COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station

New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 827-R

DISTRIBUTIONAL ANALYSIS OF PORTFOLIO CHOICE

Philip H. Dybvig

Revised January 1988

Distributional Analysis of Portfolio Choice

Philip H. Dybvig

Yale University

August, 1984

Current version January, 1988

Abstract

We compare trading in a market with receiving some particular consumption bundle, given increasing state-independent preferences and complete markets. The analysis focuses on the *distributional price* of the particular bundle. The distributional price is the price of the cheapest utility-equivalent bundle sold in the market. The distributional price is determined by the distribution functions of the outside bundle and the state price density. Simple portfolio performance measures illustrate the value of the approach. Unlike CAPM-based measures, these measures are valid even when superior information is the source of superior performance.

1. Introduction

The Capital Asset Pricing Model (CAPM) is an approach to investment analysis based on very simple assumptions. For example, mean variance analysis is implied by the following three assumptions:

1. Agents' preferences depend only on the mean and variance of consumption of a single good at a single date.

2. Agents prefer more to less, i.e., given a choice between two consumption bundles with equal variances, an agent will choose the bundle with the higher mean return.

3. The capital market equilibrium comes from our standard model of perfect markets (with no taxes, transaction costs, or information asymmetries) which may not be complete. Such a perfect market allows short sales without penalty.

From these assumptions, we can derive all the familiar features of the CAPM, including the mean variance frontier (where all portfolio choices lie), the mean variance efficiency of the market, and the security market line (which is a graphical representation of the linear relation between risk and expected return).

The CAPM remains a leading theory in finance, partly because of the simplicity of its assumptions and partly because of the power and elegance of its conclusions. Yet the CAPM cannot possibly be valid for all assets. As shown by Dybvig and Ingersoll [1982], there exist arbitrage profits if the CAPM prices all options, even if the CAPM correctly prices all primitive assets like stocks and bonds.[1] In this paper, the intent is to extend the spirit of the CAPM to complete markets, while maintaining simplicity. The result is a model with close ties to stochastic dominance. This new model, the Payoff Distribution Pricing Model (PDPM), is consistent with most existing models with complete markets.

Like the CAPM, the PDPM can also be derived from three primitive assumptions:

1. Agents' preferences depend only on the probability distribution of consumption of a single good.

2. Agents prefer more to less, i.e., given a choice between two ordered consumption bundles, an agent will always choose the bundle that is larger.

3. The market faced by an individual comes from our standard model of a perfect market (no taxes, transaction costs, or information asymmetries) that is complete over finitely many equally probable states or some atomless[2] continuum of states. Such a market allows short sales without penalty.

The first two assumptions are substantially weaker than the parallel assumptions in mean variance analysis, because they allow preferences to depend on all the moments of the distribution of consumption. These two assumptions are almost equivalent to giving each agent strictly increasing preferences in the class defined by Machina [1982].[3] Machina's class of preferences includes all von Neumann-Morgenstern utility functions. The most important restriction we place on preferences is the exclusion of state-dependent preferences. State dependence would arise implicitly if the portfolio under study does not comprise all of consumption (as if some wealth is not marketable), or would arise explicitly if market returns were correlated with health or other non-economic contributors to well-being. We are also ignoring concerns related to the existence of multiple goods or periods, although assuming separability of preferences across goods and time would allow us to apply the analysis of this paper to each good and time separately. For exposition, the text will emphasize the more familiar von Neumann-Morgenstern preferences, although all of the proofs will go through almost verbatim with increasing Machina preferences (see Appendix 1).

The third assumption is stronger than the corresponding assumption for mean variance analysis. The assumption of complete markets over a continuum of states is intended to be in the spirit of continuous-time option pricing models in which the state space is continuous and every pattern of claims across states is obtained through some trading strategy, even if it is not marketed directly. Since the term "complete markets" has meant different things to different authors, some clarification is needed. By completeness we mean that the agent can trade every claim whose value tomorrow depends on the value of any traded security tomorrow. It does not mean, for example, that the agent can trade every claim whose value tomorrow depends on other agent's private information known today. (It is consistent with Assumption 3 for other agents to have private information, since Assumption 3 says only that our agent's choice problem must be of the same form *as if* capital markets were perfect.) Also, completeness does not mean that the agent can trade every claim whose value tomorrow depends on information that will be publicly available tomorrow (but this would imply completeness). The part of the third assumption requiring equally probable states (or a continuum) is made primarily for convenience. If we assume concavity of preferences, this assumption is not needed, as discussed in Appendix 1.

Because the *probability distribution* of consumption is the focus of our analysis, the first task is to characterize how much it costs to obtain a given distribution of consumption the cheapest way possible: we refer to that cost as the *distributional price* of the distribution (to distinguish it from the ordinary *asset price*). To compute the distributional price, we use the simple observation that a minimum cost must order state-contingent payoffs inversely with the state price density (per unit probability). For agents with strictly increasing and concave von Neumann-Morgenstern preferences, this ordering can be inferred from declining marginal utility and the first order condition that marginal utility equals the state price density. The distributional price is expressed by a simple formula that depends on the payoff distribution without reference to the actual assignment of payoffs to particular states of nature. The formula can be interpreted as measure of efficiency in the spirit of the Sharpe measure. The new measure has special appeal because it does not suffer from the theoretical shortcomings of the Sharpe measure, which cannot be used reliably to evaluate market timers (see Dybvig and Ross [1985a]).

Just as there is an analog of the mean variance frontier and the Sharpe measure, there is an analog of Security Market Line (SML) analysis and deviations from the security market line. The SML analog is a sort of derivative of the distributional pricing measure as one adds the position being evaluated to a base portfolio, just as a well-behaved SML alpha can sometimes be interpreted as a gradient when one adds a position to the market portfolio (see Dybvig and Ross [1985b]).

Sample statistics for the Sharpe-like and Jensen-like measures have a flavor similar to modern "robust" data analysis, because the statistics are related to order statistics. (Formally, the estimators are L-statistics — see Shorack and Wellner [1986].) Although the measures put relatively little weight on outliers representing the best returns (as would robust estimators), they put relatively large weight on outliers representing the worst returns (in contrast to the robust estimators). This divergence from the philosophy of robust data analysis comes directly from the theory and is well founded. Robust data analysis is intended to be "exploratory," and to give a first indication of how to model a phenomenon. Here the use of a statistic is to drive some economic decision, e.g., whether to retain a portfolio manager. Large losses by an investment manager represent large damage to the client and should be weighted heavily.[4] This is also a good feature for pragmatic reasons outside the model, because a large loss may be symptomatic of other serious problems.

Perhaps more exciting than the potential empirical use of the concepts developed here is the use of the approach as a theoretical tool. In Dybvig [1988], these tools are used to measure the cost of following inefficient portfolio strategies, such as repeated portfolio insurance, that are stylized versions of strategies used by practitioners. In Dybvig and Spatt [1983], using the distributional approach led to results about an agency problem without having to rely on *ad hoc* restrictions on the sharing rules.

In Section 2, we consider some simple examples that illustrate the distributional approach, and we lay some foundations. Section 3 gives the basic results of the analysis, including the Sharpe-like measure of performance. Section 4 analyzes the distributional implications of sums and marginal changes, including our Jensen-like measure of performance. Section 5 closes the paper.

2. An Example and Some Preliminary Results

Before developing the general theory, an example will illustrate the basic ideas. Although the theory's assumptions make the most sense when there is a continuum of states, the economic ideas are easier to understand when there are finitely many equally probable states. Therefore, all examples and proofs in the text are for the discrete case, but the theorems are stated to be valid more generally. Appendix 1 summarizes the extension of the proofs to a continuum of states or to finitely many states with unequal probabilities. For the text, we assume that there are n states, each of which has probability $1/n$, and that all state prices are positive.

Our example has three equally probably states; letting π_i be the probability of state i, we have that $\pi_1 = \pi_2 = \pi_3 = 1/3$. Since markets are complete, the market must span elementary state securities for each state. By definition, an elementary state security for state i pays off one dollar if state i occurs and zero otherwise. Letting p_i be the i-th *state price*, i.e. the price of the i^{th} elementary state security, let $p_1 = 0.31$, $p_2 = 0.20$, and $p_3 = 0.39$. A riskless asset paying 1 in all states is obtained at a cost of

$$\begin{aligned} d &= 1\cdot 0.31 + 1\cdot 0.20 + 1\cdot 0.39 \\ &= 0.90. \end{aligned}$$

We call d the *discount factor*, which corresponds to the interest rate r satisfying $d = 1/(1+r)$. In the example, r is $11\ 1/9\%$.

Now we are set to consider how much it costs to purchase a given distribution of consumption. Since there are three equally probable states, the set of feasible consumption distributions is the set of *lotteries* in which each outcome has probability $1/3$. As we use the term, *lotteries* are defined without reference to which outcome comes in which state (just as in the axiomatic derivation of von Neumann–Morgenstern utility functions — see for example Luce and Raiffa [1957]). For example, the lottery giving equal chances at *10*, *20*, and *30* is the same as the lottery giving equal chances at *20*, *30*, and *10*. This lottery can be purchased in the *6* different ways we can assign the three lottery outcomes to the three states. Here are the six allocations and their costs.

Allocation	*Cost*
(10,20,30)	$0.31 \cdot 10 + 0.20 \cdot 20 + 0.39 \cdot 30 = 18.8$
(10,30,20)	$0.31 \cdot 10 + 0.20 \cdot 30 + 0.39 \cdot 20 = 16.9$
(20,10,30)	$0.31 \cdot 20 + 0.20 \cdot 10 + 0.39 \cdot 30 = 19.9$
(20,30,10)	$0.31 \cdot 10 + 0.20 \cdot 20 + 0.39 \cdot 30 = 16.1$
(30,10,20)	$0.31 \cdot 30 + 0.20 \cdot 10 + 0.39 \cdot 20 = 19.1$
(30,20,10)	$0.31 \cdot 10 + 0.20 \cdot 20 + 0.39 \cdot 30 = 17.2$

Naturally, any agent satisfying our first assumption of caring only about the distribution and our second assumption of "preferring more to less," will choose the fourth bundle. The money saved by choosing the fourth allocation could be used to buy more consumption. For example, instead of buying the bundle *(10,30,20)*, an agent could buy the bundle *(20,30,10) + (0,4,0) = (20,34,10)*. Since the agent views consumption in the three states symmetrically, this is clearly preferred to *(10,30,20)*.

In general, an example with n outcomes will have $n!$ different ways of assigning lottery outcomes to states; for example if $n = 10$ then there are *3,628,800* different assignments of outcomes to states. Of course, the listing technique for finding the optimum is impractical in general. Instead, we rely on the simple result that the cheapest way of buying a lottery orders consumption in reverse of the *state price density* $\rho_i \equiv p_i/\pi_i$. The state price density is the price per unit of probability of buying consumption in a particular state. The reason for using the state price density is that in a continuum of states, each state has probability zero and a state price of zero. The state price density is available in many asset pricing models. For example, in the lognormal Black-Scholes model with constant drift, the state price density is lognormal (see Dybvig [1988]).

Theorem 1: Suppose that markets are complete and that all states are equally probable. Then any cheapest way to achieve a lottery assigns the outcomes of the lottery to the states in reverse order of the state price density (with probability one). In the discrete case, this means that if lottery outcome c_i is chosen in state i and lottery outcome c_j is chosen in state j, then $\rho_i > \rho_j \Rightarrow c_i \le c_j$.

Proof: First note that there exists a cheapest way to achieve the lottery, since there are only finitely many ($n!$) ways to assign the lottery outcomes to the states. Now suppose that some cheapest way of assigning lottery outcomes to states is not in the reverse order as the state prices density (we will argue to a contradiction). In this case, there exist states i and j such that $\rho_i > \rho_j$ but $c_i > c_j$. Suppose we switch the lottery outcomes between states i and j. The change in cost is given by $(p_i c_j + p_j c_i) - (p_i c_i + p_j c_j) = (c_j - c_i)(p_i - p_j) = (c_i - c_j)(\rho_i - \rho_j)/n$, a negative number, implying a decrease in cost. This contradicts the supposition that the original assignment is cheapest. □

The cheapest assignment is unique if and only if the state prices and consumption amounts are distinct. If not, then reassignment of lottery outcomes among states having the same state price or the same amount of consumption does not affect the cost. Therefore, Theorem 1 gives a complete characterization of the assignments that yield the least cost.

It is useful to remember that cost-minimizing portfolio choice is an implication of utility-maximizing behavior. Consider the following choice problem.

Given $u(\cdot)$, w, n, p_1, p_2, $\ldots$, p_n, *and* $\pi_1 = \pi_2 = \ldots = \pi_n = \frac{1}{n}$,

choose c_1, c_2, $\ldots$, c_n *to*

maximize $E[u(\tilde{c})] = \sum_{i=1}^{n} \pi_i u(c_i)$

subject to $E[\tilde{\rho}\tilde{c}] = \sum_{i=1}^{n} \pi_i \rho_i c_i = \sum_{i=1}^{n} p_i c_i \leq w$.

The problem incorporates the von Neumann-Morgenstern assumption, the assumption of complete and frictionless markets without short sales restrictions, and the assumption of equally probable states. As always, we assume that the state prices are positive.

The following theorem shows that maximizing behavior is cost-minimizing, and that cost-minimization is the only restriction imposed by maximizing behavior.

Theorem 2: Assume that markets are complete (with positive state prices) and that all states are equally probable. Suppose that an assignment of consumption minimizes the cost of purchasing a lottery. Then there exists an increasing and concave von Neumann-Morgenstern utility function $u(\cdot)$ for which the assignment is the optimal choice. Conversely, if an assignment is the optimal choice for some increasing von Neumann-Morgenstern utility function $u(\cdot)$ (concave or not), then it minimizes the cost of purchasing a lottery.

Proof: We sketch the proof here, referring the reader to Dybvig and Ross [1982, Theorem 1 and Lemma 3] for omitted details.

When the c_i's are distinct, the p_i's form a non-increasing function of the c_i's (by Theorem 1). Choose any positive, non-increasing, and continuous function $g:\Re \to \Re$ that passes through the points $g(c_i) = \rho_i$, for $i = 1,\ldots,n$. Then let $u(\cdot)$ be any integral of $g(\cdot)$. Continuity of $g(\cdot)$ implies that an integral exists, and $g(\cdot)$ positive and non-increasing implies that $u(\cdot)$ is strictly increasing and concave. If the ρ_i's are distinct, $g(\cdot)$ can be chosen strictly decreasing, making $u(\cdot)$ strictly concave. This utility function satisfies the first order condition from the choice problem, because $u'(c_i) = \rho_i$. By concavity, the first order conditions imply that c_i solves the problem for this $u(.)$ when the budget constraint $w = E[\tilde{\rho}\tilde{c}]$ is satisfied.

If the c_i's are not distinct, $g(\cdot)$ will be chosen as a positive, non-increasing, and convex-valued correspondence whose integral $u(\cdot)$ is therefore strictly increasing and concave. The resulting $u(\cdot)$ will still satisfy first order conditions (in terms of subgradients), but will not generally be differentiable.

For the converse, a choice that does not minimize cost has lower expected utility than the cost-minimizing choice plus a riskless bonus equal to the difference of the costs divided by the sum of the state prices. □

3. The Payoff Distribution Pricing Model — Using Distribution Functions

In Section 2, we represented distributions without regard to state assignments as lotteries. From now on, we will represent lotteries using the associated (statistical) cumulative distribution function, which embodies all the information about a random variable that is invariant to changing state assignments.

Recall that the value of a distribution function, $F(x)$, gives the probability that the associated random variable is less than or equal to x. In our analysis, we will frequently refer to integrals involving the inverse of the distribution function. In examples with finitely many states, $F(\cdot)$ is a step function, and $F^{-1}(\cdot)$ is not defined in the ordinary way because of the discontinuity of $F(\cdot)$. Instead, for $y \in (0,1)$, we define[5]

$$F^{-1}(y) \equiv \min\{x \mid F(x) \geq y\}. \tag{1}$$

Whenever we use an inverse, we will intend this definition. (The values of $F^{-1}(1)$ and $F^{-1}(0)$ do not concern us, since we will always use $F^{-1}(\cdot)$ in integrals that ignore the value at finitely many points.) We can write the mean value of x as

$$\mu_x = \int_{\gamma=0}^{1} F^{-1}(\gamma)d\gamma. \tag{2}$$

As an aid to intuition, note that if a random variable γ is distributed uniformly on $(0,1)$, then $F^{-1}(\gamma)$ is a random variable with distribution function $F(\cdot)$.

We make a distinction between two different types of pricing operators. One prices assets and is the usual type of price or cost one would obtain from a pricing model like the CAPM, the Arbitrage Pricing Theory, or the Black-Scholes model. The new type of pricing operator prices payoff distributions. Here are formal definitions and notations for the two types of pricing operator. After the definition, Theorem 3 re-expresses Theorem 1 in terms of distribution functions.

Definition: The *asset price* of a marketed consumption stream $\tilde{c}$ is denoted by $P_A(\tilde{c};\rho) \equiv E[\tilde{\rho}\tilde{c}]$. The *distributional price* of a distribution function is the asset price of the least costly consumption stream with that distribution. The distributional price is denoted by $P_D(F_c;F_\rho) \equiv \min\{P_A(\tilde{c};\tilde{\rho})|\tilde{c} \sim F_c\}$.

Theorem 3: Suppose that markets are complete and that states are equally probable. Let $F_\rho(\cdot)$ be the distribution function of the state price density and let $F_c(\cdot)$ be the distribution function of some lottery. Then the distributional price of $F_c(\cdot)$ is given by

$$P_D(F_c;F_\rho) = \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)F_c^{-1}(1-\gamma)d\gamma. \tag{3}$$

Proof: Relabel the states so that $\rho_1 \leq \rho_2 \leq \ldots \leq \rho_n$. Then the distribution function of the state price density is given by

$$\begin{array}{lll}
F_\rho(q) = 0 & q < \rho_1 & \\
\vdots & & \\
F_\rho(q) = k/n & \rho_k \leq q < \rho_{k+1} & \text{for } k=1,\ldots,n-1 \\
\vdots & & \\
F_\rho(q) = 1 & \rho_n \leq q. &
\end{array}$$

The inverse distribution function is given by

$$F_\rho^{-1}(\gamma) = \rho_1 \qquad 0 < \gamma \le 1/n$$

$$\vdots$$

$$F_\rho^{-1}(\gamma) = \rho_k \qquad (k-1)/n < \gamma \le k/n \qquad \text{for } k=2,\ldots,n-1$$

$$\vdots$$

$$F_\rho^{-1}(\gamma) = \rho_n \qquad (n-1)/n < \gamma < 1.$$

Assigning the lottery outcomes c_i for minimum cost, in reverse order as the ρ_i's (as required by Theorem 1), $c_1 \ge c_2 \ge \ldots \ge c_n$ and consequently

$$F_c^{-1}(\gamma) = c_n \qquad 0 < \gamma \le 1/n$$

$$\vdots$$

$$F_c^{-1}(\gamma) = c_{n+1-k} \qquad (k-1)/n < \gamma \le k/n \qquad \text{for } k=2,\ldots,n-1$$

$$\vdots$$

$$F_c^{-1}(\gamma) = c_1 \qquad (n-1)/n < \gamma < 1.$$

(Note that if $\rho_i = \rho_j$ for some i and j, then the minimal cost consumption plan is not unique but can always be chosen to have $c_1 \ge c_2 \ge \ldots \ge c_n$.) Now we are ready to evaluate the integral in the statement of the theorem.

$$\begin{aligned}\int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)F_c^{-1}(1-\gamma)d\gamma &= \sum_{i=1}^{n} \int_{\gamma=(i-1)/n}^{i/n} F_\rho^{-1}(\gamma)F_c^{-1}(1-\gamma)d\gamma \\ &= \sum_{i=1}^{n} \int_{\gamma=(i-1)/n}^{i/n} \rho_i c_i \\ &= \sum_{i=1}^{n} \frac{1}{n} \rho_i c_i \\ &= E[\tilde{\rho}\tilde{c}].\end{aligned}$$

which is the asset price of the cost minimizing lottery. □

The distributional price expression (3) is the most important formula in this paper. The integral is a general formula for the expected value of the product of two *inversely ordered* random variables with distribution functions $F_\rho(\cdot)$ and $F_c(\cdot)$. Theorem 1 tells us that $P_D(F_c;F_\rho) = E[\tilde{\rho}\tilde{c}] \equiv P_A(\tilde{c};\tilde{\rho})$ if and only if $\tilde{\rho}$ and $\tilde{c}$ are inversely ordered. $P_D(F_c;F_\rho)$ is clearly symmetric in $F_c(\cdot)$ and $F_\rho(\cdot)$ (as can be proven directly by a change in variables in (3) from γ to $1-\gamma$), just as $P_A(\tilde{c};\tilde{\rho})$ is symmetric in $\tilde{c}$ and $\tilde{\rho}$. Furthermore, the integral is bi-linear in $F_c^{-1}(\cdot)$ and $F_\rho^{-1}(\cdot)$. Lemma 1 is a simple implication of this property.

Lemma 1: Let $c_2 = a + bc_1$ where $b \geq 0$. Then

$$P_D(F_{c_2};F_\rho) = \frac{a}{1+r} + bP_D(F_{c_1};F_\rho). \tag{4}$$

Furthermore, a similar result is true when we switch c and ρ. Let $\rho_2 = a + b\rho_1$ where $b \geq 0$. Then

$$P_D(F_c;F_{\rho_2}) = a\mu_c + bP_D(F_c;F_{\rho_1}). \tag{5}$$

Proof: Recall that

$$d = \frac{1}{1+r} = \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)d\gamma. \tag{6}$$

The results follow from (3) and the observation that $F_{c_2}^{-1}(\gamma) = a + bF_{c_1}^{-1}(\gamma)$ and $F_{\rho_2}^{-1}(\gamma) = a + bF_{\rho_1}^{-1}(\gamma)$. □

One implication of Lemma 1 is that we can write the minimum cost as

$$P_D(F_c;F_\rho) = \frac{\mu_c}{1+r} + P_D(F_c^*;F_\rho^*) \tag{7}$$

where F^* denotes the distribution function of a de-meaned variable. This decomposition tells us that the cost of obtaining the distribution $F_c(\cdot)$ is equal to the expected discounted value less a risk adjustment. (Since $F_c^{*-1}(1-\gamma)$ is decreasing in γ with an integral equal to zero, and $F_\rho^{*-1}(\gamma)$ is increasing and therefore puts most positive weight on negative values of $F_c^{*-1}(1-\gamma)$, it follows that the integral implicit in $P_D(F_c^*;F_\rho^*)$ is negative.) Just as $P_D(F_c;F_\rho)$ is interpreted as the product of two inversely ordered random variables, $P_D(F_c^*;F_\rho^*)$ is interpreted as the covariance of two inversely ordered random variables. Given $F_\rho^*(\cdot)$, the absolute value of $P_D(F_c^*;F_\rho^*)$ is a measure of dispersion.

Before we derive a theorem about performance measurement, it is worth discussing what it means to have superior performance. When we measure performance, we follow the tradition of comparing some particular investment strategy to the alternative of trading in a market. This comparison places substantial structure on the market (the CAPM or the PDPM), but leaves the investment strategy arbitrary. The investment strategy's payoff is just a random variable that may be defined on a state space that is larger (more refined) than the space over which the market is complete. For example, the investment strategy may include information-based trading, losses from transaction costs or fraud, or investments placed privately outside the market. This is a partial equilibrium approach in the sense that we do not analyze the equilibrium effects of information or imperfect capital markets.

We can think of this as a limiting case in which the equilibrium effects of market imperfections are small (as they are in many rational expectations models if there are few enough informed traders). This way of thinking about portfolio performance measurement is consistent with the traditional Sharpe and Jensen measures. Because we will be judging a strategy using a market (and in particular, the market's $F_\rho(\cdot)$) as a basis of comparison, we will refer to that market as the *benchmark market*.

In the CAPM, we think of a benchmark as a market portfolio and an associated riskless rate. These two summarize all that is important to an investor, since they span the efficient frontier. In the current model, the critical feature of a market is the state price density $F_\rho(\cdot)$.

Theorem 4: Let $F_c(\cdot)$ be the distribution function of consumption achieved by investing initial wealth w in the strategy to be evaluated, let $F_\rho(\cdot)$ be the distribution function of the state price density in the benchmark market, and let $v = P_D(F_c;F_\rho)$. Then,

(a) If $v > w$, there exists some increasing utility function $u(\cdot)$ for which receiving the distribution $F_c(\cdot)$ is preferred to trading in the benchmark market. (This is *superior performance* or *super-efficiency*.)

(b) If $v = w$, there exists some increasing utility function $u(\cdot)$ for which receiving the distribution $F_c(\cdot)$ is just as good as trading in the benchmark market. (This is *ordinary performance* or *efficiency*.)

(c) If $v < w$, then for *all* increasing $u(\cdot)$, receiving the distribution $F_c(\cdot)$ is dominated by trading in the benchmark market. (This is *inferior performance* or *inefficiency*.)

Proof: Case (b) follows directly from Theorems 2 and 3, which imply that $F_c(\cdot)$ is the distribution function of the optimal portfolio for some increasing utility function. In case (a), Lemma 1 plus Theorems 2 and 3 imply that $c-(v-w)/d$ is distributed as the optimal portfolio for some increasing utility function, and therefore c is preferred for that agent. In case (c), Lemma 1 plus Theorem 3 imply that $c+(w-v)/d$ is distributed as the optimal portfolio for some increasing utility function, and therefore c is dominated by that choice for all increasing utility functions. (Of course, most agents can do even better). □

Theorem 4 is closely related to stochastic dominance, because the distributional price indicates when some random variable ($\tilde{c}$) is stochastically dominated by trading in the benchmark market.

It is often convenient to work with the rate of return, $x \equiv (c-w)/w$. We now define the *dispersion* of the return.

Definition: Consider a portfolio with distribution of return $F_x(\cdot)$, being evaluated with respect to a benchmark state price density distribution function $F_\rho(\cdot)$. The *dispersion* of the return is defined by

$$\delta \equiv -d^{-1} P_D(F_x^*;F_\rho^*). \tag{8}$$

(Recall that the F^*'s are distribution functions of demeaned variables and that $d = 1/(1+r) = E[\rho]$ is the discount factor. Also, $\delta > 0$ if neither ρ nor x is constant, since we showed earlier that $P_D(F_c^*;F_\rho^*)$ is always negative.) Furthermore, from (2) the portfolio's expected return is

$$\mu_x = \int_{\gamma=0}^{1} F_x^{-1}(\gamma)d\gamma, \tag{9}$$

and its excess return is $\mu_x - r$.

Our analog of the Sharpe measure is given by the excess return $(\mu_x - r)$ of a portfolio divided by its dispersion. Theorem 5 shows that comparing this measure with the market price of dispersion classifies performance.

Theorem 5: Let $F_x(\cdot)$ be the distribution of a portfolio's return. Then,

(a) If $(\mu_x - r)/\delta > 1$, there exists some increasing utility function $u(\cdot)$ for which receiving the distribution $F_x(\cdot)$ of returns is preferred to trading in the benchmark market (*superior performance*).

(b) If $(\mu_x - r)/\delta = 1$, there exists some increasing utility function $u(\cdot)$ for which receiving the distribution $F_x(\cdot)$ of returns is just as good as trading in the benchmark market (*ordinary performance*).

(c) If $(\mu_x - r)/\delta < 1$, then for *all* increasing $u(\cdot)$, receiving the distribution $F_x(\cdot)$ of returns is dominated by trading in the benchmark market (*inferior performance*).

Proof: We want to show that conditions (a), (b), and (c) are the same as in Theorem 4. The condition from Theorem 4 that $v \gtrless w$ is equivalent to each of the following. By Theorem 3 and Lemma 1,

$$v = \mu_c d + \int_{\gamma=0}^{1} F_\rho^{*-1}(\gamma) F_c^{*-1}(1-\gamma)\, d\gamma \gtrless w.$$

Because $x \equiv (c-w)/w$, or equivalently $c = w(1+x)$, then $\mu_c = w(1+\mu_x)$ and $F_c^{*-1}(1-\gamma) = wF_x^{*-1}(1-\gamma)$. Therefore,

$$v = w(1+\mu_x)d + w\int_{\gamma=0}^{1} F_\rho^{*-1}(\gamma) F_x^{*-1}(1-\gamma)\, d\gamma \gtrless w.$$

Divide both sides by w, subtract 1, and multiply by $d^{-1} = (1+r)$ to obtain

$$(\mu_x - r) + d^{-1}\int_{\gamma=0}^{1} F_\rho^{*-1}(\gamma) F_x^{*-1}(1-\gamma)\, d\gamma \gtrless 0.$$

From the definitions of δ in (8) and the positivity of δ we have that

$$\frac{\mu_x - r}{\delta} \gtrless 1.$$

□

To this point, we have not put enough structure on the model to pin down δ, since we need to know $F_\rho(\cdot)$ to compute δ. The actual specification of $F_\rho(\cdot)$ can be motivated by almost any equilibrium model consistent with complete markets. In Dybvig [1988], it is shown how to derive $F_\rho(\cdot)$ in discrete and continuous time models in which all options are spanned by continous time trading strategies. In this paper, we give a single-period example related to the CAPM in which we can derive the form of $F_\rho(\cdot)$ (and therefore δ).

Here is the example. There is a riskless asset paying a constant return r and a risky asset ("the market") paying a return $\tilde{x}_M$ that is distributed normally with mean μ and variance σ^2. There are arbitrarily many other assets in the economy, including enough options to complete the market. To price all the assets, we assume that some agent with wealth 1 and constant absolute risk aversion (exponential utility) holds the market portfolio. By the standard result, the agent holding the market must have absolute risk aversion

$$A = \frac{\mu - r}{\sigma^2} ,$$

and therefore the utility function is $u(w) = -exp(-Aw)$, and the marginal utility is $u'(w) = Aexp(-Aw)$. The first order condition for optimally holding the market is the existence of a Lagrange multiplier λ (a constant) such that $\rho = \lambda u'(x_M) = \lambda Aexp(-Ax_M)$. Because x_M is normally distributed, this first order condition implies that ρ is constant, too. To compute the exact distribution, we need to know λ. But we know that ρ must price a riskless bond with a face of 1 correctly; i.e., we must have that

$$\begin{aligned} \frac{1}{1+r} &= E[\rho] \\ &= \lambda A e^{-A\mu + A^2\sigma^2/2} \end{aligned}$$

or

$$\log(\lambda A) = A\mu - A^2\sigma^2/2 - \log(1+r). \tag{11}$$

Therefore, we have that

$$\rho = \frac{1}{1+r} e^{-A(x_M-\mu)-A^2\sigma^2/2}, \tag{12}$$

and therefore $log(\tilde{\rho})$ is distributed normally with mean $-log(1+r)-A^2\sigma^2/2$ and variance $A^2\sigma^2$. Letting $\Phi(\cdot)$ be the unit cumulative normal distribution function, we have that

$$F_{\rho}(\rho) = \Phi\left[\frac{log(\rho)+log(1+r)+A^2\sigma^2/2}{A\sigma}\right] \tag{13}$$

and

$$F_{\rho}^{-1}(\gamma) = \frac{1}{1+r}\, exp\left[-A^2\sigma^2/2 + A\sigma\ \Phi^{-1}(\gamma)\right]. \tag{14}$$

This formula can be substituted into (3) to compute the payoff distribution price of a consumption distribution, or into (8) to compute the dispersion of a return distribution. (It is interesting to note that the measure of dispersion does not depend on the interest rate r except through the absolute risk aversion coefficient, since d^{-1} in (8) cancels $1/(1+r)$ in (14).) To obtain a sample estimate of the dispersion given i.i.d. observations on the return, we use the sample distribution function of returns.[6] This estimator is well-defined and consistent provided returns are bounded below (limited liability) and the mean return exists, even if the variance of returns is infinite.[7] Of course, this measure will suffer from many of the same *measurement* problems as other measures, for example, if the sample is too small or if returns are not i.i.d.

Once we have selected the distribution of the state price density (such as in (13) above), we can use the measure of dispersion in much the same way as we use the standard deviation. For example, we can plot sample dispersions against sample means, or we can compute the ratio of excess return to dispersion directly. According to Theorem 5, this analog of the Sharpe measure will correctly identify superior, ordinary, and inferior performance (provided we have chosen $F_{\rho}^{-1}(\cdot)$ correctly).

Unlike the Sharpe measure, the distributional measure gives the correct ordering (subject to measurement error) even in cases where superior performance is based on superior information. The failure of the Sharpe measure in the presence of information is documented by Dybvig and Ross [1985b]. The difficulty with the Sharpe measure is that information makes returns non-normal. For example, if an investor's signal and the risky asset return are joint normal, the investor's return will be the product of the normal asset return and the portfolio choice. The portfolio choice is some function of the signal, and except in degenerate cases this product will not be normally distributed. If we motivated mean variance analysis using quadratic utility, we would not have this problem, but assuming quadratic utility sacrifices monotonicity of preferences and implies absolute risk aversion is increasing. None of these problems arise in the PDPM, however, because von Neumann-Morgenstern preferences are valid for all random variables (and not, for example, not just normal random variables), including those arising from information-based trading.

4. Sums and Marginal Changes

The analysis in Section 3, like the Sharpe measure, is designed to evaluate the efficiency of the entire portfolio held by the agent. In this section, we consider measures designed to evaluate net investments. These measures consider whether investing marginally in the portfolio being evaluated would increase the distributional price of a given base portfolio. One way of evaluating marginal changes directly using the results of Section 3 is to look at our performance measure applied to the base portfolio before and after partial movement towards the portfolio being evaluated. Another approach, in the spirit of the Jensen measure, looks at the derivative of the minimum cost in the direction of a marginal movement into the new portfolio.

Generally, the choice of base portfolio matters, just as it matters which market proxy or index we use in mean-variance analysis. (We distinguish between the choice of base portfolio and the choice of benchmark market, because the two enter the analysis separately.) If we are looking at one of several managers of a pension fund, one natural base portfolio may be the portfolio representing the investments of the other managers — in other words we want to know whether the fund is better off than it would be without this manager. In other contexts, the natural choice of base portfolio is any efficient portfolio. Fortunately, as in mean-variance analysis, marginal performance does not depend on which *strictly efficient* base portfolio we choose, where strictly efficient means that $\rho_i > \rho_j \Rightarrow c_i < c_j$ rather than $\rho_i > \rho_j \Rightarrow c_i \le c_j$.[8] Marginal performance does depend on the choice of base portfolio if an *inefficient* base portfolio is used.

Now we develop our analog of the Jensen measure.

Theorem 6: The marginal change in the minimum cost, starting at the payoff $\tilde{c}$ of some (possibly inefficient) base consumption stream and moving in the direction of some net payoff $\tilde{\Delta}$ is given by

$$\frac{\partial}{\partial\xi}P(F_{c+\xi\Delta};F_\rho)\bigg|_{\xi=0^+} = \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)E[\tilde{\Delta}|c=F_c^{-1}(1-\gamma)]d\gamma \tag{15}$$

if $\tilde{c}$ takes on different values in each state, and

$$= \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)\ \lim_{\alpha\downarrow 0}\left(E[\tilde{\Delta}|c+\alpha\Delta=F_{c+\alpha\Delta}^{-1}(1-\gamma)]\right)d\gamma \tag{16}$$

in general.

If the base consumption stream $\tilde{c}$ is strictly efficient, then the expression (15) or (16) for the marginal change is equal to $E[\tilde{\Delta}\tilde{\rho}]$, which is $P_A(\tilde{\Delta};\tilde{\rho})$ if $\tilde{\Delta}$ is marketed.

Proof: For sufficiently small positive ξ and α the orderings across states of $\tilde{c}+\xi\tilde{\Delta}$ and $\tilde{c}+\alpha\tilde{\Delta}$ are the same as the ordering across states of $\tilde{c}$ (because there are finitely many states), excepting perhaps ties in $\tilde{c}$. Consequently, we have that for all ξ and α both positive and sufficiently small,

$$P(F_{c+\xi\Delta};F_\rho) = \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)\left(F_c^{-1}(1-\gamma)+\xi E\left[\tilde{\Delta}|c+\alpha\Delta=F_{c+\alpha\Delta}^{-1}(1-\gamma)\right]\right)d\gamma.$$

The result follows by differentiating with respect to ξ under the integral sign. The "limit" part of the expression in the statement of the theorem is trivial, since we have equality for *all* sufficiently small positive α.

For $\tilde{c}$ strictly efficient, ρ is a function of c and is given by $F_{\rho}^{-1}(1-F_{c}(c))$. In this case, for α sufficiently small, the right hand side of (16) is simply

$$E[\tilde{\rho}\lim_{\alpha\downarrow 0}(E[\tilde{\Delta}|c+\alpha\Delta])] = E[\tilde{\rho}\tilde{\Delta}]. \qquad \square$$

The following result characterizes when some finite movement from $\tilde{c}$ in the direction of $\tilde{\Delta}$ will increase the distributional price.

Theorem 7: The bundle $\tilde{c}+\xi\tilde{\Delta}$ represents better performance than $\tilde{c}$ (as measured by minimum cost) for some $\xi > 0$ if and only if the expression in (16) is positive.

Proof: For each pairing of the states underlying $\tilde{c}+\xi\tilde{\Delta}$ with values of ρ, the cost is affine (linear plus a constant) in ξ. The minimum cost is the minimum of affine functions and therefore concave. Concavity and Theorem 6 together imply our result. By Theorem 6, the right derivative of the distributional price of $c+\xi\Delta$ with respect to ξ at $\xi=0$ is simply the expression in (16). When this is negative, concavity of the distributional price in ξ implies the result. If the right derivative in (16) is positive, then increasing ξ increases the distributional price for sufficiently small $\xi > 0$. $\square$

Note that Theorems 6 and 7 do not assume that $\tilde{\Delta}$ is marketed. For example, $\tilde{\Delta}$ could be the result of investments by a manager who has superior information that is used effectively but who also incurs high transaction costs from churning the portfolio, netting out to an inefficient return. If

$\tilde{\Delta}$ is marketed and $\tilde{c}$ is efficient, then the measure is equivalent to valuation using state prices.

The next result puts the results of Theorems 6 and 7 in terms of returns, mimicking the form of the Jensen measure.

Theorem 8: Suppose that the distribution of returns given by $\tilde{x}$ exhibits ordinary performance in the sense of Theorem 5. Then the portfolio $\tilde{x}+\alpha(\tilde{y}-\tilde{x})$ exhibits superior performance for some $\alpha > 0$ if and only if

$$\mu_y - r - \beta_{y;x} (\mu_x - r) > 0, \tag{17}$$

where

$$\beta_{y;x} = \frac{\lim_{\alpha \downarrow 0} \int_{\gamma=0}^{1} F_\rho^{-1}(\gamma) E[\tilde{y}-\mu_y | x+\alpha(y-x) = F_{x+\alpha(y-x)}^{-1}(1-\gamma)]\, d\gamma}{P_D(F_x^*;F_\rho^*)}. \tag{18}$$

If $\tilde{x}$ is different in every state, then $\beta_{y;x}$ is written more simply as

$$\beta_{y;x} = \frac{\int_{\gamma=0}^{1} F_\rho^{-1}(\gamma) E[\tilde{y}-\mu_y | x = F_x^{-1}(1-\gamma)]\, d\gamma}{P_D(F_x^*;F_\rho^*)}. \tag{19}$$

Proof: Let $\tilde{\Delta} = w(\tilde{y}-\tilde{x})$. By Theorems 6 and 7, it suffices to show that this definition makes (17) equivalent to positivity of (15). The proof assumes that $\tilde{c}$ and $\tilde{x}$ are different in every state; the proof for the other formula is essentially the same line by line. By choice of $\tilde{\Delta}$ and because $\tilde{c} = w(1+\tilde{x})$, substitution and division by w yields that positivity of (15) is equivalent to

$$\int_{\gamma=0}^{1} F_{\rho}^{-1}(\gamma)E[\tilde{y}-\tilde{x}|x=F_{x}^{-1}(1-\gamma)]d\gamma > 0.$$

Write $\tilde{y}-\tilde{x}$ in the expectation into three terms, μ_y, $-\tilde{x}$, and $\tilde{y}-\mu_y$. By additivity of the expectation and integral operators, we have

$$\int_{\gamma=0}^{1} F_{\rho}^{-1}(\gamma)\mu_y d\gamma - \int_{\gamma=0}^{1} F_{\rho}^{-1}(\gamma)F_{x}^{-1}(1-\gamma)d\gamma + \int_{\gamma=0}^{1} F_{\rho}^{-1}(\gamma)E[\tilde{y}-\mu_y|x=F_{x}^{-1}(1-\gamma)]d\gamma > 0.$$

The first integral is $\mu_y/(1+r)$, because we can take μ_y outside the integral. The second integral is $P_D(F_x;F_\rho)$. Because x is efficient, $P_D(F_{1+x};F_\rho) = 1$, and Lemma 1 implies that $P_D(F_x;F_\rho) = -1/(1+r) + 1 = r/(1+r)$. The third integral is the numerator of $\beta_{y;x}$ in (18). By (8), Theorem 5, and the efficiency of $\tilde{x}$, the denominator of $\beta_{y;x}$ is $P_D(F_x^*;F_\rho^*) = -\delta/(1+r) = -(\mu_x-r)/(1+r)$. Therefore, the third integral is $-\beta_{y;x}(\mu_x-r)/(1+r)$. Replacing each integral by its value and multiplying by $1+r$ yields (17). □

We complete this section by asking what changes in $F_c(\cdot)$ decrease $P_A(F_c;F_\rho)$ for *all* $F_\rho(\cdot)$. The payoff distributions with lower cost than $F_c(\cdot)$ for *all* $F_\rho(\cdot)$ are those distributions that are second order stochastically dominated by $F_c(\cdot)$. (Appendix 2 provides the necessary background on stochastic dominance.) Symmetrically, the state price density distributions that price all claims lower than $F_\rho(\cdot)$ for all $F_c(\cdot)$ are those state price densities stochastically dominated by $F_\rho(\cdot)$.

Theorem 9: Consider two distributions of consumption $F_{c_1}(\cdot)$ and $F_{c_2}(\cdot)$. The distributional price of $F_{c_1}(\cdot)$ is always less than or equal to that of

$F_{c_2}(\cdot)$ for all $F_\rho(\cdot)$ if and only if $F_{c_2}(\cdot)$ second-order stochastically dominates $F_{c_1}(\cdot)$. That is,

$$P_D(F_{c_1};F_\rho) \le P_D(F_{c_2};F_\rho) \tag{20}$$

for all $F_\rho(\cdot)$, if and only if

$$\int_{\gamma=0}^{\pi} F_{c_1}^{-1}(\gamma)d\gamma \le \int_{\gamma=0}^{\pi} F_{c_2}^{-1}(\gamma)d\gamma \tag{21}$$

for all $\pi\in[0,1]$.

Similarly,

$$P_D(F_c;F_{\rho_1}) \le P_D(F_c;F_{\rho_2}) \tag{22}$$

for all $F_c(\cdot)$ if and only if

$$\int_{\gamma=0}^{\pi} F_{\rho_1}^{-1}(\gamma)d\gamma \le \int_{\gamma=0}^{\pi} F_{\rho_2}^{-1}(\gamma)d\gamma \tag{23}$$

for all $\pi\in[0,1]$.

Proof: Suppose that (20) holds for all $F_\rho(\cdot)$. Then, in particular, (20) holds for $\tilde{\rho}$ a random variable taking on 1 with probability π and ε with probability $1-\pi$, where $0 < \varepsilon < 1$. Taking the limit as $\varepsilon \downarrow 0$ and reversing the integral by changing variables from γ to $1-\gamma$ yields (21).

Suppose conversely that (21) holds for all π. Since $F_\rho^{-1}(\cdot)$ is a non-increasing step function with finitely many steps, it can be written in the form

$$F_{\rho}^{-1}(\gamma) = \sum_{i=1}^{n} \psi_i I(\gamma,\pi_i), \tag{24}$$

where the ψ_i's are positive, the π_i's lie between 0 and 1, and

$$\begin{aligned} I(\gamma,\pi) &= 1 \qquad && \text{for } \gamma \geq \pi \\ &= 0 && \text{otherwise.} \end{aligned}$$

Change the order of integration by changing variables in (20) to $1-\gamma$ and substitute in (24). This gives us the difference between the two sides of (20) as

$$\begin{aligned} \int_{\gamma=0}^{1} F_{\rho}^{-1}(1-\gamma)[F_{c_1}^{-1}(\gamma)-F_{c_2}^{-1}(\gamma)]d\gamma &= \int_{\gamma=0}^{1} \sum_{i=1}^{n} \psi_i I(1-\gamma,\pi_i)[F_{c_1}^{-1}(\gamma)-F_{c_2}^{-1}(\gamma)]d\gamma \\ &= \sum_{i=1}^{n} \psi_i \int_{\gamma=0}^{\pi_i} [F_{c_1}^{-1}(\gamma)-F_{c_2}^{-1}(\gamma)]d\gamma. \end{aligned}$$

Since the ψ_i's are positive, this last term is negative by (21).

The result for the state price distributions is true by the same argument, once we note that adding a constant to $\tilde{c}$ (to make it positive) does not affect the ordering. □

The second part of Theorem 9 (comparing $F_{\rho_1}(\gamma)$ to $F_{\rho_2}(\gamma)$) has an interesting interpretation in terms of a more highly informed agent. One possible effect of better information is to split states that were previously indistinguishable. For example, take two states with the same state prices and the same probabilities from the uninformed perspective. For an informed agent, the state prices are the same, but the probabilities are different, leaving the sum of the two probabilities the same.[9] Therefore, the state price density has effectively had mean zero noise added to it. In this example, therefore, the informed state price density is "second-order stochastically dominated" by the uniniformed density, and therefore is distributed as the uninformed density plus noise.

In fact, this property holds whenever we look at an uninformed agent and an informed agent. Given a single-agent decision problem, improved information would make any agent better off, and therefore the resultant state price density must be preferred. By Theorem 9, a preferred state price density is "second-order stochastically dominated" and therefore is distributed as the original state price density plus noise less a nonnegative variable. However, the mean (the price of the riskless asset) must be the same with and without information. Therefore the new state price density is just the original one plus added noise. In other words, in terms of $F_{\rho}^{-1}(\gamma)$, the only effect of receiving superior information on the state price density distribution is equivalent to splitting states in the sense described above.

5. Conclusion

Starting from simple assumptions, we have developed a number of tools for analyzing portfolio performance and efficiency. Collectively, we can refer to the approach described here as the *distributional approach* to analyzing portfolio problems, or as the *Payoff Distribution Pricing Model* (PDPM). The research has two broad goals. First, the PDPM provides a theoretical toolbox that may stimulate new avenues of theoretical development. Second, the PDPM provides specific techniques for measuring investment performance and testing for efficiency.

While distributional analysis has already been applied successfully in theoretical work (Dybvig and Spatt [1983] and Dybvig [1988]), much work remains on the empirical side. We can apply some the known properties of L-statistics directly to the estimators (see Shorack and Wellner [1986]), but we still require econometric analysis of the estimators in the presence of measurement errors in both the distribution of state prices and the distribution of returns. Empirical work is needed to apply the measures and to compare them to more traditional measures. Only then will it be possible to assess the empirical potential of the distributional approach.

Appendix 1 Generalizing the Results

In the text, the proofs have assumed finitely many equally probable states and monotone von Neumann-Morgenstern preferences. In this appendix, we discuss how the proofs are modified to be valid more generally. We consider three types of generalizations: (1) more general preferences, (2) unequal state probabilities, and (2) atomless continuous state spaces.

Theorem 2 is the only place where preferences appear explicitly in the proofs; other theorems that make statements about preferences build on Theorem 2. In Theorem 2, we need only a few restrictions on the class of preferences: all preferences in the class must depend only on the distribution, all preferences in the class must strictly prefer more to less, and the class must include all strictly monotone and concave von Neumann-Morgenstern utility functions. The class could be the class of strictly monotone and concave von Neumann-Morgenstern preferences, the class of strictly monotone Machina [1982] preferences, or any class nested between these two classes. (The class of strictly monotone von Neumann-Morgenstern preferences is an example of a class nested between the two.) Assuming differentiability of von Neumann-Morgenstern preferences is *not* consistent with the analysis here, because the condition for optimality would be $\rho_i > \rho_j \Rightarrow c_i < c_j$ instead of $\rho_i > \rho_j \Rightarrow c_i \le c_j$. For a discussion of the effect of differentiability and strictness of concavity on the set of efficient portfolios, see Dybvig and Ross [1982], especially Table 1 and the related discussion.

As noted in the introduction, the analysis still works when probabilities are unequal if we assume that agents are risk averse. The basic change is that we look not for the cheapest lottery but instead for the cheapest lottery that is at least as good for all increasing and concave $u(\cdot)$. The proof of Theorem 1 in the text obtains a smaller cost whenever $\rho_i > \rho_j$ and $c_i > c_j$ by swapping the consumption values in the two states. For increasing concave preferences and unequal probabilities, we swap consumption and then adjust consumption in the less probable state to make the means the same. The original distribution of consumption is distributed as the new (swapped) consumption plus noise. For example, if $\pi_i > \pi_j$, $\rho_i > \rho_j$, and $c_i > c_j$, then we change state j consumption to c_i and we change state i consumption to $(\pi_i c_i + \pi_j c_j - \pi_j c_i)/\pi_i$. This change makes any risk-averse agent better off, while reducing the cost of the consumption distribution by $\pi_j(\rho_i - \rho_j)(c_i - c_j)$. A similar swap works in the other case ($\pi_i < \pi_j$, $\rho_i > \rho_j$, and $c_i > c_j$).

The remainder of this appendix gives a sketch of how the proofs change when we have a continuum of states. While many details have been skipped, an effort has been made to cover the most important and difficult points. Here are two assumptions we will use in the general case.

Assumption A1 All utility functions are defined on a convex unbounded interval of non-negative consumption levels, and all consumption choices are non-negative random variables. In distributional terms, $F_c(0) = 0$, or equivalently $F_c^{-1}(\gamma) \geq 0$ for all $\gamma \in (0,1)$.

Assumption A2 For all state price densities $\tilde{\rho}$ and consumption patterns $\tilde{c}$ under consideration, $E(\tilde{\rho}) < \infty$ and $E(\tilde{c}) < \infty$. In distributional terms, $\int_{\gamma=0}^{1} F_\rho^{-1}(\gamma)d\gamma < \infty$ and $\int_{\gamma=0}^{1} F_c^{-1}(\gamma)d\gamma < \infty$. For the results concerning stochastic dominance, we will also want $E(\tilde{\rho}^2) < \infty$ and $E(\tilde{c}^2) < \infty$, or in distributional terms $\int_{\gamma=0}^{1} F_{\rho}^{-1}(\gamma)^2 d\gamma < \infty$ and $\int_{\gamma=0}^{1} F_{c}^{-1}(\gamma)^2 d\gamma < \infty$.

The proof of Theorem 1 in the text says that if consumption and the state price density are in the same order in two states, we can obtain the same distribution for lower cost by swapping consumption in the two states, since all states are equally probable with a positive probability. With a continuum of states, if $\tilde{\rho}$ and $\tilde{c}$ are not in reverse order with probability 1, then there exist two non-null disjoint sets of states Ω_1 and Ω_2 such that $\tilde{\rho}$ and $\tilde{c}$ are both uniformly larger in Ω_1 than in Ω_2. While Ω_1 and Ω_2 may not be equally probable, having a non-atomic state space implies that there exist equally probable non-null subsets of Ω_1 and Ω_2. By a standard measure-theoretic argument, we can switch consumption between the two subsets in a way that doesn't affect the overall distribution of consumption. As in the discrete proof, this decreases cost.

Theorem 2 is essentially the same as in the discrete case, both in the statement and in the proof. Two new wrinkles are related to closure problems at the limit when the number of states becomes infinite. One problem is that the strictly increasing utility function constructed may not be defined outside the open interval going from the bottom to the top of the support of $\tilde{c}$ — implicitly, the marginal utility may be infinite or zero outside that interval. (The utility function is defined at any endpoint that is a mass point of $\tilde{c}$, however.) The other wrinkle is that expected utility for the utility function defined this way may be infinite. The simplest solution to this problem is to redefine optimality using a shortfall criterion, as Ramsey [1928] did in growth theory. Define an optimal choice $\tilde{c}^*$ to be a choice of $\tilde{c}$ that maximizes $E[u(\tilde{c})-u(\tilde{c}^*)]$ subject to the budget constraint for $\tilde{c}$. When $E[u(\tilde{c})]$ is bounded, the new objective function is simply $E[u(\tilde{c})] - E[u(\tilde{c}^*)]$, and maximizing the new objective with respect to $\tilde{c}$ yields the same optimum as maximizing $E[u(\tilde{c})]$ with respect to $\tilde{c}$. The advantage to this new objective function is that it can be defined even when $E[u(\tilde{c})]$ and $E[u(\tilde{c}^*)]$ are both infinite. Using this more general definition of optimality, the proof of Theorem 2 goes through as is.

Theorem 3 can be proven using a sequence of finite approximations to the two distribution functions and a proof for each finite approximation that is essentially the same as the finite proof. Assumptions A1 and A2 together imply that the expression for the minimum cost given in (3) is finite. To prove this, split the integral into two integrals, one from 0 to 1/2 and one from 1/2 to 1. By dominance, the minimum cost is less than

$$\frac{1}{2}\left[F_{\rho}^{-1}(1/2)E(\tilde{c}) + F_{c}^{-1}(1/2)E(\tilde{\rho})\right] < \infty$$

For Lemma 1, finiteness of $1/(1+r)$ in (6) is just the condition $E[\tilde{\rho}] < \infty$ in A2, and it is easy to see that the proof goes through unchanged.

It is relatively straightforward to generalize the remaining proofs, most of which are direct applications of the Theorems we have already discussed.

Appendix 2 Stochastic Dominance

This appendix summarizes some familiar stochastic dominance results. The results are displayed in a tabular form which is essentially Table 1 of Ross [1971] extended to include conditions expressed in terms of F^{-1}. Most if not all of the results have appeared before in published literature. For example, see Quirk and Saposnik [1962], Hadar and Russell [1969], and Levy and Kroll [1978] for the results for the monotone (M) and monotone concave (MC) classes of utility functions. We have taken the results for the concave class (C) from Ross [1971].

We will take $F(\cdot)$ and $G(\cdot)$ to be the distribution functions of the random variables $\tilde{x}$ and $\tilde{y}$, respectively, and $\tilde{x}$ and $\tilde{y}$ are both assumed to have finite mean and variance. M is the set of nonincreasing real-valued functions on $\Re$, C is the set of concave real-valued functions on $\Re$, and MC is the intersection of these two sets. The notation $\pi(\cdot)$ is used to indicate the probability of an event. The notation $\overset{d}{=}$ indicates "is distributed as."

All conditions in a given row of Table 1 are equivalent. The rows of Table 1 correspond to dominance for all monotone utility functions (first order stochastic dominance), dominance for all concave utility functions (concave dominance), and dominance for all monotone and concave utility functions (second order stochastic dominance).

Table 1: Summary of Stochastic Dominance Results

Criterion	*Integral (F)*	*Integral* (F^{-1})	*Dist*
$\forall U(\cdot)\in M$ $E[U(\tilde{x})] \geq E[U(\tilde{y})]$	$(\forall c)$ $F(c) \leq G(c)$	$(\forall \gamma)$ $F^{-1}(\gamma) \geq G^{-1}(\gamma)$	$\tilde{y} \overset{d}{=} \tilde{x}+\tilde{z}$ $\pi(\tilde{z}\leq 0) = 1$
$\forall U(\cdot)\in C$ $E[U(\tilde{x})] \geq E[U(\tilde{y})]$	$(\forall c)$ $\int_{-\infty}^{c}[F(\tau)-G(\tau)]d\tau \leq 0$ and $\rightarrow 0$ as $c \uparrow \infty$	$(\forall \gamma)$ $\int_{0}^{\gamma}[F^{-1}(\tau)-G^{-1}(\tau)]d\tau \geq 0$ and $\rightarrow 0$ as $\gamma \uparrow 1$	$\tilde{y} \overset{d}{=} \tilde{x}+\tilde{\varepsilon}$ $E[\tilde{\varepsilon}\|x] = 0$
$\forall U(\cdot)\in MC$ $E[U(\tilde{x})] \geq E[U(\tilde{y})]$	$(\forall c)$ $\int_{-\infty}^{c}[F(\tau)-G(\tau)]d\tau \leq 0$	$(\forall \gamma)$ $\int_{0}^{\gamma}[F^{-1}(\tau)-G^{-1}(\tau)]d\tau \geq 0$	$y \overset{d}{=} \tilde{x}+\tilde{z}+\tilde{\varepsilon}$ $\pi(\tilde{z}\leq 0) = 1$ $E[\tilde{\varepsilon}\|x+z] = 0$

Bibliography

Banz, R., and Miller, M. 1978. Prices for State-Contingent Claims: Some Estimates and Applications. *Journal of Business* 51: 653–672.

Brennan, M. and Solanki, R. 1981. Optimal Portfolio Insurance. *Journal of Financial and Quantitative Analysis* 16: 279–300.

Cox, J. C., and Huang, C. 1985. A Variational Problem Arising in Financial Economics with an Application to a Portfolio Turnpike Theorem. MIT Sloan School working paper.

Cox, J. C., and Leland, H. 1982. On Dynamic Investment Strategies. *Proceedings Seminar on the Analysis of Security Prices* vol. 26 no. 2 Center for Research in Security Prices.

Dybvig, P. H. 1982. Some New Tools for Testing Market Efficiency and Measuring Mutual Fund Performance. Yale School of Management working paper.

Dybvig, P. H. 1988. Inefficient Dynamic Portfolio Strategies, or, How to Throw Away a Million Dollars in the Stock Market. *Review of Financial Studies* 1.

Dybvig, P. H., and Ross, S. A. 1982. Portfolio Efficient Sets. *Econometrica* 50: 1525–1546.

Dybvig, P. H., and Ross, S. A. 1985a. Performance Measurement Using Excess Returns Relative to a Security Market Line: Differential Information. *Journal of Finance* 40: 384–399.

Dybvig, P. H., and Ross, S. A. 1985b. The Analytics of Performance Measurement Using a Security Market Line. *Journal of Finance* 40: 401–416.

Dybvig, P. H., and Spatt, C. S. 1983. Agency and the Market for Portfolio Managers: the Principle of Preference Similarity. Yale School of Management working paper.

Luce, R. D., and Raiffa, H. 1957. *Games and Decisions*. New York: John Wiley & Sons.

Hadar, J., and Russell, W. R. 1969. Rules for Ordering Uncertain Prospects. *American Economic Review* 59: 25–34.

Levy, H., and Kroll, Y. 1978. Ordering Uncertain Options with Borrowing and Lending. *Journal of Finance* 33: 553–574.

Machina, M. 1982. 'Expected Utility' Analysis Without the Independence Axiom. *Econometrica* 50: 277–323.

Quirk, J., and Saposnik, R. 1962. Admissibility and Measurable Utility Functions. *Review of Economic Studies* 29: 140-146.

Ramsey, F. P. 1928. A Mathematical Model of Saving. *Economic Journal* 38: 543–559. Reprinted in J. Stiglitz and H. Uzawa (ed.), 1969, *Readings in the Modern Theory of Economic Growth*. Cambridge, Mass.: M.I.T. Press.

Ross, S. A. 1976. Options and Efficiency. *Quarterly Journal of Economics* 90: 75–89.

Rubinstein, M. 1976. The Valuation of Uncertain Income Streams and the Pricing of Options. *Bell Journal of Economics* 7: 407–425.

Shorack, G. R. and Wellner, J. A. 1986. *Empirical Processes with Applications to Statistics*. New York: John Wiley & Sons.

Footnotes

†I would like to acknowledge helpful comments by seminar participants at Berkeley, Cornell, MIT, Stanford, and UBC, and particularly valuable advice from Jon Ingersoll, David Pollard, Steve Ross, and an anonymous referee. I am very grateful for support received from the Batterymarch Fellowship program and from the Sloan Research Fellowship program. This paper extends and clarifies ideas originally introduced in a previous working paper, Dybvig [1982].

1. The arbitrage opportunity is the purchase of a call option on the market portfolio with a sufficiently large exercise price. Provided the market might go up enough to put the call option in the money, the call will have a *negative* price.

2. An atom is an indivisible state with positive probability. Whenever we refer to a continuous state space we will implicitly mean that there are no atoms.

3. Machina's [1982] preferences satisfy all the von Neumann-Morgenstern axioms except the independence axiom. The resulting preferences depend only on the distribution function of outcomes (as we require). The slight difference is that Machina also imposes Frechet differentiability.

4. This is obvious when von Neumann-Morgenstern preferences are concave. When there is a continuum of states, it can also be shown that a monotone agent will always be on a concave part of the utility function.

5. When F^{-1} exists in the usual sense, this definition is the same as the standard one. More generally, if random variables x_n converge to x in distribution as $n\uparrow\infty$, then the inverse distribution function in the definition converges pointwise.

6. Alternatively, we can use a slightly simpler estimator of dispersion which uses the ordinary approximation to the integral in (8) that assigns $F_{\rho}^{-1}(\cdot)$ at the midpoint to each interval along which the sample inverse distribution function of return is constant.

7. The formal properties of the estimators are available in the literature on L-statistics. Chapter 19 of Shorack and Wellner [1986] contains a rich variety of results related to consistency and asymptotic normality of L-statistics. These results apply directly given the actual distribution function of state prices and the sample distribution of consumption or returns. (For example, as a special case, if both are lognormally distributed, we have both consistency and asymptotic normality.) If we also have an estimated distribution of state price density, the derivation of the unconditional distribution of our measure requires additional analysis but appears straightforward.

8. Strict efficiency corresponds to efficiency for some agent with a differentiable von Neumann-Morgenstern utility function. See Dybvig and Ross [1982], Table 1.

9. While this makes probabilities unequal in general, this is justified if we have a continuous state space or if we restrict attention to concave preferences.