Chapter 10

ARBITRAGE, STATE PRICES AND PORTFOLIO THEORY

PHILIP H. DYBVIG

Washington University in Saint Louis

STEPHEN A. ROSS

MIT

Contents

Abstract

Neoclassical financial models provide the foundation for our understanding of finance. This chapter introduces the main ideas of neoclassical finance in a single-period context that avoids the technical difficulties of continuous-time models, but preserves the principal intuitions of the subject. The starting point of the analysis is the formulation of standard portfolio choice problems.

A central conceptual result is the Fundamental Theorem of Asset Pricing, which asserts the equivalence of absence of arbitrage, the existence of a positive linear pricing rule, and the existence of an optimum for some agent who prefers more to less. A related conceptual result is the Pricing Rule Representation Theorem, which asserts that a positive linear pricing rule can be represented as using state prices, risk-neutral expectations, or a state-price density. Different equivalent representations are useful in different contexts.

Many applied results can be derived from the first-order conditions of the portfolio choice problem. The first-order conditions say that marginal utility in each state is proportional to a consistent state-price density, where the constant of proportionality is determined by the budget constraint. If markets are complete, the implicit state-price density is uniquely determined by investment opportunities and must be the same as viewed by all agents, thus simplifying the choice problem. Solving first-order conditions for quantities gives us optimal portfolio choice, solving them for prices gives us asset pricing models, solving them for utilities gives us preferences, and solving them for probabilities gives us beliefs.

We look at two popular asset pricing models, the CAPM and the APT, as well as complete-markets pricing. In the case of the CAPM, the first-order conditions link nicely to the traditional measures of portfolio performance.

Further conceptual results include aggregation and mutual fund separation theory, both of which are useful for understanding equilibrium and asset pricing.

Keywords

1. Introduction

The modern quantitative approach to finance has its original roots in neoclassical economics. Neoclassical economics studies an idealized world in which markets work smoothly without impediments such as transaction costs, taxes, asymmetry of information, or indivisibilities. This chapter considers what we learn from single-period neoclassical models in finance. While dynamic models are becoming more and more common, single-period models contain a surprisingly large amount of the intuition and intellectual content of modern finance, and are also commonly used by investment practitioners for the construction of optimal portfolios and communication of investment results. Focusing on a single period is also consistent with an important theme. While general equilibrium theory seeks great generality and abstraction, finance has work to be done and seeks specific models with strong assumptions and definite implications that can be tested and implemented in practice.

2. Portfolio problems

In our analysis, there are two points of time, 0 and 1, with an interval of time in between during which nothing happens. At time zero, our champion (*the agent*) is making decisions that will affect the allocation of consumption between nonrandom consumption, c_0, at time 0, and random consumption $\{c_\omega\}$ across states $\omega = 1, 2, \ldots, \Omega$ revealed at time 1. At time 0 and in each state at time 1, there is a single consumption good, and therefore consumption at time 0 or in a state at time 1 is a real number. This abstraction of a single good is obviously not "true" in any literal sense, but this is not a problem, and indeed any useful theoretical model is much simpler than reality. The abstraction does, however, face us with the question of how to interpret our simple model (in this case with a single good) in a practical context that is more complex (has multiple goods). In using a single-good model, there are two usual practices: either use nominal values and measure consumption in dollars, or use real values and measure consumption in inflation-adjusted dollars. Depending on the context, one or the other can make the most sense.

Following the usual practice from general equilibrium theory of thinking of units of consumption at various times and in different states of nature as different goods, a typical consumption vector is $C \equiv \{c_0, c_1, \ldots, c_\Omega\}$, where the real number c_0 denotes consumption of the single good at time zero, and the vector $c \equiv \{c_1, \ldots, c_\Omega\}$ of real numbers $c_1, \ldots, c_\Omega$ denotes random consumption of the single good in each state $1, \ldots, \Omega$ at time 1.

If this were a typical exercise in general equilibrium theory, we would have a price vector for consumption across goods. For example, we might have the following choice problem, which is named after two great pioneers of general equilibrium theory, Kenneth Arrow and Gerard Debreu:

Problem 1: Arrow–Debreu Problem. *Choose consumptions* $C \equiv \{c_0, c_1, \ldots, c_\Omega\}$ *to maximize utility of consumption* $U(C)$ *subject to the budget constraint*

$$c_0 + \sum_{\omega=1}^{\Omega} p_\omega c_\omega = W. \tag{1}$$

Here, $U(\cdot)$ is the *utility function* that represents preferences, p is the price vector, and W is wealth, which might be replaced by the market value of an endowment. We are taking consumption at time 0 to be the numeraire, and p_ω is the price of the *Arrow–Debreu security* which is a claim to one unit of consumption at time 1 in state ω.

The first-order condition for Problem 1 is the existence of a positive Lagrangian multiplier λ (the marginal utility of wealth) such that $U'_0(c_0) = \lambda$, and for all $\omega = 1, \ldots, \Omega$,

$$U'_\omega(c_\omega) = \lambda p_\omega.$$

This is the usual result from neoclassical economics that the gradient of the utility function is proportional to prices. Specializing to the leading case in finance of time-separable von Neumann–Morgenstern preferences, named after John von Neumann and Oscar Morgenstern (1944), two great pioneers of utility theory, we have that $U(C) = v(c_0) + \sum_{\omega=1}^{\Omega} \pi_\omega u(c_\omega)$. We will take v and u to be differentiable, strictly increasing (more is preferred to less), and strictly concave (risk averse). Here, π_ω is the probability of state ω. In this case, the first-order condition is the existence of λ such that

$$v'(c_0) = \lambda, \tag{2}$$

and for all $\omega = 1, 2, \ldots, n$,

$$\pi_\omega u'(c_\omega) = \lambda p_\omega, \tag{3}$$

or equivalently

$$u'(c_\omega) = \lambda \rho_\omega, \tag{4}$$

where $\rho_\omega \equiv p_\omega/\pi_\omega$ is the *state-price density* (also called the *stochastic discount factor* or *pricing kernel*), which is a measure of priced relative scarcity in state of nature ω. Therefore, the marginal utility of consumption in a state is proportional to the relative scarcity. There is a solution if the problem is feasible, prices and probabilities are positive, the von Neumann–Morgenstern utility function is increasing and strictly concave, and there is satisfied the *Inada condition* $\lim_{c \uparrow \infty} u'(c) = 0$.[1] There are

---

[1] Proving the existence of a solution requires more assumptions in continuous-state models.

different motivations of von Neumann–Morgenstern preferences in the literature and the probabilities may be objective or subjective. What is important for us is that the von Neumann–Morgenstern utility function represents preferences in the sense that expected utility is higher for more preferred consumption patterns. [2]

Using von Neumann–Morgenstern preferences has been popular in part because of axiomatic derivations of the theory [see, for example, Herstein and Milnor (1953) or Luce and Raiffa (1957, Chapter 2)]. There is also a large literature on alternatives and extensions to von Neumann–Morgenstern preferences. For single-period models, see Knight (1921), Bewley (1988), Machina (1982), Blume, Brandenburger and Dekel (1991) and Fishburn (1988). There is an even richer set of models in multiple periods, for example, time-separable von Neumann–Morgenstern (the traditional standard), habit formation [e.g., Duesenberry (1949), Pollak (1970), Abel (1990), Constantinides (1991) and Dybvig (1995)], local substitutability over time [Hindy and Huang (1992)], interpersonal dependence [Duesenberry (1949) and Abel (1990)], preference for resolution of uncertainty [Kreps and Porteus(1978)], time preference dependent on consumption [Bergman (1985)], and general recursive utility [Epstein and Zin (1989)].

Recently, there have also been some attempts to revive the age-old idea of studying financial situations using psychological theories [like prospect theory, Kahneman and Tversky (1979)]. Unfortunately, these models do not translate well to financial markets. For example, in prospect theory framing matters, that is, the observed phenomenon of an agent making different decisions when facing *identical* decision problems described differently. However, this is an alien concept for financial economists and when they proxy for it in models they substitute something more familiar [for example, some history dependence as in Barberis, Huang and Santos (2001)]. Another problem with the psychological theories is that they tend to be isolated stories rather than a general specification, and they are often hard to generalize. For example, prospect theory says that agents put extra weight on very unlikely outcomes, but it is not at all clear what this means in a model with a continuum of states. This literature also has problems with using *ex post* explanations (positive correlations of returns are underreaction and negative correlations are overreactions) and a lack of clarity of how much is going on that cannot be explained by traditional models (and much of it can).

In actual financial markets, Arrow–Debreu securities do not trade directly, even if they can be constructed indirectly using a portfolio of *securities*. A *security* is characterized by its cash flows. This description would not be adequate for analysis of taxes, since different sources of cash flow might have very different tax treatment, but we are looking at models without taxes. For an asset like a common stock or a bond, the cash flow might be negative at time 0, from payment of the price, and positive or zero in each state at time 1, the positive amount coming from any repayment of

---

[2] Later, when we look at multiple-agent results, we will also make the neoclassical assumption of identical beliefs, which is probably most naturally motivated by common objective beliefs.

principal, dividends, coupons, or proceeds from sale of the asset. For a futures contract, the cash flow would be 0 at time 0, and the cash flow in different states at time 1 could be positive, negative, or zero, depending on news about the value of the underlying commodity. In general, we think of the negative of the initial cash flow as the price of a security. We denote by $P = \{P_1, \ldots, P_N\}$ the vector of prices of the N securities $1, \ldots, N$, and we denote by X the payoff matrix. We have that P_n is the price we pay for one unit of security n and $X_{\omega n}$ is the payoff per unit of security n at time 1 in the single state of nature ω.

With the choice of a portfolio of assets, our choice problem might become

Problem 2: First Portfolio Choice Problem. *Choose portfolio holdings $\Theta \equiv \{\Theta_1, \ldots, \Theta_N\}$ and consumptions $C \equiv \{c_0, \ldots, c_\Omega\}$ to maximize utility of consumption $U(C)$ subject to portfolio payoffs $c \equiv \{c_1, \ldots, c_\omega\} = X\Theta$ and budget constraint $c_0 + P'\Theta = W$.*

Here, Θ is the vector of portfolio weights. Time 0 consumption is the numeraire, and wealth W is now chosen in time 0 consumption units and the entire endowment is received at time 0. In the budget constraint, the term $P'\Theta$ is the cost of the portfolio holding, which is the sum across securities n of the price P_n times the number of shares or other unit Θ_n. The matrix product $X\Theta$ says that the consumption in state ω is $c_\omega = \sum_n X_{\omega n}\Theta_n$, i.e., the sum across securities n of the payoff $X_{\omega n}$ of security n in state ω, times the number of shares or other units Θ_n of security n our champion is holding.

The first-order condition for Problem 2 is the existence of a vector of shadow prices p and a Lagrangian multiplier λ such that

$$\pi_\omega u'(c_\omega) = \lambda p_\omega, \tag{5}$$

where

$$P' = pX. \tag{6}$$

The first equation is the same as in the Arrow–Debreu model, with an implicit shadow price vector in place of the given Arrow–Debreu prices. The second equation is a pricing equation that says the prices of all assets must be consistent with the shadow prices of the states. For the Arrow–Debreu model itself, the state-space tableau X is I, the identity matrix, and the price vector P is p, the vector of Arrow–Debreu state prices. For the Arrow–Debreu model, the pricing equation determines the shadow prices as equal to the state prices.

Even if the assets are not the Arrow–Debreu securities, Problem 2 may be essentially equivalent to the Arrow–Debreu model in Problem 1. In economic terms, the important feature of the Arrow–Debreu problem is that all payoff patterns are *spanned*, i.e., each potential payoff pattern can be generated at some price by some portfolio of assets. Linear algebra tells us that all payoff patterns can be generated if the payoff matrix X

has full row rank. If X has full row rank, p is determined (or over-determined) by Equation (6). If p is uniquely determined by the pricing equation (and therefore also all Arrow–Debreu assets can be purchased as portfolios of assets in the economy), we say that markets are *complete*, and for all practical purposes we are in an Arrow–Debreu world.

For the choice problem to have a solution for any agent who prefers more to less, we also need for the price of each payoff pattern to be unique (the "law of one price") and positive, or else there would be *arbitrage* (i.e., a "money pump" or a "free lunch"). If there is no arbitrage, then there is at least one vector of positive state prices p solving the pricing equation (6). There is an arbitrage if the vector of state prices is overdetermined or if all consistent vectors of state prices assign a negative or zero price to some state. The notion of absence of arbitrage is a central concept in finance, and we develop its implications more fully in the section on preference-free results.

So far, we have been stating portfolio problems in prices and quantities, as we would in general equilibrium theory. However, it is also common to describe assets in terms of *rates of return*, which are relative price changes (often expressed as percentages). The return to security n, which is the relative change in total value (including any dividends, splits, warrant issues, coupons, stock issues, and the like as well as change in the price). There is not an absolute standard of what is meant by return, in different contexts this can be the rate of return, one plus the rate of return, or the difference between two rates of return. It is necessary to figure which is intended by asking or from context. Using the notation above, the rate of return in state ω is $r_{\omega n} = (X_{\omega n} - P_n)/P_n$.[3] Often, consumption at the outset is suppressed, and we specialize to von Neumann–Morgenstern expected utility. In this case, we have the following common form of portfolio problem.

Problem 3: Portfolio Problem using Returns. *Choose portfolio proportions* $\theta \equiv \{\theta_1, \ldots, \theta_N\}$ *and consumptions* $c \equiv \{c_1, \ldots, c_\Omega\}$ *to maximize expected utility of consumption* $\sum_{\omega=1}^{\Omega} \pi_\theta u(c_\omega)$ *subject to the consumption equation* $c = W\theta'(1+r)$ *and the budget constraint* $\theta'1 = 1$.

Here, $\pi = \{\pi_1, \ldots, \pi_\Omega\}$ is a vector of state probabilities, $u(\cdot)$ is the von Neumann–Morgenstern utility function, and 1 is a vector of 1's. The dimensionality of 1 is determined implicitly from the context; here the dimensionality is the number of assets. The first-order condition for an optimum is the existence of shadow state-price density vector ρ and shadow marginal utility of wealth λ such that

$$u'(c_\omega) = \lambda \rho_\omega \tag{7}$$

---

[3] One unfortunate thing about returns is that they are not defined for contracts (like futures) that have zero price. However, this can be finessed formally by bundling a futures with a bond or other asset in defining the securities and unbundling them when interpreting the results. Bundling and unbundling does not change the underlying economics due to the linearity of consumptions and constraints in the portfolio choice problem.

and

$$1 = E[(1 + r)\rho].\tag{8}$$

These equations say that the state-price density is consistent with the marginal valuation by the agent and with pricing in the market.

As our final typical problem, let us consider a mean-variance optimization. This optimization is predicated on the assumption that investors care only about mean and variance (typically preferring more mean and less variance), so we have a utility function $V(m, v)$ in mean m and variance v. For this problem, suppose there is a risk-free asset paying a return r (although the market-level implications of mean-variance analysis can also be derived in a general model without a risky asset). In this case, portfolio proportions in the risky assets are unconstrained (need not sum to 1) because the slack can be taken up by the risk-free asset. We denote by μ the vector of mean risky asset returns and by σ the covariance matrix of risky returns. Then our champion solves the following choice problem.

Problem 4: Mean-variance optimization. *Choose portfolio proportions* $\theta \equiv \{\theta_1, \ldots, \theta_N\}$ *to maximize the mean-variance utility function* $V(r + (\mu - r\mathbf{1})' \theta, \theta' \Sigma \theta)$.

The first-order condition for the problem is

$$\mu - r\mathbf{1} = \lambda \Sigma \theta,\tag{9}$$

where θ is the optimal vector of portfolio proportions and λ is twice the marginal rate of substitution $V_v'(m, v)/V_m'(m, v)$, evaluated at $m = r + (\mu - r\mathbf{1})' \theta$ and $v = \theta' \Sigma \theta$. The first-order condition (9) says that mean excess return for each asset is proportional to the marginal contribution of volatility to the agent's optimal portfolio.

We have seen a few of the typical types of portfolio problem. There are a lot of variations. The problem might be stated in terms of excess returns (rate of return less a risk-free rate) or total return (one plus the rate of return). Or, we might constrain portfolio holdings to be positive (no short sales) or we might require consumption to be nonnegative (limited liability). Many other variations adapt the basic portfolio problem to handle institutional features not present in a neoclassical formulation, such as transaction costs, bid–ask spreads, or taxes. These extensions are very interesting, but beyond the scope of what we are doing here, which is to explore the neoclassical foundations.

3. Absence of arbitrage and preference-free results

Before considering specific solutions and applications, let us consider some general results that are useful for thinking about portfolio choice. These results are *preference-free* in the sense that they do not depend on any specific assumptions about preferences

but only depend on an assumption that agents prefer more to less. Central to this section is the notion of an arbitrage, which is a "money pump" or a "free lunch". If there is arbitrage, linearity of the neoclassical problem implies that any candidate optimum can be dominated by adding the arbitrage. As a result, no agent who prefers more to less would have an optimum if there exists arbitrage. Furthermore, this seemingly weak assumption is enough to obtain two useful theorems. The *Fundamental Theorem of Asset Pricing* says that the following are equivalent: absence of arbitrage, existence of a consistent positive linear pricing rule, and existence of an optimum for some hypothetical agent who prefers more to less. The *Pricing Rule Representation Theorem* gives different equivalent forms for the consistent positive linear pricing rule, using state prices, risk-neutral probabilities (martingale valuation), state-price density (or stochastic discount factor or pricing kernel), or an abstract positive linear operator. The results in this section are from Cox and Ross (1975), Ross (1976c, 1978b) and Dybvig and Ross (1987). The results have been formalized in continuous time by Harrison and Kreps (1979) and Harrison and Pliska (1981).

Occasionally, the theorems in this section can be applied directly to obtain an interesting result. For example, linearity of the pricing rule is enough to derive put-call parity without constructing the arbitrage. More often, the results in this section help to answer conceptual questions. For example, an option pricing formula that is derived using absence of arbitrage is *always* consistent with equilibrium, as can be seen from the Fundamental Theorem. By the Fundamental Theorem, absence of arbitrage implies there is an optimum for some hypothetical agent who prefers more to less; we can therefore construct an equilibrium in the single-agent pure exchange economy in which this agent is endowed with the optimal holding. By construction the equilibrium in this economy will have the desired pricing, and therefore any no-arbitrage pricing result is consistent with some equilibrium.

In this section, we will work in the context of Problem 2. An arbitrage is a change in the portfolio that makes all agents who prefer more to less better off. We make all such agents better off if we increase consumption sometime, and in some state of nature, and we never decrease consumption. By combining the two constraints in Problem 2, we can write the consumption C associated with any portfolio choice Θ using the stacked matrix equation

$$C = \begin{bmatrix} W \\ 0 \end{bmatrix} + \begin{bmatrix} -P' \\ X \end{bmatrix} \Theta.$$

The first row, $W - P'\Theta$, is consumption at time 0, which is wealth W less the cost of our portfolio. The remaining rows, $X\Theta$, give the random consumption across states at time 1.

Now, when we move from the portfolio choice Θ to the portfolio choice $\Theta + \eta$, the initial wealth term cancels and the change in consumption can now be written as

$$\Delta C = \begin{bmatrix} -P' \\ X \end{bmatrix} \eta.$$

This will be an arbitrage if ΔC is never negative and is positive in at least one component, which we will write as[4] $\Delta C > 0$ or

$$\begin{bmatrix} -P' \\ X \end{bmatrix} \eta > 0.$$

Some authors describe taxonomies of different types of arbitrage, having perhaps a negative price today and zero payoff tomorrow, a zero price today and a nonnegative but not identically zero payoff tomorrow, or a negative price today and a positive payoff tomorrow. These are all examples of arbitrages that are subsumed by our general formula. The important thing is that there is an increase in consumption in some state of nature at some point of time and there is never any decrease in consumption.

3.1. Fundamental theorem of asset pricing

Theorem 1: Fundamental Theorem of Asset Pricing. *The following conditions on prices P and payoffs X are equivalent:*

(i)   *Absence of arbitrage:* $(\nexists \eta) \left(\begin{bmatrix} -P' \\ X \end{bmatrix} \eta > 0 \right).$

(ii) *Existence of a consistent positive linear pricing rule (positive state prices):*
       $(\exists p \gg 0)(P' = p'X).$

(iii) *Some agent with strictly increasing preferences U has an optimum in Problem 2.*

Proof: We prove the equivalence by showing (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii), and (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): This is the most subtle part, and it follows from a separation theorem or the duality theorem from linear programming. From the definition of absence of arbitrage, we have that the sets

$$S_1 \equiv \left\{ \begin{bmatrix} -P' \\ X \end{bmatrix} \eta \mid \eta \in \mathfrak{R}^n \right\}$$

and

$$S_2 \equiv \left\{ x \in \mathfrak{R}^{\Omega + 1} \mid x > 0 \right\}$$

must be disjoint. Therefore, there is a separating hyperplane z such that $z'x = 0$ for all $x \in S_1$ and $z'x > 0$ for all $x \in S_2$. [See Karlin (1959), Theorem B3.5] Normalizing so that the first component (the shadow price of time zero consumption) is 1, we will see that p defined by $(1 \ p') = z/z_0$ is the consistent linear pricing rule we seek. Constancy

---

[4] We use the following terminology for vector inequalities: $(x \geqslant y) \Leftrightarrow (\forall i) (x_i \geqslant y_i)$, $(x > y) \Leftrightarrow ((x \geqslant y)$ & $(\exists i) (x_i > y_i))$, and $(x \gg y) \Leftrightarrow (\forall i) (x_i > y_i)$.

of zx for $x \in S_1$ implies that $(1 \ p') \begin{bmatrix} -P' \\ X \end{bmatrix} = 0$, which is to say that $P' = p'X$, i.e., p is a consistent linear pricing rule. Furthermore, $z'x$ positive for $x \in S_2$ implies $z \gg 0$ and consequently $p \gg 0$, and p is indeed the desired consistent positive linear pricing rule.

(ii) $\Rightarrow$ (iii): This part is proven by construction. Let $U(C) = (1 \ p') C$, then $\Theta = 0$ solves Problem 2. To see this, note that the objective function $U(C)$ is constant and equal to W for all Θ:

$$
\begin{aligned}
U(C) &= (1 \ p') C \\
&= (1 \ p') \left(\begin{bmatrix} W \\ 0 \end{bmatrix} + \begin{bmatrix} -P' \\ X \end{bmatrix} \Theta \right) \\
&= W + (-P' + p'X) \Theta \\
&= W.
\end{aligned}
$$

(The motivation of this construction is the observation that the existence of the consistent linear pricing rule with state prices p implies that all feasible consumptions satisfy $(1 \ p') C = W$.)

(iii) $\Rightarrow$ (i): This part is obvious, since any candidate optimum is dominated by adding the arbitrage, and therefore there can be no arbitrage if there is an optimum. More formally, adding an arbitrage implies the change of consumption $\Delta C > 0$, which implies an increase in $U(C)$. $\square$

One feature of the proof that may seem strange is the degeneracy (linearity) of the utility function whose existence is constructed. This was all that was needed for this proof, but it could also be constructed to be strictly concave, additively separable over time, and of the von Neumann–Morgenstern class for given probabilities. Assuming any of these restrictions on the class would make some parts of the theorem weaker [(iii) implies (i) and (ii)] at the same time that it makes other parts stronger [(i) or (ii) implies (iii)]. The point is that the theorem is still true if (iii) is replaced by a much more restrictive class that imposes on U any or all of strict concavity, some order of differentiability, additive separability over time, and a von Neumann–Morgenstern form with or without specifying the probabilities in advance. All of these classes are restrictive enough to rule out arbitrage, and general enough to contain a utility function that admits an optimum when there is no arbitrage.

The statement and proof of the theorem are a little more subtle if the state space is infinite-dimensional. The separation theorem is topological in nature, so we must restrict our attention to a topologically relevant subset of the nonnegative random variables. Also, we may lose the separating hyperplane theorem because the interior of the positive orthant is empty in most of these spaces (unless we use the sup-norm topology, in which case the dual is very large and includes dual vectors that do not support state prices). However, with some definition of arbitrage in limits, the economic content of the Fundamental Theorem can be maintained.

3.2. Pricing rule representation theorem

Depending on the context, there are different useful ways of representing the pricing rule. For some abstract applications (like proving put–call parity), it is easiest to use a general abstract representation as a linear operator $L(c)$ such that $c > 0 \Rightarrow L(c) > 0$. For asset pricing applications, it is often useful to use either the state-price representation we used in the Fundamental Theorem, $L(c) = \sum_{\omega} p_{\omega} c_{\omega}$, or risk-neutral probabilities, $L(c) = (1 + r^*)^{-1} E^*[c_{\omega}] = (1 + r^*)^{-1} \sum_{\omega} \pi_{\omega}^* c_{\omega}$. The intuition behind the risk-neutral representation (or martingale representation[5]) is that the price is the expected discounted value computed using a shadow risk-free rate (equal to the actual risk-free rate if there is one) and artificial *risk-neutral probabilities* π^* that assign positive probability to the same states as do the true probabilities. Risk-neutral pricing says that all investments are fair gambles once we have adjusted for time preference by discounting and for risk preference by adjusting the probabilities. The final representation using the *state-price density* (or stochastic discount factor) ρ to write $L(c) = E[\rho_{\omega} c_{\omega}] = \sum_{\omega} \pi_{\omega} \rho_{\omega} c_{\omega}$. The state price density simplifies first-order conditions of portfolio choice problems because the state-price density measures priced scarcity of consumption. The state-price density is also handy for continuous-state models in which individual states have zero state probabilities and state prices but there exists a well-defined positive ratio of the two.

Theorem 2: Pricing Rule Representation Theorem. *The consistent positive linear pricing rule can be represented equivalently using*
(i) *an abstract linear functional $L(c)$ that is positive: $(c > 0) \Rightarrow (L(c) > 0)$*
(ii) *positive state prices $p \gg 0$: $L(c) = \sum_{\omega=1}^{\Omega} p_{\omega} c_{\omega}$*
(iii) *positive risk-neutral probabilities $\pi^* \gg 0$ summing to 1 with associated shadow risk-free rate r^*: $L(c) = (1 + r^*)^{-1} E^*[c_{\omega}] \equiv (1 + r^*)^{-1} \sum_{\omega} \pi_{\omega}^* c_{\omega}$*
(iv) *positive state-price densities $\rho \gg 0$: $L(c) = E[\rho c] \equiv \sum_{\omega} \pi_{\omega} \rho_{\omega} c_{\omega}$.*

Proof: (i) $\Rightarrow$ (ii): This is the known form of a linear operator in $\mathfrak{R}^{\Omega}$; $p \gg 0$ follows from the positivity of L.

(ii) $\Rightarrow$ (iii): Note first that the shadow risk-free rate must price the riskless asset $c = 1$:

$$\sum_{\omega=1}^{\Omega} p_{\omega} 1 = (1 + r^*)^{-1} E^*[1],$$

which implies (since $E^*[1] = 1$) that $r^* = 1/p'1 - 1$. Then, matching coefficients in

$$\sum_{\omega=1}^{\Omega} p_{\omega} c_{\omega} = (1 + r^*)^{-1} \sum_{\omega} \pi_{\omega}^* c_{\omega},$$

---

[5] The reason for calling the term "martingale representation" is that using the risk-neutral probabilities makes the discounted price process a martingale, which is a stochastic process that does not increase or decrease on average.

we have that $\pi^* = p/1'p$, which sums to 1 as required and inherits positivity from p.

(iii) $\Rightarrow$ (iv): Simply let $\rho_\omega = (1 + r^*)^{-1}\pi_\omega^*$ (which is the same as p_ω/π_ω).

(iv) $\Rightarrow$ (i): immediate. $\square$

Perhaps what is most remarkable about the Fundamental Theorem and the Representation Theorem is that neither probabilities nor preferences appear in the determination of the pricing operator, beyond the initial identification of which states have nonzero probability and the assumption that more is preferred to less. It is this observation that empowers the theory of derivative asset pricing, and is, for example, the reason why the Black–Scholes option price does not depend on the mean return on the underlying stock. Preferences and beliefs are, however, in the background: in equilibrium, they would influence the price vector P and/or the payoff matrix X (or the mean return process for the Black–Scholes stock).

Although the focus of this chapter is on the single-period model, we should note that the various representations have natural multiperiod extensions. The abstract linear functional and state prices have essentially the same form, noting that cash flows now extend across time as well as states of nature and that there are also conditional versions of the formula at each date and contingency. In some models, the information set is generated by the sample path of security prices; in this case the state of nature is a sample path through the tree of potential security prices. For the state-price density in multiple periods, there is in general a state-price-density process $\{\rho_t\}$ whose relatives can be used for valuation. For example, the value at time s of receiving subsequent cash flows $c_{s+1}, c_{s+2} \ldots c_t$ is given by

$$\sum_{\tau = s+1}^{t} E_s\left[\frac{\rho_\tau}{\rho_s} c_\tau\right], \tag{10}$$

where $E_s[\cdot]$ denotes expectation conditional on information available at time s. Basically, this follows from iterated expectations and defining ρ_t as a cumulative product of single-period ρ's. Similarly, we can write risk-neutral valuation as

$$P_s = E_s^*\left[\frac{P_\tau}{(1 + r_s^*)(1 + r_{s+1}^*) \ldots (1 + r_\tau^*)}\right]. \tag{11}$$

Note that unless the risk-free rate is nonrandom, we cannot take the discount factors out of the expectation.[6] This is because of the way that the law of iterated expectations works. For example, consider the value V_0 at time 0 of the cash flow in time 2.

$$\begin{aligned}
V_0 &= (1 + r_1^*)^{-1} E_0^*[V_1] \\
&= (1 + r_1^*)^{-1} E_0^*[(1 + r_2^*)^{-1} E_1^*[c_2]] \\
&= (1 + r_1^*)^{-1} E_0^*[(1 + r_2^*)^{-1} c_2]. \tag{12}
\end{aligned}$$

---

[6] It would be possible to treat the whole time period from s to τ as a single period and apply the pricing result to that large period in which case the discounting would be at the appropriate $(\tau - s)$-period rate. The problem with this is that the risk-neutral probabilities would be different for each pair of dates, which is unnecessarily cumbersome.

Now, $(1 + r_1^*)^{-1}$ is outside the expectation (as could be $(1 + r_{s+1}^*)^{-1}$ in Equation (11), but $(1 + r_2^*)^{-1}$ cannot come outside the expectation unless it is nonrandom.[7] So, it is best to remember that when interest rates are stochastic, discounting for risk-neutral valuation should use the rolled-over spot rate, within the expectation.

4. Various analyses: Arrow–Debreu world

The portfolio problem is the starting point of a lot of types of analysis in finance. Here are some implications that can be drawn from portfolio problems (usually through the first-order conditions):
• optimal portfolio choice (asset allocation or stock selection)
• portfolio efficiency
• aggregation and market-level implications
• asset pricing and performance measurement
• payoff distribution pricing
• recovery or estimation of preferences
• inference of expectations
We can think of many of these distinctions as a question of what we are solving for when we look at the first-order conditions. In optimal portfolio choice and its aggregation, we are solving for the portfolio choice given the preferences and beliefs about returns. In asset pricing, we are computing the prices (or restrictions on expected returns) given preferences, beliefs about payoffs, and the optimal choice (which is itself often derived using an aggregation result). In recovery, we derive preferences from beliefs and idealized observations about portfolio choice, e.g. at all wealth levels. Estimation of preferences is similar, but works with noisy observations of demand at a finite set of data points and uses a restriction in the functional form or smoothing in the statistical procedure to identify preferences. And, inference of expectations derives probability beliefs from preferences, prices, and the (observed) optimal demand. In this section, we illustrate the various analyses in the case of an Arrow–Debreu world.

Analysis of the complete-markets model has been developed by many people over a period of time. Some of the more important works include some of the original work on competitive equilibrium such as Arrow and Debreu (1954), Debreu (1959) and Arrow and Hahn (1971), as well as some early work specific to security markets such as Arrow (1964), Rubinstein (1976), Ross (1976b), Banz and Miller (1978) and Breeden and Litzenberger (1978). There are also a lot of papers set in

---

[7] In the special case in which c_2 is uncorrelated with $(1 + r_2^*)^{-1}$ (or in multiple periods if cash flows are all independent of shadow interest rate moves), we can take the expected discount factor outside the expectation. In this case, we can use the multiperiod riskfree discount bond rate for discounting a simple expected final. However, in general, it is best to remember the general formula (11) with the rates in the denominator inside the expectation.

multiple periods that contributed to the finance of complete markets; although not strictly within the scope of this chapter, we mention just a few here: Black and Scholes (1973), Merton (1971, 1973), Cox, Ross and Rubinstein (1979) and Breeden (1979).

4.1. Optimal portfolio choice

The optimal portfolio choice is the choice of consumptions $(c_0, c_1, \ldots, c_\Omega)$ and Lagrange multiplier λ to solve the budget constraint (1) and the first-order conditions (2) and (3). If the inverse $I(\cdot)$ of $u'(\cdot)$ and the inverse $J(\cdot)$ of $v'(\cdot)$ are both known analytically, then finding the optimum can be done using a one-dimensional monotone search for λ such that $J(\lambda) + \sum_{\omega=1}^{\Omega} p_\omega I(\lambda p_\omega / \pi_\omega) = W$. In some special cases, we can solve the optimization analytically. For logarithmic utility, $v(c) = \log(c)$ and $u(c) = \delta \log(c)$ for some $\delta > 0$, optimal consumption is given by $c_0 = W/(1 + \delta)$ and $c_\omega = \pi_\omega W \delta / ((1 + \delta) p_\omega)$ (for $\omega = 1, \ldots, \Omega$). The portfolio choice can also be solved analytically for quadratic utility.

4.2. Efficient portfolios

Efficient portfolios are the ones that are chosen by some agent in a given class of utility functions. For the Arrow–Debreu problem, we might take the class of utility functions to be the class of differentiable, increasing and strictly concave time-separable von Neumann–Morgenstern utility functions $U(c) = v(c_0) + \sum_{\omega=1}^{\Omega} \pi_\omega u(c_\omega)$.[8] Since $u(\cdot)$ is increasing and strictly concave, $(c_\omega > c_{\omega'}) \Leftrightarrow u'(c_\omega) < u'(c_{\omega'})$. Consequently, the first-order condition (4) implies that $(c_\omega > c_{\omega'}) \Leftrightarrow (\rho_\omega < \rho_{\omega'})$. Since the state-price density $\rho_\omega \equiv p_\omega / \pi_\omega$ is a measure of priced social scarcity in state ω, this says that we consume less in states in which consumption is more expensive. This necessary condition for efficiency is also sufficient; if consumption reverses the order across states of the state-price density, then it is easy to construct a utility function that satisfies the first-order conditions. Formally,

Theorem 3: Arrow–Debreu Portfolio Efficiency. *Consider a complete-markets world (in which agents solve Problem 1) in which state prices and probabilities are all strictly positive, and let $\mathcal{U}$ be the class of differentiable, increasing and strictly concave time-separable von Neumann–Morgenstern utility functions of the form $U(c) = v(c_0) + \sum_{\omega=1}^{\Omega} \pi_\omega u(c_\omega)$. Then there exists a utility function in the class $\mathcal{U}$ that chooses the consumption vector c satisfying the budget constraint if and only if consumptions at time 1 are in the opposite order as the state-price densities, i.e., $(\forall \omega, \omega' \in \{1, \ldots, \Omega\})((c_\omega > c_{\omega'}) \Leftrightarrow (\rho_\omega < \rho_{\omega'}))$.*

---

[8] A non-time-separable version would be of the form $U(c) = \sum_{\omega=1}^{\Omega} \pi_\omega u(c_0, c_\omega)$.

Proof: The "only if" part follows directly from the first-order condition and concavity as noted in the paragraph above. For the "if" part, we are given a consumption vector with the appropriate ordering and we will construct a utility function that will choose it and satisfy the first-order condition with $\lambda = 1$. For this, choose $v(c) = \exp(-(c - c_0))$ (so that $v'(c_0) = 1$ as required by Equation 2), and choose $u'(c)$ to be any strictly positive and strictly decreasing function satisfying $u'(c_\omega) = \rho_\omega$ for all $\omega = \{1, 2, \dots, \Omega\}$, for example, by "connecting the dots" (with appropriate treatment past the endpoints) in the graph of ρ_ω as a function of c_ω. Integrating this function yields a utility function $u(\cdot)$ such that the von Neumann–Morgenstern utility function satisfies the first-order conditions, and by concavity this first-order solution is a solution. $\square$

Friendly warning. *There are many notions of efficiency in finance: Pareto efficiency, informational efficiency, and the portfolio efficiency we have mentioned are three leading examples. A common mistake in heuristic arguments is to assume incorrectly that one sense of efficiency necessarily implies another.*

4.3. Aggregation

Aggregation results typically show what features of individual portfolio choice are preserved at the market level. Many asset pricing results follow from aggregation and the first-order conditions. The most common type of aggregation result is the efficiency of the market portfolio. For most classes of preferences we consider, the efficient set is unchanged by rescaling wealth, and consequently the market portfolio is always efficient if and only if the efficient set is convex. This is because the market portfolio is a rescaled version of the individual portfolios. (If the portfolios are written in terms of proportions, no rescaling is needed). When the market portfolio is efficient, then we can invert the first-order condition for the hypothetical agent who holds the market portfolio to obtain the pricing rule.

In the Arrow–Debreu world, the market portfolio is always efficient. This is because the ordering across states is preserved when we sum individual portfolio choices to form the market portfolio. Consider agents $m = 1, \dots, M$ with felicity functions $v^1(\cdot), \dots, v^M(\cdot)$ and $u^1(\cdot), \dots, u^M(\cdot)$ and optimal consumptions $C^{1*}, \dots, C^{M*}$. The following results are close relatives of standard results in general equilibrium theory.

Theorem 4: Aggregation Theorem. *In a pure exchange equilibrium in a complete market,*
(i)   all agents order time 1 consumption in the same order across states,
(ii) aggregate time 1 consumption is in the same order across states,
(iii) equilibrium is Pareto optimal, and
(iv) there is a time separable von Neumann–Morgenstern utility function that would choose optimally aggregate consumption.

Proof: (i) and (ii) Immediate, given Theorem 3.

(iii) Let λ^m be the Lagrangian multiplier at the optimum in the first order condition in agent m's decision problem. Consider the problem of maximizing the linear social welfare function with weights λ^m, namely

$$\sum_{n=1}^{N} \lambda^n \left(v^n(c_0) + \sum_{\omega=1}^{\Omega} u^n(c_\omega^n) \right).$$

It is easy to verify from the first-order conditions from the equilibrium consumptions that they solve this problem too. This is a concave optimization, so the first-order conditions are sufficient, and since the welfare weights are positive the solution must be Pareto optimal (or else a Pareto improvement would increase the objective function).

(iv) Define $v^A(c) \equiv \max_{c^{n's}} \sum_{n=1}^{N} \lambda^n v^n(c_n)$ to be the first-period aggregate felicity function and define $u^A(c) \equiv \max_{c^{n's}} \sum_{n=1}^{\Omega} \lambda^n u^n(c_n)$ to be the second-period aggregate felicity function. Then the utility function $v^A(c_0) + E[u^A(c_\omega)]$ is a time-separable von Neumann–Morgenstern utility function that would choose the market's aggregate consumption, since the objective function is the same as for the social welfare problem described under the proof of (iii). $\square$

There is a different perspective that gives an alternative proof of the existence of a represenatitive agent (iv). The existence of a representative agent follows from the convexity of the set of efficient portfolios derived earlier. The main condition we require to have this work is that the efficient set of portfolio proportions is the same at all wealth levels, which is true here and typically of the cases we consider.

4.4. Asset pricing

Asset pricing gets its name from valuation of cash flows, although asset pricing formulas may be expressed in several different ways, for example as a formula explaining expected returns across assets or as a moment condition satisfied by returns that can be tested econometrically. Let $v^A(\cdot)$ and $u^A(\cdot)$ represent the preferences of the hypothetical agent who holds aggregate consumption, as guaranteed by the aggregation theorem 4. Then we can solve the first-order conditions (2) and (3) to compute $p_\omega = \pi_\omega u^{A'}(c_\omega^A)/v^{A'}(c_0^A)$ and therefore the time-0 valuation of the time-1 cash flow vector $\{c_1, \ldots, c_\Omega\}$ is

$$
\begin{aligned}
L(c_1, \ldots, c_\Omega) &= \sum_{\omega=1}^{\Omega} \pi_\omega \frac{u^{A'}(c_\omega^A)}{v^{A'}(c_0^A)} c_\omega \\
&= E\left[\frac{u^{A'}(c_\omega^A)}{v^{A'}(c_0^A)} c_\omega \right].
\end{aligned}
\tag{13}
$$

This formula (with state-price density $\rho_\omega = u^{A'}(c_\omega^A)/v^{A'}(c_0^A)$) is the right one for pricing assets, but asset pricing equations are more often expressed as explanations of mean

returns across assets or as moment conditions satisfied by returns. Defining the rate of return (the relative value change) for some asset as $r_\omega \equiv (c_\omega - P)/P$ where c_ω is the asset value in state ω and P is the asset's price. Letting r_f be the risk-free rate of return (or the riskless interest rate), which must be

$$r_f = \frac{1}{E[u^{A'}(c_\omega^A)/v^{A'}(c_0^A)]}, \tag{14}$$

we have that Equation (13) implies

$$E[r_\omega] = r_f + (1 + r_f)\,\mathrm{cov}\left(\frac{u^{A'}(c_\omega^A)}{v^{A'}(c_0^A)}, r_\omega\right), \tag{15}$$

so that the risk premium (the excess of expected return over the risk-free rate) is proportional to covariance of return with the state-price density. This is the representation of asset pricing in terms of expected returns, and is also the so-called consumption-capital asset pricing model (CCAPM) that is more commonly studied in a multiperiod setting.

Either of the pricing relations could be used as moment conditions in an asset-pricing test, but it is more common to use the moment condition

$$1 = E\left[\frac{u^{A'}(c_\omega^A)}{v^{A'}(c_0^A)}(1 + r_\omega)\right], \tag{16}$$

to test the CCAPM. This same equations characterize pricing for just about all the pricing models (perhaps with optimal consumption for some agent in place of aggregate consumption). Recall that the first-order conditions are just about the same whether markets are complete or incomplete. The main difference is that the state prices are shadow prices (Lagrangian multipliers) when markets are incomplete, but actual asset prices in complete markets. Either way, the first-order conditions are consistent with the same asset pricing equations.

4.5. Payoff distribution pricing

For von Neumann–Morgenstern preferences (expected utility theory) and more general Machina preferences, preferences depend only on distributions of returns and payoffs and do not depend on the specific states in which those returns are realized. Consider, for example, a simple example with three equally probable states, $\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$. Suppose that an individual has to choose one of the following payoff vectors for consumption at time 1: $c_1' = (1, 2, 2)$, $c_2' = (2, 1, 2)$, and $c_3' = (2, 2, 1)$. These three consumption patterns have the same distribution of consumption, giving consumption of 1 with probability $\frac{1}{3}$ and consumption of 2 with probability $\frac{2}{3}$. Therefore, an agent with von Neumann–Morgenstern preferences or more general Machina preferences

would find all these consumption vectors are equally attractive. However they do not all cost the same unless the state-price density (and in the example, the state price) is the same in all states. However, having the state-price density the same in all states is a risk-neutral world – all consumption bundles priced at their expected value – which is not very interesting since all risk-averse agents would choose a riskless investment. [9] In general, we expect the state-price density to be highest of states of social scarcity, when the market is down or the economy is in recession, since buying consumption in states of scarcity is a form of insurance.

Suppose that the state-price vector is $p' = (.3, .2, .4)$. Then the prices of the bundles can be computed as $p'c^a = .3 \times 1 + .2 \times 2 + .4 \times 2 = 1.5$, $p'c^b = .3 \times 2 + .2 \times 1 + .4 \times 2 = 1.6$, and $p'c^c = .3 \times 2 + .2 \times 2 + .4 \times 1 = 1.4$. The cheapest consumption pattern is c^c, which places the larger consumption in the cheap states and the smallest consumption in the most expensive state. This gives us a very useful cash-value measure of the inefficiency of the other strategies. An agent will save $1.5 - 1.4 = 0.1$ cash up front by choosing c^c up front instead of c^a or $1.6 - 1.4 = 0.2$ cash up front by choosing c^c instead of c^b. Therefore, we can interpret 0.1 as a lower bound on the amount of inefficiency in c^a, since any agent would pay that amount to swap to c^c and perhaps more to swap to something better. The only assumption we need for this result is that the agent has preferences (such as von Neumann–Morgenstern preferences or Machina preferences) that care only about the distribution of consumption and not the identity of the particular states in which different parts of the distribution are realized.

The general result is based on the "deep theoretical insight" that you should "buy more when it is cheaper". This means that efficient consumption is decreasing in the state-price density. We can compute the (lower bound on the) inefficiency of the portfolio by reording its consumption in reverse order as the state-price density and computing the decline in cost. The *payoff distributional price* of a consumption pattern is the price of getting the same distribution the cheapest possible way (in reverse order as the state-price density).

There is a nice general formula for the distributional price. Let $F_c(\cdot)$ be the cumulative distribution function of consumption and let $\iota_c(\cdot)$ be its inverse. Similarly, let $F_\rho(\cdot)$ be the cumulative distribution function of the state-price density and let $\iota_\rho(\cdot)$ be its inverse. Let c^* be the efficient consumption pattern with distribution function $F_c(\cdot)$. Then the distributional price of the consumption pattern can be written as

$$E[c^*\rho] = \int_{z=0}^{1} \iota_c(z) \, \iota_\rho(1-z) \, \mathrm{d}z. \tag{17}$$

In this expression, z has units of probability and labels the states ordered in reverse of the state-price density, $\iota_\rho(1-z)$ is state-price density in state z, and $\iota_c(z)$ is the optimal

---

[9] This is different from there existing a change of probability that gives risk-neutral pricing. In a risk-neutral world, the actual probabilities are also risk-neutral probabilities.

consumption c^* in state z. This formula is simplest to understand for a continuous state space, but also makes sense for finitely many equally-probable states as in the example, provided we define the inverse distribution function at mass points in the natural way.

The original analysis of Payoff Distribution pricing for complete frictionless markets was presented by Dybvig (1988a,b). Payoff Distribution pricing can also be used in a model with incomplete markets or frictions, as developed by Jouini and Kallal (2001), but that analysis is beyond the scope of this chapter.

5. Capital asset pricing model (CAPM)

The *Capital Asset Pricing Model (CAPM)* is an asset-pricing model based on equilibrium with agents having mean-variance preferences (as in Problem 4). It is based on the mean-variance analysis pioneered by Markowitz (1952, 1959) and Tobin (1958), and was extended to an equilibrium model by Sharpe (1964) and Lintner (1965). Even though there are many more modern pricing models, the CAPM is still the most important. This model gives us most of our basic intuitions about the trade-off between risk and return, about how market risk is priced, and about how idiosyncratic risk is not priced. The CAPM is also widely used in practice, not only in the derivation of optimal portfolios but also in the *ex post* assessment of performance. Sometimes people still refer to mean-variance analysis by the term *Modern Portfolio Theory* without intending a joke, even though we are approaching its 50th anniversary.

In theoretical work, the mean-variance preferences assumed in Problem 4 are usually motivated by joint normality of returns (a restriction on beliefs) or by a restriction on preferences (a quadratic von Neumann–Morgenstern utility function). When returns are jointly normal, so are portfolio returns, so the entire distribution of a portfolio's return (and therefore utility that depends only on distribution) is determined by the mean and variance. For quadratic utility, there is an algebraic relation between expected utility and mean and variance. Letting $u(c) = k_1 + k_2 c - k_3 c^2$,

$$
\begin{aligned}
E[u(c)] &= k_1 + k_2 E[c] - k_3 E[c^2] \\
&= k_1 + k_2 E[c] - k_3 (\text{var}(c) + (E[c])^2),
\end{aligned}
\tag{18}
$$

which depends on the preferences parameters k_1, k_2, and k_3 and the mean and variance of c and not on other features of the distribution (such as skewness or kurtosis). Neither assumption is literally true, but we must remember that models must be simpler than the world if they are to be useful.

You may wonder why we need to motivate the representation of preferences by the utility function $V(m, v)$, since it may seem very intuitive to write down preferences for risk an return directly. However, it is actually a little strange to assume that these preferences apply to all random variables. For example, if there is a trade-off between risk and return (so the agent cares about risk), then there should exist $m_1 > m_2$ and $v_1 > v_2 > 0$ such that the agent $V(m_1, v_1) < V(m_2, v_2)$ and the agent would turn
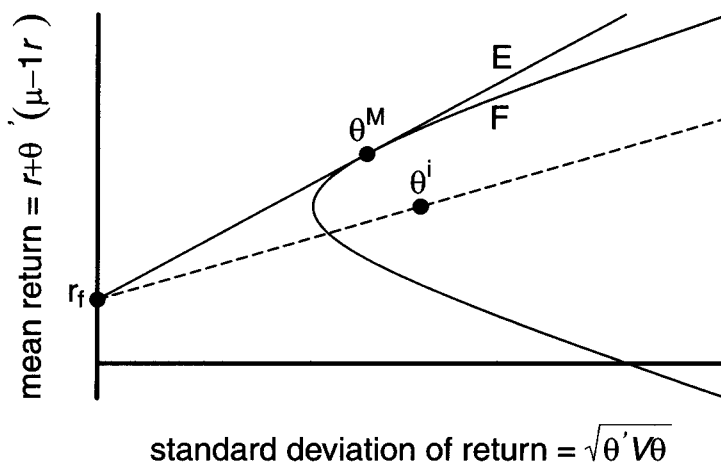
Fig. 1. The efficient frontier in means and standard deviations

down the higher return because of the higher risk. However, it is easy to construct random variables x_1 and x_2 with $x_1 > x_2$ that have means m_1 and m_2 and variances v_1 and v_2. In other words, a non-trivial mean-variance utility function (that does not simply maximize the mean) cannot always prefer more to less.

The two typical motivations of mean-variance preferences have different resolutions of this conundrum. Quadratic utility does not prefer more to less, so there is no inconsistency. This is not a nice feature of quadratic utility but it may not be a fatal problem either. Multivariate normality does not define preferences for all random variables, and in particular the random variables that generate the paradox are not available. When using any model, we need to think about whether the unrealistic features of the model are important for the application at hand.

Many important features of the CAPM are illustrated by Figures 1 and 2. In Figure 1, F is the *efficient frontier of risky asset returns* in means and standard deviations. Other feasible portfolios of risky assets will plot to the right of F and will not be chosen by any agent who can choose only among the risky assets and prefers less risk at a given mean. And, agents who choose higher mean given the standard deviation, will only choose risky portfolios on the upper branch of F, which is called the *positively efficient frontier of risky assets*. When the risk-free asset r_f is always available, all agents preferring a higher mean at a given standard deviation will choose a portfolio along the frontier E.[10] One important feature in either case is *two-fund separation*, namely, that the entire frontier F or E is spanned by two portfolios, which can be chosen to be any portfolios at two distinct points on the frontier. This is called a

---

[10] For agents who prefer less risk at a given mean but may not prefer a higher mean at a given level of risk, there is another branch of E below that is the reflection of its continuation to the left of the axis.

"mutual fund separation" result because we can *separate* the portfolio choice problem into two stages: first find two "mutual funds" (portfolios) spanning the efficient frontier (which can be chosen independently of preferences) and then find the mixture of the two funds appropriate for the particular preferences. For a typical agent who prefers more to less and prefers to avoid risk, preferences are increasing up and to the left in Figure 1. A more risk-averse agent will choose a portfolio on the lower left part of the frontier, with low return but low risk, and a less risk-averse agent will choose a portfolio on the upper right part of the frontier, accepting higher risk in exchange for higher return.

Figure 1 also illustrates the *Sharpe ratio* [Sharpe (1966)], which is used for performance measurement. The line through the riskless asset r_f and the market portfolio θ^M has a slope in Figure 1 that is larger than the slope for any inefficient portfolio such as θ^i. The slope of the line through a particular portfolio is the *Sharpe ratio* for the particular portfolio. The Sharpe ratio is largest for an efficient portfolio and the shortfall below that amount is the measure of inefficiency for any other portfolio. (An even greater Sharpe ratio would be possible if the efficient proxy is inefficient in sample or if we are considering a portfolio, say from an informed trading strategy, that is not a fixed portfolio of the assets.) In practice, due to random sampling error, even an efficient portfolio will have a measured Sharpe ratio that is not the largest value. When stock returns are Gaussian, there is an important connection between the measured Sharpe ratio of the market portfolio and the likelihood ratio test of the CAPM [Gibbons, Ross and Shanken (1989)].

Figure 2 shows the security market line, which quantifies the relation between risk and return in the CAPM. Risk is measured using the beta coefficient, which is the slope coefficient of a linear regression of the asset's return on the market's return. If the CAPM is true, all assets and portfolios will plot on the *Security Market Line* (SML) that goes through the risk-free asset r_f and the market portfolio of risky assets θ^M. In practice, measured asset returns are affected by random sampling error; if the CAPM is true it is entirely random whether a portfolio will plot above or below the security market line *ex post*. The use of beta as the appropriate measure of risk tells us that investors are rewarded for taking on market risk (correlated with market returns) not taking on idiosyncratic risk (uncorrelated with the market).

If the security market line tells us how much of a reward is justified for a given amount of risk, it makes intuitive sense that deviations from the security market line can be used to measure superior or inferior performance. This is the intuition behind the *Treynor Index* and *Jensen's alpha* [Treynor (1965) and Jensen (1969)]. For example, in Figure 2, Jensen's alpha for θ^s is $\alpha^s > 0$, indicating superior performance, and Jensen's alpha for θ^u is $\alpha^u < 0$, indicating underperformance. Unfortunately, any formal motivation for using Jensen's alpha must come from outside the CAPM, since if the CAPM is true then the expected value of Jensen's alpha is zero and the realized value is purely random. Theoretical models that incorporate superior performance from information-gathering have given mixed results on the value of using the security market line for measuring performance: a superior performer with
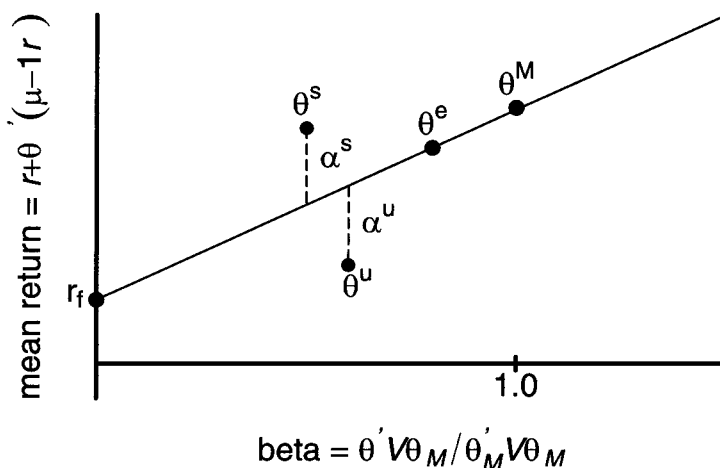
Fig. 2. The security market line connecting risk and return.

security-specific information will have a positive Jensen's alpha, but for market timing a superior performer may have a negative Jensen's alpha and may even plot inside the efficient frontier for static strategies [Mayers and Rice (1979) and Dybvig and Ross (1985)]. The Treynor Index is the slope of the line through the evaluated portfolio and the risk-free asset in the security market line diagram Figure 2. Performance is determined by comparing a portfolio's Treynor Index to that of the market; a larger Treynor Index indicates better performance. The Treynor index will indicate superior or inferior performance compared to the market the same as the Jensen measure. However, the ordering of superior or inferior performers can be different because the Treynor measure is adjusted for leverage.

The main results of the CAPM can be derived from the first-order condition (9). The first-order condition for agent n is

$$\mu - r\mathbf{1} = \lambda^n \Sigma \theta^n, \tag{19}$$

where $\lambda^n = 2V^{n\prime}_v(m, v)/V^{n\prime}_m(m, v)$, evaluated at the optimum $m = r + (\mu - r\mathbf{1})'\theta^n$ and $v = \theta^{n\prime} \Sigma \theta^n$. Now, the market portfolio is the wealth-weighted average of all agents' portfolios,

$$\theta^M = \frac{\sum_{n=1}^{N} w^n \theta^n}{\sum_{n=1}^{N} w^n}, \tag{20}$$

and consequently we have the wealth-weighted average of the first-order conditions

$$\mu - r\mathbf{1} = \lambda^M \Sigma \theta^M, \tag{21}$$

where

$$\lambda^M = \frac{\sum_{n=1}^{N} w^n \lambda^n}{\sum_{n=1}^{N} w^n}. \tag{22}$$

We can plug in the market portfolio to solve for λ^M and we obtain

$$\mu - r\mathbf{1} = \frac{\Sigma\theta^M}{\theta^{M\prime}\Sigma\theta^M}(\mu^M - r), \tag{23}$$

where $\mu^M \equiv \theta^{M\prime}\mu$ is the mean return on the market portfolio of risky assets.

Applying Equation (23) to obtain the expected excess return of a portfolio θ of risky assets (with $\theta'\mathbf{1} = 1$ since a portfolio of risky assets does not include any holdings of the risk-free asset), we have that [11]

$$\begin{aligned}
\theta'\mu - r &= \lambda^M \theta'\Sigma\theta^M \\
&= \beta^\theta(\mu^M - r),
\end{aligned} \tag{24}$$

where β^θ is the portfolio's beta, which is the slope coefficient of a regression of the returns of the portfolio θ's return on the market return,

$$\beta^\theta \equiv \frac{\theta'\Sigma\theta^M}{\theta^{M\prime}\Sigma\theta^M}. \tag{25}$$

The SML equation we plotted in Figure 2 is Equation (24).

For a portfolio θ, Jensen's alpha is given by

$$\theta'\mu - r - \beta^\theta(\mu^M - r), \tag{26}$$

its Treynor index is

$$\frac{\theta'\mu - r}{\beta^\theta}, \tag{27}$$

and its Sharpe ratio is

$$\frac{\theta'\mu - r}{\sqrt{\theta'\Sigma\theta}}. \tag{28}$$

The portfolios encountered in practice are actively managed and the formulas for these performance measures would be more complex than for the simple fixed mix

---

[11] We looked at the simpler case in the text, but the same pricing result holds for a portfolio including a holding in the risky asset. In this case, the expected return on the portfolio is $\theta'\mu + (1 - \theta'\mathbf{1})r$ and the expected excess return is $\theta'\mu + (1 - \theta'\mathbf{1})r - r = \theta'\mu - \theta'\mathbf{1}r$.

of assets θ. However, the concepts are unchanged with the natural adaptations, e.g., replacing $\theta'\mu$ by the sample mean return on the portfolio and replacing $\beta^\theta = \theta'\Sigma\theta^M/\theta^{M'}\Sigma\theta^M$ by the estimated slope from the regression of the portfolio return on the market return.

6. Mutual fund separation theory

The general portfolio problem for arbitrary preferences and distributions is sufficiently rich to allow for nearly any sort of qualitative behavior [see Hart (1975) for negative results or Cass and Stiglitz (1972) for positive results in special cases]. In an effort to simplify this problem and obtain results that allow for aggregation so that the general behavior of the market can be understood in terms of the primitive properties of risk aversion and of the underlying distributions a collection of results known as separation results have been developed.

Mutual Fund Separation is the separation of portfolio choice into two stages. The first stage is the selection of small set of "mutual funds" (portfolios) among which choice is to be made, and the second stage is the selection of an allocation to the mutual funds. We have "k-fund separation" for a particular class of distributions and a particular class of utility functions if for each joint return distribution in the class there exist k funds that can be used in the two-step procedure while making agents with utility in the class and any wealth level just as well off as the choosing in the whole market. The important restriction is that the choice of the funds is done once for the entire class of utility functions. In the literature, there are two general approaches: one approach [Hakansson (1969) and Cass and Stiglitz (1970)] restricts utility functions and has relatively unrestricted distributions, while the other approach [Ross (1978a)] restricts distributions and has relatively unrestricted utility functions. Either approach is useful for deriving asset pricing results because, for example, if individual investors hold mixtures of two funds, then the market portfolio must be a mixture of the same two funds.

6.1. Preference approach

The preference approach focuses on classes of special utility functions. Many of the results involve utility functions that have properties of homotheticity or invariance. It is important that we require the same funds to work for each utility function at all wealth levels, since this avoids "accidental" cases such as a set containing any two utility functions over returns. Analysis in this section will use Problem 3, in some cases adding the assumption that one of the assets is riskless.

First, we consider one-fund separation, which requires all portfolio choices to lie in a ray. Given the budget constraint, this implies that the portfolio choice is just proportional to wealth. For this to happen at all prices, the preferences have to be

homothetic. And, given the von Neumann–Morgenstern restriction, this is equivalent to either logarithmic utility, $u(c) = \log(c)$, or power utility, $u(c) = c^{1-R}/(1-R)$.

Theorem 5: One-fund separation from preferences. *The following are equivalent properties of a nonempty class U of utility functions:*

(1) For each joint distribution of security returns there exists a single portfolio θ, such that every $u \in U$ is just as well off choosing a multiple of θ as choosing from the entire market.

(2) The class U consists of a single utility function (up to an affine transform that leaves preferences unchanged) of the form $u(c) = \log(c)$ or $u(c) = c^{1-R}/(1-R)$.

Proof: $(2) \Rightarrow (1)$ Let u be the single utility function in U. The objective function in terms of portfolio proportions is $E[u(w\theta'r)]$. In the log case, this is $E[\log(w\theta'r)] = \log(w) + E[\log(\theta'r)]$, and maximizing the objective is the same as maximizing the second term which does not depend on w. In the power case, the objective is $E[(w\theta'r)^{1-R}/(1-R)] = w^{1-R}E[(\theta'r)^{1-R}/(1-R)]$, and maximizing the objective is the same as maximizing the second factor which does not depend on w. In either case, choosing the proportions that work at one wealth level gives a portfolio in proportions that will be optimal at all wealth levels.

$(1) \Rightarrow (2)$ Suppose u is an element of the class U. Then, the first-order condition for an optimum implies

$$E[(1 + r - \gamma)\, u'(W\theta'(1 + r))] = 0, \tag{29}$$

where $\gamma = \lambda/E[u'(W\theta'(1 + r))]$. In general, ρ must satisfy Equation (8) and may vary with W, but for complete markets ρ is uniquely determined by Equation (8) and may be taken as given. For the same portfolio weights θ to be optimal for all W, it follows that the derivative of the first-order condition is zero and for complete markets we have

$$E[(1 + r - \gamma)\, \theta'(1 + r)\, u''(W\theta'(1 + r))] = 0. \tag{30}$$

Now, one-fund separation implies that in all complete markets Equation (29) implies Equation (30), but the only way this can always be true is if everywhere

$$cu''(c) = -Ru'(c), \tag{31}$$

where $R = 1$ implies logarithmic utility and any other $R \geqslant 0$ implies power utility. ($R \geqslant 0$ corresponds to a convex utility function.) And all utility functions in the class U must correspond to the same preferences or else it is easy to construct a 2-state counterexample. $\square$

The utility functions in the theorem comprise the *Constant Relative Risk Aversion* (CRRA) class for which the Arrow–Pratt coefficient of relative risk aversion

$-cu''(c)/u'(c)$ is a constant. [See Arrow (1965) and Pratt (1964, 1976).[12]] Other special utility functions lead to two-fund separation if there is a riskless asset. The *Constant Absolute Risk Aversion* (CARA) class of utility functions of the form $u(c) = -\exp(-Ac)/A$ for which the Arrow–Pratt coefficient of absolute risk aversion $-u''(c)/u'(c)$ is constant leads to a special two-fund separation result in which the risky portfolio holding is constant and only the investment in the risk-free asset is changing as wealth changes. When there is a riskless asset, there is also two-fund separation in the larger *Linear Risk Tolerance* (LRT) class which encompasses the other two classes as well as wealth-translated relative risk aversion preferences of the form $u(c) = \log(c - c_0)$ or $u(c) = (c - c_0)^{1-R}/(1 - R)$. The linear risk tolerance class is defined by the risk tolerance $-u'(c)/u''(c)$ having the linear form $a(c - c_0)$. We can include in this class the satiated utility functions of the form $-(c - c_0)^{1-R}/(1 - R)$ defined for $c \leq c_0$ (and is typically extended to $c > c_0$ in the obvious way in the quadratic case $R = -1$). With quadratic utility, we have a special result of two-fund separation even without a risk-free asset due to linearity of marginal utility. In these results, all utility functions in the class U must have the same power (or absolute risk aversion coefficient for exponential utility) but can have different translates c_0 (but exponential utility is unchanged under translation). For details and proofs, see Cass and Stiglitz (1970).

6.2. Beliefs

We have already seen one case of separation based on beliefs, which is in mean-variance analysis motivated by multivariate normality, as discussed in the section on the CAPM. Mean-variance preferences can also be derived from more general transformed spherically distributed preferences discussed by Chamberlain (1983).[13] We turn now to a strictly more general class, the separating distributions of Ross (1978a).

The central intuition behind the separating distributions is that risk-averse agents will not choose to take on risk without any reward. This is the same intuition as in mean-variance analysis, but it is somewhat more subtle because risk can no longer be characterized by variance for general concave von Neumann–Morgenstern preferences. The appropriate definition of risk is related to Jensen's inequality, which says that for any convex function $f(\cdot)$ and any random variable x, $E[f(x)] \geq f(E[x])$, with strict equality if $f(\cdot)$ is strictly concave and x is not (almost surely) constant. A risk-averse von Neumann–Morgenstern utility function $u(\cdot)$ is concave (so that $-u(\cdot)$ is convex), and consequently for any random consumption c, $E[u(c)] \leq u(E[c])$, with strict inequality for strictly concave u and nonconstant c. More importantly for portfolio choice problems, we can use Jensen's inequality and the law of iterated expectations

[12] Some other ways of comparing risk aversion are given by Kihlström, Romer and Williams (1981) and Ross (1981).

[13] Another special case of one-fund separation is the symmetric case of Samuelson (1967).

to conclude that adding conditional-mean-zero noise makes a risk-averse agent worse off. That gives us the following useful result:

Lemma 1. *If $E[\varepsilon|c] = 0$ and u is concave, then*

$$E[u(c + \varepsilon)] \leqslant E[u(c)]. \tag{32}$$

Proof:

$$\begin{aligned} E[u(c + \varepsilon)] &= E[E[u(c + \varepsilon)|c]], \\ &\leqslant E[u(E[c + \varepsilon|c])], \tag{33} \\ &= E[u(c)], \end{aligned}$$

by Jensen's inequality and the law of iterated expectations. $\square$

In fact, it can be shown that one random variable is dominated by another with the same mean for all concave utility functions if and only if the first has the same distribution as the second plus conditional-mean-zero noise. This is one of the results of the theory of *Stochastic Dominance*, which was pioneered by Quirk and Saposnik (1962) and Hadar and Russell (1969) and was popularized by Rothschild and Stiglitz (1970, 1971).

The basic idea behind the separating distributions is that there are k funds (e.g., 2 funds for 2-fund separation) such that everything else is equal to some portfolio of the k funds, plus conditional-mean-zero noise. Formally, we have

Theorem 6. *Consider a world with k funds that are portfolios with weights $y^1, \ldots, y^k$ summing $(\forall j)\mathbf{1}'y^j = 1$ (or in vector notation, $\mathbf{1}'y = \mathbf{1}'$).*[14] *Further assume that returns on each asset i can be written as*

$$r_i = \sum_{j=1}^{k} \beta_{ij} y^{j'} r + \varepsilon_i, \tag{34}$$

(i.e., $r = \beta'y'r + \varepsilon$), where $\sum_{j=1}^{k} \beta_{ij} = 1$ (i.e., $\mathbf{1}\beta = \mathbf{1}$) and for all linear combination η of the fund returns, ε is conditional-mean-zero noise:

$$E[\varepsilon|\eta'\beta r] = \mathbf{0}. \tag{35}$$

Then any agent with increasing and concave von Neumann–Morgenstern preferences will be just as happy choosing a portfolio of the k funds, as choosing from the

---

[14] As discussed earlier, the dimension of $\mathbf{1}$ is determined by context; in $\mathbf{1}'y = \mathbf{1}'$, the first occurrence of $\mathbf{1}$ is $k \times 1$ and the second occurrence of $\mathbf{1}$ is $n \times 1$.

entire market. More formally, for each monotone and concave u and for each feasible portfolio θ with $1'\theta = 1$, there exists another portfolio η with $1'\eta$ such that $Eu(\eta'r) \geqslant Eu(\theta'r)$.

Proof: Consider any portfolio θ with $1'\theta = 1$. Then

$$
\begin{aligned}
\theta'r &= \theta'(\beta y'r + \varepsilon), \\
&= \theta'\beta y'r + \theta'\varepsilon.
\end{aligned}
\tag{36}
$$

But $y\beta'\theta$ is a valid portfolio because $1'y\beta'\theta = 1'\beta'\theta = 1'\theta = 1$. And, the second term is conditional-mean-zero noise. Therefore, by Lemma 1 all agents with concave preferences would be at least as happy to switch from θ to the portfolio $y\beta'\theta$, which is a portfolio of the k funds (with weights $\beta'\theta$). $\square$

In the case of 1- and 2-fund separating distributions, the characterization is necessary as well as sufficient [see Ross (1978a)].

In the CAPM derived using multivariate normality, it is easy to show that the SML implies that each mean-variance inefficient portfolio has a payoff equal to the payoff of the efficient portfolio with the same mean plus conditional-mean-zero noise. Given that the mean-variance frontier is spanned by two portfolios, we see that the CAPM with multivariate normality is indeed in the class of 2-fund separating distributions.

7. Arbitrage pricing theory (APT)

The Arbitrage Pricing Theory, which was introduced in Ross (1976a,c), is a model of security pricing that generalizes the pricing relation in the CAPM and also builds on the intuition of the separating distributions. First, we start with a *factor model* of returns of the sort studied in statistics:

$$
r = \mu + f\beta + \varepsilon,
\tag{37}
$$

where μ is a vector of mean returns (unrestricted at the moment but to be restricted by the theory), f is a vector of factor returns, of dimensionality much less than r, β is a matrix of factor loadings, and ε is a vector of uncorrelated idiosyncratic noise terms. We can represent the restriction to the factor model by writing the covariance matrix as

$$
\text{var}(r) = \beta\beta' + D,
\tag{38}
$$

where we have assumed an orthonormal set of factors with the identity matrix as covariance matrix (without loss of generality because we can always work with a linear transformation), and where D is a diagonal matrix which is the covariance matrix of the vector of security-specific noise terms ε. The factor model is a useful restriction for

empirical work on security returns: given that typically we have many securities for the number of time periods, the full covariance matrix is not identified but a sufficiently low dimensional factor model has many fewer parameters and can be estimated.

One intuition of the APT is that idiosyncratic risk is not very important economically and should not be priced. Another intuition of the APT is that compensation for risk should be linear or else there will be arbitrage. For example, if there is a single factor and two assets have different exposures to the factor (betas), excess return must be proportional to the risk exposure. Suppose the compensation per unit risk is larger for the asset with a larger risk exposure. Then a portfolio mixture of the risk-free asset and the high-risk asset will have the same risk exposure as the low-risk asset but a higher expected return, and combining a long position in the mixture with a short position in the low-risk asset gives a pure profit. This profit will be riskless in the absence of idiosyncratic risk; it will be profitable for some agents if idiosyncratic risk is diversifiable. Conversely, if the compensation per unit risk is larger for the asset with the lower risk exposure, the other asset can be dominated by a combination of a long position in the less risky asset with a short (borrowing) position in the risk-free asset.

The main consequence of the APT is a pricing equation that looks like a multifactor version of the CAPM equation:

$$\mu = r_f \mathbf{1} + \Gamma\beta. \tag{39}$$

Here, r_f is the risk-free rate and Γ is the vector of factor risk premia. There are several approaches to motivating this APT pricing equation; see for example Ross (1976a,c), Dybvig (1983), or Grinblatt and Titman (1983).

The APT shares important features of the CAPM: the value of diversification, compensation for taking on systematic risk, and no compensation for taking on idiosyncratic risk. The main difference is that there may be multiple factors, and that the priced factors are the common factors that appear in many securities and not necessarily just the market factor.

8. Conclusion

On reflection it is surprising that even our simplest context of a single-period neoclassical model of investments has such a rich theoretical development. We have hit on many of the highlights but even so we cannot claim to an exhaustive review of all that is known.

References

Abel, A.B. (1990), "Asset prices under habit formation and catching up with the Joneses", American Economic Review 80:38–42.

Arrow, K., and G. Debreu (1954), "Existence of an equilibrium for a competitive economy", Econometrica 22:265–290.

Arrow, K., and F. Hahn (1971), General Competitive Analysis (Holden-Day, San Francisco).

Arrow, K.J. (1964), "The role of securities in the optimal allocation of risk-bearing", Review of Economic Studies 31:91–96.

Arrow, K.J. (1965), "Aspects of the theory of risk-bearing", Yrjo Jahnsson Lectures (Yrjö Jahnssonin Säätiö, Helsinki).

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

Barberis, N., M. Huang and T. Santos (2001), "Prospect theory and asset prices", Quarterly Journal of Economics 116:1–53.

Bergman, Y. (1985), "Time preference and capital asset pricing models", Journal of Financial Economics 14:145–159.

Bewley, T. (1988), "Knightian uncertainty", Nancy Schwartz Lecture (Northwestern MEDS, Evanston).

Black, F., and M. Scholes (1973), "The pricing of options and corporate liabilities", Journal of Political Economy 81:637–654.

Blume, L., A. Brandenburger and E. Dekel (1991), "Lexicographic probabilities and choice under uncertainty", Econometrica 59:61–79.

Breeden, D.T. (1979), "An intertemporal asset pricing model with stochastic consumption and investment opportunities", Journal of Financial Economics 7:265–296.

Breeden, D.T., and R.H. Litzenberger (1978), "Prices of state-contingent claims implicit in option prices", Journal of Business 51:621–651.

Cass, D., and J.E. Stiglitz (1970), "The structure of investor preferences and separability in portfolio allocation: a contribution to the pure theory of mutual funds", Journal of Economoic Theory 2: 122–160.

Cass, D., and J.E. Stiglitz (1972), "Risk aversion and wealth effects on portfolios with many assets", Review of Economic Studies 39:331–354.

Chamberlain, G. (1983), "A characterization of the distributions that imply mean-variance utility functions", Journal of Economic Theory 29:185–201.

Constantinides, G. (1991), "Habit formation: a resolution of the equity premium puzzle", Journal of Political Economy 98:519–543.

Cox, J.C., and S.A. Ross (1975), "A survey of some new results in financial option pricing theory", Journal of Finance 31:383–402.

Cox, J.C., S.A. Ross and M. Rubinstein (1979), "Option pricing: a simplified approach", Journal of Financial Economics 7:229–263.

Debreu, G. (1959), Theory of Value: An Axiomatic Analysis of Economic Equilibrium (Yale University Press, New Haven).

Duesenberry, J.S. (1949), Income, Saving, and the Theory of Consumer Behavior (Harvard University Press, Cambridge, MA).

Dybvig, P.H. (1983), "An explicit bound on individual assets' deviations from APT pricing in a finite economy", Journal of Financial Economics 12:483–496.

Dybvig, P.H. (1988a), "Distributional analysis of portfolio choice", Journal of Business 61:369–393.

Dybvig, P.H. (1988b), "Inefficient dynamic portfolio strategies, or how to throw away a million dollars in the stock market", Review of Financial Studies 1:67–88.

Dybvig, P.H. (1995), "Duesenberry's ratcheting of consumption: optimal dynamic consumption and investment given intolerance for any decline in standard of living", Review of Economic Studies 62:287–313.

Dybvig, P.H., and S.A. Ross (1985), "Differential information and performance measurement using a security market line", Journal of Finance 40:383–399.

Dybvig, P.H., and S.A. Ross (1987), "Arbitrage", in: J. Eatwell, M. Milgate and P. Neuman, eds., The New Palgrave: a Dictionary of Economics (Macmillan, London) pp. 100–106.

Epstein, L.G., and S.E. Zin (1989), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework", Econometrica 57:937–969.

Fishburn, P. (1988), Nonlinear Preference and Utility Theory (John Hopkins, Baltimore).

Gibbons, M.R., S.A. Ross and J. Shanken (1989), "A test of the efficiency of a given portfolio", Econometrica 57:1121–1152.

Grinblatt, M., and S. Titman (1983), "Factor pricing in a finite economy", Journal of Financial Economics 12:497–507.

Hadar, J., and W.R. Russell (1969), "Rules for ordering uncertain prospects", American Economic Review 59:25–34.

Hakansson, N.H. (1969), "Risk disposition and the separation property in portfolio selection", Journal of Financial and Quantitative Analysis 4:401–416.

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

Harrison, J.M., and S. Pliska (1981), "Martingales and stochastic integrals in the theory of continuous trading", Stochastic Processes and Their Applications 11:215–260.

Hart, O.D. (1975), "Some negative results on the existence of comparative statics results in portfolio theory", The Review of Economic Studies 42:615–621.

Herstein, I.N., and J. Milnor (1953), "An axiomatic approach to measurable utility", Econometrica 21:291–297.

Hindy, A., and C. Huang (1992), "On intertemporal preferences for uncertain consumption: a continuous time approach", Econometrica 60:781–801.

Jensen, M.C. (1969), "Risk, the pricing of capital assets, and the evaluation of investment portfolios", Journal of Business 42:167–247.

Jouini, E., and H. Kallal (2001), "Efficient trading strategies in the presence of market frictions", Review of Financial Studies 14:343–369.

Kahneman, D., and A. Tversky (1979), "Prospect theory: an amalysis of decision under risk", Econometrica 47:263–292.

Karlin, S. (1959), Mathematical Methods and Theory in Games, Programming, and Economics (Addison-Wesley, Reading, MA).

Kihlström, R.E., D. Romer and S. Williams (1981), "Risk aversion with random initial wealth", Econometrica 49:911–920.

Knight, F.H. (1921), Risk, Uncertainty, and Profit (Houghton Mifflin, New York).

Kreps, D., and E. Porteus (1978), "Temporal resolution of uncertainty and dynamic choice theory", Econometrica 46:185–200.

Lintner, J. (1965), "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", Review of Economics and Statistics 47:13–37.

Luce, R.D., and H. Raiffa (1957), Games and Decisions (Wiley, New York).

Machina, M.J. (1982), " 'Expected Utility' Analysis without the Independence Axiom", Econometrica 50:277–324.

Markowitz, H. (1952), "Portfolio selection", Journal of Finance 7:77–91.

Markowitz, H. (1959), Portfolio Selection, Efficient Diversification of Investments (Wiley, New York).

Mayers, D., and E.M. Rice (1979), "Measuring portfolio performance and the empirical content of asset pricing models", Journal of Financial Economics 7:3–28.

Merton, R.C. (1971), "Optimal consumption and portfolio rules in a continuous-time model", Journal of Economic Theory 3:373–413.

Merton, R.C. (1973), "An intertemporal capital asset pricing model", Econometrica 41:867–887.

Pollak, R.A. (1970), "Habit formation and dynamic demand functions", Journal of Political Economy 78:745–763.

Pratt, J.W. (1964), "Risk aversion in the small and the large", Econometrica 32:122–136.

Pratt, J.W. (1976), "Risk aversion in the small and the large (erratum)", Econometrica 55:420.

Quirk, J.P., and R. Saposnik (1962), "Admissibility and measurable utility functions", Review of Economic Studies 29:140–146.

Ross, S.A. (1976a), "The arbitrage theory of capital asset pricing", Journal of Economic Theory 13:341–360.

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

Ross, S.A. (1976c), "Return, risk, and arbitrage", in: I. Friend and J. Bicksler, eds., Risk and Return in Finance 1 (Ballinger, Cambridge, MA) pp. 189–218.

Ross, S.A. (1978a), "Mutual fund separation in financial theory – the separating distributions", Journal of Economic Theory 17:254–286.

Ross, S.A. (1978b), "A simple approach to the valuation of risky streams", Journal of Business 51: 453–475.

Ross, S.A. (1981), "Some stronger measures of risk aversion in the small and the large", Econometrica 49:621–638.

Rothschild, M., and J.E. Stiglitz (1970), "Increasing risk: I. A definition", Journal of Economic Theory 2:225–243. See also "Addendum to 'Increasing risk: I. A definition' ", Journal of Economic Theory 5:306.

Rothschild, M., and J.E. Stiglitz (1971), "Increasing risk: II. Its economic consequences", Journal of Economic Theory 3:66–84.

Rubinstein, M. (1976), "The valuation of uncertain income streams and the pricing of options", Bell Journal of Economics 7:407–425.

Samuelson, P.A. (1967), "General proof that diversification pays", Journal of Financial and Quantitative Analysis 2:1–13.

Sharpe, W.F. (1964), "Capital asset prices: a theory of market equilibrium under conditions of risk", Journal of Finance 19:425–442.

Sharpe, W.F. (1966), "Mutual fund performance", Journal of Business 39:119–138.

Tobin, J. (1958), "Liquidity preference as behavior towards risk", Review of Economic Studies 25:65–86.

Treynor, J. (1965), "How to rate management of investment funds", Harvard Business Review 43:63–75.

von Neumann, J., and O. Morgenstern (1944), Theory of Games and Economic Behavior (Princeton University Press, Princeton, NJ).