1]$. That is, the (common) judgement of more rewarding or more risky relationships between any two strategies should not depend on any individual investor's capital allocation decision. The other is called the *properness axiom*, which says that the perceived gain function g and loss function ℓ should satisfy $0 < g'(r_0) < \infty$ and $0 < \ell'(r_0) < \infty$ at the point of the risk-free rate (the neutral return). Together with some other more standard axioms, the inherent reward and risk defined above are the only measures (ratio scales) that follow, i.e., unique up to a positive multiplicative transformation. Thus, the preferences of inherently rational individuals differ only in their wealth and reward-risk utilities.

The risk premium of strategy x will be denoted $\mu(x) = E(x(r) - r_0) = u(x) - v(x)$, which by the definitions of u and v can also be viewed as the strategy's excess inherent reward (after subtracting inherent risk from the reward).

Let $\Psi_0 = \{V \in \Psi : U = wu, D = wv;\ (u, v) \text{ satisfying (8)-(9)}\}$ denote the set of inherently rational preferences, i.e., $V \in \Psi_0$ if and only if $V(U, D) = V(wu, wv)$. As will be shown in the next section, the determination of security prices and capital market equilibrium can be derived easily from inherent rationality. In the subsequent sections, it will further be demonstrated that even if investors have different gain and loss functions in their assessment of investment reward and risk their choices often can be made *as though* they are inherently rational.

Definition 3 (inherent reward-to-risk ratio) *The inherent reward to risk ratio of any strategy* x, $z : x \in \Omega \rightarrow z(x) \in R_+$, *is defined as*

$$z(x) = \begin{cases} \infty & \text{if } v(x) = 0,\ u(x) > 0 \\ u(x)/v(x) & \text{if } v(x) \neq 0 \\ 1 & \text{if } v(x) = 0,\ u(x) = 0. \end{cases} \tag{10}$$

In the special case where $v(x) = 0$, the strategy has no inherent risk. In the absence of risk-free arbitrage opportunities, any strategy with zero inherent risk must also have zero inherent reward, i.e., $u(x) = 0$, and it follows by definition that $z(x) = 1$. In the case where x indeed represents a *risk-free arbitrage strategy*, $z(x) = \infty$.

The next concept plays an important role in the subsequent analyses.

Definition 4 (perfect co-domain) *For any* $x \in \Omega$ *and* $y \in \Omega$*, let the subset of* $\mathbb{M}$

$$COD(x, y) = \mathbb{M} \cap \{r : (x(r) - r_0)(y(r) - r_0) \geq 0\}$$

be called the (reward-risk) co-domain between x *and* y*. Then,* x *and* y *are said to have perfect co-domain on* $\mathbb{M}$ *under probability distribution* P *if and only if* $P(COD(x, y)) = 1$*. That is, the excess returns of the two strategies have the same sign almost surely (a.s.), where zero is treated as both positive and negative signs.*

Alternatively, it is equivalent to state that two strategies have *imperfect* co-domain on $\mathbb{M}$ under probability distribution P if there exists a subset of $\mathbb{M}$ with strictly positive probability measure on which the excess returns of the two strategies have opposite signs. That is,

$$P(\mathbb{M} \cap \{r : (x(r) - r_0)(y(r) - r_0) < 0\}) > 0. \tag{11}$$

The next lemma show the diversification effect among strategies that do not share perfect co-domain.

Lemma 1 (sub-additivity) *Let any $x \in \Omega$, $y \in \Omega$, and any $a \in [0, 1]$ be given. Then*

$$u(ax + (1-a)y) \leq au(x) + (1-a)u(y) \tag{12}$$

$$v(ax + (1-a)y) \leq av(x) + (1-a)v(y) \tag{13}$$

For $a \in (0, 1)$, the strict inequalities hold in (12) and (13) simultaneously if and only if (11) holds. In other words, the inherent reward and risk measures satisfy strict sub-additivity between strategies that have imperfect co-domain.

Proof: See Appendix.

Lemma 2 (quasi-concavity) *Let any $x \in \Omega$ and $y \in \Omega$ be given such that $z(x) \leq z(y) < \infty$. Then $z(ax + (1-a)y) \geq z(x)$ for all $a \in [0, 1]$. The inequality holds strictly for $a \in (0, 1)$ if and only if $z(x) < z(y)$ or $P(COD(x, y)) < 1$. In other words, the inherent reward-to-risk ratio is quasi-concave on Ω and strictly quasi-concave on the set of strategies that do not have perfect co-domains among each other.*

Proof: Let any $x \in \Omega$ and $y \in \Omega$ and any real number $a \in (0, 1)$ be given such that $z(x) \leq z(y) < \infty$. If $v(x) = 0$, then $x = r_0$ (a.s.) and it is trivially true that $z(ax+(1-a)y) = z(y) \geq z(x)$ with strict inequality if and only if $z(y) > z(x) = 1$, or, equivalently in this case, $P(COD(x, y)) < 1$.

Otherwise, since $z(x) - 1 = \mu(x)/v(x)$ and $\mu(ax + (1-a)y) = a\mu(x) + (1-a)\mu(y)$,

$$\begin{aligned} z(ax + (1-a)y) &= \frac{a\mu(x) + (1-a)\mu(y)}{v(ax + (1-a)y)} + 1 \\ &\geq \frac{a\mu(x) + (1-a)\mu(y)}{av(x) + (1-a)v(y)} + 1 \qquad (14) \\ &= \frac{au(x) + (1-a)u(y)}{av(x) + (1-a)v(y)} \geq \frac{u(x)}{v(x)} = z(x) \qquad (15) \end{aligned}$$

From Lemma 1, the inequality in (14) holds strictly if and only if $P(COD(x, y)) < 1$. The inequality in (15) holds strictly if and only if $z(x) < z(y)$, as a result of a simple mathematical property. □

Definition 5 (inherent dominance) *Let any two strategies x and y from Ω be given. Then y is said to inherently dominates x if and only if $z(y) \geq z(x)$.*

The inherent dominance relationship has the desirable properties of being complete (provided only that the means exist for all strategies in Ω), transitive, and continuous (Zou, 2000a). Moreover, it is a necessary condition for the criterion of first-order stochastic dominance (e.g., Fishburn, 1977; Bawa, 1976, 1978; and Zou, 2000a). Thus if strategy $y \in \Omega$ is preferred to strategy $x \in \Omega$ by *all* investors who prefer more wealth to less, then it is *necessary* that $z(y) \geq z(x)$. It will be shown later that for investors whose gain- and loss-functions are weakly concave, the relationship $z(y) > z(x)$ is also *sufficient* for them to prefer y to x – provided that the capital market is quasi-complete.

Definition 6 (inherent efficiency) *The strategy space Ω is said to be (inherently) efficient if there exists a strategy $y \in \Omega$ such that $z(x) \leq z(y) < \infty$ for all $x \in \Omega$. The set $\mathbb{E}(\Omega) \subseteq \Omega$ will be called the (inherent) efficiency set of strategies in Ω, defined as*

$$\mathbb{E}(\Omega) = \{y \in \Omega : \forall x \in \Omega, z(x) \leq z(y) < \infty\}.$$

Note that inherent efficiency is defined requiring no arbitrage. It is difficult to call a market efficient (by any criteria) if there are arbitrage opportunities. Any market involving arbitrage opportunities is transient in nature; consequently, it is not of very much theoretical interest *per ce* except its straightforward implication. Another reason for ruling out this case is that most of the analysis in this paper does require that there is no arbitrage. Also note that since $r_0 \in \Omega$ and $z(r_0) = 1$, $z(y) \geq 1$ for all $y \in \mathbb{E}(\Omega)$.[16] It will be assumed thoughout the paper that $z(y) = \overline{z} > 1$ for $y \in \mathbb{E}(\Omega)$.

In what follows, the notation $\sim$ means "equivalent" in the sense that two strategies have the same inherent reward and inherent risk.

Theorem 3 *Suppose $y \in \mathbb{E}(\Omega)$ and let any $x \in \Omega$ be given. Then the following conditions hold for $x \in \mathbb{E}(\Omega)$:*

1. *If there exists $\overline{y} \in \mathbb{E}(\Omega)$ such that $P(\mathbb{N}(\overline{y})) = 0$ and x and y have perfect co-domain, and only if x and y have perfect co-domain;*
2. *if and only if for all $a > 0$ such that $r_0 + a(x - r_0) \in \Omega$, $r_0 + a(x - r_0) \in \mathbb{E}(\Omega)$;*
3. *if and only if $ax + (1 - a)y \in \mathbb{E}(\Omega)$ for all $a \in [0, 1]$;*
4. *if and only if there exists $a \in [0, 1]$ such that $x \sim r_0 + a(y - r_0)$ or $y \sim ar_0 + (1 - a)x$.*

Proof: Suppose $y \in \mathbb{E}(\Omega)$.

1. Sufficiency will follow later from Corollary 1. Necessity follows from Lemma 1 since $z(x) = z(y)$ only if x and y have perfect co-domain.

2. If for all $a > 0$ such that $r_0 + a(x - r_0) \in \Omega$, $r_0 + a(x - r_0) \in \mathbb{E}(\Omega)$, choosing $a = 1$ yields $x \in \mathbb{E}(\Omega)$. If $x \in \mathbb{E}(\Omega)$ and $x_a = r_0 + a(x - r_0) \in \Omega$, then $u(x_a) = au(x)$ and $v(x_a) = av(x)$ so that $z(x_a) = z(x)$.

3. By Lemma 2, if $z(x) = z(y)$ then $z(ax + (1 - a)y) \geq z(y)$ for all $a \in [0, 1]$ so that $ax + (1 - a)y \in \mathbb{E}(\Omega)$. Sufficiency is trivially true.

4. If there exists $a \in [0,1]$ such that $u(x) = au(y)$ and $v(x) = av(y)$, or $u(y) = au(x)$ and $v(y) = av(x)$, then $z(x) = z(y)$. If $z(x) = z(y)$ then let $a = u(x)/u(y)$ or $a = u(y)/u(x)$ such that $a \in [0,1]$. It follows that $x \sim r_0 + a(y - r_0)$ or $y \sim ar_0 + (1-a)x$. $\square$

Theorem 3 is useful for the subsequent analysis and is important in its own implications. Since whether two strategies have perfect co-domain is easy to verify, Theorem 3 dramatically simplifies the task for identifying efficient strategies.

Example 2 *(inherently efficient and inefficient strategies)*

See Figures 1–4. Suppose $y \in \Omega$ is an efficient strategy. Then the positive linear combination of y and the risk-free asset is efficient (Figure 1). A bull spread of y can be designed that is also efficient. This can be achieved with a long position in y plus a self-financed reversal, or simply a long and a short call on y with properly chosen exercise prices. As shown in Figure 2, such a bull spread has a perfect co-domain with y. However, the simple covered call (a long position in y plus a short call option on y with the proceeds of the short call invested in the risk-free asset) or protective put (a long position in y plus a long put option on y purchased with risk-free borrowing) are not efficient since they do not have perfect co-domain with y. For the same reason, a straddle strategy (a long call and a long put option on y) is not efficient. The inherent efficiency of many other popular strategies can be examined easily in the same fashion.

For ease of writing, let $E(\cdot|\Delta)$ denote the expectation operator conditional on $r \in \Delta \subseteq \mathbb{M}$ and let $E(\cdot,\Delta) = E(\cdot|\Delta)P(\Delta)$. The inherent reward and risk of $x \in \Omega$ computed on the subset $\Delta \subseteq \mathbb{M}$ can then be denoted as $u(x,\Delta) = E(x - r_0, \Delta \cap \mathbb{G}(x))$ and $v(x,\Delta) = E(r_0 - x, \Delta \cap \mathbb{L}(x))$ respectively. Note that it holds for all $\Delta \subseteq \mathbb{M}$ that the risk premium $\mu(x,\Delta) = u(x,\Delta) - v(x,\Delta)$.

Theorem 4 *There exists $y \in \mathbb{E}(\Omega)$ if and only if for all $x \in \Omega$,*

$$v(x, \mathbb{N}(y)) - u(x, \mathbb{N}(y))z(y) \le \mu(x, \mathbb{G}(y)) + \mu(x, \mathbb{L}(y))z(y) \le v(x, \mathbb{N}(y))z(y) - u(x, \mathbb{N}(y)) \quad (16)$$

Proof: Straightforward by substituting 1 for $g'(\cdot)$ and $\ell'(\cdot)$ in Theorem 2. □

Example 3 *(Pricing ranges of contingent claims).*

Consider a market Ω in which asset $A \in \mathbb{E}(\Omega)$ and B is risk free. Their gross returns are given as follows in each of the three equally probable states

State	Prob	$r(A)$	$r(B)$
1	$\frac{1}{3}$	3	1
2	$\frac{1}{3}$	1	1
3	$\frac{1}{3}$	0.5	1

What is the price (range) of a call option on A with strike price equal to 2? The inherent reward and risk of A are $u(A) = (1/3)(3 - 1) = (2/3)$ and $v(A) = (1/3)(1 - 0.5) = 1/6$. Thus $z(A) = 4$. Since $A \in \mathbb{E}(\Omega)$ the price (range) of this option must satisfy (16). The three states in this example can be clearly represented by the three different returns on A. Let $c(r)$ denote the return of this option as a function of $r = r(A)$. Then $c(3) = 1/C$, $c(1) = c(0.5) = 0$. From the left inequality in (16), $(1/3)(c(3) - 1) + (1/3)(0 - 1)4 \ge (1/3)(1 - 0) - 0 \Rightarrow c(3) \ge 6$. And from the right inequality in (16) $(1/3)(c(3) - 1) + (1/3)(0 - 1)4 \le (1/3)(1 - 0)4 - 0 \Rightarrow c(3) \le 9$.Thus, the price of this call option $C \in [1/9, 1/6]$.

It is instructive to verify this conclusion by contradiction. First suppose that $c(3) = 9 + \epsilon$. Then $u(C) = (1/3)(9 + \epsilon - 1)$, $v(C) = (2/3)(1 - 0)$, and $z(C) = (1/3)(9 + \epsilon - 1)/(2/3) = 4 + \frac{1}{2}\epsilon$. To be consistent with $A \in \mathbb{E}(\Omega)$, it follows that $\epsilon \le 0$. Next suppose $c(3) = 6 - \epsilon$. Then consider a portfolio with some short positions in the call. To avoid negative wealth, the maximum short position (for $\epsilon \geq 0$) in the call is 0.5 or 50% of one's capital (so that in state 1 the portfolio return ≥ 0). This gives $u(1.5A - 0.5C) = (1/3)(1.5 \times 3 - 0.5(6 - \epsilon) - 1) + (1/3)(1.5 - 0 - 1)$, $v((1+a)A - aC) = (1/3)(1 - 1.5 \times 0.5)$, and $z(C) = 4 + 2\epsilon$. Again, to be consistent with $A \in \mathbb{E}(\Omega)$ it follows that $\epsilon \leq 0$. □

Example 4 *(Exact pricing of contingent claims).*

Now suppose asset $A \in \mathbb{E}(\Omega)$ satisfies $P(\mathbb{N}(A)) = 0$. Its gross returns are as follows (four possible states).

State	Prob	Payoff A	Payoff B
1	0.3	2	1
2	0.2	1.2	1
3	0.1	0.8	1
4	0.4	0.5	1

What is the exact price of a call option on A with strike price equal to 1? And of a put option on A with strike price equal to 1?

The inherent reward and risk of A are $u(A) = 0.3(2 - 1) + 0.2(1.2 - 1) = .34$, $v(A) = 0.1(1 - 0.8) + 0.4(1 - 0.5) = .22$. Thus $z(A) = 1.5455$. Since $P(\mathbb{N}(A)) = 0$, the two option prices can be determined exactly.

$0.3((1/C) - 1) + 0.2((0.2)/C - 1) - 1.5455[0.1(1 - 0) + 0.4(1 - 0)] = 0 \Rightarrow C = .26714$

$0.3(1 - 0) + 0.2(1 - 0) - 1.5455[0.1((0.2)/P - 1) + 0.4((0.5)/P - 1)] = 0 \Rightarrow P = .26715.$

This confirms also the put-call parity. □

4 APPROXIMATE ARBITRAGE AND SECURITY PRICES

This section examines further implications of inherent efficiency for security prices. The pricing theories presented below are purely abstract and do not rely on any specifications of investor preferences, thanks to the equivalence between inherent efficiency and the absence of (approximate) arbitrage.

Comparing Examples 3 and 4 it is obvious that the choice of the benchmark strategy $y \in \mathbb{E}(\Omega)$ matters for determining prices of other strategies. Most desirable, of course, is to have an inherently efficient strategy y with $y(r) \neq r_0$ almost surely, so that an exact pricing relationship will be established for all securities. In Section 5, it will be shown that the market portfolio $r_m \in \mathbb{E}(\Omega)$ so that it suffices to assume that $P(\mathbb{N}(r_m)) = 0$.

Assumption 3 *If $\mathbb{E}(\Omega) \neq \emptyset$, then there exists $\overline{y} \in \mathbb{E}(\Omega)$, called the efficient benchmark, with $P(\mathbb{N}(\overline{y})) = 0$. The gain- and loss-domain of $\overline{y}$ will be denoted $G = \mathbb{G}(\overline{y})$ and $L = \mathbb{L}(\overline{y})$ respectively, so that $G \cup L = \mathbb{M}$ (a.s.). In particular, if the market portfolio $r_m \in \mathbb{E}(\Omega)$ then $P(\mathbb{N}(r_m)) = 0$.*

An immediate implication of Assumption 3 is that for all $y \in \mathbb{E}(\Omega)$, $\mathbb{G}(y) \subseteq G$ and $\mathbb{L}(y) \subseteq L$. This follows from Theorem 3 (condition 1) that for any $y \in \mathbb{E}(\Omega)$ it is necessary that y and $\overline{y}$ share perfect co-domain; consequently, $u(y) = \mu(y, G)$ and $v(y) = -\mu(y, L)$. For convenience, let the inherent reward and risk of any $y \in \mathbb{E}(\Omega)$, conditional on G and L (rather than on $\mathbb{G}(y)$ and $\mathbb{L}(y)$), be denoted as $\widehat{u}(y) = \mu(y|G)$ and $\widehat{v}(y) = -\mu(x|L)$. For the efficient benchmark, let $\overline{z} = \overline{u}/\overline{v} = u(\overline{y})/v(\overline{y})$ and $\widehat{z} = \widehat{u}/\widehat{v}$, where $\widehat{u} = u(\overline{y}|G)$ and $\widehat{v} = v(\ \overline{y}|L)$.

4.1 Pricing models in the general market Ω

In the current context, approximate arbitrage means a situation in which an investor, with maximum expected loss (inherent risk) of 1 dollar, can arbitrarily increase the expected payoff (say, higher than 1 trillion dollars) via some feasible strategy.

Definition 7 (approximate arbitrage) *The capital market Ω allows approximate arbitrage if and only if for any large number $M > 0$, there exists a strategy $x \in \Omega$ whose inherent reward-to-risk ratio $z(x) > M$.*

Approximate arbitrage opportunities will be hard to find, of course.[17] The next lemma is straightforward by definitions of inherent efficiency and approximate arbitrage.

Lemma 3 *There exists no approximate arbitrage in Ω if and only if $\mathbb{E}(\Omega) \neq \emptyset$.*

In order to examine the prices (or fair costs) of securities or strategies, note that $x = p(x)/p^0(x) \in \Omega$ if $p^0(x)$ and $p(x)$ are the non-normalized prices of the strategy at time 0 and time 1 respectively.

Theorem 5 *Under Assumption 3 there is no approximate arbitrage in Ω if and only if for all $x \in \Omega$,*

$$p^0(x) = \delta_1 E(p(x)|G) + \delta_2 E(p(x)|L) \tag{17}$$

where

$$\delta_1 = \frac{1}{(1+\widehat{z})r_0}, \quad \delta_2 = \frac{\widehat{z}}{(1+\widehat{z})r_0}. \tag{18}$$

Proof: Under Assumption 3, the pricing relationships in (16) simplifies to

$$\mu(x|G) + \mu(x|L)\widehat{z} = 0 \tag{19}$$

for all $x \in \Omega$. Substituting $p(x)/p^0(x)$ for x in (19) yields

$$E(p(x) - p^0(x)r_0|G) = \widehat{z}E(p^0(x)r_0 - p(x)|L) \tag{20}$$

Taking out the constant $p^0(x)r_0$ from the expectation yields

$$E(p(x)|G) + \widehat{z}E(p(x)|L) = p^0(x)r_0(1 + \widehat{z}) \tag{21}$$

Rearranging terms yields the expression in (17), with $\delta_1$and δ_2 defined as in (18). □

Theorem 5 reveals some important insights in the pricing of risky assets. It says that in the absence of approximate arbitrage, all strategies (despite their structure of random payoffs) can be priced as the sum of two present values. There are two *common* factors, δ_1 and δ_2, for discounting all strategies' conditional expected values on G and L respectively.

These common discount factors have a clear meaning of their own, where δ_1 is the price of a security that pays 1 dollar conditional on $r \in G$ and 0 otherwise, and δ_2 the price of a security that pays 1 dollar conditional on $r \in L$ and 0 otherwise. For the reason to be shown later that the market portfolio $r_m \in \mathbb{E}(\Omega)$, δ_1 will be called the *up-market discount factor*, and δ_2 the *down-market discount factor*. Together they form a *(conditional) pricing kernel* (δ_1, δ_2) that is unique even if the capital market is not complete. As a record, the price of a zero-coupon risk-free bond that pays 1 dollar at time 1 is correctly priced according to (17):

$$\delta_1 E(1|G) + \delta_2 E(1|L) = \delta_1 + \delta_2 = \frac{1}{r_0}. \tag{22}$$

For their theoretical and practical importance, securities that pay a fixed amount contingent on the underlying asset or portfolio being above or below an exercise price and zero otherwise will be called the *flat options*, formally defined as follows.

Definition 8 *Security $y_c(x, x_0)$ is called a flat-call option on the underlying asset x with exercise price x_0 if $y_c = 1$ for $x \geq x_0$ and $y_c = 0$ for $x < x_0$. Security $y_p(x, x_0)$ is called a flat-put option on the underlying asset x with exercise price x_0 if $y_c = 0$ for $x \geq x_0$ and $y_c = 1$ for $x < x_0$.*

Note that (22) is a simple "put-call" parity for the flat options. The price of a flat-call on the market (with strike price r_0), for instance, is δ_1 and that of a flat-put is δ_2.

Since the gross returns of the strategy is denoted $x(r)$, (17) has an equivalent expression in terms of the gross returns:

$$1 = \delta_1 E(x|G) + \delta_2 E(x|L) \qquad \forall x \in \Omega \tag{23}$$

or equivalently, using (22),

$$\delta_1 \mu(x|G) = -\delta_2 \mu(x|L) \quad \forall x \in \Omega. \tag{24}$$

Interpreting $\mu(x|G)$ and $\mu(x|L)$ as the strategy's *up-market potential* and *down-market potential* respectively, equation (24) says that the absolute present values of the two potentials should be equal. Their signs depend on the correlation of the strategy's returns with that of the efficient benchmark, but they must be always opposite to each other. If the correlation is positive, then it is likely that $\mu(x|G) \geq 0$ and $\mu(x|L) \leq 0$; otherwise, the opposite is likely to be true. In the special case where the strategy is uncorrelated with the benchmark, then $\mu(x|G) = \mu(x|L) = 0$ or equivalently, $E(x) = r^0$. Since the present value of up-market (resp. down-market) potentials of all strategies are discounted by the same δ_1 (resp. δ_2), it follows that the ratio of the up-market potential over the down-market potential of every risky strategy is the same $\delta_2/\delta_1 = \widehat{z}$.

Since equation (23) holds also for the benchmark $\overline{y}$, using (22) the discount factors δ_1 and δ_2 may also be expressed as functions of $\widehat{u}$ and $\widehat{v}$:

$$\delta_1 = \frac{\widehat{v}}{(\widehat{u} + \widehat{v}) r_0}, \quad \delta_2 = \frac{\widehat{u}}{(\widehat{u} + \widehat{v}) r_0} \tag{25}$$

It follows that

$$\frac{-\mu(x|L)}{\mu(x|G)-\mu(x|L)} = \frac{\widehat{v}}{\widehat{u}+\widehat{v}} = \delta_1 r_0 \qquad \forall x \in \Omega \tag{26}$$

$$\frac{\mu(x|G)}{\mu(x|G)-\mu(x|L)} = \frac{\widehat{u}}{\widehat{u}+\widehat{v}} = \delta_2 r_0 \qquad \forall x \in \Omega \tag{27}$$

From (26) and (27) it follows that for any given $y \in \mathbb{E}(\Omega)$, a parameter $\beta(x,y)$ can be defined for all $x \in \Omega$ as follows provided there is no (approximate) arbitrage.

$$\begin{aligned}\beta(x,y) &= \frac{\mu(x|G)}{\widehat{u}(y)} = \frac{-\mu(x|L)}{\widehat{v}(y)} = \frac{\mu(x,G)}{u(y)} = \frac{-\mu(x,L)}{v(y)} \\ &= \frac{\mu(x|G)-\mu(x|L)}{\widehat{u}(y)+\widehat{v}(y)} = \frac{\mu(x,G)-\mu(x,L)}{u(y)+v(y)}\end{aligned} \tag{28}$$

The next corollary follows immediately.

Corollary 1 *Under Assumption 3 there is no approximate arbitrage in* Ω *if and only if for all* $x \in \Omega$ *and* $y \in \mathbb{E}(\Omega)$

$$E(x-r_0|G) = \beta(x,y)E(y-r_0|G) \tag{29}$$

$$E(x-r_0|L) = \beta(x,y)E(y-r_0|L) \tag{30}$$

or equivalently

$$E(x-r_0,G) = \beta(x,y)E(y-r_0,G) \tag{31}$$

$$E(x-r_0,L) = \beta(x,y)E(y-r_0,L) \tag{32}$$

Dividing (31) by (32) one derives that if x shares perfect co-domain with any $y \in \mathbb{E}(\Omega)$ than $x \in \mathbb{E}(\Omega)$. This concludes the sufficiency part of the proof of Theorem 3 (condition 1).

It is interesting to observe that the six ratio expressions in the definition of $\beta(x,y)$ in (28), although required to be equal to avoid approximate arbitrage, are very different in their economic interpretations. Some are related to the up-market potentials, some down-market potentials, and some the spread of these potentials of the strategies. By Corollary 1, a higher β means that the strategy tends to realize a higher gain when the benchmark goes up and a higher loss when the benchmark goes down. If β is negative, it suggests that the strategy tends to realize gains when the benchmark realizes losses and losses when the benchmark realizes gains (e.g., a put option on the benchmark). Corollary 1 thus predicts that for any investment strategy, its up-market potential or down-market potential are directly related to its β. Note also that $\beta(x, y) = 0$ for a strategy x if and only if $E(x) = r^0$, and that $\beta(y, y) = 1$.

To further understand the meanings of β suppose that the market portfolio $r_m \in \mathbb{E}(\Omega)$ and consider the primitive assets $\{r_i\}$ that make up the market, of which $\beta(r_i, m)$ can be re-written as follows in terms of the market's inherent reward $u(m)$ and inherent risk $v(m)$.

$$\beta(r_i, m) = \frac{\mu(r_i, G)}{u(m)} = \frac{-\mu(r_i, L)}{v(m)}. \tag{33}$$

The terms $\mu(r_i, G)$ and $-\mu(r_i, L)$ can be interpreted as the asset's *share of the market inherent reward* and *share of the market inherent risk* respectively. This interpretation comes from the fact that the inherent reward and risk of the market are equal to the weighted average of the shares of reward and risk of its component primitive assets. That is,

$$u(m) = \sum_{i=1}^{m} \frac{m_i}{m} \mu(r_i, G), \quad v(m) = -\sum_{i=1}^{m} \frac{m_i}{m} \mu(r_i, L), \quad \text{where} \quad \sum_{i=1}^{m} \frac{m_i}{m} = 1. \tag{34}$$

Thus, when the market is inherently efficient all the primitive asset prices are determined by their share of the market inherent reward and risk. Alternatively, $\beta(r_i, m)$ can also be interpreted as the share of the asset's *market spread:* $u(m) + v(m)$, from the expression $\beta(r_i, m) = (\mu(r_i, G) - \mu(r_i, L))/(u(m) + v(m))$. These possible interpretations of β show that the common interpretation of β in the mean-variance CAPM as the asset's systematic risk alone is rather limited. The new insights here are that all assets in equilibrium or in the absence of arbitrage must be priced such that their up-market and down-market potentials have the same proportion in the upside potential and downside potential of the market. This proportion is measured by the asset's β, which can be interpreted either as the asset's *systematic reward* $\mu(r_i, G)/u(m)$, *systematic risk* $-\mu(r_i, L)/v(m)$, or *systematic spread* $(\mu(r_i, G) - \mu(r_i, L))/(u(m) + v(m))$. The inherent pricing theory in this section shows that all these measures are the same.

Clearly, the predictions of Corollary 1 are stronger than the existing beta-models for asset pricing. It separately predicts securities' up-market potentials and their down-market potentials. As a result, the pricing models presented here have more empirical contents, being easier to be falsified than other existing models. Especially, the pricing relationships in (29) and (30) depend only on the conditional expectations (on market being up, i.e., $r_m \geq r_0$ and on market being down, i.e., $r_m \leq r_0$). Thus they require a weaker assumption on investors' beliefs for equilibrium than the assumption of homogeneous beliefs adopted in deriving CAPM. Similar to the security market line (SML) of CAPM, (29) and (30) give two separate lines for pricing the up-market and down-market potentials of all the securities. I call these lines the *up-market line* and the *down-market line* (see Figure 5). The next corollary shows how Corollary 1 implies the single-SML beta-models.

Corollary 2 *Under Assumption 3, if $y \in \mathbb{E}(\Omega)$ then for all $x \in \Omega$*

$$u(x) = v(x) + \beta(x, y)(u(y) - v(y)) \tag{35}$$

or, equivalently

$$E(x) = r_0 + \beta(x, y)(E(y) - r_0). \tag{36}$$

where $\beta(x, y)$ is defined as in (28).

Proof: Adding up (31) and (32) yields directly the desired results. □

Corollary 2 involves a weaker prediction than Corollary 1 since it concerns only the combined values of $u(x) - v(x)$. This result has been first derived by Bawa and Lindenberg (1977), who also showed that if all return distributions are jointly normal then $\beta(x, y) = cov(x, y)/var(y)$. In other words, the mean-variance CAPM is a special case of Corollary 2, which holds when all security returns are joint-normally distributed and investors have homogeneous beliefs.

The equivalence between (35) and (36) follows from $u(x) - v(x) = E(x) - r_0$. However, these two expressions have different interpretations. The term $u(x) - v(x)$ can be interpreted as the *inherent reward minus its inherent risk* of the investment strategy, whereas the term $E(x) - r_0$ is what conventionally termed the risk premium of the strategy without specifying what the risk is. Thus, $u(x) - v(x)$ is more specific in its meaning about the reward and risk that determine the risk premium of the strategy. Interpreting $E(x) - r_0$ as the *inherent risk premium* of strategy x, the equations (35) and (36) state a linear relationship between the inherent risk premium on any strategy with that of an inherently efficient benchmark.

The next corollary concludes this subsection for the finite strategy spaces.

Corollary 3 *If the strategy space contains only a finite distinct strategies, i.e.,* $\Omega = \{x_1, x_2, ..., x_K\}$*, then the term "approximate arbitrage" can be replaced by "arbitrage" in Theorem 5 and Corollaries 1 – 2.*

Proof: Arbitrage is defined as a strategy x with $z(x) = \infty$. Since there are only a finite number of strategies, $\max(z(x) : x \in \Omega) < \infty$ if and only if there is no arbitrage in Ω. □

4.2 State price density for tradable and non-tradable states

This subsection extends the preceding results to a quasi-complete market. The point is that even if a security is not tradable, it can have a perceived price. The quasi-complete market requires that the perceived prices of non-tradable securities are consistent with that of the tradable securities. Let $\psi_j(\cdot)r_0$ denote the (subjective) state-price density on $\mathbb{M}$ as perceived by investor $j \in J$, defined for all tradable and non-tradable states. And define

$$\widetilde{\Omega}_j = \{x : r \in \mathbb{M} \to x(r) \in [0, \overline{M}); r_0 \int_{\mathbb{M}} x(r)\psi_j(r)dF_j(r) = 1, \int_{\mathbb{M}} \psi_j(r)dF_j(r) = \frac{1}{r_0}\}.$$

as the "would-be" complete market as perceived by the investor. In other words, the investor believes that the non-tradable securities if made tradable will be priced according to $\psi_j(\cdot)$.

Definition 9 (quasi-completeness) *The capital market Ω is said to be quasi-complete if there exists $y \in \mathbb{E}(\Omega)$ such that $z(y) \geq z(x)$ for all $x \in \widetilde{\Omega}_j$ and for all $j \in J$.*

Implied in this definition is that if the market is quasi-complete and there is no approximate arbitrage then every investor believes that prices of the tradable securities in Ω will not be affected by the opening-up of any new market for currently non-tradable (derivative) securities. It is easy to note that if a market is complete, then it is quasi-complete. But the reverse is not true. It is possible in a quasi-complete market that some investors would be better off if some non-existent (derivative) securities could be made feasible. An example is gambling in a place where gambles are forbidden. In general, the existing security prices in the absence of approximate arbitrage would not be affected by any new security markets as long as their *aggregate* supply and demand do not change. So the quasi-complete market is weaker than the complete market assumption.

The purpose of introducing the concept of quasi-completeness is to develop a *unique* pricing model that unifies the existing models under complete and incomplete markets. One of the major advantages of the complete market assumption is its implication for a unique pricing kernel that determines the prices of all securities under no arbitrage. However, the complete-market assumption is clearly unrealistic – if only by casual observation that the number of tradable securities in practice are far less than the number of possible states.[18] Alternatively, one can view the complete capital market as an ultimate goal (or limit) of an exchange economy as it moves towards higher efficiency, and form some opinion about the degree of efficiency (or completeness) of the current market. If a market is *not* quasi-complete then it is possible to create new financial products that improve further inherent efficiency of the current market. Since it will be shown that inherent efficiency is a necessary condition for individual optimality (Theorems 7 and 8), this would mean to make everyone better off. The profit potential of such new products would be huge, and it is thus at least for some good reason to assume that such profit opportunities are fully exploited in a competitive market, i.e., that the market is quasi-complete.

Moreover, an ability to reliably price new securities in an incomplete market is of both theoretical and practical interest. This typically involves the use of existing security prices. One popular approach, for instance, is to derive a risk-neutral distribution from the traded option prices that could serve to price other securities (e.g., Rubinstein, 1994). Implicit in such approaches, however, is the assumption of a (quasi-)complete market – for else the derived risk-neutral distributions might not be justified. No theory to date, to my knowledge, has been able to present a *unique* pricing formula under *incomplete* market conditions, consistent with no arbitrage, that can be used to price new securities as well as the existing securities. The assumption of a quasi-complete market is sufficient to achieve this goal.

Let f denote a generalized (subjective) density function of r on $\mathbb{M}$, so that $f(r)$ is the probability that r will occur (for the discrete case) or the density function (for the continuous case) of r on $\mathbb{M}$. The conditional density will be denoted as $f(r|\Delta) = f(r)/P(\Delta)$ for $P(\Delta) > 0$.

Theorem 6 (state price density) *Let Assumption 3 hold and assume that Ω is quasi-complete.*

Then the (subjectively believed) state price density $\psi(r)r_0$ on $G \cup L$ is determined by

$$\psi(r) = \begin{cases} \delta_1 f(r|G) & \text{if } r \in G, \\ \delta_2 f(r|L) & \text{if } r \in L, \end{cases} \tag{37}$$

and $\psi(r_0) = 0$ for the discrete case (since $f(r_0) = 0$) or it can be quite arbitrarily defined for the continous case (where the density $f(r_0) > 0$) as any positive finite number, say, $\psi(r_0) \in [\lim_{r\downarrow r_0} \delta_1 f(r|G), \lim_{r\uparrow r_0} \delta_2 f(r|G)]$.

Proof: It is straightforward to verify that Theorem 5 implies (37). It is also clear that $\psi(r) \geq 0$ with strict inequality for r such that $f(r) > 0$. Further,

$$\int_{\mathbb{M}} \psi(r)dr = \delta_1 \int_G f(r|G)dr + \delta_2 \int_L f(r|L)dr = \delta_1 + \delta_2 = \frac{1}{r_0}$$

Thus $\psi(r)r_0$ is a state-price density. The price $\psi(r_0)$ is immaterial since $P(\{r : r = r_0\}) = 0.$[19] □

Example 5 *(State prices in a quasi-complete market).*

Consider an economy with two traded assets A and B, both priced at 1 dollar, and four possible states, with equal probability. The payoffs of the two assets are given as follows

State	Prob	Payoff A	Payoff B
1	$\frac{1}{4}$	3	1
2	$\frac{1}{4}$	2	1
3	$\frac{1}{4}$	$\frac{1}{2}$	1
4	$\frac{1}{4}$	0	1

Suppose the market is quasi-complete. What is the state-price density? It is clear that $A \in E(\widetilde{\Omega})$ with $\widehat{z} = 2$. Which yields $\delta_1 = 1/3$ and $\delta_2 = 2/3$. It follows that $(\psi_1, \psi_2, \psi_3, \psi_4) = (1/6, 1/6, 1/3, 1/3)$, which can be used to price any new securities in this market. □

5 CAPITAL MARKET EQUILIBRIUM

I proceed now to the equilibrium analysis of the capital market. Although the previous results hold for subjectively believed probability distributions (P_j) and apply to individually perceived security prices ($\psi_j(r)$), results in this section may require some homogeneity in individual beliefs. But let us first define the market equilibrium.

Definition 10 (capital market equilibrium) *Let Assumptions 1-3 hold. Then the capital market is said to be in equilibrium given the feasible strategy space Ω, investor beliefs $\{P_j, j \in J\}$ and investor preferences $\{V_j \in \Psi\}$ if there is no approximate arbitrage in Ω and there exists a vector of feasible strategies $\mathbf{x}^* = (x_1^*, x_2^*, ..., x_J^*)$, $x_j^* \in \Omega$ such that*

1. *for all $V_j \in \Psi$, $j \in J$, and $x \in \Omega$,*

$$V_j(U_j(x_j^*, w_j), D_j(x_j^*, w_j)) \geq V_j(U_j(x, w_j), D_j(x, w_j)) \tag{38}$$

   with U_j and D_j defined as in (1) and (2);

2. *for all $r \in \mathbb{M}$,*

$$\sum_{j \in J} x_j^*(r) w_j = \sum_{i \in I} m_i r_i = m r_m(r) \tag{39}$$

Condition 1 says that x_j^* is the optimal strategy for investor j. Condition 2 has several implications. First, the sum of all the individual wealth at time 0 are equal to the total market value at time 0. This is because taking conditional expectation on both side of (39) on G and L, adding them up, and using (23) yields

$$\sum w_j = \sum_{j \in J} [\delta_1 E(x_j^*|G) + \delta_2 E(x_j^*|L)] w_j = m[\delta_1 E(r_m|G) + \delta_2 E(r_m|L)] = m$$

Second, the sum of all individuals' wealth at time 1 are equal to the total market value at time 1 for whatever realization of $r \in \mathbb{M}$. This condition is imposed so that the market clears at time 1 as well. It follows from these implications that the net aggregate supply and demand of the risk-free asset and derivative securities are zero at time 0, and that the aggregate profits and losses on all the derivative contracts sum up to zero at time 1. Less transparent, yet also true, is the implication of Condition 2 that the supply and demand of each individual primitive security be cleared at time 0. To see this, let Δr_i be a partial marginal change of r_i. From (39), the partial marginal changes in the market value and in the aggregate investor wealth are equal, implying $m_i = \sum[\Delta x_j^*(r)/\Delta r_i]w_j$. Since this must hold for all levels of r_i, the total holdings of security i under the equilibrium condition (39) must be m_i.[20]

Assumption 4 *Investors have homogeneous beliefs about the conditional density function $f(\cdot|G)$ and $f(\cdot|L)$ on $\mathbb{M}$, i.e.,*

$$f_j(r|G) = f(r|G), \text{ and } f_j(r|L) = f(r|L), \text{ for all } j \in J$$

Investors can thus differ in their general opinion about the market as a whole, e.g., in their assessment about $P(G)$ and $P(L)$. Some can be more optimistic or bullish (assessing a higher $P(G)$) and some more pessimistic or bearish (assessing a higher $P(L)$).

Theorem 7 *Under Assumptions 1 – 4 and assuming that Ω is inherently efficient and complete ($\Omega = \Omega_\infty$), the market is in equilibrium if and only if (i) all strategies are priced by the state-price density given in (37) and (ii) there exist $x_j^* \in E(\Omega_\infty)$ for all $j \in J$ that satisfy (38) and (39).*

Proof: See Appendix.

By Theorem 7, all investors would prefer a higher inherent reward-to-risk ratio even when their subjective perceptions of reward and risk differ. This validates the claim made earlier that all investors would behave *as though* they perceive reward and risk inherently. The next theorem shows that the assumption of a complete market is not necessary and that the same insight holds for a quasi-complete market where a special flat option on the market portfolio can be traded.

Theorem 8 *Let Assumptions 1 – 4 hold. Further assume that (i) the market Ω is quasi-complete, (ii) for all investors their gain- and loss-functions are weakly concave, and (iii) the flat-call option $y_c(r_m, r_0)$ on the market portfolio is feasible in Ω. Then the market is in equilibrium if and only if (i) $r_m \in \mathbb{E}(\Omega)$, (ii) all tradable and non-tradable securities are priced by the state-price density given in (37), and (iii) there exist $x_j^* \in \mathbb{E}(\Omega)$ for all $j \in J$ that satisfy (38) and (39). In particular, it is optimal for investors having a strictly concave gain function to hold a flat–call, those having a strictly concave loss function to hold a flat put, on the market portfolio.*[21]

Proof: *Necessity*: Suppose the market is in equilibrium. Let us start with a fictitious assumption (to be shown superfluous shortly) that the market is complete. Then from Theorem 7 all securities are priced by (37) and the optimal strategies $x_j^* \in \mathbb{E}(\Omega)$ for all $j \in J$. Since (39) implies that r_m is a finite convex combination of $x_j^* \in \mathbb{E}(\Omega)$, from Theorem 3 it follows that $r_m \in \mathbb{E}(\Omega)$.

Now assume that g and ℓ are concave functions. Let $x = x^+ y_c(r_m, r_0) + x^- y_p(r_m, r_0) \in \mathbb{E}(\Omega)$ be defined as a combination of a flat-call and a flat-put option on any efficient strategy x^*, where

$$\begin{aligned} x^+ &= \int_G x^*(r)dF(r) = E(x^*, G) \\ x^- &= \int_L x^*(r)dF(r) = E(x^*, L) \end{aligned}$$

The cost of x is the same as x^* since

$$\begin{aligned} \int_G x^+ \psi(r)dr &= \frac{\delta_1}{P(G)} \int_G x^*(r)dF(r) = \int_G x^*(r)\psi(r)dr \\ \int_L x^- \psi(r)dr &= \frac{\delta_2}{P(G)} \int_L x^*(r)dF(r) = \int_L x^*(r)\psi(r)dr \end{aligned}$$

However, by Jensen's inequality

$$\begin{aligned} U(x,w)-U(x^*,w) &= \int_G [g(wx^+)-g(wx^*(r))]dF(r) = P(G)g(E(wx^*)) - E(g(wx^*),G) \geq 0 \\ D(x,w)-D(x^*,w) &= \int_L [\ell(wx^*(r))-\ell(wx^-)]dF(r) = E(\ell(wx^*),L) - P(L)\ell(E(wx^*)) \leq 0 \end{aligned}$$

with strict inequality for strictly concave g or ℓ. Thus, by Theorem 1, strategy x is (weakly) preferred to x^* by investors having concave g and/or ℓ. In other words, there is no loss of generality for all investors with (weak) concave gain- and loss-functions to restrict their attention to strategies in Ω, or more precisely, in $\mathbb{E}(\Omega)$. The fictitious assumption of a complete market is thus not necessary. The *sufficiency* part of the proof is by straightforward verification. □

For the next assumption, let $u_{\max} = \sup\{u(x) : x \in \mathbb{E}(\Omega)\}$ and $v_{\max} = \overline{z}u_{\max}$.

Assumption 5 *All investors are inherently rational ($V_j \in \Psi_0$). For all investors $j \in J$, their degree of risk aversion ρ_j satisfies $\rho_j(0,0) < \overline{z} < \rho_j(w_j u_{\max}, w_j v_{\max}) < \infty$. Moreover, for all $(w_j u, w_j v)$ on R^2_{++},*

1. *non-increasing risk aversion w.r.t. reward, i.e., $\partial\rho_j/\partial u \leq 0$;*
2. *non-decreasing risk aversion w.r.t. risk, i.e., $\partial\rho_j/\partial v \geq 0$;*
3. *increasing risk aversion w.r.t. wealth, i.e., $\partial\rho_j/\partial w_j > 0$.*

Properties 1 and 2 in Assumption 5 are fairly intuitive. They say that as the upside potential increases and/or the downside risk decreases one becomes more tolerant to risk. Property 3 says that more wealth leads to more risk aversion. A simple example is the quadratic reward-risk utility

$$V = U - (D + \frac{\kappa}{2}D^2) = wu - wv - \frac{\kappa}{2}(wv)^2, \quad \kappa > 0$$

of which $\rho(wu, wv) = 1 + \kappa D = 1 + \kappa wv$. It satisfies Properties 1-3 in Assumption 5 trivially. In the next lemma the subscript j is dropped for brevity.

Lemma 4 *Under Assumptions 1 – 5, for all* $V \in \Psi_0$ *there exists* $x \in \mathbb{E}(\Omega)$ *that maximizes* V *on* Ω*. Moreover, for any* x *that maximizes* V *on* Ω *it is necessary and sufficient that* $\rho(wu(x), wv(x)) = z(x)$.

Proof: By Theorem 1, there is no loss of generality to restrict attention to $\mathbb{E}(\Omega)$ for finding optimal strategies that maximize $V \in \Psi_0$. Under Assumption 5 there exists $y \in \mathbb{E}(\Omega)$ such that $\overline{z} < \rho_j(w_j u(y), w_j v(y))$ for all j. For any $x \in \mathbb{E}(\Omega)$ satisfying $u(x) \leq u(y)$, then, there exists $a \in (0, 1]$ such that $x \sim (1-a)r_0 + ay \in \mathbb{E}(\Omega)$ (i.e., $V(wu(x), wv(x)) = V(wau(y), wav(y)$; see Theorem 3, condition 2). It will be shown that the rest of the $x \in \mathbb{E}(\Omega)$ with $u(x) > u(y)$, if any, are suboptimal. Thus the investor's choice problem reduces to

$$\max_{a \in (0,1]} V(wau(y), wav(y))$$

The unconstrained first-order condition is

$$\frac{\partial V}{\partial a} = \frac{\partial V}{\partial U} wu(y) + \frac{\partial V}{\partial D} wv(y) = 0 \tag{40}$$

or equivalently,

$$\rho(wu(x), wv(x)) = \rho(wau(y), wav(y)) = z(y) = z(x) = \overline{z}. \tag{41}$$

The existence of an unconstrained (interior) optimal $a \in (0, 1)$ follows from $u(x) > v(x)$ and the assumption that $\partial \rho_j / \partial w_j > 0$, so that for $a \in (0, 1)$

$$\rho(0, 0) < \rho(wau(y), wav(y)) = \overline{z} < \rho(wu(y), wv(y)). \tag{42}$$

The strict concavity of V in u and v ensures that the first-order condition is also sufficient. . $\square$

For the existence of equilibrium, some restrictions on the market return and investor preferences are needed. Let $\underline{r}_m \geq 0$ denote the lower bound of r_m so that $P(\{r : r_m(r) < \underline{r}_m\}) = 0$.

Then $\overline{a} = r_0/(r_0 - \underline{r}_m)$ is the maximum leverage on the market portfolio without violating the nonnegative-wealth constraint.

Theorem 9 (separation) *Let Assumptions 1 –5 hold, and assume that* $\underline{r}_m > 0$ *and for all* $j \in J$, $\overline{z} < \rho_j(w_j\overline{a}u(m), w_j\overline{a}v(m))$. *Then the capital market is in equilibrium if there exists a vector* $(a_1^*, a_2^*, ..., a_J^*)$, $a_j^* \in [0, \overline{a}]$ *for all* $j \in J$, *such that* $(1 - a_j)r_0 + a_j r_m \in \mathbb{E}(\Omega)$ *and satisfies*

$$\rho_j(w_j a_j^* u(m), w_j a_j^* v(m)) = z(m), \tag{43}$$

$$\sum_{j \in J} a_j^* w_j = m. \tag{44}$$

Proof: Suppose (43) and (44) hold. Then by Lemma 4, $x_j^* = (1 - a_j^*)r_0 + a_j^* r_m \in \mathbb{E}(\Omega)$ satisfies (38). It is also straightforward to verify that substituting x_j^* in (44) yields (39). □

Analogous to the mean-variance analysis (e.g., Tobin, 1958), this theorem is called the separation theorem for its implication that the investment decisions by inherently rational investors can be separated into two tasks : a common task of finding an efficient strategy $y \in \mathbb{E}(\Omega)$, and an individual task of choosing the optimal capital allocation a such that each $V \in \Psi_0$ is maximized (see Figure 6).[22]

Theorem 10 (existence of equilibrium) *Under Assumptions 1 –5, there exists a capital market equilibrium which is unique in that for any level of market risk* $v(m)$, *there is a unique market risk premium* $z(m) = u(m)/v(m)$ *such that all securities are uniquely priced according to Theorem 5.*

Proof: Dropping m and j for brevity, consider the necessary and sufficient condition for individual optimality

$$\rho(wau, wav) - \frac{u}{v} = 0 \tag{45}$$

From (42), for any levels v there exists $u > v$ such that a solution $a(u, v, w)$ exists satisfying (45). Further, $\partial\rho/\partial w > 0$ is equivalent to $\partial\rho/\partial a > 0$. It follows that

$$\begin{aligned} \frac{\partial a}{\partial u} &= (\frac{1}{v} - \frac{\partial \rho}{\partial u})/\frac{\partial \rho}{\partial a} > 0 \\ \frac{\partial a}{\partial v} &= -(\frac{\partial \rho}{\partial v} + \frac{u}{v^2})/\frac{\partial \rho}{\partial a} < 0 \end{aligned}$$

In other words, $a(u, v, w)$ strictly increases in u or in $z = u/v$, and strictly decreases in v.

Now, for fixed wealth levels, let $h(u, v) = \sum_{j \in J} a_j^*(u, v, w_j)w_j - m$ where $a_j^*(u, v, w_j)$ is an implicit function satisfying (45). The function h is continuous and strictly increasing in u for any given level of $v > 0$. When $u = v$ there is no demand for the risky assets since $\rho(wu, wv) \geq 1$ as assumed, in which case h is negative. As u increases and becomes sufficiently large, h will become positive because $\partial\rho_j/\partial u \leq 0$ implies that $\rho_j(w_j u, w_j v) < u/v$ as u becomes sufficiently large, i.e., all investors would like to borrow and invest more than 100% of their capital in the market portfolio. From the mean value theory, then, there exists a unique u or $z = u/v$ for any level of v such that $h(u, v) = 0$. □

This simple existence and uniqueness theorem, evidently, circumvents difficulties in deriving the existence and uniqueness equilibrium results in the mean-variance paradigm (e.g., Allingham, 1991; and Nielsen, 1988, 1990).[23] I close this section with a brief look at *Pareto efficiency.*

Definition 11 (Pareto efficiency) *The vector of strategies* $x^* = (x_1^*, x_2^*, ..., x_J^*)$, $x_j^* \in \Omega$, *is Pareto optimal if there is no feasible strategies* $x = (x_1, x_2, ..., x_J)$, $x_j \in \Omega$, *such that* $V_j(x_j, w_j) \geq V_j(x_j^*, w_j)$ *for all* $j \in J$ *with at least one strict inequality. If there is no feasible strategies* $x = (x_1, x_2, ..., x_J)$, $x_j \in \Omega_\infty$, *such that* $V_j(x_j, w_j) \geq V_j(x_j^*, w_j)$ *for all* $j \in J$ *with at least one strict inequality, then* x^* *is fully Pareto optimal and the market* Ω *is said to have attained Full Pareto Efficiency.*

The next theorem shows the implication of quasi-complete market for full Pareto efficiency (see, e.g., Amershi, 1985, for a general result on full Pareto efficiency). The proof is straightforward hence omitted.

Theorem 11 *Let the assumptions in Theorem 7 or in Theorem 8 hold, and assume that the market is in equilibrium. Then Ω attains Full Pareto Efficiency.*

6 CONCLUDING REMARKS

The contributions of this paper may be summerized in terms of two attributes: one of unifying and generalizing the known results, and one of presenting new insights and tools for analyzing investment strategies and security markets.

Formulated in the dichotomous-choice paradigm of Zou (2000a), individual preferences are more generally represented by their perceptions of reward and risk, and by their reward-risk utility. Since this new paradigm encompasses the expected utility models and other legitimate reward-risk utility models, the results obtained in this paper also apply to these special cases.

The paper re-enforces some important insights that are already known from the modern portfolio theory, such as the effect of risk diversification, the (inherent rather than mean-variance) efficiency of the market in equilibrium, the separation theorem, the role of "beta" in the determination of asset prices, the equilibrium implication for Pareto efficiency, etc., except that these results are derived here more generally and rigorously, being consistent with no-arbitrage and the first-order stochastic dominance (thus overcoming shortcomings of the mean-variance model).

The paper also presents many new insights. Most important, perhaps, is the implications of the dichotomous pricing model (Theorem 5) that (i) security returns in equilibrium or in the absence of approximate arbitrage can be analyzed separately in terms of their up-market and down-market potentials (see Figure 5); (ii) these potentials should be in the same proportion, measured by the security's beta, to those of the market; and (iii) all securities' prices can be expressed as the sum of two discounted values: the up-market expected value and the down-market expected value, and there are two common factors (δ_1 and δ_2) for discounting the expected values of all securities respectively. In deriving these results as a condition for market equilibrium, only homogenous beliefs on the *conditional* (on market being up or down) distributions of security returns are assumed. Thus the predictions hold even when investors differ in their opinions about the probability that the market will be up or down.

Another important insight is the implication of quasi-complete market and flat options for a unique state-price density (Theorem 8). This state-price density has also two parts, respectively in constant proportion to the conditional density of the primitive states (on the gain-domain and the loss-domain of the market). It is able to price both traded and non-traded securities, consistent with no-arbitrage, and with market equilibrium when investors have (weakly) concave gain- and loss-functions. This result unifies the existing (two-date) models of complete market and incomplete market into a single consistent framework.

Many issues remain to be investigated, however. I list here just a few thoughts (some reflected in Zou, 2000b) to conclude the paper. What are the inter-temporal properties of inherent reward and risk? Are the main results here extendable to the multiple-period investment (and consumption) environment? How to test the predictions of the model empirically? How to incorporate multiple sources of risk in the current model? How to incorporate more general heterogeneous information and/or beliefs? And so on...

APPENDIX

Proof of Theorem 2: Assume that there exists $x^* \in \Omega$ such that $V(x,w) \leq V(x^*,w)$ for all $x \in \Omega$, which implies by Theorem 1 that $Z(x,w) \leq Z(x^*,w) < \infty$ for all $x \in \Omega$. Thus for any strategy $x \in \Omega$ and any real number $a \in (-\epsilon, 1)$ where ϵ is a sufficiently small positive number such that the portfolio $\pi = ax + (1-a)x^* \in \Omega$,[24]

$$Z(x^*,w) \geq Z(ax + (1-a)x^*, w), \qquad \forall a \in (-\epsilon, 1).$$

the return on portfolio at time 1 is $\pi(r) = ax(r) + (1-a)x^*(r)$ if the realized state is r. By definition,

$$\begin{aligned} U(\pi,w) &= \int_{\mathbb{G}(\pi)} [g(awx(r) + (1-a)wx^*(r)) - g(wr_0)]dF(r) \\ D(\pi,w) &= \int_{\mathbb{L}(\pi)} [\ell(wr_0) - \ell(awx(r) - (1-a)wx^*(r))]dF(r). \end{aligned}$$

It follows from x^* being optimal that $\partial Z(\pi,w)/\partial a \leq 0$ at $a = 0_+$ and $\partial Z(\pi,w)/\partial a \geq 0$ at $a = 0_-$. Equivalently,

$$\frac{\partial U(\pi,w)}{\partial a} D(x^*,w) - \frac{\partial D(\pi,w)}{\partial a} U(x^*,w) \leq 0 \text{ at } a = 0_+ \tag{46}$$

$$\frac{\partial U(\pi,w)}{\partial a} D(x^*,w) - \frac{\partial D(\pi,w)}{\partial a} U(x^*,w) \geq 0 \text{ at } a = 0_- \tag{47}$$

Since

$$\begin{aligned} & U(\pi,w) - U(x^*,w) \\ = & \int_{\mathbb{G}(\pi)} [g(w\pi(r)) - g(wr_0)]dF(r) - \int_{\mathbb{G}(x^*)} [g(wx^*(r)) - g(wr_0)]dF(r) \\ = & \int_{\mathbb{G}(x^*)} [g(w\pi(r)) - g(wx^*(r))]dF(r) + \int_{\mathbb{G}(\pi)\cap\mathbb{L}(x^*)} [g(w\pi(r)) - g(wr_0)]dF(r), \end{aligned}$$

taking derivative from the right at $a = 0$ yields

$$\frac{\partial U(\pi,w)}{\partial a}|_{a=0_+} = \lim_{a\to 0_+} \frac{1}{a}[\int_{\mathbb{G}(x^*)} [g(wx^*(r) + aw(x(r) - x^*(r))) - g(wx^*(r))]dF$$

$$+\int_{\mathbb{M}\cap\{r:a(x(r)-x^*(r))\geq r_0-x^*(r)\geq 0\}}[g(wx^*(r)+aw(x(r)-x^*(r)))-g(wr_0)]dF]$$
$$= \int_{\mathbb{G}(x^*)} g'(wx^*(r))w(x(r)-x^*(r))dF+\int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} g'(wr_0)w(x(r)-r_0)dF$$

Similarly,

$$\begin{aligned}
\frac{\partial U(\pi,w)}{\partial a}|_{a=0_-} &= \int_{\mathbb{G}(x^*)} g'(wx^*(r))w(x(r)-x^*(r))dF+\int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} g'(wr_0)w(x(r)-r_0)dF \\
\frac{\partial D(\pi,w)}{\partial a}|_{a=0_+} &= \int_{\mathbb{L}(x^*)} \ell'(wx^*(r))w(x^*(r)-x(r))dF+\int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)w(r_0-x(r))dF \\
\frac{\partial D(\pi,w)}{\partial a}|_{a=0_-} &= \int_{\mathbb{L}(x^*)} \ell'(wx^*(r))w(x^*(r)-x(r))dF+\int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)w(r_0-x(r))dF.
\end{aligned}$$

Substituting the above into (46) and (47) yields

$$\begin{aligned}
& \int_{\mathbb{G}(x^*)} g'(wx^*(r))w(x(r)-x^*(r))dF+\int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} g'(wr_0)w(x(r)-r_0)dF \\
\leq\ & Z(x^*,w)[\int_{\mathbb{L}(x^*)} \ell'(wx^*(r))w(x^*(r)-x(r))dF+\int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)w(r_0-x(r))dF] \quad (48) \\
& \int_{\mathbb{G}(x^*)} g'(wx^*(r))w(x(r)-x^*(r))dF+\int_{\mathbb{L}(x)\cap\mathbb{N}(x^*)} g'(wr_0)w(x(r)-r_0)dF \\
\geq\ & Z(x^*,w)[\int_{\mathbb{L}(x^*)} \ell'(wx^*(r))w(x^*(r)-x(r))dF+\int_{\mathbb{G}(x)\cap\mathbb{N}(x^*)} \ell'(wr_0)w(r_0-x(r))dF]. \quad (49)
\end{aligned}$$

In particular, for $x=r_0$, (48) and (49) collapse to an equation:

$$\int_{\mathbb{G}(x^*)} g'(wx^*(r))w(r_0-x^*(r))dF(r) = Z(x^*,w)\int_{\mathbb{L}(x^*)} \ell'(wx^*(r))w(x^*(r)-r_0)dF(r). \quad (50)$$

Subtracting (50) from (48) and (49) yields then the statement of the theorem. □

Proof of Lemma 1: I only prove the statement concerning v since the one concerning u is analogous. Let any $x\in\Omega$ and $y\in\Omega$ and any real number $a\in(0,1)$ be given and consider the inherent risk of portfolio $\pi=ax+(1-a)y\in\Omega$:

$$v(\pi) = \int_{\mathbb{L}(\pi)}[a(r_0-x(r))+(1-a)(r_0-y(r))]dF(r)$$

$$\begin{aligned}
&= a[\int_{\mathbb{L}(\pi)\cap\mathbb{L}(x)} (r_0 - x(r))dF + \int_{\mathbb{L}(\pi)\cap\mathbb{G}(x)} (r_0 - x(r))dF] && (51)\\
&\quad +(1-a)[\int_{\mathbb{L}(\pi)\cap\mathbb{L}(y)} (r_0 - y(r))dF + \int_{\mathbb{L}(\pi)\cap\mathbb{G}(x)} (r_0 - y(r))dF] && (52)\\
&\le \int_{\mathbb{L}(\pi)\cap\mathbb{L}(x)} a(r_0 - x(r))dF + \int_{\mathbb{L}(\pi)\cap\mathbb{L}(y)} (1-a)(r_0 - y(r))dF && (53)\\
&\le av(x) + (1-a)v(y). && (54)
\end{aligned}$$

Thus the first part of the lemma is proved. The following statements are in the sense of *almost everywhere* with respect to the probability distribution P, although not repeated.

Sufficiency: Suppose $P(COD(x, y)) = 1$, then

$$x(r) > r_0 \Rightarrow y(r) \ge r_0 \Rightarrow \pi(r) > r_0, \text{ thus } P(\mathbb{L}(\pi) \cap \mathbb{G}(x)) = 0,$$

which implies that the second terms in (51) and (51) are zero. Similarly, the second term in (52) is zero. It follows that the strict equalities in (53) holds. Further,

$$\begin{aligned}
x(r) &\le r_0 \Leftrightarrow y(r) \le r_0 \Rightarrow \pi(r) \le r_0,\\
\text{thus } P(\{r &: \pi(r) \le r_0, x(r) \le r_0\}) = P(\{r : x(r) \le r_0\}) = P(\{r : y(r) \le r_0\})
\end{aligned}$$

It follows that the strict equality in (54) holds.

Necessity: Suppose $P(COD(x, y)) < 1$. Then there exists a subset $\Delta \subseteq \mathbb{M}$, $P(\Delta) > 0$, on which the excess returns of the two strategies have opposite signs. Without loss of generality suppose that $x(r) > r_0$ and $y(r) < r_0$ on Δ.

Now suppose contrary to what will be proved that the equalities in both in (53) and (54) hold, which means that $P(\mathbb{L}(\pi) \cap \mathbb{G}(x)) = 0$ by inspecting the second terms in (51) and (51). This would imply, however, that almost everywhere on $\mathbb{M}$

$$x(r) > r_0 \Rightarrow \pi(r) \ge r_0 \Rightarrow a(x(r) - r_0) \ge -(1-a)(y(r) - r_0). \quad (55)$$

Moreover, equality in (54) implies that almost everywhere on $\mathbb{M}$

$$y(r) < r_0 \Rightarrow \pi(r) < r_0 \Rightarrow a(x(r) - r_0) < -(1-a)(y(r) - r_0). \tag{56}$$

Clearly, the inequalities in (55) and (56) contradict each other on Δ. As a result, at least one of the inequalities in (53) and (54) must hold strictly. $\square$

Proof of Theorem 7: Let any $V \in \Psi$ be given. If investment strategy $x^* \in \Omega$ is a solution to $\mathbb{P}$, then the relationships in (7) of Theorem 2 hold for all $x \in \Omega$. Let strategy x whose time-1 payoff is defined as follows.

$$\begin{aligned} x(r) &= x^*(r) \text{ on } G \cap \mathbb{G}(x^*) \text{ or } L \cap \mathbb{L}(x^*) \\ &= r_0 \ \text{ otherwise} \end{aligned}$$

Clearly, if the cost of x is less than or equal to 1 then $x \in \mathbb{E}(\Omega)$. This will be verified shortly to be true. For the moment, however, it is convenient to assume that x costs 1 so that the conditions in (7) can be used. Substituting $x^*(r)$ and $x(r)$ into (7) yields then, respectively,

$$\int_{\mathbb{G}(x^*)} g'(wx^*(r))(x^*(r) - r_0)dF(r) = Z(x^*, w) \int_{\mathbb{L}(x^*)} \ell'(wx^*(r))(r_0 - x^*(r))dF(r) \tag{57}$$

and

$$\int_{\mathbb{G}(x^*)} g'(wx^*(r))(x(r) - r_0)dF(r) = Z(x^*, w) \int_{\mathbb{L}(x^*)} \ell'(wx^*(r))(r_0 - x(r))dF(r) \tag{58}$$

Subtracting (58) from (57), however, yields

$$\int_{L\cap\mathbb{G}(x^*)} g'(wx^*(r))(x^*(r) - r_0)dF(r) = Z(x^*, w) \int_{G\cap\mathbb{L}(x^*)} \ell'(wx^*(r))(r_0 - x^*(r))dF(r), \tag{59}$$

from which it follows that $P(L \cap \mathbb{G}(x^*)) > 0$ if and only if $P(G \cap \mathbb{L}(x^*)) > 0$. If $P(L \cap \mathbb{G}(x^*)) = P(G \cap \mathbb{L}(x^*)) = 0$ then $x^* \in \mathbb{E}(\Omega)$ by Theorem 3, and the theorem is proved. If otherwise, by (37) the cost difference between the two strategies

$$
\begin{aligned}
& \int_{\mathbb{M}} (x^*(r) - x(r))\psi(r)dr \\
= \; & \int_{G\cap\mathbb{L}(x^*)} (x^*(r) - r_0)\delta_1 f(r|G)dr + \int_{L\cap\mathbb{G}(x^*)} (x^*(r) - r_0)\delta_2 f(r|L)dr \\
= \; & \frac{\delta_1}{P(G)}[\overline{z}\int_{L\cap\mathbb{G}(x^*)} (x^*(r) - r_0)f(r)dr - \int_{G\cap\mathbb{L}(x^*)} (r_0 - x^*(r))f(r)dr] \\
> \; & \frac{\delta_1}{P(G)}[\int_{L\cap\mathbb{G}(x^*)} g'(x^*(r))(x^*(r) - r_0)f(r)dr - Z(x^*, w)\int_{G\cap\mathbb{L}(x^*)} \ell'(x^*(r))(r_0 - x^*(r))f(r)dr] \\
= \; & 0
\end{aligned}
$$

where the strict inequality comes from $\overline{z} > 1$, $Z(x^*, w) \geq Z(r_0, w) = 1$, and Assumption 1 ($g' \leq \ell'$). This contradicts x^* being optimal, however, since raising x to some $x + \epsilon$ on $\mathbb{G}(x^*)$ and $\mathbb{L}(x^*)$ will change the "=" in (58) into ">" implying that x^* is suboptimal. As a result, x^* must have perfect co-domain with x. By Theorem 3 again, $x^* \in \mathbb{E}(\Omega)$. □

REFERENCES

Allingham, M., "Existence Theorems in the CAPM." *Econometrica*, 59, pp. 1169-1174.

Amershi, Amin H., "A Complete Analysis of Full Pareto Efficiency in Financial Markets for Arbitrary Preferences." *The Journal of Finance*, Vol. XL, no. 4, September 1985, pp. 1235-1243.

Arrow, K.J., "The Role of Securities in the Optimal Allocation of Risk-bearing." The Review of Economic Studies, Vol. 31, no. 2, April, 1964, pp. 91-96.

Arrow, K. J. *Essays in the Theory of Risk Bearing.* Markham Publishing Company, Chicago, 1971.

Artzner, Philippe, Freddy Delbaen, and Jean-Marc Eber, "Coherent Measure of Risk." *Journal of Mathematical Finance*, vol. 9, no. 3, July 1999, 203-228.

Bawa, Vijay S. "Optimal Rules for Ordering Uncertain prospects." *Journal of Financial Economics*, 2, 1975, pp. 95-121.

Bawa, Vijay S. "Admissible Portfolios For All Individuals." *The Journal of Finance*, vol. XXXI, no. 4, September 1976, pp. 1169-1183.

Bawa, Vijay S. and Lindenberg, Eric B. "Capital market Equilibrium in A Mean-Lower Partial moment Framework." *Journal of Financial Economics*, 5, 1977, pp. 189-200.

Bawa, Vijay S. "Safety-First, Stochastic Dominance, and Optimal Portfolio Choice." *Journal of Financial and Quantitative Analysis*, June 1978, pp. 255-269.

Bell, David E., and Fishburn, Peter C., "Utility Functions of Wealth." *Journal of Risk and Uncertainty*, 20, 2000. pp. 5-44

Benartzi, Shlomo, and Richard H. Thaler. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, February, 1997, pp. 73-92.

Bernardo, Antonio E., and Olivier Ledoit, "Gain, Loss, and Asset Pricing." *Journal of Political Economy*, vol. 108, no. 1, 2000, pp. 144-172.

Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81, no. 3, May/June 1973, pp. 637-54.

Brennan M. J. and Solanki, R. "Optimal portfolio Insurance" *Journal of Financial and Quantitative Analysis*, vol. XVI, no. 3, September 1981, pp. 279-300.

Campbell, John Y., "Asset Pricing at the Millennium." *The Journal of Finance*, Vol. LV, No. 4, August 2000, pp. 1515-1567.

Cass, David and Stiglitz, Joseph E. "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, 1970, pp. 122-160.

Clark, Stephen A., "Arbitrage Approximation Theory." *Journal of Mathematical Economics*, 33, 2000, pp. 167-181.

Dana, R.A., "Existence, Uniqueness and Determinacy of Equilibrium in CAPM with a riskless asset." *Journal of Mathematical Economics*, 32, no. 2, 1999, pp. 167-175.

Debreu, G. "Theory of Value: An Axiomatic Analysis of General Equilibrium." *New Haven: Yale University Press*, 1959.

Diamond, P., and J. E. Stiglitz. "Increase in Risk and in Risk Aversion." *Journal of Economic Theory*, 8, 1974, pp. 337-360.

Domar, E.V., and R.A. Musgrave. "Proportional Income Taxation and Risk-Taking." *Quarterly Journal of Economics*, vol. 58, May, 1944, pp. 389-422.

Dybvig, Philip H. "Distributional Analysis of Portfolio Choice." *Journal of Business*, vol. 61, no. 3, 1988, pp. 369-393.

Dybvig, Philip H. "Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market." *Review of Financial Studies*, Vol. 1, no. 1, (1988), pp. 67-88.

Dybvig, Philip H. and Jr. Ingersoll, Jonathan E. "Mean-Variance Theory in Complete Markets." *Journal of Business*, vol. 55, no. 2, 1982, pp. 233-251.

Dybvig, Philip H. and Ross, Stephen. "Portfolio Efficient Sets." *Econometrica*, vol. 50, XL, no. 6, Nov. 1982, pp. 1525-1546.

Dybvig, Philip H. and Ross, Stephen A. "Short Sales Restrictions and Kinks on the Mean Variance Frontier." *The Journal of Finance*, vol. 39, no. 1, March 1984, pp. 239-244.

Dybvig, Philip H. and Ross, Stephen A. "Differential Information and Performance Measurement Using a Security Market Line." *The Journal of Finance*, vol. XL, no. 2, June 1985a, pp. 383-399.

Dybvig, Philip H. and Ross, Stephen. "The Analytics of Performance Measurement Using a Security market Line." *The Journal of Finance*, vol. XL, no. 2, June 1985b, pp. 401-416.

Dybvig, Philip H. and Huang, Chi-fu. "Nonnegative Wealth, Absence of Arbitrage, and Feasible Consumption Plans." *Review of Financial Studies*, vol. 1, no. 4, Winter 1988, pp. 377-401.

Ellsberg, D. "Risk, Ambiguity and the Savage Axioms,", *Quarterly Journal of Economics*, 75, pp. 643-669.

Elton, Edwin J. "Expected Return, Realized Return, and Asset Pricing Tests." *The Journal of Finance*, vol. LIV, no. 4, August 1999, pp. 1199-1220.

Fishburn, Peter C. "Mean-Risk Analysis with Risk Associated with Below-Target Returns" *The American Economic Review*, vol. 67, no. 2, March 1977, pp. 116-126.

Friedman, M., and L. J. Savage. "The Utility Analysis of Choices Involving Risk." *Journal of Political Economy*, 56, 1948, pp. 279-304.

Hadar, J., and W. Russell. "Rules for Ordering Uncertain Prospects." *American Economic Review*, 59, 1969, pp. 25-34.

Hagen, O. "Towards a Positive Theory of Preference Under Risk." *Expected Utility Hypothesis and the Allais Paradox* ed. By M. Allais and O. hagen. Dordrecht, Holland: D. Reidel Publishing Company, 1979, pp. 271-302.

Hakansson, Nils H., "Welfare Aspects of Options and Supershares." *The Journal of Finance*, Vol. XXXIII, no. 3, June 1978, pp. 759-776.

Hakansson, Nils H., "Changes in the Financial Market: Welfare and Price Effects and the Basic Theorems of Value Conservation." *The Journal of Finance*, Vol. XXXVII, no. 4, Sept. 1982, pp. 977-1004.

Harlow, W.V., and Ramesh K.S. Rao. 1989. "Asset Pricing in a Generalized Mean-lower Partial Moment Framework: Theory and Evidence." *Journal of Financial and Quantitative Analysis*, (September):285-311.

Harlow, W.V. "Asset Allocation in a Downside-risk Framework." *Financial Analysts Journal*, vol. 47, no. 5, September/October 1991, pp. 28-40.

Hart, Oliver D., "On the Existence of Equilibrium in a Securities Model." *Journal of Economic Theory*, 9, 1974, pp. 293-311.

Herstein, I. N., and J. Milnor, "An Axiomatic Approach to Measurable Utility." *Econometrica*, 21, 1953, pp. 291-297.

Hicks, J.R., *Value and Capital.* Oxford University Press, Fair Lawn, N.J., 1939.

Kahneman, Daniel, and Amos Twersky. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, vol. 47, no.2, March 1979, pp. 263-291.

Knight, F.H., *Risk, Uncertainty and Profit.* New York: Houghton Mifflin 1921 (*Reprints of Economic Classics*, New York, 1964).

Kreps, D.M., "A Representation Theorem for 'Preference for Flexibility'." *Econometrica*, 47, 1979, pp. 565-577.

Kreps, D.M., "Arbitrage and Equilibrium in Economies with Infinitely many Commodities." *Journal of Mathematical Economics* 8, 1981, pp.15-35.

Leland, Hayne E. "Who Should Buy Portfolio Insurance?" *The Journal of Finance*, vol.XXXV, no. 2, may 1980, pp. 581-594.

Levy, Haim, and Yoram Kroll. "Stochastic Dominance with Riskless Assets." *Management Science*, December 1976, pp. 743-777.

Lintner, J. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics*, 47, Feb. 1965a, 13-37.

Lintner, J. "Security Prices, Risk, and Maximal Gains from Diversification." *Journal of Finance*, 20, Dec. 1965b, pp. 587-615.

MacCrimmon, K. R., and S. Larsson. "Utility Theory: Axioms versus 'Paradoxes'." *Expected Utility Hypothesis and the Allais Paradox* ed. by M. Allais and O. Hagen. Dordrecht, Holland: D. Reidel Publishing Company, 1979.

Machina, M. J. "Expected Utility Analysis without the Independence Axiom." *Econometrica*,50, 277-323.

Malinvaud, E. "Note on von Neumann-Morgenstern's Strong Independence Axiom." *Econometrica*, 20, 1952, 679.

Markowitz, Harry. “Portfolio Selection.” *Journal of Finance* 7, no. 1, 1952, pp. 77-91.

Markowitz, Harry. *Portfolio Selection: Efficient Diversification of Investments*, New York: Wiley, 1959.

Marschak, J. “Rational Behavior, Uncertain Prospects, and Measurable Utility.” *Econometrica*, 18, 1950, pp. 111-141. (“Errata,” *Econometrica*, 18, 1950, 312.)

Mehra, Rajnish and Prescott, Edwards C. “The Equity Premium-A Puzzle.” *Journal of Monetary Economics*, 15, 1985, pp. 145-161.

Merton, Robert C. *Continuous-Time Finance*, Blackwell Publishers Inc. MA, USA. 1997.

Mondigliani, Franco, and M.H. Miller. “The Cost of Capital, Corporation Finance and the Theory of Investment.” *American Economic Review*, 48, June 1958, pp. 261-297.

Morgan, J.P. *RiskMetrics Monitor*, 4th Quarter, 1995.

Morrison, D. G. “On the Consistency of Preferences in Allais’s Paradox.” *Behavioral Science*, 12,1967, pp. 373-383.

Moskowitz, H. “Effects of Problem Representation and Feedback on Rational Behavior in Allais and Morlat-Type problems.” *Decision Science*, 5, 1974, pp. 225-242.

Mossin, Jan. “Equilibrium in a Capital Asset Market.” *Econometrica*, 34, 1966, pp. 768-783.

Nielsen, L.T., “Uniqueness of Equilibrium in the Classical Capital Asset Pricing Model.” *Journal of Financial and Quantitative Analysis*, 23, 1988, pp. 329-336.

Nielsen, L.T., “Existence of Equilibrium in CAPM.” *Journal of Economic Theory*, 52, 1990, pp. 223-231.

Peleg, B., and Yaari, M.E., "A Price Characterization of Efficient Random Variables." *Econometrica*, 43, no. 2, March 1975, pp. 283-292.

Porter, R. Burr. "Semivariance and Stochastic Dominance: A Comparison." *American Economic Review*, vol. 64, no. 1, March 1974, pp. 200-204.

Pratt, J. W. "Risk Aversion in the Small and in the Large." *Econometrica*, 32, 1964, pp. 122-136.

Pratt, John W. and Zeckhauser, Richard J. "Proper Risk Aversion." *Econometrica*, vol. 55, no. 1, January 1987, pp. 143-154.

Ross, Stephen A. "Options and Efficiency." *Quarterly Journal of Economics*, 90, 1976, pp. 75-89.

Ross, Stephen A. "Mutual Fund Separation in Financial Theory-The Separating Distributions." *Journal of Economic Theory*, 17, 1978, pp. 254-286.

Rothschild, M., and J. Stiglitz. "Increasing Risk: I. A Definition." *Journal of Economic Theory*, 2, 1970, pp. 225-243. ("Addendum," *Journal of Economic Theory*, 5, 1972, 306.)

Rothschild, M., and J. Stiglitz. "Some Further Results on the Measurement of Inequality." *journal of Economic Theory*, 6, 1973, pp. 188-204.

Roy, A.D. "Safety-First and Holding of Assets." *Econometrica*, Vol. 20, 1952. pp. 431-449.

Rubinstein, Mark. "An Aggregation Theorem for Securities markets." *journal of Financial Economics*, 1, 1974, pp. 225-244.

Rubinstein, Mark. "Implied Binomial Trees." *The Journal of Finance*, Vol. LXIX, no. 3, 1994, pp. 771-818.

Samuelson, P. A. "Utility, Preference, and Probability." abstract of a paper given before the conference on "Les Fondements et Applications de la Theorie du Risque en Econometrie." Paris, May, 1952. Reprinted as Chapter 13 in *Collected Scientific Papers of Paul A. Samuelson*, Volume I ed. by J. Stiglitz. Cambridge. Mass.: MIT Press, 1966.

Samuelson, Paul A. "The Fundamental Approximation Theorem of Portfolio Analysis in terms of Means, Variances and Higher Moments." *Review of Economics Studies*, 37, 1970, pp. 537-42.

Savage, L.J., *The Foundations of Statistics*, New York: John Wiley & Sons, 1954. (New York: Dover Publications, 1972).

Sharpe, William F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance*, 19, 1964, pp. 423-439.

Sharpe, William F. "Mutual Fund Performace." *Journal of Business*, January 1966, pp. 119-138.

Shelley, Marjorie K. "Gain/Loss Asymmetry in Risky Intertemporal Choice" *Organizational Behavior and Human Decision Processes*, 59, 1994, pp. 124-159.

Starmer, Chris, "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk." *Journal of Economic Literature*, Vol. XXXVIII, June 2000, pp. 332-382.

Stoll, Hans R., "Friction." *The Journal of Finance*, Vol. LV, No. 4, August 2000, pp. 1479-1514.

Stone, B.K., "A General Class of Three Parameter Risk Measures." *Journal of Finance*, vol. 28, June 1973, pp. 675-685.

Tobin, J. "Liquidity Preference as Behavior Toward Risk." *Review of Economic Studies*, 25, 1958, pp. 65-85.

Tversky, A. "Utility Theory and Additivity Analysis of Risky Choices." *journal of Experimental Psychology*, 75, 1967a, pp. 27-36.

Tversky, A. "Additivity, Utility, and Subjective Probability." *Journal of Mathematical Psychology*, 4, 1967b, pp. 175-201.

Tversky, A. "Intransitivity of Preferences." *Psychological Review*, 76, 1969, pp. 31-48.

Tversky, A. "A Critique of Expected Utility Theory: Descriptive and Normative Considerations." *Erkenntnis*, 9, 1975, pp. 163-173.

Von Neumann, J. and O. Morgenstern "Theory of games and Economic Behavior." *Princeton: Princeton University Press*, 1944. Second Edition, 1947. Third Edition, 1953.

Williams, Joseph T. "Capital Asset Prices with Heterogeneous Beliefs." *Journal of Financial Economics*, 5, 1977, pp. 219-239.

Zou, Liang. "Inherent Reward and Risk (Part I): Towards a Universal Paradigm for Investment Analysis." Tinbergen Institute Discussion Paper Series, 2000a, TI 2000 050/2. Revised under new entitle: "The Dichotomous Theory of Choice under Risk." Working Paper, Finance Group, University of Amsterdam.

Zou, Liang. "Inherent Reward and Risk (Part II): Dynamic Investment Strategies and Pricing of Contingent Claims." Working Paper, Finance Group, University of Amsterdam, 2000b.

FOOTNOTES

1. Part of this paper was written while I was visiting Hong Kong Baptist University in the fall of 1999. The generous support of my host, the helpful comments by Rod Aya, Henk Koster, Kin Lam, Wei Li, colleagues and seminar participants at University of Amsterdam, Hong Kong Baptist University, University of Hong Kong, Tinbergen Institute and, in particular, the detailed remarks by Laixiang Sun, Florian Wagener and Peter Wakker on earlier versions, are gratefully acknowledged. All errors remain to be my own, of course.

2. The reader need not have prior knowledge of this theory, a summary of which is presented in Sections 2 and 3 below.

3. For a recent survey and other references, see Campbell (2000).

4. Recently, Bernardo and Ledoit (2000) present a pricing model that has some structural similarity with the inherent-model to be developed here. Their gain-loss ratio, however, is computed using risk-adjusted probabilities of some given benchmark portfolio whereas I work with probabilities that are actually believed by the investors. Bernardo and Ledoit obtain relationships between bounds of the pricing kernel and of the gain-loss ratio, assuming a discrete state space. In the absence of a theory of choice, though, these bounds are quite arbitrary. In contrast, the pricing models to be presented here are built within a rigorous theoretical framework and offer much sharper (e.g., a precise prediction of security prices) and more general (e.g., on a more general state space) results.

5. For a recent survey of utility functions for wealth, see Bell and Fishburn (2000).

6. In Zou (2000b) I show how this two-date model can be readily extended to a multi-period

model with path consistency.

7. These payoffs might also be interpreted as a consequence of some dynamic trading strategies pre-programed to mimic a targeted payoff structure at time 1. However, since this interpretation requires further assumptions that are unnecessary in this static setting (e.g., smooth security price processes and the ability of the traders to commit to such mimicking strategies over time), I choose to interpret such payoffs as a result of the derivative contracts.

8. It is difficult to construe the concept of negative wealth. The problem lies in the willingness of the creditor to treat the debtor's negative wealth as his asset. If the debtor has the ability to render service or defer payment of his debt, then his current wealth should reflect the value of his service or the present value of his future wealth. Even when a creditor enjoys the pure psycho-pleasure of "consuming" his debtor's negative wealth, such pleasure has value that can be used to reduce the debtor's negative wealth. After depleting all such possibilities the remaining liabilities, if any, has to be written off from the creditor's asset side of the balance sheet. Analytically, it suffices to assume that $r_i \in [0, M]$ almost surely, i.e., the probability that $r_i > M$ is zero. Note also that the analysis is easier without the non-negative wealth constraint. The purpose here is not to simplify, but to show how the thorny issue of negative wealth can be easily resolved in the present analytical framework. See, e.g., Dybvig and Huang (1988) on nonnegative wealth and its implications.

9. Alternatively, one can assume that the brokerage firms or the clearing houses are vigilant enough so that an account is automatically liquidated as soon as its value falls to zero. From the viewpoint of dynamic rebalancing, this is equivalent for the investor to buying a put option at the start so that he never risks bankruptcy.

10. Implicit in this definition is that every strategy must have an initial cost to establish, and that the cost is finite. Investors with zero wealth, for instance, will not be allowed to trade risky assets – for else, under risk, negative wealth may occur at time 1. This restriction, however, does not preclude the consideration of derivative securities that have a zero cost, since it is easy to combine such derivatives with a risk-free asset and analyze the combined portfolio as a strategy. Similarly, if a derivative contract has a negative value, then investors on the other side of the contract must see it having a positive value.

11. See, e.g., Leland (1980) and Brennan and Solanki (1981) for similar modelling of strategies as functions of the underlying asset. Existing models usually consider only one single risky asset as the underlying, whereas here x is a function of a n-dimensional vector.

12. Note that assuming an upper bound for all the investment returns does not preclude possible arbitrage strategies. An arbitrage strategy does not have to mean that with 1 dollar investment one could make *instantly* arbitrarily high profits. It is more properly defined as a strategy that could realize an above-normal risk-free return (e.g., higher than the risk-free interest rate), and it must able to be (infinitely) repeated for (infinitely) high profits.This assumption, however, does preclude many interesting distribution functions, such as normal distributions. In my opinion, it should not be considered as a limitation of the theory. Instead, since the total wealth of the world economy is bounded at any given point in time, it is more realistic and rigorous to preclude the unbounded return distributions. The inconsistency between short-sales and unbounded stock returns can thus be avoided. Justification for using unbounded distributions to approximate reality can be relegated to applications of the theory.

13. This level of wealth, in nominal terms, may depend on the current prices of the securities as w_j may be interpreted as the total value of the investor's current portfolio based on market prices. However, since the prices of the primitive securities are normalized to 1 and the numbers of the aggregate supply of the primitive securities are assumed to be fixed, w_j is a constant for each investor at time 0.

14. In the mean-variance analysis there is a similar problem that kinks may occur on the mean-variance frontier. Dybvig (1984) shows that this happens only under some rare conditions, and with a short-sales restriction. Note that the inequalities here in (7) apply to subjectively believed security prices (or returns) and the optimal strategy can differ across investors.

15. The reason to call such investors "inherently rational" is that examples can be constructed in which investors who do not agree with the inherent measures of reward and risk may subject themselves to "mispricing to their disadvantage"; in other words, they may create arbitrage opportunities for the others. This subject is not persued here, however.

16. Efficiency criterions vary, of course, with the individual objectives. In the expected utility paradigm, for instance, see Peleg and Yaari (1975) on efficiency sets of security prices, and Dybvig and Ross (1982) on efficiency sets of portfolio decisions.

17. See Clark (2000) for more on approximate arbitrage in a model of infinitely many commodities proposed by Kreps (1981).

18. Even if in theory a market can be spanned efficiently with only supershares (Hakanssen 1978, 1982) or with only options on an efficient portfolio, (e.g., Ross, 1976), this requires a complete spectrum of supershares or options with a full range of exercise prices that we do not observe. Also, even if in theory strategies can be continuously rebalanced to mimic the payoff of any

option (portfolios), this requires unrealistic assumptions that there is no transaction costs and that the stock returns follow a continuous stochastic process (e.g., Black and Scholes, 1973; Merton, 1997). Fortunately, these concerns become superfluous in a quasi-complete market provided that a single "flat option" on the market portfolio can be traded (with a weakly concave restriction on investors' perception of reward and risk, thus allowing weakly concave von Neumann-Morgenstern utilities).

19. The reader may wonder if it is possible to approximately arbitrage by buying the up-market states $[r_0, r_0 + \varepsilon] \subset G$ and short selling the down-market states $[r_0 - \varepsilon, r_0] \subset L$ for infinitesimally small ε. The answer is no. Provided $\varepsilon > 0$, it is easy to veryfy that such a strategy (with the rest in cash) yields an inherent reward-to-risk ratio that is less or equal to $z(y)$, $y \in \mathbb{E}(\Omega)$. On the other hand, there is a bound ($\varepsilon \geq \varepsilon_0 > 0$) on the maximum short positions taken in such a strategy before incurring hedging costs under limited liability (e.g., see Example 1).

20. Requiring market to clear at both times ensures consistency with the nonnegative-wealth constraint. In the equilibrium studies of market models it is often only required that the market clears at time 0 (see, e.g., Hart, 1974; Allingham, 1991; Dana, 1999; and Nielsen, 1988, 1990).

21. There is no need to introduce the flat-put option since it can be constructed using the risk-free asset and the flat-call only by the put-call parity for flat options (22). It is used here for convenience.

22. For other separation theories in the framework of expected utility or in more restricted classes of return distributions see, e.g., Cass and Stiglitz (1982), Ross (1978), and Rubinstein (1974).

23. An unverified conjecture is that this existence theorem could be extented to a quasi-complete market with general perceptions of reward and risk, provided that the total wealth of the inherently rational investors is sufficiently large so that they can absorb the flat-option effect. It is obvious that an equilibrium cannot exist if all investors prefer flat options, since flat options assign fixed payoffs whereas the market returns are generically random.

24. Since there is a uniform upper bound for the payoffs of all the perception on which the strategy $x \in \Omega$ is defined, a short position in x with sufficiently small exposure $\alpha > -\epsilon$ is also a feasible strategy in Ω, i.e., without risking losing more than 100% almost surely.

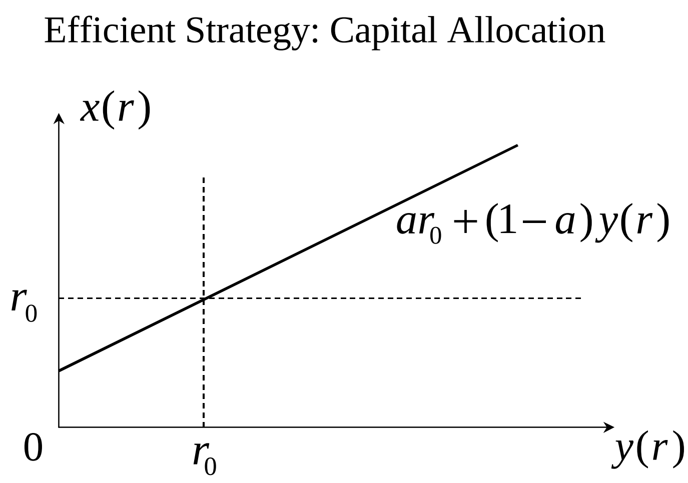

Figure 1: A linear combination of y and a risk-free asset is inherently efficient.

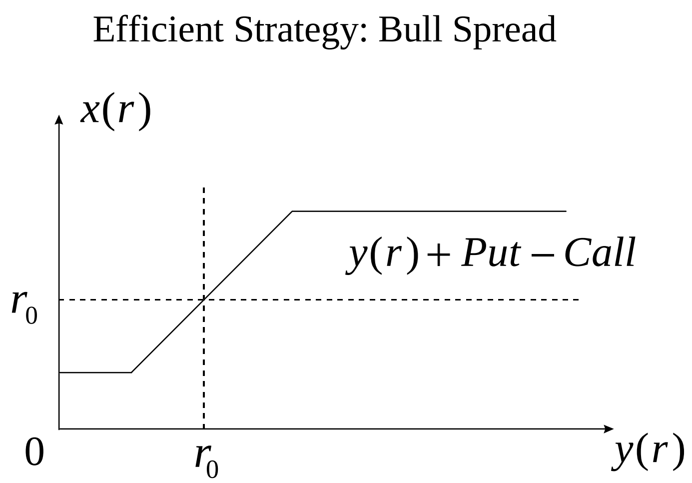

Figure 2: A bull spread can be inherently efficient provided it is constructed to have perfect co-domain with y.

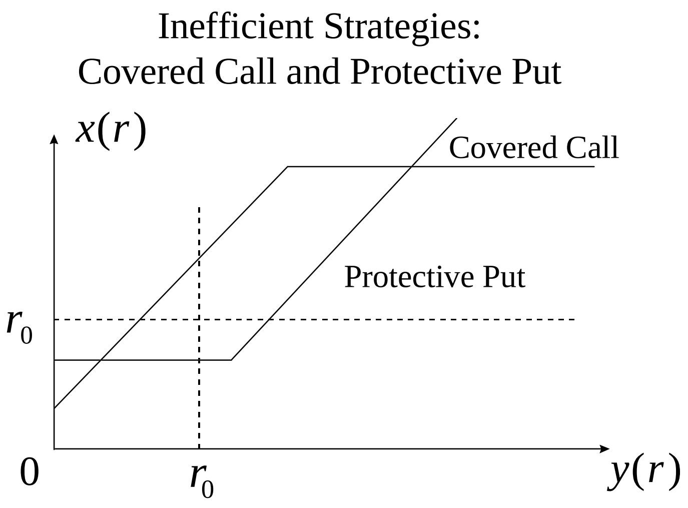

Figure 3: Both the covered call strategy and the protective put strategy as depicted are inefficient because they do not have perfect co-domain with y.

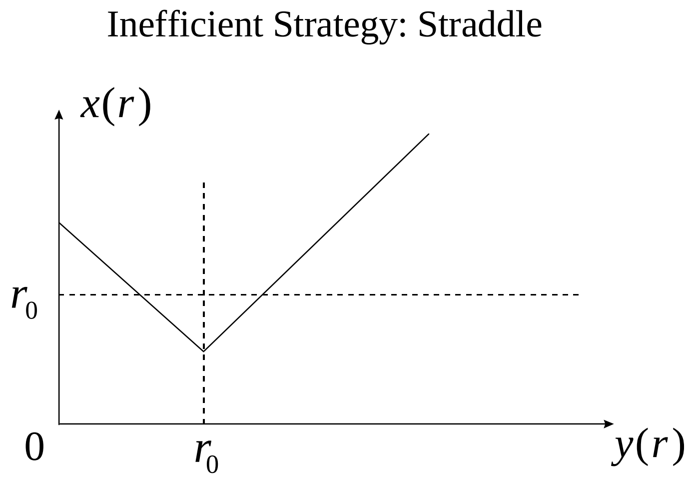

Figure 4: A straddle cannot be efficient because it never can be made to have perfect co-domain with y.

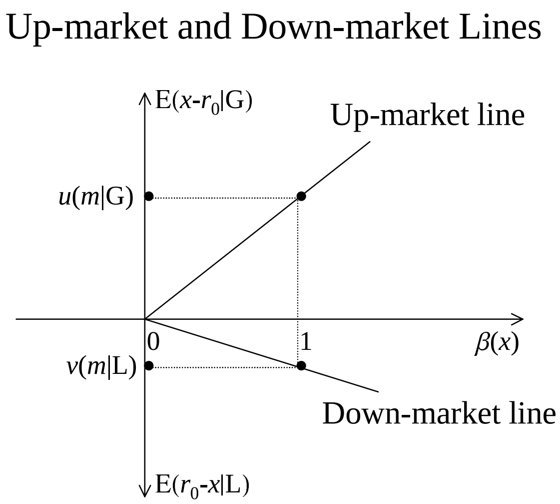

Figure 5: The dichotomous pricing model predicts that in equilibrium every security's up-market potential and down-market potential are linear functions of its beta. Thus the up-market (resp. down-market) potentials of all securities should fall on the up-market (resp. down-market) line. Note that these lines can be extended to the left to allow for negative betas, which are not depicted in the figure.

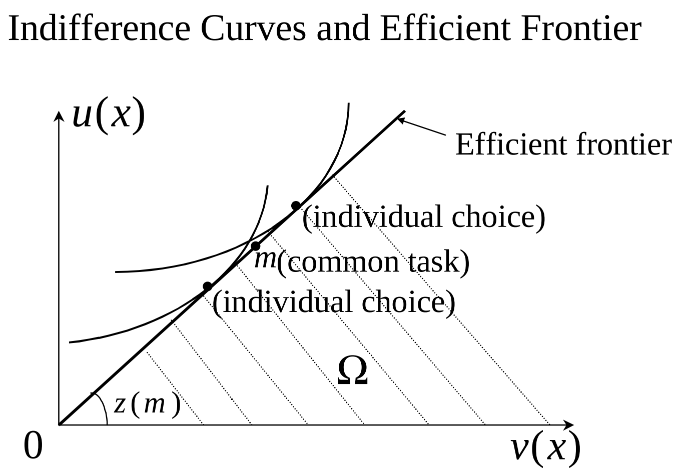

Figure 6: If all investors share common perception of inherent reward and inherent risk, then maximizing the inherent reward-to-risk ratio z is a common task. In equilibrium the market portfolio m offers the highest such ratio $z(m)$ and individuals choose optimal strategies along the efficient frontier.