http://www.jstor.org

THE ROLE OF CONDITIONING INFORMATION IN DEDUCING TESTABLE RESTRICTIONS IMPLIED BY DYNAMIC ASSET PRICING MODELS[1]

By Lars Peter Hansen and Scott F. Richard

The purpose of this paper is to investigate testable implications of equilibrium asset pricing models. We derive a general representation for asset prices that displays the role of conditioning information. This representation is then used to examine restrictions implied by asset pricing models on the unconditional moments of asset payoffs and prices. In particular, we analyze the effect of information omission on the mean-variance frontier of one-period returns on portfolios of securities. Also, we deduce an information extension of equilibrium pricing functions that is useful in deriving restrictions on the unconditional moments of payoffs and prices.

KEYWORDS: Asset pricing, conditioning information, mean-variance analysis, Hilbert spaces.

INTRODUCTION

IN A COMPETITIVE EQUILIBRIUM MODEL of asset markets, prices are found endogenously as a consequence of the aggregation of the decisions of economic agents. Equilibrium prices are determined by a pricing function that maps uncertain payoffs in the future into prices today. Alternative models of asset prices imply alternative pricing functions. Two models that imply the same pricing function are observationally indistinguishable using payoff and price data from asset markets. Hence we can index observationally equivalent classes of asset pricing models by their implied pricing functions. Data from sources other than asset markets are required to discriminate among models within an equivalence class.

To understand the contribution of this paper, it is convenient to think of the analysis of asset pricing models as proceeding in two steps. The first step is to derive alternative pricing functions from more primitive assumptions on the underlying economic environment, e.g., preferences, endowments, and the technological opportunities for production. The second step is to deduce the restrictions that these alternative pricing functions imply for the population moments of time series data on asset payoffs and prices. While both steps are important, our paper contributes to the second step of this analysis.

Conditioning information is a crucial ingredient in our analysis of pricing functions. In most intertemporal models, information accumulates over time and traders in asset markets form portfolios contingent on information available at the time trades are made. The accumulated information becomes imbedded in

asset prices that clear the competitive markets. (For example, see the equilibrium models of Brock (1980), Cox, Ingersoll, and Ross (1985), Lucas (1978), or Richard and Sundaresan (1981).) Consequently, a pricing function maps payoffs modeled as random variables into prices that are also modeled as random variables, but are constrained to be in the information set of traders at the time prices are quoted. From the vantage point of an outside observer or econometrician there is no useful sense in which payoffs are random and prices are not. Instead each of these classes of random variables are presumed to satisfy different informational constraints. A distinctive feature of our analysis is that the range of the pricing function is the collection of random variables in the information set of traders at the time the portfolio decisions are made.

In this paper we consider implications of asset pricing functions for population moments of asset payoffs and prices. Although theoretical models assume traders make calculations conditioned on information available at the time of trading, it is often convenient for an econometrician to deduce empirical restrictions that do not depend on this conditioning information. For this reason, we consider the restrictions on population or unconditional moments implied by theoretical models of asset prices. Such moments are limits of time series averages of observable data for many data generation processes.

In Section 1 we give a motivational example of a class of data generation processes for portfolio payoffs and prices which guarantee that time series averages converge (almost surely) to population moments. These data generation processes may be consistent with a variety of asset pricing models. On the other hand, there are other interesting data generation processes to which our analysis applies.

The formal analysis in this paper proceeds as follows. In Section 2 we presume the existence of pricing functions that satisfy two important properties. The first property is value-additivity, i.e., the price of a portfolio is the portfolio of the prices. The second property is continuity, i.e., if a sequence of payoffs goes to zero, then the sequence of its prices goes to zero. Other authors have analyzed asset pricing functions that satisfy these two properties. See Rubinstein (1976), Ross (1977, 1978), Harrison and Kreps (1979), and Chamberlain and Rothschild (1983). The main contribution of this section is to show that alternative asset pricing functions that embody conditioning information can be represented using alternative random variables among the collection of payoffs from portfolios. This result follows from a conditional counterpart to the Riesz representation of a linear functional on a Hilbert space. We obtain this representation by extending conventional Hilbert space analysis to accommodate conditioning information. This extension is required for the analysis in Sections 3 and 4 and for other applications as well (Eichenbaum, Hansen, and Richard (1984)).

In Section 3 we discuss the role of conditioning information in determining the mean-variance implications of asset pricing models. Chamberlain and Rothschild (1983) showed that the mean-variance frontier implied by a pricing function can be characterized in terms of two payoffs, one of which is the payoff used in representing the pricing function. The first step in our analysis shows that the

same result is true when conditioning information is introduced. We then take the analysis one step further and characterize the mean-variance frontier calculated without using conditioning information. In particular, we display the relation between the conditional and unconditional mean-variance frontiers. This relation is important because standard approaches to testing mean-variance implications examine unconditional means, variances, and covariances of returns.

All of the analysis in Section 3 uses pricing functions that map into the space of random variables in the conditioning information set of traders. The analysis in Section 4 proceeds along a somewhat different vein. Pricing functions are used to construct linear functionals defined on a standard Hilbert space of portfolio payoffs. These linear functionals map into the space of real numbers and are used to deduce testable restrictions in terms of unconditional moments. The empirical analyses conducted by Hansen and Singleton (1982) and Brown and Gibbons (1985) can be viewed as using this approach.

Finally, Section 5 contains our conclusions and an Appendix is included that contains proofs of the lemmas and theorems presented in the text.

1. DATA GENERATION

In this section we describe an important class of data generation processes to which our theoretical analysis applies. We include this description to help in relating our theoretical analysis to empirical analyses of asset market data that use summary statistics calculated by taking time series averages. The analysis in later sections applies to many data generation processes that are not in the class described here. Consequently, the discussion in this section is meant to be illustrative rather than exhaustive.

Let (Ω, F, Pr) define a probability space where Ω is a set of sample points, F is a sigma algebra of subsets of Ω, and Pr is a probability measure. It is convenient to introduce a deterministic law of motion governing the evolution of states of the world over time. Let S be a measurable, measure-preserving transformation mapping Ω into itself. This transformation defines the temporal evolution of states over time in the sense that if ω is the state at time zero, then $S^t(\omega)$ is the state of the world at time t, where S^t is defined to mean the transformation S applied t times in succession.

Even though the law of motion is deterministic, the true state will not be observed directly. For instance, suppose x is a measurable function (random vector) mapping Ω into k-dimensional Euclidean space that determines a vector of observations as a function of the underlying state. If ω is the state of the world at time zero, then $x[S^t(\omega)]$ is the corresponding observation vector at time t. Hence, we define a vector stochastic process $\{x_t: t = 1, 2, \ldots\}$ via

$$(1.1) \quad x_t(\omega) = x[S^t(\omega)],$$

which determines a sequence of time series observations for each initial state of the world. In general, the observation vector $x(\omega)$ will not reveal ω. Thus even

though we have a deterministic law of motion governing the state of the world, x_t will not in general be perfectly forecastable given $x_{t-1}, x_{t-2}, \ldots, x_1$.

Since S is assumed to be measure-preserving, for any random vector x the stochastic process constructed via (1.1) is strictly stationary. Such stochastic processes satisfy a law of large numbers in the sense that $\{(1/T) \sum_{t=1}^{T} x_t : T = 1, 2, \ldots\}$ converges almost surely as long as x has a finite first moment. In general the limit random vector is $E(x | F^*)$ where F^* is the sigma algebra of invariant sets of the transformation S (Brieman (1968, p. 113) or Doob (1953, p. 465)). When these invariant sets all have probability measure zero or one, we say that S is ergodic. In this case, $E(x | F^*) = E(x)$ so that time series averages converge almost surely to the average taken across states of the world using the measure Pr. For simplicity we assume that S is ergodic in our analysis (although much of our analysis also applies more generally with expectations conditioned on F^* replacing unconditional expectations).

In an environment such as this, an econometrician can learn about moments of random vectors by calculating time series averages. For the purpose of this paper, we will adopt the simplifying assumption that for any given state ω, the entire infinite sequence of observations $\{x_t(\omega) : t = 1, 2, \ldots\}$ is available to the econometrician. Thus we abstract from some important issues pertaining to statistical inference.

As is standard in rational expectations models, we presume that economic agents in this environment make decisions at every date knowing the true probability distribution over states of the world. In general, their information will be limited so that the true state of the world is not revealed to them at any point in time. To define the common set of information available to all consumers at time t, let G denote a subsigma algebra of F. We interpret G as the information at time zero. Define

(1.2) $G_t = \{A_t : A_t = S^{-t}(A) \text{ for some } A \text{ in } G\}$,   for $t = 1, 2, \ldots$.

Then G_t is the sigma algebra defining the information available to economic agents at date t. We assume that G_1 contains G so that $\{G_t : t = 1, 2, \ldots\}$ is nondecreasing. Finally, for each $t = 1, 2, \ldots$, let I_t denote the set of all random variables that are measurable with respect to G_t and let I denote the set of all random variables that are measurable with respect to G.

At any date t economic agents make portfolio decisions based on their current information. Asset prices are assumed to be measurable with respect to G_t and hence observed by all consumers. A one-period security purchased at time t has a payoff at time $t+1$. The payoff for this security is assumed to be in the information set of economic agents at the date of payoff ($t+1$). Let p denote a random variable that is in I_1 and is used to define a sequence of such payoffs. The corresponding payoff at time $t+1$ is given by

(1.3) $p_{t+1}(\omega) = p[S^t(\omega)]$.

Let $\pi(p)$ denote the time zero price of the payoff p at time one so that $\pi(p)$ is a random variable in I. We assume that for any payoff sequence that is generated

via (1.3) for some p (which may be limited to an admissible set of payoffs),

$$(1.4) \qquad \pi_t(p_{t+1})(\omega) = \pi(p)[S^t(\omega)],$$

where $\pi_t(p_{t+1})$ is the time t price of p_{t+1}.

In this paper we will study the pricing of one-period securities in the initial time period zero. As long as payoff sequences are generated via (1.3) and the prices of those payoffs evolve via (1.4), an initial period analysis will extend to subsequent time periods via the transformation S. Furthermore, any payoff sequence $\{p_{t+1}: t = 1, 2, \ldots\}$ or price sequence $\{\pi_t(p_{t+1}): t = 1, 2, \ldots\}$ will be a strictly stationary and ergodic stochastic process whose moments can be calculated by taking time series averages. Hence, it will be convenient to deduce implications of asset pricing models in terms of unconditional moments. Although our analysis focuses on one-period securities, it could be extended easily to accommodate multi-period securities.

2. PRELIMINARY MATHEMATICAL ANALYSIS

In this section we develop the mathematics that underlies our analysis in Sections 3 and 4. These same tools turn out to be useful in analyzing other problems as well. We will use notation that is compatible with that used in Section 1. The transformation S, however, will not play a role in the analysis in this section.

Harrison and Kreps (1979) use Hilbert space methods to represent pricing functions and Chamberlain (1983) and Chamberlain and Rothschild (1983) deduce mean-variance implications for returns in a Hilbert space setting. We follow their lead except that we make one extension. For these other authors, the range of the pricing function at the date of initial trading, analogous to π (introduced in Section 1) is the set of real numbers. As part of their analysis of continuous time diffusion models, Harrison and Kreps show that prices can be represented as conditional expectations after the initial trading date when new information has been revealed. We show this conditional expectations representation applies to more general stochastic processes and we allow conditioning at the initial trading date. Hence the range of π in our analysis is the set of random variables, I, that are measurable with respect to a subsigma algebra, G, of F.

The purpose of this section of the paper is to show how to recast Hilbert spaces methods to take account of conditioning information. In particular, we will represent π as a conditional analogue to a linear functional mapping random variables (payoffs on portfolios of assets) that are measurable with respect to a subsigma algebra, G_1, of F into prices that are measurable with respect to G. As in Section 1, we assume that G_1 contains G.

First, we introduce some notation. In Loeve (1978) it is shown that for any random variable p, $E(p^2|G)$ is always well-defined almost surely although it may equal infinity with positive probability. Let

$$(2.1) \qquad P^+ = \{p \text{ in } I_1: E(p^2|G) < \infty\},$$

where I_1 is the set of random variable that are measurable with respect to G_1.[2] The set P^+ satisfies several nice properties that will be described subsequently. To describe these properties, it is convenient to introduce a conditional counterpart to an inner product on P^+. For any p_1 and p_2 in P^+, let

$$(2.2) \qquad \langle p_1 | p_2 \rangle_G = E(p_1 p_2 | G).$$

A conditional version of the Cauchy–Schwarz Inequality guarantees that the right-hand side of (2.2) is well-defined and finite.[3] In accordance with (2.2), we define a conditional norm to be

$$(2.3) \qquad \|p\|_G = [\langle p | p \rangle_G]^{1/2}.$$

Both the conditional inner product $\langle \cdot | \cdot \rangle_G$ and the conditional norm $\|\cdot\|_G$ map into I. To define convergence using these constructs, we must use a notion of convergence of random variables that are measurable with respect to G. We use convergence in probability in the following definitions of conditional Cauchy sequences and conditional convergent sequences.

DEFINITION 2.1: A sequence $\{p_j : j = 1, 2, \ldots\}$ in P^+ converges conditionally to p_0 if for any $\varepsilon > 0$, $\lim_{j \to \infty} \Pr\{\|p_j - p_0\|_G > \varepsilon\} = 0$.

DEFINITION 2.2: A sequence $\{p_j : j = 1, 2, \ldots\}$ in P^+ is conditionally Cauchy if for any $\varepsilon > 0$, $\lim_{j,k \to \infty} \Pr\{\|p_j - p_k\|_G > \varepsilon\} = 0$.

Next, we introduce a subset P of P^+ that is restricted to satisfy conditional counterparts to linearity and completeness.

DEFINITION 2.3: A set P is a conditional linear subspace of P^+ if for any w_1 and w_2 that are in I and any p_1 and p_2 in P, $w_1 p_1 + w_2 p_2$ is in P.

DEFINITION 2.4: A set P is conditionally complete if every conditional Cauchy sequence in P is conditionally convergent to some element in P.

It is easily verified that the set P^+ is conditionally linear. This follows from the conditional Minkowski Inequality and from the fact that

$$(2.4) \qquad \|wp\|_G = |w| \cdot \|p\|_G.$$

In the Appendix we show that P^+ is conditionally complete.

---

[2] Throughout this paper, equality and inequality relations between random variables are only required to hold with probability one. We will adopt the usual convention of viewing random variables in an equivalence class, that are equal with probability one and satisfy appropriate measurability restrictions, as being the same random variable.

[3] To see this, let $p^+ = \max(0, p)$ and $p^- = \max(0, -p)$. Then $(p_1 p_2)^+ = p_1^+ p_2^+ + p_1^- p_2^-$ and $(p_1 p_2)^- = p_1^+ p_2^- + p_1^- p_2^+$. Conditional versions of the Minkowski, Holder, and Cauchy–Schwarz Inequalities are valid (Loeve (1978, p. 14)). The conditional Cauchy–Schwarz Inequality implies that $p_1^+ p_2^+$, $p_1^- p_2^-$, $p_1^+ p_2^-$, and $p_1^- p_2^+$, all have finite expectations conditioned on G. Hence $(p_1 p_2)^+$ and $(p_1 p_2)^-$ have finite expectations conditioned on G. Following Loeve (1978), we define

$$E(p_1 p_2 | G) = E[(p_1 p_2)^+ | G] - E[(p_1 p_2)^- | G].$$

In our analysis we view P as a set of payoffs at time one from portfolios purchased at time zero. One possibility is for P to equal P^+ in which case P contains (among other things) all event contingent claims based on information in I_1. We allow P to be a proper subspace of P^+ to accommodate situations in which either the econometrician omits assets from his analysis or the markets open to economic agents are incomplete. (Markets can be incomplete, even if the set of payoffs P is conditionally complete in the mathematical sense of Definition 2.4.)

ASSUMPTION 2.1: *P is a conditionally complete linear subspace of P^+.*

The restriction that P be a conditional linear space from the standpoint of economic agents can be motivated by the fact that economic agents are allowed to adjust their portfolios based on information that is available at the purchase date of the portfolios.

Next, we introduce a pricing function π that maps the set of payoffs P into I. We impose two assumptions on the pricing function π and the set of payoffs P. The first assumption is termed value-additivity and requires that the price of a conditional linear combination of payoffs be equal to the corresponding linear combination of prices for the individual payoffs.

ASSUMPTION 2.2: *For any p_1 and p_2 in P and any w_1 and w_2 in I, $\pi(w_1 p_1 + w_2 p_2) = w_1 \pi(p_1) + w_2 \pi(p_2)$.*

The second assumption requires that the pricing function be conditionally continuous at the zero payoff.

ASSUMPTION 2.3: *If $\{p_j : j = 1, 2, \ldots\}$ is a sequence of payoffs in P that converges conditionally to zero, then for any $\varepsilon > 0$, $\lim_{j \to \infty} \Pr\{|\pi(p_j)| > \varepsilon\} = 0$.*

Assumption 2.3 specifies the sense in which small payoffs have small prices. Equivalently, it stipulates the sense in which value-additivity holds for an infinite series of payoffs: If $\{\sum_{k=1}^{j} p_k : j = 1, 2, \ldots\}$ converges conditionally to p_0, then $\{\sum_{k=1}^{j} \pi(p_k) : j = 1, 2, \ldots\}$ converges in probability to $\pi(p_0)$. Assumptions 2.2 and 2.3 require that π be the conditional analogue to a continuous linear functional on a Hilbert space.

Next we consider a different motivation for Assumption 2.3. When there is free disposal of the good used to denominate the payoffs, prices of payoffs that are nonnegative with probability one will have equilibrium prices that are non-negative. In this case Assumption 2.3 can be omitted as long as π is defined on P^+.

LEMMA 2.1: *Suppose $P = P^+$, Assumption 2.2 is satisfied, and for any $p \geqslant 0$ in P^+, $\pi(p) \geqslant 0$. Then Assumption 2.3 is satisfied.*

In light of Lemma 2.1, the continuity requirement is implied whenever π can be extended from P to P^+ in such a way as to preserve nonnegativity and assign finite prices to all payoffs in that extension.[4] The assignment of finite prices to all elements in P^+ appears to be an important restriction, the plausibility of which will depend on the particular asset pricing model under investigation.

For our analysis, it is convenient to rule out certain nontrivial pricing functions.

ASSUMPTION 2.4: *There exists a payoff p_0 in P for which* $\Pr\{\pi(p_0) = 0\} = 0$.

A sufficient condition for Assumption 2.4 to be valid is that the pricing function π be strictly positive on P^+:

LEMMA 2.2: *Suppose $P = P^+$, Assumption 2.2 is satisfied, and* $\Pr\{\pi(p) > 0\} = 1$ *for any p in P^+ for which* $\Pr\{p > 0\} = 1$. *Then Assumptions 2.3 and 2.4 are satisfied.*

An example of a pricing function that satisfies Assumptions 2.2 and 2.3 is $\pi(p) = \langle p \mid p^* \rangle_G$ for some p^* in P. It turns out that all pricing functions have such a representation. This result is an implication of the conditional analogue to the Riesz Representation Theorem from the theory of Hilbert spaces.

THEOREM 2.1: *Suppose Assumptions 2.1–2.4 are satisfied. Then there exists a unique payoff p^* in P that satisfies $\pi(p) = \langle p \mid p^* \rangle_G$ for all p in P. Furthermore,* $\Pr\{\|p^*\|_G > 0\} = 1$.[5]

Theorem 2.1 shows that each asset pricing function is uniquely indexed by a benchmark payoff, p^*, in P.

There is a particularly convenient interpretation of p^* when π has no arbitrage opportunities on P^+.

DEFINITION 2.4: *A pricing function π has no arbitrage opportunities on P if for any payoff p in P for which* $\Pr\{p \geq 0\} = 1$, $\Pr(\{\pi(p) \leq 0\} \cap \{p > 0\}) = 0$.

An interpretation of no-arbitrage is that nonnegative payoffs that are positive with positive probability conditioned on G have positive prices.[6] This restriction on π is the conditional counterpart to the no-arbitrage assumption used by Ross (1978).

---

[4] Kreps (1981) gives necessary and sufficient conditions for such an extension when the pricing function is a (unconditional) linear functional.

[5] Theorem 2.1 would also be true if the notion of convergence in Assumptions 2.1 and 2.3 was convergence almost surely instead of convergence in probability. We use convergence in probability because it simplifies our proofs. Furthermore, a version of Theorem 2.1 remains valid under a weakened version of Assumption 2.4. See Hansen and Richard (1985).

[6] To see that no-arbitrage can be defined equivalently using conditional probabilities, let 1_p denote that random variable that is one when p is positive and zero otherwise. Then

$$\Pr(\{\pi(p) \leq 0\} \cap \{p > 0\}) = \int_{\{\pi(p) \leq 0\}} 1_p \, d\Pr = \int_{\{\pi(p) \leq 0\}} E(1_p \mid G) \, d\Pr,$$

as long as $\Pr(\{\pi(p) \leq 0\}) > 0$. Therefore, $\Pr(\{\pi(p) \leq 0\} \cap \{p > 0\}) = 0$ if, and only if, $\Pr(\{\pi(p) \leq 0\} \cap \{E(1_p \mid G) > 0\}) = 0$.

LEMMA 2.3: *Suppose $P = P^{+}$ and Assumption 2.2 is satisfied. Then π has no arbitrage opportunities on P^{+} if, and only if, $\Pr \{ p^{*} > 0 \} = 1$.*

In light of Lemma 2.3, when π has no arbitrage opportunities on P^{+}, p^{*} will be strictly positive with probability one. In this case we can interpret p^{*} as a measure of the equilibrium intertemporal marginal rate of substitution of the numeraire good used to denominate prices for the numeraire good used to denominate payoffs. Harrison and Kreps (1979) use p^{*} to define a new probability measure such that asset prices can be calculated as conditional expectations. In our analysis we will continue to use the original probability measure.

In the next section we use the results in this section to study the impact of omitting conditioning information when studying asset pricing models using time series data.

3. IMPLICATIONS FOR OMITTING INFORMATION

Chamberlain and Rothschild (1983) and Chamberlain (1983) use Hilbert space theory to characterize restrictions of asset pricing models in terms of means and variances of returns. In light of the mathematical results given in Section 2, it is not surprising that much of their analysis will carry over to a conditional Hilbert space setting. We show that this is indeed true. Then we proceed to study the impact of omitting conditioning information in calculating mean-variance frontiers for infinite dimensional spaces. In particular we compare the mean-variance frontier conditioned on G to the unconditional mean-variance frontier.

Consider a set of payoffs P and a pricing function π that satisfies Assumptions 2.1–2.4. We define two level sets of π that are central to our analysis:

$$(3.1) \qquad R = \{ p \text{ in } P \colon \pi(p) = 1 \}$$

and

$$(3.2) \qquad Z = \{ p \text{ in } P \colon \pi(p) = 0 \}.$$

The set R is the collection of all payoffs with unit prices, i.e., the set of all returns. The set Z consists of all payoffs with zero prices. Since P is a conditional linear subspace and π is a conditional linear functional, the zero payoff is in Z. Assumption 2.4 guarantees that the payoff $p_0 / \pi(p_0)$ is in R. Hence, both Z and R are not empty.

In Section 2 we showed that π could be represented as

$$(3.3) \qquad \pi(p) = \langle p \mid p^{*} \rangle_{G} \quad \text{for all } p \text{ in } P.$$

Furthermore, we showed when the pricing function is nontrivial (i.e., when Assumption 2.4 is satisfied), $\Pr \{ \pi(p^{*}) = 0 \} = 0$ so that

$$(3.4) \qquad r^{*} = p^{*} / \pi(p^{*})$$

is a well-defined return in R.

LEMMA 3.1: *Suppose that (P, π) satisfies Assumptions 2.1–2.4. Then*

(i) $\langle r^* | z \rangle_G = 0$,   *for all z in Z;*

(ii) $\langle r^* | r^* \rangle_G \leqslant \langle r | r \rangle_G$ *for all r in R.*

Lemma 3.1 shows that r^* is the minimum conditional second moment return. This implication carries over when unconditional second moments are used in place of conditional second moments. By the Law of Iterated Expectations,

(3.5) $E(p^2) = E[\langle p | p \rangle_G]$ for any p in P

where the expectation in (3.5) may be infinite. Implication (ii) of Lemma 3.1 and relation (3.5) together imply that

(3.6) $E(r^{*2}) \leqslant E(r^2)$ for any r in R.

Either side of (3.6) may be infinite. When the left-hand side of (3.6) is infinite, then (3.6) is interpreted to imply that the right-hand side of (3.6) also is infinite.

The difference between two returns in R is always in Z. Hence singling out any return in R as a benchmark, all other returns can be expressed as the sum of that benchmark return and an element in Z. When r^* is used as the benchmark, we obtain

(3.7) $R = \{r: r = r^* + z \text{ for some } z \text{ in } Z\}$.

This is a particularly convenient representation of R since r^* is conditionally orthogonal to Z [implication (i) of Lemma 3.1].

The set Z is itself a conditionally complete linear subspace of P^+. Furthermore, the conditional expectation operator $E(\cdot | G)$ defines a conditional analogue to a continuous linear functional on Z. Hence, we have the result:

LEMMA 3.2: *If (P, π) satisfies Assumptions 2.1–2.3, then $[Z, E(\cdot | G)]$ satisfies Assumptions 2.1–2.3.*

We impose the following additional restriction:

ASSUMPTION 3.1: *There exists a payoff z_0 in Z for which $\Pr\{E(z_0 | G) = 0\} = 0$.*

An equivalent statement of this assumption is that $[Z, E(\cdot | G)]$ satisfies Assumption 2.4. Assumption 3.1 requires that the conditional expectation operator and the pricing function disagree somewhere on Z (with probability one). This means that prices are inconsistent with risk neutrality. Assumption 3.1 is satisfied, for instance, when the variance of p^* conditioned on G is different from zero with probability one and P is P^+.

With the addition of Assumption 3.1, Theorem 2.1 implies that there exists a unique payoff z^* in Z for which

(3.8) $\langle z | z^* \rangle_G = E(z | G)$ for all z in Z.

Theorem 2.1 also shows that $\langle z^* | z^* \rangle_G = E(z^{*2} | G) = E(z^* | G)$ is positive with probability one. Notice that

(3.9) $\quad \mathrm{Var}\,(z^* | G) = E(z^{*2} | G) - E(z^* | G)^2 = E(z^* | G)(1 - E(z^* | G)) \geqslant 0$

implies that with probability one

(3.10) $\quad E(z^* | G) = \langle z^* | z^* \rangle_G \leqslant 1.$

The random variable z^* defines one (conditional) dimension of Z. All remaining dimensions that are orthogonal to z^* conditioned on G must have conditional mean zero since (3.8) is satisfied. Hence, the set Z can be represented as

(3.11) $\quad Z = \{z: z = wz^* + n \text{ for some } w \text{ in } I \text{ and some } n \text{ in } N\}$

where N is the set

(3.12) $\quad N = \{z \text{ in } Z: E(z | G) = 0\}.$

Combining (3.7) and (3.11) gives the representation

(3.13) $\quad R = \{r: r = r^* + wz^* + n \text{ for some } w \text{ in } I \text{ and some } n \text{ in } N\}.$

Next, we consider solutions to the mean-variance problem conditioned on G.

PROBLEM 3.1: Minimize $\langle r | r \rangle_G$ for r in R subject to the constraint $E(r | G) = w$ for some w in I.

Notice that the objective function for Problem 3.1 is not real-valued but is a random variable in I. Hence, this objective only induces a partial ordering on P. Nevertheless, as we show in the subsequent analysis, Problem 3.1 does have a solution. Also, notice that $\langle r | r \rangle_G$ in Problem 3.1 could be replaced by the variance of r conditioned on G without altering the solution. Any return that solves Problem 3.1 for some w in I is said to be on the mean-variance frontier conditioned on G.

Problem 3.1 can be solved conveniently using the decomposition of returns given in (3.13). Let

(3.14) $\quad w^* = [w - E(r^* | G)] / E(z^* | G).$

Then r in R satisfies $E(r | G) = w$ if, and only if,

(3.15) $\quad r = r^* + w^* z^* + n$

for w^* given by (3.14) and some n in N. To solve Problem 3.1, n in (3.15) is set to zero since the components on the right-hand side of (3.15) are orthogonal conditioned on G. We have proved the following lemma.

LEMMA 3.3: *If (P, π) satisfies Assumptions 2.1–2.4 and $[Z, E(\cdot | G)]$ satisfies Assumption 3.1, then $r_w = r^* + w^* z^*$ is the solution to Problem 3.1 for w^* given by (3.14).*

The characterization of the mean-variance frontier given in Lemma 3.3 is the conditional counterpart to the characterization of the unconditional mean-variance frontier given by Chamberlain and Rothschild (1983). (Other than conditioning there is a minor difference between our approach and that taken by Chamberlain and Rothschild (1983). Our decomposition is conditionally orthogonal while theirs is not; this is a matter of convenience.) A conditional two-fund theorem can be obtained as an immediate implication of the characterization given in Lemma 3.3 (Hansen and Richard (1984)).

In Section 1 we described a set of circumstances in which unconditional moments of returns could be estimated using time series data. In these circumstances, sample means and covariances calculated using time series data converge almost surely to unconditional means and covariances. The goal of this section is to study the impact of omitting the conditioning information when calculating the mean-variance frontier. To this end we confine our attention to the set of payoffs with finite unconditional second moments:

$$(3.16) \qquad P^* = \{ p \text{ in } P \colon E(p^2) < \infty \}.$$

On this set we define an unconditional inner product

$$(3.17) \qquad \langle p_1 | p_2 \rangle = E(p_1 p_2) \quad \text{for } p_1 \text{ and } p_2 \text{ in } P^*.$$

(Alternatively (3.17) can be viewed as a specialization of the conditional inner product (2.2) when G is the trivial sigma-algebra containing only Ω and the null set.) In addition, we consider the unconditional counterparts to the sets R, Z, and N. Let

$$(3.18) \qquad R^* = R \cap P^*;$$

$$(3.19) \qquad Z^* = Z \cap P^*;$$

and

$$(3.20) \qquad N^* = \{ z \text{ in } Z^* \colon Ez = 0 \}.$$

We study the relation between solutions to Problem 3.1 and solutions to the following problem.

PROBLEM 3.2: Minimize $\langle r | r \rangle$ for r in R^* subject to the constraint that $E(r) = c$ for some real number c.

We are interested in circumstances in which this problem is not vacuous. For this reason, we make the following assumption.

ASSUMPTION 3.2: *There exists a return r_0 in R^*.*

Given Assumption 3.2, r_0 and hence r^* are always in R^*. Taking the unconditional expectation of (3.10) gives $\langle z^* | z^* \rangle \leq 1$ so z^* is in Z^*. Also, the zero random variable is always in Z^* and N^*. Therefore, these three sets are not empty.

Any solution to Problem 3.2 for some real number c is said to be on the unconditional mean-variance frontier. Our strategy for calculating this frontier is very similar to the one used to calculate the mean-variance frontier conditioned on G. Applying the Law of Iterated Expectations to (3.8) gives

$$(3.21) \quad \langle z \,|\, z^* \rangle = E(z) \quad \text{for all } z \text{ in } Z^*.$$

Hence, the random variable z^* is orthogonal to N^* (unconditionally) so that elements in Z^* can be represented as the sum of a scalar multiple of z^* and an element in N^*. Consequently, R^* can be represented as

$$(3.22) \quad R^* = \{r: r = r^* + c^* z^* + n^* \text{ for some real number } c^*$$
$$\text{and some } n^* \text{ in } N^*\}.$$

Let

$$(3.23) \quad c^* = [c - E(r^*)]/E(z^*).$$

Then r in R^* has expectation c if, and only if,

$$(3.24) \quad r = r^* + c^* z^* + n^*$$

for some n^* in N^* and c^* given in (3.23). Problem 3.2 is then solved by setting n^* to zero. We have proved the following lemma.

LEMMA 3.4: *Suppose that (P, π) satisfies Assumptions 2.1–2.4 and 3.2, and $[Z, E(\cdot \,|\, G)]$ satisfies Assumption 3.1. Then the solution to Problem 3.2 is $r_c = r^* + c^* z^*$ for c^* given by (3.23).*

Taken together, Lemmas 3.3 and 3.4 display the impact of omitting conditioning information when evaluating whether a return is on the mean-variance frontier. Notice that any return that is on the unconditional frontier must also be on the conditional frontier. In other words, r_c given in Lemma 3.4 also solves Problem 3.1 for w equal to $E(r_c \,|\, G)$. The converse is not true, however. A return may be on the mean-variance frontier conditioned on G and not be on the unconditional frontier. For instance, consider r_w as given by Lemma 3.3. Then as long as w^* as given by (3.14) is not equal to a constant with probability one, r_w will not be on the unconditional mean-variance frontier. A risk-free return can be used to illustrate this fact.

When P contains a unit payoff and π has no arbitrage opportunities on P, R will contain a return that is risk-free. This return is given by

$$(3.25) \quad r^f = 1/\langle p^* \,|\, 1 \rangle_G$$
$$= \langle r^* \,|\, r^* \rangle_G / \langle r^* \,|\, 1 \rangle_G.$$

Notice that the conditional inner product of any payoff in P with a unit payoff is just the conditional expectation of that payoff. In this case, the variable z^* used in representing the mean-variance frontier is the conditional residual of regressing 1 on r^*. Hence

$$(3.26) \quad z^* = 1 - \langle r^* \,|\, 1 \rangle_G \, r^* / \langle r^* \,|\, r^* \rangle_G.$$

It is straightforward to verify that z^* satisfies (3.8). The risk-free return can be expressed as

$$(3.27) \qquad r^f = r^* + r^f z^*.$$

This means that the risk-free return will be on the unconditional mean-variance frontier if, and only if, r^f equals a constant with probability one.

The implications of the static CAPM are typically portrayed using a single-beta representation of the risk-return tradeoff. We call a return r_β in R a *reference return* for a single-beta representation conditioned on G if $\Pr\{\text{Var}(r_\beta \mid G) = 0\} = 0$ and

$$(3.28) \qquad E(r \mid G) - \alpha = \frac{\text{Cov}[r_\beta, r \mid G]}{\text{Var}[r_\beta \mid G]} [E(r_\beta \mid G) - \alpha]$$

for all r in R, where the random variable α is in I. In general, any return in R that is conditionally uncorrelated with r_β has a conditional mean equal to α. In particular, if a risk-free return r^f is in I, then $r^f = \alpha$ since $\text{Cov}[r_\beta, r^f \mid G] = 0$.

Our next result establishes the conditions under which being on the conditional mean-variance frontier is equivalent to being a reference return for a conditional single-beta representation. It is the conditional counterpart of Roll's (1977) Corollary 6.

LEMMA 3.5: *Suppose (P, π) satisfies Assumption 2.1–2.4, π has no arbitrage opportunities on P, and $[Z, E(\cdot \mid G)]$ satisfies Assumption 3.1. Then r_β is a reference return for a single-beta representation conditioned on G if, and only if, $r_\beta = r^* + w^* z^*$ where w^* is in I and*

$$(3.29) \qquad \Pr\{w^* = E(r^* \mid G)/(1 - E(z^* \mid G))\} = 0.$$

The no-arbitrage restriction on π ensures that $(1 - E(z^* \mid G))$ is not zero with probability one. Condition (3.29) guarantees that the probability that r_β is equal to the minimum conditional variance return is zero.

We call a return r_β a reference return for an unconditional single-beta representation if $\text{Var}(r_\beta) > 0$ and the unconditional counterpart to (3.28) is satisfied for all returns in R^*, where α is a real number. Notice that the unconditional version of (3.28) defines a set of restrictions across the means of returns and the population regression coefficients of returns on r_β. There is also an unconditional version of Lemma 3.5.

COROLLARY 3.1: *Suppose the assumptions of Lemma 3.5 and Assumption 3.2 are satisfied. Then r_β is a reference return for an unconditional single-beta representation if, and only if, $r_\beta = r^* + c^* z^*$ where c^* is a constant and*

$$(3.30) \qquad c^* \neq E(r^*)/(1 - E(z^*)).$$

The standard approach to testing whether a return is on the mean-variance frontier is to test whether a return satisfies a single-beta representation using regression techniques. Corollary 3.1 gives a defense for this procedure. Under the class of data generating processes given in Section 1, the unconditional first and second moments of return processes constructed using elements in R^* will be time invariant and can be estimated using time series averages. In such circumstances, standard methods can be used to estimate unconditional means of returns and the slope coefficients for regressions of returns on a hypothetical reference return.[7] Such methods are appropriate for testing whether returns are on the unconditional mean-variance frontier, but are not necessarily appropriate for testing whether returns are on the conditional mean-variance frontier. Omitting returns from the analysis, which is often required for the empirical analysis to be tractable, will not in general restore the validity of the regression methodology for testing whether returns are on the mean-variance frontier conditioned on G. Hence an econometrician is not permitted to ignore the conditioning information when testing whether returns are on the conditional mean-variance frontier.

The analysis in this section has implications for the empirical evaluation of particular asset pricing models and for the performance evaluation of portfolio managers. For example, the static Capital Asset Pricing Model of Sharpe (1964), Lintner (1965), and Mossin (1966) can be interpreted to imply that the return on the aggregate wealth portfolio is a reference return for a conditional single-beta representation, while Breeden's (1979) continuous time asset pricing model implies that the return on the aggregate consumption portfolio is. But neither of these portfolios will necessarily be a reference return for an unconditional single-beta representation, (e.g., see Cornell (1981)). Along the second vein, our results are consistent with the finding of Dybvig and Ross (1985) that when returns from a managed-portfolio are found to be unconditionally mean-variance inefficient, they still may be conditionally mean-variance efficient. The analysis in this section gives a characterization of all returns that stay on the mean-variance frontier when conditioning information is omitted.

4. AN ALTERNATIVE PRICING FUNCTION

In Section 3 we showed that the mean-variance implications of asset pricing models are sensitive to the omission of conditioning information. In this section we suggest a method for deducing implications that are expressable in terms of unconditional moments. Recall that unconditional moments can be estimated consistently using time series data when the time series are generated by processes like those described in Section 1.

In this section we restrict our attention to payoffs with finite unconditional second moments, i.e., payoffs in P^*. Conditioning information was essential to

---

[7] The standard methods for inference are often inappropriate, however, because they assume the time series process for returns are independent and identically distributed. These methods can be modified using the asymptotic distribution theory in Hansen (1982).

the analysis in Section 3 because the pricing function π mapped into I. Consequently, some payoffs in P^* have prices that are not constant with probability one. Our approach in this section is to define a new pricing function π^* that is constructed from π but assigns prices that are real numbers. For this construction to be valid, we require that the payoff p^*, used in representing π, be in P^*.

ASSUMPTION 4.1: $\langle p^* \mid p^* \rangle < \infty$.

Assumption 4.1 will not be satisfied for all choices of numeraire for quoting asset prices. More precisely, let w be a random variable in I for which $\Pr\{w > 0\} = 1$. Then a pricing function π^- with a different numeraire is given by $\pi^-(p) = \pi(p)/w$. This pricing function has a benchmark payoff p^*/w. If p^* satisfies Assumption 4.1, then it is not necessarily the case that p^*/w will satisfy Assumption 4.1. On the other hand, if p^* does not satisfy Assumption 4.1 then one can always find a random variable w such that p^*/w will satisfy Assumption 4.1. For instance, let $w = \| p^* \|_G$. Then $E[(p^*/w)^2 \mid G] = 1$ implying that p^*/w has a finite unconditional second moment. Therefore, requiring a pricing function to satisfy this assumption restricts the choice of numeraire.

We define π^* to be

$$(4.1) \qquad \pi^*(p) = E[\pi(p)].$$

Notice that π^* maps into the real numbers. This function behaves like the pricing function π when the trivial sigma algebra is used for the conditioning information set. The only random variables that are measurable with respect to the trivial sigma algebra are constant over all sample points. We refer to prices calculated using π^* as pseudo-prices.

THEOREM 4.1: Suppose (P, π) satisfies Assumptions 2.1–2.4 and 4.1. Then (P^*, π^*) satisfies Assumptions 2.1–2.4 for G given by the trivial sigma algebra. Furthermore, if π has no arbitrage opportunities on P, then π^* has no arbitrage opportunities on P^*.

Elements in Z^* and R^* will be assigned the same prices by π and π^*, but many other payoffs in P^* will be assigned distinct prices by these functions. Notice that some payoffs in P^* that are not in R^* will satisfy $\pi^*(p) = 1$. For example, suppose there is a unit payoff in P such that the riskfree return, $r_f = 1/E(p^* \mid I)$, is in R^*, but is not a constant. We can use the pseudo-pricing function π^* to construct a pseudo-riskfree return $r_f^* = 1/E(p^*)$, which is a constant, but is not in R^* and is not equal to r_f. Similarly, some payoffs in P^* that are not in Z^* will satisfy $\pi^*(z) = 0$. Consequently, the set of pseudo-returns can be larger than R^* and the set of payoffs with pseudo-prices that are zero can be larger than Z^*.

In addition to assigning prices that are real numbers, the pseudo-pricing function π^* has other nice properties. For instance, this function is constructed

so that the inner product representation

$$(4.2) \qquad \pi^*(p) = E[\langle p \,|\, p^* \rangle_G]$$
$$= \langle p \,|\, p^* \rangle$$

is valid. Hence, p^* can be used to represent π^* as well as π. This feature is attractive for the following reason. Suppose time series data are available on p^* as implied by a particular asset pricing model. In addition, suppose time series data are available on asset payoffs and prices. Under the data generation processes described in Section 1, both $E[\pi(p)]$ and $E(pp^*)$ can be estimated for any p in P^*. Relations (4.1) and (4.2) imply that these two quantities should be the same. Consequently, the asset pricing model can be tested by checking whether sample counterparts to $E[\pi(p)]$ and $E(pp^*)$ could be equal after accounting for estimation error. This procedure can be viewed as an interpretation of the econometric approach suggested by Hansen and Singleton (1982) for testing intertemporal asset pricing models.

Notice that payoffs can be constructed that are conditional linear combinations of some initial collection of payoffs as long as the coefficients are measurable with respect to G. The prices (using π as the pricing function) of the resulting payoffs are just the conditional linear combinations of the corresponding prices (again using π) of the initial collection of payoffs. Hence, an analyst is given great flexibility in constructing payoffs and prices to be used in this procedure. The conditional weights used in forming payoffs and prices correspond to the instrumental variables in the Hansen–Singleton analysis. Hansen and Singleton show formally how to estimate preference parameters and test restrictions using time series versions of instrumental variables methods.

A question emerges as to whether information is lost in testing the implications using π^* instead of π. One way to think about this problem is to ask when can two pricing functions, say π and π^+, defined on P imply the same pseudo pricing function on P^*. Suppose these two pricing functions have the representations

$$(4.3) \qquad \pi(p) = \langle p \,|\, p^* \rangle_G \quad \text{for all } p \text{ in } P$$

and

$$(4.4) \qquad \pi^+(p) = \langle p \,|\, p^+ \rangle_G \quad \text{for all } p \text{ in } P,$$

where both p^* and p^+ are in P. In addition, suppose p^* and p^+ are in P^* and π and π^+ imply the same pseudo pricing function:

$$(4.5) \qquad \pi^*(p) = E[\pi(p)]$$
$$= E[\pi^+(p)]$$

for all p in P^*. Then

$$(4.6) \qquad \langle p^* - p^+ \,|\, p^* - p^+ \rangle = 0$$

since $p^* - p^+$ is in P^*. Hence, p^* is equal to p^+ with probability one implying that π and π^+ are the same pricing function. This analysis shows that the

conditioning down approach used in constructing π^* can discriminate among distinct pricing functions defined on P.

Testing whether

$$(4.7) \qquad \langle p \,|\, p^* \rangle = E[\pi(p)] \quad \text{for all } p \text{ in } P^*$$

may be an arduous task. Recall that P is a conditional linear space that will be infinite dimensional in many, if not most circumstances. Even if P is generated by taking all conditional linear combinations of a finite dimensional collection of primitive payoffs, P^* may not be a finite dimensional unconditional linear space because of the presence of conditioning information in I. If relation (4.7) is tested for only a subset of assets, then restrictions are lost. For instance, let P^- be an unconditionally complete linear subspace of P^*. Then it is possible to construct distinct pricing functions on P^* that agree on P^-. This claim follows from the analysis in Harrison and Kreps (1979) and Kreps (1981). They show that when an extension of a pricing function exists that preserves no-arbitrage, this extension is not necessarily unique. Consequently, distinct pricing functions on P may imply the same pseudo pricing function when restricted to P^-.

In summary, models of asset prices are indexed by their pricing function defined on P. We have shown that conditioning down *per se* does not prevent one from distinguishing between models of asset prices. (Of course, two alternative models of asset prices that imply the same pricing functions are indistinguishable in our analysis.) On the other hand, the omission of payoffs can result in statistical tests that are not powerful against particular families of alternative models. Nonetheless, even with a subset of assets, some discrimination is possible.

5. DISCUSSION AND CONCLUSION

Dynamic models that posit environments with uncertainty and, in particular, dynamic models of asset pricing require specifications of the changing information available to economic agents when making their consumption and investment decisions. Hence, it is convenient, if not necessary, to view both payoffs and prices as random variables which inherit their randomness from the underlying state of the economy. For this reason, we introduce conditioning information explicitly and establish conditions under which equilibrium pricing functions can be viewed as conditional linear functionals with conditional inner product representations. These conditional inner product representations are useful in a variety of applications, two of which are illustrated in Sections 3 and 4 of this paper. Both illustrations analyze approaches to testing asset pricing models using unconditional moments of payoffs and prices.

One approach to deducing testable restrictions of asset pricing models that was not analyzed in this paper is Ross' (1976) Arbitrage Pricing Theory (APT). Comments similar to those made in Section 3 also apply to the APT. Ross assumes that a specified set of primitive asset returns have a factor structure. He shows that if the number of primitive assets becomes arbitrarily large, then to avoid arbitrage there must be a linear relationship between the expected asset returns

and their factor loadings. As Shanken (1983) points out, the factor structure is not robust to the formation of portfolios. Consequently, the analysis is sensitive to the specification of the primitive securities. Chamberlain and Rothschild (1983) circumvent Shanken's objection by making a weaker assumption that only a fixed number of eigenvalues of the covariance matrix of returns are unbounded as the number of assets increases. They derive conclusions analogous to those obtained by Ross without specifying a factor structure on the set of primitive returns.

One way to obtain a dynamic version of the APT is to assume that the restrictions on the covariance matrices of returns apply to the conditional covariance matrices. A factor structure or a restriction on the number of unbounded eigenvalues is not, however, robust to conditioning information.[8] Consequently, unconditional covariance matrices could not be used directly in such an approach. An alternative approach, suggested by Stambaugh (1983) and Rothschild (1985), is to assume that the restrictions apply directly to the unconditional covariance matrices of returns. The restriction of a fixed number of unbounded eigenvalues will not, however, be robust to portfolio formation because portfolios can be constructed using weights that are in the conditioning information set of economic agents. Furthermore, it is more difficult to relate the factors to the underlying state variables in the economy using an unconditional representation instead of a conditional representation.

Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, IL 60637, U.S.A.

and

Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA 15213, U.S.A.

APPENDIX

In this Appendix we establish conditional versions of several of the important Hilbert-space theorems. We also prove all of the lemmas and theorems discussed in the text.

We first prove a conditional version of the Riesz–Fischer Theorem which shows that L^2 is complete.

THEOREM A.1: *If $P^+ = \{p \text{ in } I_1: E(p^2 | G) < \infty\}$, then P^+ is conditionally complete.*

To prove Theorem A.1 we need a lemma which shows that if $\{p_j: j = 1, 2, \ldots\}$ is conditionally Cauchy in P^+, then this sequence is Cauchy in probability. Let $\varepsilon > 0$ and define the sets

$$A_{jk}(\varepsilon) = \{|p_j - p_k| > \varepsilon\},$$

and

$$B_{jk}(\varepsilon) = \{\|p_j - p_k\|_G > \varepsilon\}.$$

---

[8] In an earlier version of this paper (Hansen and Richard (1984)) we have an example with an infinite number of consumers where the endowments satisfy a conditional factor structure. In this example a dynamic, conditional version of the APT can be found. There is, however, no unconditional factor structure so that no unconditional dynamic version of the APT exists.

LEMMA A.1: *If* $\{p_j : j = 1, 2, \ldots\}$ *is a conditional Cauchy sequence, then for any* $\varepsilon > 0$ *there exists an* $M(\varepsilon)$ *such that for* j, $k \geqslant M(\varepsilon)$, $\Pr\{A_{jk}(\varepsilon)\} < \varepsilon$.

PROOF: Let $B_{jk}^c(\varepsilon)$ denote the complement of $B_{jk}(\varepsilon)$, and let $\delta = \min\{\varepsilon/2, (\varepsilon^3/2)^{1/2}\}$. Since $\{p_j : j = 1, 2, \ldots\}$ is conditionally Cauchy, there exists an $M(\varepsilon)$ such that for j, $k \geqslant M(\varepsilon)$, $\Pr\{B_{jk}(\delta)\} < \delta$. But on $B_{jk}^c(\delta)$, $E[(p_j - p_k)^2 | G] \leqslant \delta^2$ so that by the definition of conditional expectation,

$$\delta^2 \geqslant \int_{B_{jk}^c(\delta)} E[(p_j - p_k)^2 | G] \, d\Pr = \int_{B_{jk}^c(\delta)} (p_j - p_k)^2 \, d\Pr.$$

Partition $B_{jk}^c(\delta)$ into the two sets $C_1 = B_{jk}^c(\delta) \cap A_{jk}^c(\varepsilon)$ and $C_2 = B_{jk}^c(\delta) \cap A_{jk}(\varepsilon)$. Hence for j, $k \geqslant M(\varepsilon)$

$$\delta^2 \geqslant \int_{B_{jk}^c(\delta)} (p_j - p_k)^2 \, d\Pr = \int_{C_1} (p_j - p_k)^2 \, d\Pr + \int_{C_2} (p_j - p_k)^2 \, d\Pr$$

$$\geqslant \int_{C_2} (p_j - p_k)^2 \, d\Pr \geqslant \varepsilon^2 \Pr\{C_2\},$$

since the integral over C_1 is nonnegative and on C_2, $(p_j - p_k)^2 > \varepsilon^2$. Now for j, $k \geqslant M(\varepsilon)$

$$\Pr\{A_{jk}(\varepsilon)\} = \Pr\{B_{jk}^c(\delta) \cap A_{jk}(\varepsilon)\} + \Pr\{B_{jk}(\delta) \cap A_{jk}(\varepsilon)\}$$

$$\leqslant \Pr\{C_2\} + \Pr\{B_{jk}(\delta)\}$$

$$< \delta^2/\varepsilon^2 + \delta < \varepsilon. \qquad\qquad Q.E.D.$$

PROOF OF THEOREM A.1: Suppose $\{p_j : j = 1, 2, \ldots\}$ is a conditional Cauchy sequence in P^+. First, we show that there is a subsequence $\{p_{j(k)} : k = 1, 2, \ldots\}$ which converges conditionally to p_0 in P^+. Then, we show that the original sequence converges conditionally to p_0.

From Lemma A.1 and the fact that $\{p_j : j = 1, 2, \ldots\}$ is conditionally Cauchy, we know that we can construct a subsequence $\{p_{j(k)} : k = 1, 2, \ldots\}$ which satisfies

$$a_k = \Pr\{|p_{j(k+1)} - p_{j(k)}| > 2^{-k}\} < 2^{-k}$$

and

$$b_k = \Pr\{\|p_{j(k+1)} - p_{j(k)}\|_G > 2^{-k}\} < 2^{-k}$$

for $k = 1, 2, \ldots$. Let $s_k = p_{j(k+1)} - p_{j(k)}$. Since $\sum_{k=1}^{\infty} a_k < \infty$ and $\sum_{k=1}^{\infty} b_k < \infty$, the Borel-Cantelli Lemma guarantees that the following infinite series converge almost surely:

$$(A.1) \qquad \sum_{k=1}^{\infty} |s_k| < \infty$$

and

$$(A.2) \qquad \sum_{k=1}^{\infty} \|s_k\|_G < \infty.$$

We now construct p_0 and show it is in P^+. Define p_0 by

$$p_0 = \sum_{k=1}^{\infty} s_k + p_{j(1)}.$$

Relation (A.1) guarantees that p_0 is well-defined as an almost sure limit so that p_0 is in I_1. By the conditional version of the Minkowski Inequality and (A.2)

$$\|p_0\|_G \leqslant \sum_{k=1}^{\infty} \|s_k\|_G + \|p_{j(1)}\|_G < \infty,$$

so p_0 is in P^+.

To show that $\{p_j\}$ converges to p_0 conditioned on G, note that

$$\lim_{k \to \infty} \|p_{j(k)} - p_0\|_G = \lim_{k \to \infty} \left\| \sum_{i=k}^{\infty} s_i \right\|_G$$

$$\leqslant \lim_{k \to \infty} \sum_{i=k}^{\infty} \|s_i\|_G = 0 \quad \text{almost surely,}$$

by the conditional version of the Minkowski Inequality and (A.2). Since convergence almost surely implies convergence in probability, $\{p_{j(k)}: k = 1, 2, \ldots\}$ converges conditionally to p_0.

Finally, we show that the original sequence $\{p_k: k = 1, 2, \ldots\}$ converges conditionally to p_0. By the conditional version of the Minkowski Inequality,

$$\| p_k - p_0 \|_G \leq \| p_k - p_{j(k)} \|_G + \| p_{j(k)} - p_0 \|_G.$$

Since $\{p_k: k = 1, 2, \ldots\}$ is conditionally Cauchy and $j(k) \geq k$, for any $\varepsilon > 0$

$$\lim_{k \to \infty} \Pr \{ \| p_k - p_{j(k)} \|_G > \varepsilon/2 \} = 0.$$

We have already shown that

$$\lim_{k \to \infty} \Pr \{ \| p_{j(k)} - p_0 \|_G > \varepsilon/2 \} = 0.$$

Therefore,

$$\lim_{k \to \infty} \Pr \{ \| p_k - p_0 \|_G > \varepsilon \} = 0. \qquad \qquad \qquad Q.E.D.$$

PROOF OF LEMMA 2.1: The proof is by contradiction. Suppose $\{p_j: j = 1, 2, \ldots\}$ converges conditionally to zero and that there exists an $\varepsilon > 0$ such that

$$\limsup_{j \to \infty} \Pr \{ | \pi(p_j) | > \varepsilon \} > 0.$$

Let $q_j = p_j^+ + p_j^-$ where $p_j^+ = \max\{0, p_j\}$ and $p_j^- = \max\{0, -p_j\}$. Then by linearity and positivity, $\pi(q_j) = \pi(p_j^+) + \pi(p_j^-) \geq |\pi(p_j)|$. Hence

$$(A.3) \qquad \limsup_{j \to \infty} \Pr \{ \pi(q_j) > \varepsilon \} > 0.$$

Furthermore, $\| p_j \|_G = \| q_j \|_G$ so that $\{q_j: j = 1, 2, \ldots\}$ converges conditionally to zero.

Next we show that $\{\pi(q_j): j = 1, 2, \ldots\}$ converges in probability to zero which contradicts (A.3). Theorem 4.1.5 in Chung (1974) gives a metric on the space of random variables that induces a notion of convergence that is equivalent to convergence in probability. Hence, we can establish convergence in probability to zero by showing that every subsequence contains a further subsequence that converges in probability to zero. Let $\{q_j^*: j = 1, 2, \ldots\}$ be any subsequence of $\{q_j: j = 1, 2, \ldots\}$. Then there exists a subsequence $\{q_{j(k)}^*: k = 1, 2, \ldots\}$ of $\{q_j^*: j = 1, 2, \ldots\}$ such that

$$a_k = \Pr \{ | q_{j(k)}^* |_G > 2^{-k} \} < 2^{-k},$$

$$b_k = \Pr \{ \| q_{j(k)}^* \|_G > 2^{-k} \} < 2^{-k},$$

since $\{q_j^*: j = 1, 2, \ldots\}$ converges conditionally to be zero. Using logic that parallels closely the proof of Theorem A.1, the infinite sum

$$\sum_{k=1}^{\infty} q_{j(k)}^* = q_0^*$$

converges almost surely and conditionally to q_0^*, which is in P^+. Since $q_{j(k)}^*$ is nonnegative, $\pi[q_{j(k)}^*]$ is nonnegative for all $k > 0$. Hence, for any $l > 0$

$$\sum_{k=1}^{l} \pi[q_{j(k)}^*] \leq \pi(q_0^*).$$

Therefore, the infinite sum

$$\sum_{k=1}^{\infty} \pi[q_{j(k)}^*]$$

converges almost surely implying that $\{\pi[q_{j(k)}^*]: k = 1, 2, \ldots\}$ converges almost surely to zero. Since convergence almost surely implies convergence in probability, $\{\pi[q_{j(k)}^*]: k = 1, 2, \ldots\}$ converges in probability to zero. $\qquad Q.E.D.$

PROOF OF LEMMA 2.2: Suppose p in P satisfies $\Pr \{ p \geq 0 \} = 1$. Let $p_n = p + 2^{-n}$ for $n = 1, 2, 3, \ldots$ so that $\pi(p_n) = \pi(p) + 2^{-n}\pi(1) > 0$ with probability one since π is strictly positive. Letting $n \to \infty$ shows that $\pi(p) \geq 0$ with probability one. Hence π is positive and Lemma 2.1 implies Assumption 2.3 is satisfied. To establish Assumption 2.4, consider the unit payoff on Ω, so strict positivity implies $\Pr(\{\pi(1) > 0\}) = 1.$ $\qquad Q.E.D.$

Our next theorem establishes the conditional version of the Classical Projection Theorem for Hilbert spaces. First we prove three preliminary lemmas.

LEMMA A.2: *Let H be a conditional linear subspace of P^+, and let p be an arbitrary element in P^+. If there is an h_0 in H such that*

(A.4) $\| p - h_0 \|_G \leqslant \| p - h \|_G$

for all h in H, then h_0 is unique. A necessary and sufficient condition that h_0 in H be the unique element of H satisfying (A.4) is that the error element $p - h_0$ be conditionally orthogonal to H.

PROOF: We show first that if h_0 satisfies (A.4), $p - h_0$ is conditionally orthogonal to H. Let h be any element in H and let

$$w = \begin{cases} 0 & \text{if } \|h\|_G = 0, \\ \langle p - h_0 \,|\, h \rangle_G / \|h\|_G^2 & \text{otherwise.} \end{cases}$$

Notice that w is in I so that

$$\| p - h_0 - wh \|_G^2 = \| p - h_0 \|_G^2 + w^2 \|h\|_G^2 - 2w \langle p - h_0 \,|\, h \rangle_G$$
$$= \| p - h_0 \|_G^2 - w^2 \|h\|_G^2 .$$

Since h_0 satisfies (A.4) and $h_0 + wh$ is in H, w must be zero with probability one. But w is zero only when $\langle p - h_0 \,|\, h \rangle_G$ is zero. Since our choice of h in H is arbitrary, $p - h_0$ is conditionally orthogonal to H.

We now show that if $p - h_0$ is conditionally orthogonal to H, then h_0 is the unique minimizing element. For any h in H,

$$\| p - h \|_G^2 = \| p - h_0 + h_0 - h \|_G^2 = \| p - h_0 \|_G^2 + \| h_0 - h \|_G^2 .$$

Hence if $h_0 \neq h$ on a set of positive probability, then $\| p - h \|_G > \| p - h_0 \|_G$ on that set.   Q.E.D.

Our next two lemmas use the space $L^2 = \{ p \text{ in } P^+ : \| p \| < \infty \}$ where $\| p \| = (Ep^2)^{1/2}$. Mean-square convergence on this space is defined using the unconditional norm $\| \cdot \|$. It is a well-known result that L^2 is complete.

LEMMA A.3: *Suppose $\{ p_j : j = 1, 2, \ldots \}$ converges in mean-square to p_0 in L^2. Then this sequence also converges conditionally to p_0.*

PROOF: Note that

$$\lim_{j \to \infty} E[\| p_j - p_0 \|_G^2] = \lim_{j \to \infty} \| p_j - p_0 \|^2 = 0 .$$

Therefore, $\{ \| p_j - p_0 \|_G^2 : j = 1, 2, \ldots \}$ converges in probability to zero implying that $\{ p_j : j = 1, 2, \ldots \}$ converges conditionally to zero.   Q.E.D.

LEMMA A.4: *Suppose H is a conditionally complete linear subspace of P^+. Then $H^* = \{ h \text{ in } H : \|h\| < \infty \}$ is a (mean-square) closed linear subspace of L^2.*

PROOF: The set H^* is a subspace of L^2 because H is a conditional linear subspace of P^+ and unconditional linear combinations of random variables with finite second moments themselves have finite second moments. To prove that H^* is closed, let $\{ h_j : j = 1, 2, \ldots \}$ be any mean-square Cauchy sequence of H^*. Since L^2 is complete, this sequence converges in mean-square to an element h_0 in L^2. Lemma A.3 and the conditional closure of H guarantee that h_0 is in H^*.   Q.E.D.

We now prove a conditional analogue to the Classical Projection Theorem.

THEOREM A.2: (Conditional Projection Theorem): *Suppose H is a conditionally complete linear subspace of P^+. Corresponding to any element p in P^+, there is a unique element h_0 in H such that $\| p - h_0 \|_G \leqslant \| p - h \|_G$ for all h in H. Furthermore, a necessary and sufficient condition that h_0 in H be the unique minimizing element is that $p - h_0$ be conditionally orthogonal to H.*

PROOF: The uniqueness and orthogonality have been established in Lemma A.2. We now establish the existence of a minimizing element. This is accomplished by examining a subset of P^+ for which unconditional second moments are finite. We apply the Classical Projection Theorem directly to this subset and then we extend this solution to all of P^+.

Let

$$H^* = \{h \text{ in } H: \|h\| < \infty\}.$$

Then Lemma A.4 implies that H^* is a closed (in mean-square) linear subspace of L^2. We calculate the conditional projection of p in P^+ onto H in terms of an unconditional projection onto H^*. There are two cases. In the first case, $\|p\| < \infty$. By the Classical Projection Theorem there exists an h_0 in H^* which satisfies

$$E(\|p - h_0\|_G^2) \leq E(\|p - h\|_G^2) \quad \text{for all } h \text{ in } H^*.$$

We next prove by contradiction that

$$\|p - h_0\|_G \leq \|p - h\|_G \quad \text{for all } h \text{ in } H,$$

which establishes the theorem for this case. Suppose there exists an h in H such that

(A.5) $\|p - h\|_G < \|p - h_0\|_G$

on a set A in G, where $\Pr(A) > 0$. Let

$$\hat{h} = \begin{cases} h \text{ on } A, \\ h_0 \text{ on } A^c. \end{cases}$$

Since A is in G, $\hat{h}$ is a conditional linear combination of h and h_0 where the coefficients are the indicator function of A and A^c, respectively. Hence $\hat{h}$ is in H. Inequality (A.5) implies that

(A.6) $E(\|p - \hat{h}\|_G^2) < E(\|p - h_0\|_G^2) < \infty.$

Therefore, $\hat{h}$ is in H^* and satisfies (A.6), contradicting the assumption that h_0 is minimizing in H^*.

In the second case, $Ep^2 = \infty$ so we must transform p in order to use the result of case one. Let $\tilde{p} = p/v$, where v in I is defined by

$$v = \begin{cases} \|p\|_G & \text{when } \|p\|_G > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Now $\|\tilde{p}\|_G \leq 1$, implying $\tilde{p}$ is in L^2 since $E(\|\tilde{p}\|_G^2) \leq 1$. From case one we know there exists an $\tilde{h}_0$ in H^* such that

$$\|\tilde{p} - \tilde{h}_0\|_G \leq \|\tilde{p} - h\|_G \quad \text{for all } h \text{ in } H.$$

Let $h_0 = v\tilde{h}_0$ so that

$$\|p - h_0\|_G = v\|\tilde{p} - \tilde{h}_0\|_G \leq v\|\tilde{p} - h/v\|_G = \|p - h\|_G,$$

for all h in H, since h/v is in H.   Q.E.D.

Next we prove a conditional version of the Riesz–Frechet Representation Theorem. We begin with a preliminary lemma.

LEMMA A.5: *Suppose* (P, π) *satisfies Assumptions 2.1-2.3. Then* $Z = \{p \text{ in } P: \pi(p) = 0\}$ *is a conditionally complete linear subspace or* P.

PROOF: It is an immediate consequence of the conditional linearity of π that Z is a conditional linear subspace of P. To prove that Z is conditionally complete let $\{z_j: j = 1, 2, \ldots\}$ be a conditional Cauchy sequence in Z. Since P is conditionally complete this sequence converges conditionally to an element p_0 in P. Furthermore, $\{z_j - p_0: j = 1, 2, \ldots\}$ converges conditionally to zero implying that $\{\pi(z_j - p_0): j = 1, 2, \ldots\}$ converges in probability to zero. However,

$$\pi(z_j - p_0) = \pi(z_j) - \pi(p_0) = -\pi(p_0).$$

Consequently, $\pi(p_0)$ is zero and p_0 is in Z. Therefore, Z is conditionally complete.   Q.E.D.

PROOF OF THEOREM 2.1: Let p_0 satisfy Assumption 2.4 and let z_0 be its conditional projection onto Z. Then $\pi(p_0) = \pi(p_0 - z_0)$ since $\pi(z_0)$ is zero. Define the benchmark return r^* by

(A.7) $r^* = (p_0 - z_0)/\pi(p_0)$.

Notice that r^* is conditionally orthogonal to Z. Given any p in P, $p - \pi(p)r^*$ is in Z since $\pi[p - \pi(p)r^*] = \pi(p) - \pi(p)\pi(r^*) = 0$. Hence, $\langle p - \pi(p)r^* | r^* \rangle_G = 0$ or

(A.8) $\langle p | r^* \rangle_G = \pi(p)\|r^*\|_G^2$.

We digress to show by contradiction that $\Pr\{\|r^*\|_G = 0\} = 0$. Suppose $\|r^*\|_G = 0$ on a set A where $\Pr(A) > 0$. Hence on A, $r^* = 0$. Define a zero payoff

$$z = \begin{cases} r^* & \text{on } A, \\ 0 & \text{otherwise.} \end{cases}$$

By Assumption 2.2 $\pi(z) = 0$, but by (A.7) $\pi(z) = 1$ on A, establishing the contradiction.
Returning to the main proof, divide (A.8) by $\|r^*\|_G^2$ and define

$p^* = r^*/\|r^*\|_G^2$

to get

$\pi(p) = \langle p | p^* \rangle_G$.

The element p^* is clearly unique since if p' is any element of P for which $\pi(p) = \langle p | p' \rangle_G$ for all p, we have $\langle p^* | p^* \rangle_G = \langle p' | p^* \rangle_G = \langle p' | p' \rangle_G$ so that $\|p^* - p'\|_G = 0$. Since $\|p^*\|_G = 1/\|r^*\|_G$, it follows that $\Pr\{\|p^*\|_G > 0\} = 1$.   Q.E.D.

PROOF OF LEMMA 2.3: First we show that $\Pr\{p^* > 0\} = 1$ implies no arbitrage. Let p in P^+ be such that $\Pr\{p \geq 0\} = 1$. Since $\pi(p) = E(pp^* | I)$, $\Pr\{\pi(p) \geq 0\} = 1$. If $\Pr\{\pi(p) = 0\} = 0$, then the result follows. Otherwise $\Pr\{\pi(p) = 0\} > 0$, so that

$$0 = \int_{\{\pi(p)=0\}} \pi(p)\, d\Pr = \int_{\{\pi(p)=0\}} p^* p\, d\Pr.$$

Since $\Pr\{p^* > 0\} = 1$, we must have $\Pr(\{p = 0\} \cap \{\pi(p) = 0\}) = \Pr(\{\pi(p) = 0\})$, implying $\Pr(\{\pi(p) = 0\} \cap \{p > 0\}) = 0$. The result follows from the fact that $\Pr(\{\pi(p) \leq 0\} \cap \{p > 0\}) = \Pr(\{\pi(p) = 0\} \cap \{p > 0\}) = 0$.
Next we show that no arbitrage on P^+ implies $\Pr\{p^* > 0\} = 1$. Suppose to the contrary that $\Pr(A) > 0$ where $A = \{p^* \leq 0\}$. The indicator function, $1_A \geq 0$, is in P^+. We know that $\pi(1_A) \leq 0$ since $\pi(1_A) = \langle p^* | 1_A \rangle_G$. Therefore, $\Pr(\{\pi(1_A) \leq 0\} \cap \{1_A > 0\}) = \Pr\{1_A > 0\} = \Pr(A) > 0$, establishing the required contradiction.   Q.E.D.

PROOF OF LEMMA 3.1: This proof follows directly from Theorem 2.1 and its proof. In the proof of Theorem 2.1, $r^* = p^*/\pi(p^*)$ was constructed to be conditionally orthogonal to Z. It remains to show that r^* satisfies (ii). For any r in R, $r - r^*$ is in Z so that

$\langle r^* | r \rangle_G = \langle r^* | r^* \rangle_G = \|r^*\|_G^2$.

Consequently,

$$0 \leq \|r - r^*\|_G^2 = \|r\|_G^2 - 2\langle r^* | r \rangle_G + \|r^*\|_G^2$$

$$= \|r\|_G^2 - \|r^*\|_G^2,$$

establishing that (ii) is satisfied.   Q.E.D.

PROOF OF LEMMA 3.2: We have already shown Z is a conditionally complete subspace of P; hence Assumption 2.1 is satisfied. Assumption 2.2 follows directly from the conditional linearity of the conditional expectation operator. Finally Assumption 2.3 follows from the conditional version of the Cauchy-Schwarz Inequality since $|E(z | G)| \leq \|z\|_G$.   Q.E.D.

PROOF OF LEMMA 3.5: The proof of the "if" part proceeds in three steps. The first step is to show that $\Pr\{1 - E(z^*|G) = 0\} = \Pr\{\text{Var}(z^*|G) = 0\} = 0$. Note that $\text{Var}(z^*|G) = E(z^{*2}|G) - E(z^*|G)^2 = E(z^*|G)(1 - E(z^*|G))$ is zero if, and only if, $1 - E(z^*|G)$ is zero since $E(z^*|G)$ is positive with probability one by Theorem 2.1. We can establish a contradiction by supposing that $\Pr\{\text{Var}(z^*|G) = 0\} > 0$. Define

$$\hat{z} = \begin{cases} 0 & \text{if Var}(z^*|G) > 0, \\ z^* & \text{if Var}(z^*|G) = 0, \end{cases}$$

so that $\hat{z}$ is in Z. Note that $\text{Var}(\hat{z}|G) = 0$ implying $\hat{z}$ is in I so that $E(\hat{z}|G) = \hat{z}$. Also when $\text{Var}(z^*|G) = 0$, $E(z^*|G) = 1$ implying

$$\hat{z} = \begin{cases} 0 & \text{if Var}(z^*|G) > 0, \\ 1 & \text{if Var}(z^*|G) = 0. \end{cases}$$

Since $\pi(\hat{z}) = 0$ this contradicts no-arbitrage.

Suppose $r_\beta = r^* + w^* z^*$ where w^* is in I, $\Pr\{w^* = v\} = 0$, and $v = E(r^*|G)/[1 - E(z^*|G)]$. The second step is to prove that $\Pr\{\text{var}(r_\beta|G) = 0\} = 0$. Let $r_v = r^* + vz^*$. By direct computation we find that

$$\text{Var}(r_\beta|G) - \text{Var}(r_v|G) = \text{Var}(z^*|G)(w^* - v)^2,$$

which is positive with probability one by step one.

The final step is to show that r_β is a reference return for a single-beta representation conditioned on G. Define

(A.9) $$\alpha = \frac{[E(r^{*2}|G) - E(r^*|G)w^*]E(z^*|G)}{(v - w^*)\text{Var}(z^*|G)}.$$

By step one, the denominator of (A.9) is nonzero with probability one. Let r be any return in R. Then $r = r^* + wz^* + n$ for some w in I and some n in N. Using the conditional orthogonality of r^*, z^*, and n, we find that

$$E(r_\beta|G) = E(r^*|G) + w^* E(z^*|G);$$

$$E(r|G) = E(r^*|G) + wE(z^*|G);$$

$$\text{Var}(r_\beta|G) = \text{Var}(r^*|G) + w^{*2}\text{Var}(z^*|G) - 2w^* E(r^*|G)E(z^*|G);$$

and

$$\text{Cov}(r_\beta, r|G) = \text{Var}(r^*|G) + ww^* \text{Var}(z^*|G) - (w + w^*)E(r^*|G)E(z^*|G).$$

It is straightforward but tedious to verify by direct substitution that (3.28) holds.

The proof of the "only if" part of Lemma 3.5 also requires three steps. First we show that $\Pr\{E(r_\beta|G) = \alpha\} = 0$. Note that $r_0 = r_\beta + z^*/E(z^*|G)$ is a return in R and that $E(r_0|G) = E(r_\beta|G) + 1$. This in turn implies that

$$1 + E(r_\beta|G) - \alpha = E(r_0|G) - \alpha = \frac{\text{Cov}(r_\beta, r_0|G)}{\text{Var}(r_\beta|G)}[E(r_\beta|G) - \alpha].$$

Therefore,

(A.10) $$1 = \left[\frac{\text{Cov}(r_\beta, r_0|G)}{\text{Var}(r_\beta|G)} - 1\right][E(r_\beta|G) - \alpha],$$

implying that $\Pr\{E(r_\beta|G) = \alpha\} = 0$.

The second step is to show that $r_\beta = r^* + w^* z^*$, for some w^* in I. Since r_β is a return, it can be represented as $r_\beta = r^* + w^* z^* + n$, for some w^* in I and some n in N. Let $r_w = r^* + w^* z^*$. Since $E(r_\beta|G) = E(r_w|G)$ and r_w is in R, part one and (3.28) imply that $E(r_\beta^2|G) = E(r_w r_\beta|G)$. Since r^*, z^*, and n are conditionally orthogonal, $E(n^2|G) = 0$ implying that $\Pr\{n = 0\} = 1$.

The final step is to prove that w^* satisfies (3.29). Let $A = \{w^* = v\}$ and suppose to the contrary that $\Pr(A) > 0$. It is easily calculated that on A, $\text{Cov}(r_0, r_\beta|G)/\text{Var}(r_\beta|G) = 1$:

$$\text{Cov}(r_0, r_\beta|G) = \text{Var}(r_\beta|G) + \text{Cov}(z^*, r_\beta|G)/E(z^*|G)$$

$$= \text{Var}(r_\beta|G) + vE(z^{*2}|G)/E(z^*|G) - [E(r^*|G) + vE(z^*|G)]$$

$$= \text{Var}(r_\beta|G) > 0.$$

Hence, $\text{Cov}(r_0, r_\beta|G)/\text{Var}(r_\beta|G) = 1$ on A. This contradicts (A.10). Q.E.D.

PROOF OF COROLLARY 3.1: Let $P^- = \{p \text{ in } P^*: \pi(p) = c \text{ for some real number } c\}$, where P^* is defined in (3.16). Clearly R^*, Z^*, and N^* are subsets of P^- and r^* is in P^-. We will prove that (P^-, π) satisfies the requirements of Lemma 3.5 for G equal to the trivial sigma algebra. If $\{p_j: j = 1, 2, \ldots\}$ is a mean-square Cauchy sequence in P^-, then it converges in mean-square to an element p_0 in P^*. Lemma A.3 shows that $\{p_j: j = 1, 2, \ldots\}$ converges conditionally to p_0. Hence, $\{\pi(p_j): j = 1, 2, \ldots\}$ converges in probability to $\pi(p_0)$. Since $\pi(p_j) = c_j$ for some real number c_j, $\pi(p_0) = c_0$ for some real number c_0. Consequently, p_0 is in P^- proving that P^- is closed in mean-square. Also, an (unconditional) linear combination of payoffs with constant prices has a constant price. Therefore, P^- satisfies Assumption 2.1. The pair (P^-, π) satisfies Assumption 2.2 trivially. Suppose $\{p_j: j = 1, 2, \ldots\}$ in P^- converges in mean-square to zero. Lemma A.3 implies that this sequence converges conditionally to zero. Since (P, π) satisfies Assumption 2.3, $\{\pi(p_j): j = 1, 2, \ldots\}$ converges in probability to zero. However, $\pi(p_j) = c_j$ for some real number c_j. Therefore, $\{c_j: j = 1, 2, \ldots\}$ converges to zero and (P^-, π) satisfies Assumption 2.3. Since r^* is in P^-, (P^-, π) satisfies Assumption 2.4. If π has no arbitrage opportunities on P, then π has no arbitrage opportunities on a subset of P. Finally, z^* is in P^- and $E(z^*) > 0$. Hence, $[z^*, E(\cdot)]$ satisfies Assumption 3.1 unconditionally.   Q.E.D.

PROOF OF THEOREM 4.1: We show in turn that (P^*, π^*) satisfies Assumptions 2.1 through 2.4 when G is replaced by the trivial sigma algebra. Lemma A.4 shows that P^* satisfies Assumption 2.1. It follows from the linearity of the expectations operator that (P^*, π^*) satisfies Assumption 2.2. The Cauchy-Schwarz Inequality implies that $|\pi^*(p)| \leq \|p^*\| \|p\|$. By applying this inequality, it can be verified that (P^*, π^*) satisfies Assumption 2.3. Theorem 2.1 guarantees that $\pi^*(p^*) = E[\langle p^* | p^* \rangle_G] > 0$ so that (P^*, π^*) satisfies Assumption 2.4.

Finally, suppose π has no arbitrage opportunities on P. Let $p \geq 0$ and let $A = \{p > 0\}$ have positive probability. Then $\pi^*(p) = E[\pi(p)] = \int_A \pi(p) \, d \Pr > 0$.   Q.E.D.

REFERENCES

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

BREIMAN, L. (1968): *Probability*. Reading, Massachusetts: Addison-Wesley.

BROCK, W. A. (1980): "Asset Pricing in an Economy with Production: A 'Selective' Survey of Recent Work on Asset-Pricing Models," in *Dynamic Optimization and Mathematical Economics*, ed. by P.-T. Lin. New York: Plenum Press.

BROWN, D. P., AND M. R. GIBBONS (1985): "A Simple Econometric Approach for Utility-Based Asset Pricing Models," *Journal of Finance*, 40, 359–381.

CORNELL, B. (1981): "The Consumption Based Asset Pricing Model: A Note on Potential Tests and Applications," *Journal of Financial Economics*, 9, 103–108.

CHAMBERLAIN, G. (1983): "Funds, Factors, and Diversification in Arbitrage Pricing Models," *Econometrica*, 51, 1305–1323.

CHAMBERLAIN, G., AND M. ROTHSCHILD (1983): "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51, 1281–1304.

CHUNG, K. L. (1974): *A Course in Probability Theory*, 2nd Ed. New York: Academic Press.

COX, J. C., J. E. INGERSOLL, AND S. A. ROSS (1985): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 51, 363–384.

DOOB, J. L. (1953): *Stochastic Processes*. New York: John Wiley.

DYBVIG, P., AND S. A. ROSS (1985): "Differential Information and Performance Measurement Using a Security Market Line," *Journal of Finance*, 40, 383–400.

EICHENBAUM, M., L. P. HANSEN, AND S. F. RICHARD (1984): "The Dynamic Equilibrium Pricing of Durable Consumption Goods," Carnegie-Mellon Working Paper.

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

HANSEN, L. P., AND S. F. RICHARD (1984): "A General Approach for Deducing Testable Restrictions Implied by Asset Pricing Models," Carnegie-Mellon University Working Paper.

——— (1985): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," N.O.R.C. Discussion Paper 85-13.

HANSEN, L. P., AND K. J. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

HARRISON, J., AND D. KREPS (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381–408.

KREPS, D. (1981): "Arbitrage and Equilibrium in Economies with Infinitely Many Commodities," *Journal of Mathematical Economics*, 15–35.

LINTNER, J. (1965): "Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 13–37.

LOEVE, M. (1978): *Probability Theory*, Vol. 2, 4th Edition. New York: Springer-Verlag.

LUCAS, R. E., JR. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1426–1446.

MOSSIN, J. (1968): "Equilibrium in a Capital Asset Market," *Econometrica*, 34, 768–783.

RICHARD, S. F., AND M. SUNDARESAN (1981): "A Continuous Time Equilibrium Model of Forward Prices and Futures Prices in a Multigood Economy," *Journal of Financial Economics*, 9, 347–371.

ROLL, R. (1977): "A Critique of the Asset Pricing Theory's Tests: Part I," *Journal of Financial Economics*, 4, 129–176.

ROSS, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

—— (1977): "Return, Risk, and Arbitrage," in *Risk and Return in Finance*, Vol. 1., ed. by I. Friend and J. L. Bicksler. Cambridge, Massachusetts: Ballinger Publishing Co., 189–218.

—— (1978): "A Simple Approach to the Valuation of Risky Streams," *Journal of Business*, 3, 453–476.

ROTHSCHILD, M. (1985): "Asset Pricing Theory," Working Paper, Department of Economics, University of California–San Diego.

RUBINSTEIN, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics and Management Science*, 7, 407–425.

SHANKEN, J. (1982): "The Arbitrage Pricing Theory: Is it Testable?" *Journal of Finance*, 37, 1129–1140.

SHARPE, W. F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 429–442.

STAMBAUGH, R. F. (1983): "Arbitrage Pricing with Information," *Journal of Financial Economics*, 12, 357–369.