



Implications of Security Market Data for Models of Dynamic Economies

Lars Peter Hansen; Ravi Jagannathan

The Journal of Political Economy, Vol. 99, No. 2 (Apr., 1991), 225-262.

Stable URL:

<http://links.jstor.org/sici?sici=0022-3808%28199104%2999%3A2%3C225%3AIOSMDF%3E2.0.CO%3B2-L>

The Journal of Political Economy is currently published by The University of Chicago Press.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

Implications of Security Market Data for Models of Dynamic Economies

Lars Peter Hansen

*University of Chicago, National Bureau of Economic Research, and
National Opinion Research Center*

Ravi Jagannathan

University of Minnesota and Federal Reserve Bank of Minneapolis

We show how to use security market data to restrict the admissible region for means and standard deviations of intertemporal marginal rates of substitution (IMRSs) of consumers. Our approach (i) is non-parametric and applies to a rich class of models of dynamic economies, (ii) characterizes the duality between the mean–standard deviation frontier for IMRSs and the familiar mean–standard deviation frontier for asset returns, and (iii) exploits the restriction that IMRSs are positive random variables. The region provides a convenient summary of the sense in which asset market data are anomalous from the vantage point of intertemporal asset pricing theory.

I. Introduction

In this paper we investigate the implications of asset market data for a rich class of models of dynamic economies. The models within this

Hansen gratefully acknowledges support from the National Science Foundation and Jagannathan from the Institute for Financial Studies, Carlson School of Management. Part of this research was completed while Hansen was a visiting scholar at Stanford University. Helpful comments were made by Darrell Duffie, Wayne Ferson, Larry Glosten, Allan Kleidon, Robert Litterman, Robert E. Lucas, Jr., Masao Ogaki, Scott Richard, José Scheinkman, Karl Snow, Alex Taber, Grace Tsiang, Larry Weiss, and especially John Cochrane and John Heaton. Karl Snow provided valuable assistance with the computations.

[*Journal of Political Economy*, 1991, vol. 99, no. 2]

© 1991 by The University of Chicago. All rights reserved. 0022-3808/91/9902-0003\$01.50

class differ with respect to the heterogeneity of consumers' preferences, the span of the payoffs on tradable securities, and the role of money in the acquisition of consumption goods. In spite of these differences, a common implication of these models is that the equilibrium price of a future payoff on any traded security can be represented as the expectation (conditioned on current information) of the product of the payoff and an appropriately interpreted intertemporal marginal rate of substitution (IMRS) of any consumer (see, e.g., LeRoy 1973; Rubinstein 1976; Lucas 1978; Breeden 1979; Harrison and Kreps 1979; Hansen and Richard 1987). This representation is a generalization of the familiar tenet from price theory that prices should equal marginal rates of substitution. To apply this principle to models of asset pricing, securities are viewed as claims to a numeraire good indexed by future states of the world.

If price data were available from a complete set of security markets, the IMRSs of all consumers could be inferred from Arrow-Debreu prices. However, economic agents may not trade in a complete set of contingent-claims markets. Furthermore, it may be practical for an econometrician to use data on only a small array of securities. Because of these limitations, asset market data alone are typically not sufficient to identify IMRSs.

One approach that has been used extensively is to identify IMRSs by restricting them to be parametric functions of data observed by an econometrician (see, e.g., Hansen and Singleton 1982; Brown and Gibbons 1985; Epstein and Zin, this issue). This approach imposes potentially stringent limits on the class of admissible asset pricing models and then tests whether the particular parameterizations are consistent with the observed asset market data.

While this parametric approach has yielded interesting insights into the empirical plausibility of particular families of models, the approach proposed in this paper goes to another extreme. We purposely enlarge the class of asset pricing models under investigation by imposing as little structure as possible on the admissible class of models. In doing so we eliminate most of the testable implications except possibly the law of one price (portfolios with the same payoffs have the same price) and the absence of arbitrage opportunities (non-negative payoffs that are positive with positive probability have positive prices). Although we are not able to identify the IMRSs fully, we can extract information about them. When IMRSs are constant, portfolio payoffs with the same price must also have the same mean. Thus the existence of portfolios of securities with the same price but distinct expected payoffs implies that IMRSs must vary. We exploit this observation to derive greatest lower bounds on the standard devi-

ations of IMRSs, that is, volatility bounds. These bounds are expressed most conveniently as regions of admissible mean–standard deviation pairs for the IMRSs.

The existence of volatility bounds on IMRSs was originally noted by Shiller (1982) (see also Hansen 1982). His goal was to construct a diagnostic for a particular family of asset pricing models that is insensitive to the alignment of the data. The volatility implications he deduced for IMRSs used only two asset returns and, even for the two-asset case, are weaker than those reported here.

Our reasons for examining volatility bounds are somewhat different from Shiller's. First, our nonparametric approach can serve as a useful complement to the parametric approach that is prevalent in the literature. In particular, it can assist in understanding better why particular models are rejected on the basis of statistical tests: does the parameterization admit too little variability in the IMRSs? Second, our approach provides a common set of diagnostics for a potentially large class of asset pricing models. These diagnostics can also be used to evaluate models in which IMRSs are parameterized as functions of observables as well as models for which moments can be computed from characterizations of the stochastic equilibria. Third, our approach allows us to determine which asset market data sets present the most stringent restrictions for IMRSs and consequently the most startling implications for dynamic economic models. It allows us to make these comparisons without having to focus on a parametric family of such models.

To illustrate these points, we provide an alternative characterization of the so-called *equity premium puzzle* (see, e.g., Mehra and Prescott 1985). In contrast to other characterizations, ours does not depend either on a Markov chain approximation with a small number of states or on a narrow class of asset valuation models. Figure 1 reports a restricted region for the means and standard deviations of IMRSs implied by the annual (1891–1985) time-series data on stocks and bonds used by Campbell and Shiller (1988). The shaded region gives the admissible pairs of means and standard deviations for IMRSs. As benchmarks, we also report time-series sample means and standard deviations for IMRSs implied by a representative consumer model with commonly used period utility functions of the form

$$U(c) = \frac{c^{\gamma+1} - 1}{\gamma + 1}$$

for negative values of γ . For this specification of preferences, the IMRS can be measured by forming a consumption ratio for two different points in time, raising it to the power γ and discounting. For

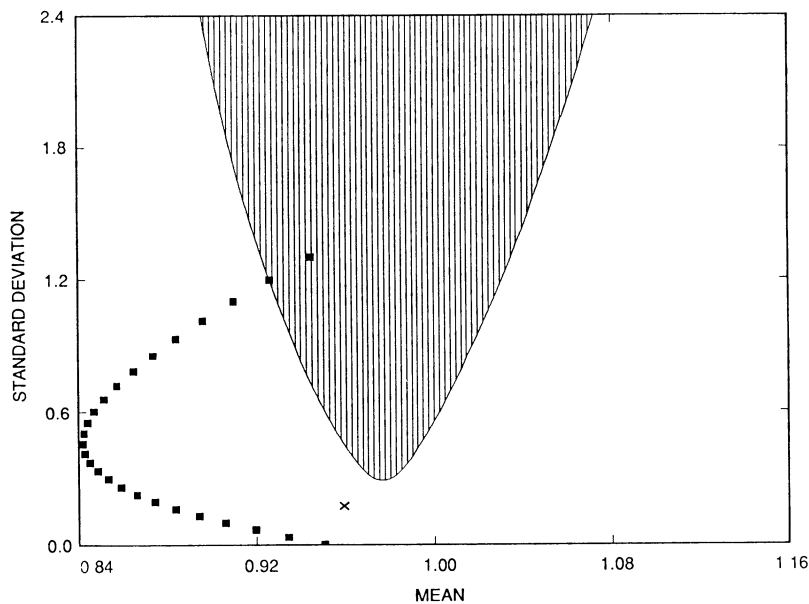


FIG. 1.—IMRS frontier computed using annual data

illustrative purposes, the annual subjective discount factor is taken to be .95. The “boxes” in the figure represent mean–standard deviation pairs for alternative values of γ ranging from zero to -30 . As $|\gamma|$ increases, the volatility of the IMRS increases but the effect on the mean of the IMRS is not uniform. Initially the mean decreases but subsequently increases so that for large $|\gamma|$ the boxes are in the admissible shaded region.

Our strategy for constructing regions such as that reported in figure 1 is to construct minimum variance random variables with pre-specified means that are related to asset payoffs and prices in the same manner as the IMRSs. We refer to such random variables as being on the *mean–standard deviation frontier* for IMRSs. In Section III, we construct these frontier random variables ignoring the fact that IMRSs must be positive. In this case the minimum variance random variables are simply linear combinations of the asset payoffs translated by a constant. As a by-product of this construction, we relate our analysis to two commonly used empirical paradigms in finance: mean-variance analysis and linear factor pricing. More precisely, we characterize the *duality* between the mean–standard deviation frontier for IMRSs and the familiar mean–standard deviation frontier for asset payoffs. This analysis reveals that asset payoffs on

the mean–standard deviation frontier are sufficient to generate the mean–standard deviation frontier for IMRSs. Hence the dimensionality reduction techniques used in linear factor pricing models can be exploited to derive a region like that reported in figure 1.

In Section IV, we modify the analysis of Section III by incorporating the restriction that IMRSs are positive random variables. For prespecified means, we construct nonnegative random variables that behave like IMRSs and have minimum variances. These random variables are not necessarily linear functions of the payoffs but instead can be interpreted as European call and put options on portfolios of these payoffs. In contrast to the analysis in Section III, for some prespecified means there may not be any nonnegative random variables with finite second moments that behave like IMRSs. While the approach of this section yields more restrictive (and therefore more informative) volatility bounds, these sharper bounds are harder to compute.

In Section V, we illustrate the results in Sections III and IV by displaying volatility bounds computed using alternative data sets and generating mean–standard deviation pairs for alternative parametric models of IMRSs. Among other things, we use these bounds to help assess the plausibility of some parametric models of asset prices.

II. A General Model of Asset Pricing

In this section we present a general model of asset valuation. Consider an environment in which multiple consumers trade in securities markets. The preferences and information sets of these consumers may be heterogeneous. We fix both the trading period (say time 0) and the time period for the receipts of the asset payoffs (say time $\tau > 0$). Let I^j denote the information set of consumer j at time 0, and $I = \cap I^j$, where the intersection is taken over the consumers in the economy who trade securities. The prices of securities traded at date 0 are presumed to be in the individual information set I^j of individual j for each j and hence in I . Let P denote a set of portfolio payoffs of the numeraire good at time τ that are traded at time 0. Since the prices of the portfolio payoffs are in I , we represent these prices as a function π_I mapping P into I . Hence $\pi_I(p)$ is the price at time 0 of a portfolio that will pay p units of a numeraire good at a future date τ .

Consumers are presumed to solve optimal portfolio problems in determining their asset holdings. This imposes restrictions relating marginal rates of substitution to asset payoffs and prices. To see this, let mu_0^j and mu_τ^j denote the equilibrium marginal utilities of consumer j in terms of the numeraire consumption good at dates 0 and

τ , respectively. In equilibrium the marginal utility–scaled price must equal the expected marginal utility–scaled payoff conditioned on I^j :

$$mu_0^j \pi_I(p) = E(mu_\tau^j p | I^j) \quad \text{for all } p \text{ in } P. \quad (1)$$

As long as consumer j is not satiated at time 0, $mu_0^j > 0$ and we can divide both sides of (1) by mu_0^j , which yields $\pi_I(p) = E(pm^j | I^j)$ for all p in P , where $m^j = mu_\tau^j / mu_0^j$ is the IMRS of consumer j . Since asset prices are presumed to be observed by all consumers, it follows from the law of iterated expectations that

$$\pi_I(p) = E(pm^j | I) \quad \text{for all } p \text{ in } P. \quad (2)$$

In a world with common information sets and complete markets, marginal rates of substitution are equated across consumers ($m^j = m$ for all j). In such a world, P can be chosen to be sufficiently large that the common IMRS is uniquely determined by (2). In general, (2) does not uniquely determine m^j . As we shall see, however, (2) does restrict the unconditional moments of m^j even when markets are incomplete.¹ Since the restrictions we derive apply to all the individual IMRSs, to simplify notation we drop the j superscript on m .

We now give a more complete description of P and the associated asset pricing function π_I . We do not require that P contain all the portfolio payoffs that are traded by consumers. Omitting payoffs will, however, weaken the implications for m . As a matter of convenience, we consider the case in which there is an n -dimensional vector \mathbf{x} of asset payoffs at date τ . The time 0 prices of these assets can also be represented as an n -dimensional vector, say \mathbf{q} , and pricing relation (2) can be expressed as

$$\mathbf{q} = E(\mathbf{x}m | I). \quad (3)$$

We are interested in the implications of (3) for the IMRS m . To investigate this relation empirically, we must have some way to replicate observations on payoffs, prices, and information over time. As in Hansen and Richard (1987), we imagine an environment in which relation (3) is replicated over time. In other words, there is a composite process $\{(m_t, \mathbf{x}_t, \mathbf{q}_t)\}$ and a sequence of information sets $\{I_t\}$ that satisfy a version of (3) for all t . Econometricians seeking to study this economy are presumed to have data on a finite record $(\mathbf{x}_t, \mathbf{q}_t)$, for $t = 1, 2, \dots, T$, and the composite process $\{(m_t, \mathbf{x}_t, \mathbf{q}_t)\}$ is presumed to be sufficiently regular that a time-series version of a law of large numbers applies. Thus sample moments formed from the finite rec-

¹ In the case of incomplete markets, formula (2) abstracts from the possible existence of short sale constraints.

ords of data converge to population counterparts as the sample size T becomes large. Even though asset prices are determined τ periods prior to the realization of the asset payoffs, from the vantage point of econometricians, we model $\{\mathbf{q}_t\}$ as a stochastic process to accommodate possible variation over time in the asset prices. In what follows we use the unconditional expectation operator E to represent the limit points of the time-series averages of the sample moments.²

We now impose restrictions on m , \mathbf{x} , and \mathbf{q} , which are expressed in terms of unconditional expectations.

ASSUMPTION 1. $E|m|^2 < \infty$, $E|\mathbf{x}|^2 < \infty$, $E\mathbf{x}\mathbf{x}'$ is nonsingular, and $E|\mathbf{q}| < \infty$.

The restriction that the second-moment matrix of \mathbf{x} is nonsingular is made as a matter of convenience to rule out cases in which the entries of \mathbf{x} are linearly dependent. Among other things, this guarantees that the law of one price holds trivially for linear combinations of \mathbf{x} . If the moment restrictions imposed on \mathbf{x} and \mathbf{q} are not satisfied for an original vector of assets, then it is often possible to scale the payoffs and prices so that these restrictions are satisfied. A special case of such scaling occurs when all the payoffs are constructed to have a unit price as in the case of measured returns to holding securities between time 0 and time τ .

Applying the law of iterated expectations to the pricing relation (3) results in the following restriction.

RESTRICTION 1. $E\mathbf{q} = E\mathbf{x}m$.

We focus on the unconditional moment restriction 1 instead of the conditional moment restriction (3) because it is typically easier to estimate unconditional moments than conditional moments. Restriction 1, however, is in general weaker than (3). Gallant, Hansen, and Tauchen (1990) show how to extend some of the analysis in this paper by exploiting characterizations of the moments of \mathbf{x} conditioned on (possibly a subset of) I .

As long as consumers are not satiated in the numeraire consumption good at time τ , the IMRS should be strictly positive.

RESTRICTION 2. $m > 0$.

Restriction 2 is sufficient to imply the absence of arbitrage opportunities. That is, the restriction guarantees that nonnegative payoffs that are strictly positive with positive probability conditioned on I have positive prices. In the next two sections we explore the implica-

² Our use of the unconditional expectation operator in this context is justified formally when the time series converges appropriately to a stochastic steady state and is ergodic. In this case, unconditional expectations are computed using the stationary distribution. For processes that are asymptotically stationary but not ergodic, the limit points can often be represented as conditional expectations in which the conditioning occurs on the invariant sets for the approximating stationary stochastic process.

tions that restrictions 1 and 2 have for the mean and standard deviation of m .

So far, we have treated the case in which only a finite vector of asset payoffs and prices is investigated. In our subsequent analysis, it will be convenient to extend the pricing function and its unconditional expectation to the linear span of \mathbf{x} . Define $P \equiv \{\mathbf{c} \cdot \mathbf{x} : \mathbf{c} \in \mathbb{R}^n\}$.³ In Section IV, we shall also consider derivative claims formed by taking particular nonlinear functions of payoffs in P . In light of assumption 1, each portfolio payoff in P has a finite second moment. With this in mind we define a norm on P to be $\|p\| \equiv [E(p^2)]^{1/2}$. Notice that the standard deviation of a portfolio payoff p , denoted $\sigma(p)$, is given by $\|p - Ep\|$.

Since the portfolio payoffs in \mathbf{x} are linearly independent, for each p in P there is a unique \mathbf{c} in \mathbb{R}^n for which p is equal to $\mathbf{c} \cdot \mathbf{x}$. We extend the pricing function so that the prices of these payoffs are given by the corresponding linear combinations of \mathbf{q} : $\pi_I(\mathbf{c} \cdot \mathbf{x}) \equiv \mathbf{c} \cdot \mathbf{q}$. As required, π_I maps P into I . Notice that π_I is constructed so that (3) extends to the linear span \mathbf{x} : $\pi_I(p) = E(pm|I)$ for all p in P .

It is also of interest to define a functional π mapping portfolio prices into the *expected value* of the prices: $\pi(p) \equiv E\pi_I(p)$. Hence π maps P linearly into the real line \mathbb{R} . Again the law of iterated expectations implies that

$$\pi(p) = E(mp) \quad \text{for all } p \text{ in } P. \quad (4)$$

It is straightforward to show that restriction (4) is equivalent to restriction 1.

III. Implications of Restriction 1

In this section we characterize the volatility restrictions for m as implied by restriction 1. In subsection *A* we suppose that there is a unit payoff in P , while in subsection *B* we consider the more common case in which such a payoff is not included in P . Finally, in subsection *C* we describe how existing empirical methodologies in finance can be used to characterize these volatility restrictions.

A. Riskless Payoff

Suppose that P contains a payoff that is equal to one with probability one. In deriving implications for the volatility of m , we should first

³ Notice that a larger set P of portfolio payoffs could be constructed by following more complex trading strategies in which the vector \mathbf{c} is replaced by a vector of random variables in I . The theoretical analysis in Hansen and Richard (1987) is designed to accommodate this case as is the econometric analysis in Gallant et al. (1990). We focus on the linear span of \mathbf{x} for pedagogical convenience and empirical tractability.

construct a random variable m^* in P that satisfies restriction 1. This amounts to finding a vector α_o in \mathbb{R}^n such that

$$E\mathbf{x}\mathbf{x}'\alpha_o = E\mathbf{q}, \tag{5}$$

where $m^* = \mathbf{x} \cdot \alpha_o$. Solving (5) for α_o gives $\alpha_o = (E\mathbf{x}\mathbf{x}')^{-1}E\mathbf{q}$. Notice that α_o depends on the second moment of \mathbf{x} and the first moment of \mathbf{q} . Hence m^* can be constructed from asset market data.

Consider any other random variable m satisfying restriction 1. Since P contains a unit payoff, $Em = \pi(1) = Em^*$. Consequently, all random variables m that satisfy restriction 1 have the same mean, and this mean is equal to the expected price of a unit payoff. Also, $E[\mathbf{x}(m - m^*)] = 0$ because both m and m^* satisfy restriction 1. In other words, the discrepancy between m and m^* is orthogonal to the random vector \mathbf{x} . Since m^* is in P , m^* is the least-squares projection of m onto P and

$$\sigma^2(m) = \sigma^2(m^*) + \sigma^2(m - m^*).$$

Therefore, we have the following relations:

$$\sigma(m) \geq \sigma(m^*), \quad Em^* = Em. \tag{6}$$

The volatility bound in (6) is as sharp as possible because m^* satisfies restriction 1 by construction.

B. No Riskless Payoff

Next we consider the more usual case in which P does not contain a unit payoff. It turns out that much of the previous analysis can be exploited in analyzing this case. Let \mathbf{x}^a denote the $(n + 1)$ -dimensional random vector formed by augmenting \mathbf{x} with a unit payoff. Since $E\mathbf{x}\mathbf{x}'$ is nonsingular and no linear combination of \mathbf{x} is equal to one with probability one, $E\mathbf{x}^a\mathbf{x}^{a'}$ is also nonsingular. We build an augmented payoff space P^a containing a unit payoff by using \mathbf{x}^a in place of \mathbf{x} .

To apply the analysis in Section IIIA, we must assign a number v to $\pi(1)$, which is the expected price of a unit payoff. Such price data may not be available, and for this reason we examine implications for an array of hypothetical expected prices. Let v be any candidate for $\pi(1)$ and π_v the corresponding extension of π from P to P^a . We then replicate the analysis in subsection A to construct a random variable m_v in P^a such that

$$E\mathbf{x}m_v = E\mathbf{q}, \quad Em_v = v. \tag{7}$$

The counterpart to volatility bound (6) is

$$\sigma(m) \geq \sigma(m_v) \tag{8}$$

for any random variable m that satisfies restriction 1 and has mean v . This volatility bound is as sharp as possible because, by construction, m_v satisfies restriction 1 and has mean v .

We replicate the construction of m_v for all real numbers v and generate an indexed collection $\{m_v; v \in \mathbb{R}\}$ of random variables, each of which satisfies restriction 1. This collection is of interest because for any m satisfying restriction 1, the ordered pair $[Em, \sigma(m)]$ is in the region

$$S \equiv \{(v, w) \text{ in } \mathbb{R}^2: w \geq \sigma(m_v)\}. \quad (9)$$

This region summarizes the volatility implications for m implied by restriction 1. We refer to the boundary of S as being the mean–standard deviation frontier for IMRSs, and we refer to members of the set $\{m_v; v \in \mathbb{R}\}$ as being on this frontier.

It is of interest to derive an expression for $\sigma(m_v)$ that is both easy to compute and easy to interpret. The moment conditions in (7) can be rewritten as terms of the covariance of m and \mathbf{x} :

$$E[(\mathbf{x} - E\mathbf{x})(m_v - v)] = E\mathbf{q} - vE\mathbf{x}. \quad (10)$$

Now

$$m_v = (\mathbf{x} - E\mathbf{x})' \beta_v + v \quad (11)$$

for some β_v in \mathbb{R}^n because m_v is a linear combination of a unit payoff and the entries of \mathbf{x} and Em_v is v . Substituting (11) into (10) and solving for β_v give

$$\beta_v = \Sigma^{-1}(E\mathbf{q} - vE\mathbf{x}),$$

where Σ is the covariance matrix of \mathbf{x} . It follows that

$$\sigma(m_v) = [(E\mathbf{q} - vE\mathbf{x})' \Sigma^{-1} (E\mathbf{q} - vE\mathbf{x})]^{1/2}. \quad (12)$$

Notice that for a given v , $\sigma(m_v)$ depends only on the means of \mathbf{q} and \mathbf{x} and the covariance matrix of \mathbf{x} .

The standard deviation bound given in (12) has the following interpretation. Consider a risk-neutral valuation of the asset payoffs in which m is set to a constant value v for all states of the world. In this case the means of the prices should be proportional to the means of the asset payoffs with proportionality factor v . The bound in (12) is the square root of a quadratic form in the vector of deviations of the observed average prices from the average risk-neutral prices. For a fixed Σ , larger deviations from risk-neutral pricing imply larger bounds on the volatility of m . Shanken (1987) derived a related bound on the pricing error induced by using error-ridden proxies in computing the valuation of asset payoffs. When a constant v is used as a

proxy for m , the bound in (12) can be viewed as a special case of Shanken's bound (see his proposition 1, pp. 93–94).

C. Relation to Empirical Models of Asset Prices

In this subsection we derive the relation between the mean–standard deviation frontier for m and the mean-variance frontier for asset returns. This latter frontier is the focal point of the static capital asset pricing model. The link we deduce between the two frontiers provides an alternative interpretation of the volatility bounds for m . We then describe how linear factor restrictions as imposed in Ross's (1976a) arbitrage pricing model (see also Chamberlain 1983; Chamberlain and Rothschild 1983; Connor 1984) can be used to characterize the mean–standard deviation frontier for m .

Define

$$R \equiv \{p \text{ in } P: \pi(p) = 1\}. \tag{13}$$

When the vector q is not random, R is the collection of (gross) returns on portfolios in P . More generally, R contains all the payoffs in P with *expected prices* that are equal to one.

Consider first the case in which P contains a unit payoff and $\pi(1)$ is different from zero. Then $1/\pi(1)$ is in R . A second payoff in R is $r^* \equiv m^*/\pi(m^*)$. Note that $\pi(m^*) = E[(m^*)^2]$, and hence

$$\|r^*\| = \frac{\|m^*\|}{\|m^*\|^2} = \frac{1}{\|m^*\|}. \tag{14}$$

Furthermore, Hansen and Richard (1987) established that r^* is the payoff in R that has the smallest norm (second moment). Consequently, r^* is the solution to the following optimization problem:

$$\underset{r \text{ in } R}{\text{minimize}} \sigma(r) \quad \text{subject to } Er = \mu$$

when μ is set equal to Er^* . Therefore, m^* is proportional to a particular payoff on the mean–standard deviation frontier for R .

To relate the bound for $\sigma(m)$ given in (8) to the slope of the mean–standard deviation frontier for R , note that

$$\frac{\sigma(m)}{Em} \geq \frac{\sigma(m^*)}{Em^*} = \frac{\sigma(r^*)\|m^*\|^2}{Em^*} = \frac{\sigma(r^*)}{Er^*}. \tag{15}$$

Recall that the second moment of a random variable r satisfies $E(r^2) = \sigma(r)^2 + (Er)^2$. Since P contains a unit payoff, the mean–standard deviation frontier for R is a cone with apex at $[0, 1/\pi(1)]$ and axis parallel to the horizontal axis. In order for r^* to be the minimum second-moment payoff in R , the ordered pair $[\sigma(r^*), Er^*]$ must occur

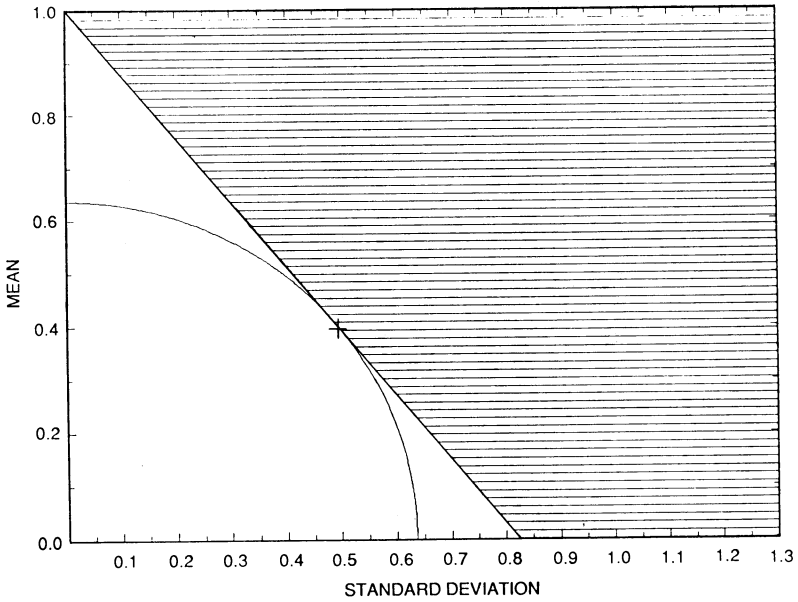


FIG. 2.—Minimum second-moment payoff in R

at the tangency of a circle with center $(0, 0)$ and the lower (inefficient) portion of the mean–standard deviation frontier for R . This tangency point is depicted in figure 2. Since the lower portion of the frontier is a ray from $[0, 1/\pi(1)]$ through $[\sigma(r^*), Er^*]$, the slope of this ray is the *Sharpe ratio* of the payoff r^* , $\{Er^* - [1/\pi(1)]\}/\sigma(r^*)$, and the slope of the circle with center $(0, 0)$ that passes through $[\sigma(r^*), Er^*]$ is $-\sigma(r^*)/Er^*$. Therefore,

$$\frac{\sigma(r^*)}{Er^*} = \frac{[1/\pi(1)] - Er^*}{\sigma(r^*)}. \quad (16)$$

In light of (15) and (16), the bound on the ratio $\sigma(m)/Em$ is given by the absolute value of the slope of the mean–standard deviation frontier for R .⁴ These relations demonstrate the precise sense in which a

⁴ An alternative way to derive the bound is as follows. Consider any payoff z with an expected price equal to zero. By restriction 1, $Ezm = 0$. It follows from the covariance decomposition and the Cauchy-Schwarz inequality that $\sigma(m)/Em \geq |Ez|/\sigma(z)$. The sharpest bound on $\sigma(m)/Em$ is given by the zero expected price payoff z with the largest Sharpe ratio $|Ez|/\sigma(z)$. As is well known from financial economics, the largest Sharpe ratio is equal to the absolute value of the slope of the mean–standard deviation frontier when R contains a riskless payoff.

steep slope of a mean–standard deviation frontier for asset payoffs can imply a potentially dramatic bound on the volatility of m .

Next we consider the case in which P does not contain a unit payoff and hence R does not contain an unconditionally riskless payoff. We follow the strategy used in subsection B by augmenting \mathbf{x} with a unit payoff and assigning this payoff an expected price v . This results in an expansion of R to R_v , where $1/v$ is now in R_v . Let r_v^* denote the payoff in R_v with the smallest second moment. Since r_v^* is on the mean–standard deviation frontier for R_v , it is well known from static capital asset pricing theory that r_v^* is a linear combination (with coefficients that sum to one) of $1/v$ and any other distinct return on the mean–standard deviation frontier for R_v . As long as $1/v$ is not equal to the mean of the minimum variance payoff in R , we can find a payoff r_v that is on the mean–standard deviation frontier for both R and R_v . Also, for each v the variable m_v is proportional to r_v^* . Therefore, with one exception, for each random variable m_v on the mean–standard deviation frontier for IMRSs, there is a corresponding payoff r_v on the mean–standard deviation frontier for R such that m_v is a linear combination of r_v and a unit payoff. In this sense the mean–standard deviation frontier for IMRSs can be thought of as the *dual* of the mean–standard deviation frontier for R . The exceptional case occurs when $1/v$ is the mean of the minimum variance payoff in R . In this case, m_v is a linear combination of a unit payoff and a payoff that is on the mean–standard deviation frontier for the space of payoffs with expected prices equal to zero.

The impact of augmenting R with $1/v$ can be seen graphically by passing a ray from the point $(0, 1/v)$ through a tangent point on the mean–standard deviation frontier for R . One side of the mean–standard deviation frontier for the augmented set R_v is given by this tangent ray, and the other is a reflection about a horizontal ray from $(0, 1/v)$. This construction is displayed graphically in figure 3. In the special case in which $1/v$ is the mean of the minimum variance payoff in R , it is not possible to draw a tangent line to the mean–standard deviation frontier of R from the point $(0, 1/v)$. Instead the frontier for R_v is given by the two asymptotes.

Once the frontier for the augmented set R_v is obtained, the construction illustrated in figure 2 can be mimicked using R_v in place of R . Thus for any m with mean v that satisfies restriction 1,

$$\frac{\sigma(m)}{Em} \geq \frac{\sigma(m_v)}{v} = \frac{\sigma(r_v^*)}{Er_v^*} = \frac{[1/\pi(1)] - Er_v^*}{\sigma(r_v^*)}. \tag{17}$$

The relations in (17) show the connection between the volatility bound on m 's with mean v and the slope of the mean–standard devia-

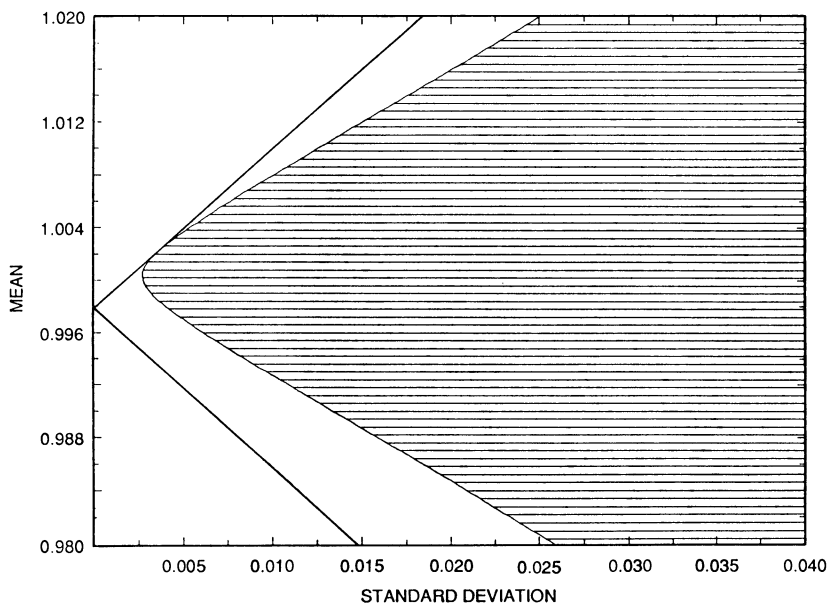


FIG. 3.—Mean-standard deviation frontiers for R and R_v

tion frontier for R_v . A steeper slope of the frontier for R_v implies a correspondingly sharper volatility bound for m .⁵

Since the mean-standard deviation frontier for R is known to have a two-fund characterization, the preceding results show that the mean-standard deviation frontier for m can be represented using two distinct frontier payoffs in R . For the general class of asset pricing models considered in this paper, there is no prediction that particular payoffs in R , say the returns on the wealth portfolios of consumers, are mean-variance efficient. Thus without additional restrictions, there is no guidance on how to reduce, a priori, a potentially large collection of portfolio payoffs into a small collection used in a time-series analysis.

One ad hoc approach that is often used to reduce the dimensionality of the collection of payoffs is factor analysis as employed in empirical arbitrage pricing models (see, e.g., Connor and Korajczyk 1988; Lehmann and Modest 1988). Suppose that P is generated by a se-

⁵ Using conditioning information in clever ways can sharpen the volatility bounds on m by increasing the maximum Sharpe ratio of the payoffs in R_v . For example, Breen, Glosten, and Jagannathan (1989) show that information in Treasury bill returns can be used to construct a portfolio that has the same average return as the value-weighted index of New York Stock Exchange securities but is only half as variable.

quence $\{p_j\}$, where

$$p_j = \gamma_j \cdot \mathbf{f} + e_j \quad (18)$$

and \mathbf{f} is a vector of common factors for all the payoffs.⁶ Often, the factors \mathbf{f} are in an appropriately defined span of $\{p_j\}$. Hence, it follows from the law of one price that there exists a unique vector $\pi(\mathbf{f})$ of hypothetical expected prices for the factor payoffs. One possible strategy for deducing volatility bounds on m is to use the extensive collection of payoffs $\{p_j\}$ (or possibly a subset of it) to identify the first two moments of \mathbf{f} and the expected price vector $\pi(\mathbf{f})$. A region S then could be constructed from these factor moments and prices using formula (12).

In general, information is lost in going from the larger space P to the smaller space F of linear combinations of factors. Tests of factor models of asset pricing examine whether the pricing relation

$$\pi(p_j) = \gamma_j \cdot \pi(\mathbf{f}) \quad (19)$$

holds at least approximately. When (19) holds exactly, the regions S generated by P and F coincide. Therefore, if asset payoffs can be priced in terms of a small number of factors \mathbf{f} , there is no loss to constructing the region S from F instead of the larger space P .⁷

As argued in Hansen and Richard (1987), an unconditional factor decomposition as in (18) may not be very appealing when economic agents can use conditioning information in I to make investments. If the factor decomposition (18) is conditioned on an information set I and γ_j is a vector of random variables in I , a reduction in payoffs is more complicated but still feasible.

IV. Implications of Restriction 2

In Section III, we showed how to construct minimum variance random variables that satisfy restriction 1. These random variables may be negative with positive probability and hence may fail to satisfy

⁶ Although our derivation of the volatility bounds for m assumed that the payoff space P is finite-dimensional, this restriction was made for pedagogical convenience. In fact the duality relation between the mean–standard deviation frontiers for m 's that satisfy restriction 1 and for payoffs in R extends to environments in which P is generated by an infinite number of payoffs, say by $\{p_j\}$.

⁷ In contrast to factor analytic approaches, Huberman and Kandel (1987) test whether the dimensionality of P can be reduced to a prespecified *observed* subset of security returns, namely three size-based portfolios of New York Stock Exchange securities. In this case, F can be constructed using these three returns. Huberman and Kandel find, however, that this construction of F is not adequate to span the mean–standard deviation frontier for the original P constructed using 33 size-sorted portfolios. Hence in this case the dimensionality reduction from P to F will result in weaker implications for m .

restriction 2.⁸ As long as we limit ourselves to candidate IMRSs that are translations of payoffs in P , it may not be possible to ensure that frontier random variables are strictly positive or, for that matter, nonnegative.

In this section we initially replace restriction 2 by a weaker requirement that m be nonnegative. We then construct minimum variance candidates for m among the class of nonnegative random variables satisfying restriction 1. It turns out that these minimum variance random variables can be interpreted as either European call or put options on payoffs in P . Recall that when the payoff on the underlying portfolio is p and the strike price is k , a European call option entitles an investor to the payoff $\max\{p - k, 0\}$ and a put option to $\max\{k - p, 0\}$. These payoffs are clearly nonnegative, but they may be nonlinear functions of \mathbf{x} . The resulting volatility bounds for nonnegative random variables satisfying restriction 1 also apply when the random variables are restricted to be strictly positive (satisfy restriction 2). However, in this case the lower bounds may only be approximated rather than attained.

This section is divided into three subsections. In subsection *A* we suppose that there is a unit payoff in P , while in subsection *B* we consider the more common case in which such a payoff is not included in P . Finally, in subsection *C* we discuss the close connection between our analysis and work by Harrison and Kreps (1979) and Kreps (1981) on the viability of equilibrium pricing functions consistent with the absence of arbitrage opportunities.

A. Riskless Payoff

First consider the case in which there is a unit payoff in P . For each p in P , let p^+ denote $\max\{p, 0\}$. Note that for any p' in P and any nonnegative strike price k that is proportional to the unit payoff, the payoffs $p' - k$ and $k - p'$ are also in P . Therefore, the collection of all random variables p^+ for some p in P includes the payoffs on European call and put options with constant strike prices.

Suppose that we weaken restriction 2 to the requirement that m be nonnegative. By construction, all derivative claims of the form p^+ for payoffs p in P are nonnegative. It turns out that the minimum variance nonnegative random variable \bar{m} satisfying restriction 1 is given by such a derivative claim. Hence we are led to the problem of finding

⁸ As Dybvig and Ingersoll (1982) have pointed out, naive use of m^* to compute (expected) prices of contingent claims may lead to assignment of negative (expected) prices to some positive payoffs and hence to the appearance of an arbitrage opportunity.

a vector α_0 in \mathbb{R}^n such that

$$E[\mathbf{x}(\mathbf{x}'\alpha_0)^+] = E\mathbf{q}, \tag{20}$$

where $\tilde{m} = (\mathbf{x}'\alpha_0)^+$. In what follows we shall first show that \tilde{m} , when it exists, has the smallest variance among all nonnegative random variables m satisfying restriction 1. We then discuss the existence and computation of a solution to (20).

To show that \tilde{m} has the smallest variance, consider any other nonnegative random variable m satisfying restriction 1. Clearly, $E\mathbf{x}m = E\mathbf{x}\tilde{m}$. Exploiting the nonnegativity of m , we have that

$$E\tilde{m}m \geq \alpha_0'E\mathbf{x}m = \alpha_0'E\mathbf{x}\tilde{m} = E[(\tilde{m})^2]. \tag{21}$$

It follows from the Cauchy-Schwarz inequality that $\|m\| \geq \|\tilde{m}\|$. Since P contains a unit payoff, both m and \tilde{m} must have the same mean. Therefore, we have the following relations:⁹

$$\sigma(m) \geq \sigma(\tilde{m}), \quad Em = E\tilde{m}. \tag{22}$$

Next we ask whether the volatility bound in (22) can be sharpened by requiring m to be strictly positive instead of nonnegative. If \tilde{m} is strictly positive (with probability one), then clearly the answer is no. This can occur only when \tilde{m} coincides with m^* computed in Section IIIA. Consider the case in which m^* is not strictly positive with probability one, and let m be any random variable satisfying restrictions 1 and 2. Then \tilde{m} is zero with positive probability, and it follows from (21) that

$$0 < \|m - \tilde{m}\|^2 = \|m\|^2 - 2E\tilde{m}m + \|\tilde{m}\|^2 \leq \|m\|^2 - \|\tilde{m}\|^2.$$

Therefore, at the very least the weak inequality (\geq) in (22) is replaced by the strong inequality ($>$). In fact, no further improvements are possible. To see this form,

$$m_j = \left(1 - \frac{1}{j}\right)\tilde{m} + \left(\frac{1}{j}\right)m. \tag{23}$$

Then m_j is strictly positive and $\{\sigma(m_j)\}$ converges to $\sigma(\tilde{m})$. Therefore, $\sigma(\tilde{m})$ is in fact the greatest lower bound for $\sigma(m)$ when m is restricted to satisfy restrictions 1 and 2.

⁹ An alternative way to deduce these bounds is to exploit the fact that when payoffs on calls and puts are included in the analysis, the space of admissible payoffs is essentially complete (see Ross 1976*b*; Breeden and Litzenberger 1978; Arditti and John 1980; Green and Jarrow 1987). If the prices of all such payoffs were available, the counterpart to the random variable m^* in Sec. III would be strictly positive. Although this extensive collection of option price data is typically not available, we can follow Merton (1973) and use lower bounds on option prices to obtain a lower bound on the volatility of m .

Equation system (20) is nonlinear in the parameter vector α_o , and its solution cannot necessarily be represented in terms of matrix manipulations. There is a closely related optimization problem whose solution may be easier to compute. This problem entails finding a payoff in R whose truncation has the smallest second moment:

$$\min_{r \text{ in } R} \|r^+\|^2. \quad (24)$$

In Appendix A we show that (24) has a solution, although this solution may not be unique. Furthermore, a necessary and sufficient condition for \bar{r} to be a solution to (24) is

$$E(\bar{r}^+ z) = 0 \quad \text{for all } z \text{ in } P \text{ such that } \pi(z) = 0. \quad (25)$$

We can think of (25) as being the first-order condition for optimization problem (24).

It turns out that we can construct a solution to (20) by scaling \bar{r}^+ appropriately. Let

$$\bar{m} = \frac{\bar{r}^+}{\|\bar{r}^+\|^2}. \quad (26)$$

This scaling is permissible because $\|\bar{r}^+\|$ must be strictly positive as long as there exists at least one random variable m satisfying restrictions 1 and 2. To see this, suppose to the contrary that $\|\bar{r}^+\|$ is zero. Then $-\bar{r}$ is a nonnegative payoff with a strictly negative expected price. Such a payoff is inconsistent with restriction 2 because it implies that there exists an arbitrage opportunity.

Clearly \bar{m} as given by (26) can be represented as $(\alpha'_o \mathbf{x})^+$ for some α_o in \mathbb{R}^n . To verify that α_o solves (20), we must show that \bar{m} as given by (26) satisfies restriction 1. Let p be any payoff in P and form the payoff $z = p - \pi(p)\bar{r}$. Note that $\pi(z) = 0$ because $\pi(\bar{r}) = 1$. It follows from first-order condition (25) that

$$\begin{aligned} 0 &= E\bar{m}z = E\bar{m}p - \pi(p)E\bar{m}\bar{r} \\ &= E\bar{m}p - \pi(p) \frac{E[\bar{r}^+ \bar{r}]}{\|\bar{r}^+\|^2} \\ &= E\bar{m}p - \pi(p). \end{aligned}$$

Thus \bar{m} satisfies restriction 1 as required.

This construction of \bar{m} parallels a similar construction reported in Hansen and Richard (1987) and in Section III. If one ignores restriction 2, one way to construct the random variable m^* , which has minimum variance among the class of random variables satisfying restriction 1, is to compute the minimum second-moment payoff, r^* , in R and divide it by its second moment, $\|r^*\|^2$. We have just demonstrated

that a similar strategy works for constructing a random variable \tilde{m} that attains the volatility bound among the class of nonnegative random variables satisfying restriction 1. Instead of computing the minimum second-moment payoff in R , we calculate the minimum *truncated* second-moment payoff, \tilde{r} , in R . To form \tilde{m} , the truncation of this payoff, \tilde{r}^+ , is divided by the second moment of its truncation, $\|\tilde{r}^+\|^2$. Whereas $\|m^*\|$ is given by $1/\|r^*\|$, $\|\tilde{m}\|$ is given by $1/\|\tilde{r}^+\|$. Since truncating a random variable reduces its norm, as required, \tilde{m} has a larger second moment than m^* . The difference in the two norms reflects the incremental contribution of restriction 2 for the volatility bound on m .

One advantage to solving optimization problem (24) instead of solving directly the nonlinear equation system (20) is that optimization problem (24) has a convex objective function $\|r^+\|^2$ and a convex constraint set R so that numerical solutions are quite feasible to obtain. Although \tilde{r} is not necessarily unique, its truncation \tilde{r}^+ is (see App. A). A sufficient condition for \tilde{r} to be unique, which is often satisfied in practice, is that no two payoffs in R have the same truncation.

B. No Riskless Payoff

Consider the more common case in which P does not contain a unit payoff. As in Section IIIB, augment \mathbf{x} with a unit payoff and form an augmented payoff space P^a . Similarly, assign alternative strictly positive numbers v for $\pi(1)$ and extend π from P to P^a . Let R_v be the augmented set of payoffs with expected prices equal to one when $\pi(1)$ is assigned v . The counterpart to equation (20) is not guaranteed to have a solution, however. It turns out that there are additional limits on the admissible choices of v consistent with restriction 2.

To investigate these limits, we study the counterparts to optimization problem (24) using the augmented space of payoffs R_v in place of R . Define

$$\delta_v \equiv \inf_{r \text{ in } R_v} \|r^+\|^2. \tag{27}$$

When δ_v is positive, the bound on $\|m\|^2$ among the class of nonnegative random variables satisfying restrictions 1 and 2 with mean v is $1/\delta_v$. However, particular choices of v may result in δ_v being zero and hence $1/\delta_v$ being infinite. For instance, when there is a portfolio payoff p in P such that p is less than or equal to one with probability one and v is strictly less than $\pi(p)$, δ_v is zero. This is true because the random variable $(1 - p)/[v - \pi(p)]$ is in R_v and is less than or equal to zero with probability one. Consequently, the norm of its truncation is zero.

As noted by Merton (1973), Cox, Ross, and Rubinstein (1979), Harrison and Kreps (1979), and Kreps (1981), it is possible to obtain arbitrage bounds on the admissible (expected) prices that can be assigned to payoffs not in P . In the case of a unit payoff, the upper and lower bounds are given by

$$\bar{\pi}(1) \equiv \begin{cases} \inf\{\pi(p): p \geq 1\} & \text{if } \{p \text{ in } P: p \geq 1\} \text{ is not empty} \\ +\infty & \text{otherwise,} \end{cases}$$

$$\underline{\pi}(1) \equiv \sup\{\pi(p): p \leq 1\},$$

respectively.¹⁰ Since the zero payoff is in P , $\underline{\pi}(1)$ is always nonnegative. The arbitrage bounds $\underline{\pi}(1)$ and $\bar{\pi}(1)$ determine the range of admissible values of Em that are compatible with m being a nonnegative random variable. Clearly, $\{v: \delta_v > 0\}$ must be a subset of the interval $[\underline{\pi}(1), \bar{\pi}(1)]$. In fact, the interiors of these sets coincide (see App. A).

When δ_v is strictly positive, there exists a minimum variance, nonnegative random variable with mean v that satisfies restriction 1 (see lemma A4 in App. A). Let this random variable be denoted \tilde{m}_v . The corresponding volatility bound is

$$\sigma(m) \geq \sigma(\tilde{m}_v), \quad (28)$$

and the family of random variables, $\{\tilde{m}_v: \delta_v > 0\}$, comprises the mean–standard deviation frontier for nonnegative random variables satisfying restriction 1. Thus the counterpart to the region S given in (9) is

$$S^+ \equiv \{(v, w): \delta_v > 0 \text{ and } w \geq \sigma(\tilde{m}_v)\}. \quad (29)$$

The set S^+ is convex. To see this consider two values of v for which δ_v is strictly positive, say $v(l) \leq v(u)$. Form convex combinations of the random variables $\tilde{m}_{v(l)}$ and $\tilde{m}_{v(u)}$. These convex combinations are nonnegative random variables that also satisfy restriction 1. Recall that the mean of a convex combination of random variables is equal to the convex combination of the means, and by the triangle inequality, the standard deviation of a convex combination is less than or equal to the convex combination of the standard deviations. While convex combinations of $\tilde{m}_{v(l)}$ and $\tilde{m}_{v(u)}$ are not necessarily on the mean–standard deviation frontier, the ordered pairs of their means and standard deviations must be in S^+ . This is sufficient for S^+ to be convex.

¹⁰ The characterization reported in Harrison and Kreps (1979) and Kreps (1981) is somewhat more complicated because they allow the counterpart to the space P to be infinite-dimensional.

Next we consider the incremental contribution of requiring m to be strictly positive as in restriction 2. It is shown in Appendix A that Em must be in the open interval $(\underline{\pi}(1), \overline{\pi}(1))$ (see lemma A6). Hence one effect of the imposition of strict positivity is that endpoints of the interval $\{v: \delta_v > 0\}$ are eliminated.

For v 's in $(\underline{\pi}(1), \overline{\pi}(1))$, \tilde{m}_v can be interpreted as either a call or put option on a payoff in P . More precisely, we obtain the counterpart to the result in Section IVA that $\tilde{m}_v = (\tilde{p}_v - k)^+$, where \tilde{p}_v is a portfolio payoff in P and k is in \mathbb{R} . When k is nonnegative, \tilde{m}_v is a call option on a portfolio with payoff \tilde{p}_v and strike price k ; when k is negative, \tilde{m}_v is a put option on a portfolio with payoff $-\tilde{p}_v$ and strike price $-k$. Therefore, for any v in $(\underline{\pi}(1), \overline{\pi}(1))$, the counterpart to equation (20) has a solution.

As in subsection A, this gives us a simple check of the incremental impact of positivity on the volatility bounds given in (28). For any v in $(\underline{\pi}(1), \overline{\pi}(1))$ such that \tilde{m}_v is strictly positive, the bound in (28) cannot be improved by restricting m to be strictly positive. This can occur only when \tilde{m}_v coincides with m_v calculated in Section IIIB. On the other hand, for any v for which m_v is not strictly positive, the weak inequality (\geq) in (28) is replaced by a strong inequality ($>$).

Even though S^+ may be a proper subset of S , the region S is still of interest for a variety of reasons. First, S is easier to use in practice because a characterization of S^+ may require that a nonquadratic optimization problem be solved for each value of v . Second, the lower boundaries of S^+ and S coincide for values of v for which m_v is nonnegative. Consequently, it is advantageous to characterize S as a first step in characterizing S^+ and then check for nonnegativity of m_v . Finally, even for values of v for which m_v is negative with positive probability, the coefficients on \mathbf{x}^a given in representation (11), when scaled appropriately, can be used as starting values for a numerical search routine used in computing δ_v .

In figure 4 we report plots of the regions S and S^+ for the same financial data set that was used to generate figure 1. The region S^+ is shaded, and the lower boundary of the region S is given by the dashed line below S^+ . While the lower boundaries of these regions coincide for points closest to the horizontal axis, they diverge for other points. The divergence between the boundaries is greater when the volatility bounds are more restrictive. Recall that in generating the lower boundary of S , we constructed random variables m_v with mean v that satisfy restriction 1 and are linear combinations of \mathbf{x}^a . When these random variables have large standard deviations relative to their means, it is not surprising that they are negative with high probability. As a result, the positivity restriction 2 often has more bite when $\sigma(m_v)/v$ is larger.

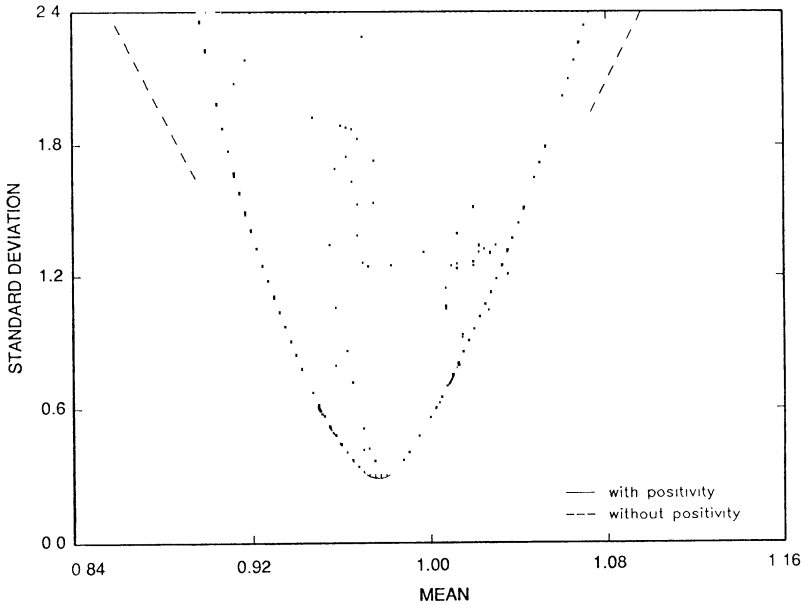


FIG. 4.—IMRS frontier with and without positivity imposed

As is true for S , the dimensionality of P can sometimes be reduced prior to the construction of S^+ . Suppose that members of P have factor decompositions of the form $p = \gamma \cdot \mathbf{f} + e$, where \mathbf{f} is a vector of common factors. Suppose further that the idiosyncratic components of the payoffs satisfy $E(e|\mathbf{f}) = 0$ and $\pi(e) = 0$. Hence we have exact factor pricing and each payoff p in P is a mean-preserving spread of a payoff $\gamma \cdot \mathbf{f}$ with the same price. Consequently, $\|p^+\|^2 \geq \|(\gamma \cdot \mathbf{f})^+\|^2$ because the function $[(p)^+]^2$ of p is convex. Hence when one solves (24) or (27), it suffices to restrict attention to linear combinations of the factors with expected prices equal to one. While it is evident how to use this reduction when the factors are observed, unobserved factors are problematic because it may be difficult to compute or estimate $\|(\gamma \cdot \mathbf{f})^+\|^2$ for arbitrary vectors γ . Because of the truncation of $(\gamma \cdot \mathbf{f})$, calculating $\|(\gamma \cdot \mathbf{f})^+\|^2$ requires knowledge of the entire probability distribution of \mathbf{f} , whereas typical factor analytic procedures identify only the first two moments of \mathbf{f} .

C. Viability of Equilibrium Pricing Functions and Arbitrage Pricing

The analysis in this section is intimately connected to general treatments of pricing derivative claims (see, e.g., Ross 1978; Harrison and

Kreps 1979; Kreps 1981). Among other things, Harrison and Kreps and Kreps consider the following question. Given a set of payoffs on primitive securities and the prices of those securities, when is it possible to extend the pricing function to a larger collection of payoffs in such a way as to preserve no arbitrage? As emphasized by Kreps, this experiment should not be construed as introducing new markets in an economy that might alter the resulting competitive equilibrium allocations. It is merely a hypothetical extension leaving intact the (expected) prices of the payoffs in P . When such an extension is possible, Harrison and Kreps and Kreps refer to the pricing function as being *viable*.

Throughout the analysis in this section, we have presumed that the family of m 's that satisfy restrictions 1 and 2 is not empty. Clearly this is sufficient to eliminate arbitrage opportunities on P . Rather than assume that this family is not empty, an alternative starting point is to verify that no arbitrage opportunities exist on P and then to appeal to theorem 3 in Kreps (1981) to show that π can be extended from P to the collection L^2 of all random variables that are (Borel measurable) functions of \mathbf{x} and have finite second moments.¹¹ The existence of an m satisfying restrictions 1 and 2 then follows from the Riesz representation theorem applied to L^2 (see also lemma 2.3 in Hansen and Richard [1987]).

V. Illustrations and Discussion

We now illustrate our analysis with alternative parametric models of m and alternative data sets on asset payoffs and prices. The model of m described in the Introduction and used to generate figure 1 assumed that consumers' preferences are separable over time and states of the world. In subsection *A* we investigate the impact on m of relaxing time separability. In subsection *B* we focus on logarithmic risk preferences but do not require these preferences to be state separable. Finally, in subsection *C* we describe the implications of price data on short-term Treasury bills for IMRSs and comment briefly on the implications for monetary models.

A. Preferences That Are Not Time Separable

Consider the following stylized version of a model with time nonseparabilities in preferences. As in the Introduction, we use a time- and

¹¹ In addition to a *no-arbitrage* restriction, Kreps also imposed a *no-free-lunch* restriction on (P, π) . As demonstrated by Clark (1990), this extra restriction is not needed when P is a closed subspace of an L^2 space, as is true in our analysis.

state-separable specification of preferences for *consumption services* with a power utility function

$$E \sum_{t=0}^{\infty} \lambda^t \frac{s_t^{\gamma+1} - 1}{\gamma + 1}, \quad (30)$$

except now s_t depends on measured consumption in the current period and one previous period:

$$s_t = c_t + \theta c_{t-1}. \quad (31)$$

More general versions of this model have been investigated by Dunn and Singleton (1986), Eichenbaum, Hansen, and Singleton (1988), Gallant and Tauchen (1989), and Eichenbaum and Hansen (1990). We shall proceed as if there is a single representative consumer. As noted by Wilson (1968) and Rubinstein (1974), this assumption can be relaxed when θ is zero, the consumption allocations are consistent with the existence of complete contingent-claims markets, and all consumers have the same preferences. This aggregation result also applies more generally, say when θ is different from zero, as long as there are, in effect, complete markets in consumption services (see Eichenbaum, Hansen, and Richard 1987). When θ is positive, consumption generates positive services in the current as well as in one subsequent time period. In this case there is *intertemporal substitution* in generating consumption services from consumption goods. More precisely, there is a durable component to consumption that depreciates fully after one time period. Alternatively, when θ is negative, there is *intertemporal complementarity* in generating consumption services from consumption goods. Put somewhat differently, the term $-\theta c_{t-1}$ is a component of current-period consumption that reflects either *committed consumption* from the previous time period or *habit persistence*. Sundaresan (1989), Constantinides (1990), and Novales (1990) have argued that habit persistence may be important in explaining the relation between asset market data and economic aggregates.

For these forms of time nonseparabilities, the marginal utility of consumption is

$$mu_{\tau} = (s_{\tau})^{\gamma} + \lambda \theta E[(s_{\tau+1})^{\gamma} | I_{\tau}].$$

The IMRS between time 0 and time τ is the corresponding ratio of marginal utilities scaled by λ^{τ} . Constructing m requires computation of the conditional expectation $E[(s_{\tau+1})^{\gamma} | I_{\tau}]$ except in the special case in which θ is zero.

To illustrate what impact positive and negative values of θ have for the volatility of m , we report calculations from Gallant et al. (1990).

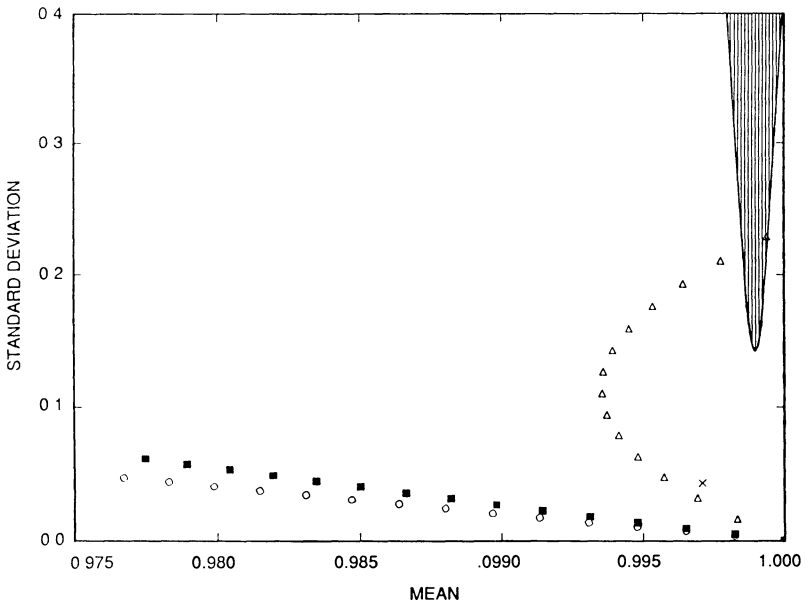


FIG. 5.—IMRS frontier computed using monthly data

For these calculations the ratio c_t/c_{t-1} is a component of a Markov process with a stochastic law of motion estimated by Gallant and Tauchen (1989) for monthly data on the consumption of nondurables and services (for more details, see Gallant and Tauchen [1989] and Gallant et al. [1990]). The estimated law of motion was then used to compute $E[(s_{t+1})^\gamma | I_t]$ required in forming a time series for m . Sample means and standard deviations were calculated for m 's implied by alternative values of γ and θ .¹²

For this illustration we let $\theta = -.5$, $\theta = 0$, and $\theta = .5$. The results are reported in figure 5. The boxes are used to denote mean–standard deviation pairs for $\theta = 0$, the circles for $\theta = .5$, and the triangles for $\theta = -.5$. For each choice of θ , we let γ range from zero to -14 with decrements of minus one. In all cases the subjective discount factor λ is set to one. Smaller values of λ decrease proportionately the mean and standard deviation of m . When $\gamma = 0$, m is one for all choices of θ . In this case, $[E(m), \sigma(m)] = (1, 0)$.

Consider first the case in which $\theta = 0$. Increasing $|\gamma|$ magnifies the volatility of m but initially reduces its mean. Extrapolated much fur-

¹² Note that the calculations of mean–standard deviation pairs for m when $\theta = 0$ do not exploit the Markov specification estimated by Gallant and Tauchen (1989) and are consequently more robust.

ther, the curve (indexed by γ) does not turn around until $|\gamma|$ is in the vicinity of 100; afterward, increasing $|\gamma|$ enlarges the mean of m . The initial decline in the mean of m reflects the dominant role of positive growth rates in consumption. For extremely large values of $|\gamma|$, observations with negative growth rates in consumption come to dominate the sample mean, eventually resulting in a change of slope of the curve. In comparing the curves denoted by boxes in figures 1 and 5, recall that the long annual time series used to generate figure 1 contains negative growth rate observations on consumption during the depression. The absence of *bad* events in the monthly data set is responsible for the fact that the curve (indexed by γ) does not turn until the magnitude of γ is substantial.¹³

Consider next the case in which $\theta = .5$. Not surprisingly, introducing this local durability into preferences reduces the volatility of m . The quantitative effect of this smoothing does not appear to be very substantial, however. The curves for $\theta = .5$ and $\theta = 0$ are similar for the range of γ 's that are plotted. Hence there is little adverse effect on the volatility of m to introducing durability by setting $\theta = .5$.

Finally, consider the case in which $\theta = -.5$. This intertemporal complementarity has the anticipated impact of increasing the volatility of m for a given value of γ . This effect is quite dramatic as indicated in figure 5. Furthermore, the value of $|\gamma|$ at which the curve turns is reduced dramatically. For $\theta = -.5$ the turning point for γ is in the vicinity of -7 , and the initial decline in the mean of m is much less dramatic.

We now compare the three curves, which describe alternative mean–standard deviation pairs for parametric models of m , to a region S^+ generated using monthly data on asset payoffs and prices. The asset market data are the same as those used by Hansen and Singleton (1982) except that data revisions were incorporated and more recent data points were included. The resulting time period is 1959:3–1986:12. The first two asset payoffs are the 1-month real return on Treasury bills and the 1-month real value-weighted return on the New York Stock Exchange. Six additional time series of asset payoffs were constructed using these data by scaling the original two payoffs and prices by the one-period lagged returns and the one-period lag in the consumption ratio. For the range of hypothetical means considered, the region S described in Section III was essentially the same as the region S^+ described in Section IV.

¹³ The *sample* volatility of m may be substantially lower than the *population* volatility if consumers anticipate that extremely bad events can occur with small probability when such events do not occur in the sample. Reitz (1988) argued that this phenomenon could explain the equity premium puzzle.

For the specification of preferences with $\theta = 0$, larger values of $|\gamma|$ initially make the mean–standard deviation pair for m further from the S^+ region because of the adverse effect on the mean of m . This is consistent with the fact that Hansen and Singleton (1982) found point estimates for γ that were close to zero but substantial evidence against the overidentifying restrictions. As emphasized by Singleton (1990), estimates of the discount factor λ are often greater than one when bond returns are included in the analysis. For a fixed γ , enlarging λ has the desired effect of increasing proportionately the mean and standard deviation of m .

From the vantage point of figure 5, the case for intertemporal complementarities in preferences is appealing. For a given value of γ , a negative value of θ increases both the mean and the standard deviation of m . However, it is quite possible for m to have a mean and standard deviation in S^+ and not satisfy restriction 1. In other words, for a given parametric specification of m , requiring $[E(m), \sigma(m)]$ to be in S^+ does not exhaust the testable implication of restriction 1. As emphasized by Gallant et al. (1990), there is substantial statistical evidence that the resulting m 's violate restriction 1. In fact, empirical studies that use similar data and preference specifications, such as Dunn and Singleton (1986), Eichenbaum, Hansen, and Singleton (1988), and Eichenbaum and Hansen (1990), typically find parameter estimates that reflect intertemporal substitution ($\theta > 0$), although they find statistical evidence against the resulting parametric model of m .

B. Logarithmic Risk Preferences

In (30) and (31), suppose that θ is zero and γ is minus one. In this case, preferences are logarithmic. As noted by Rubinstein (1976), m is equal to the reciprocal of the return on the wealth portfolio of the representative consumer between time 0 and time τ (see also Brown and Gibbons 1985). Epstein and Zin (this issue) showed that this same conclusion applies to a parametric class of recursive preferences that are not state separable so long as the risk preferences remain logarithmic. Whereas in the state-separable case the return on the wealth portfolio is equal to the discounted consumption ratio, this exact relation no longer applies when state separability is relaxed. Nevertheless, the return on the wealth portfolio can still be used as a valid measure of m .

For this reason we have included a “cross” in figures 1 and 5. In figure 1, this cross denotes the sample mean–standard deviation pair for the reciprocal of the measured annual return on the Standard and Poor's 500 stock price index, and in figure 5 it represents the sample mean–standard deviation pair for the reciprocal of the mea-

sured monthly value-weighted return on the New York Stock Exchange. In both cases the means are near the points in S^+ that are closest to the horizontal axis, but the mean–standard deviation pair is outside of S^+ . However, after a sampling error is taken, there is very little evidence against the null hypothesis that this model of m satisfies restriction 1.¹⁴

C. Treasury Bill Data and Monetary Models

We also calculated the regions S and S^+ using monthly data on 3-month holding period returns on Treasury bills. The holding period returns were constructed using bond prices on 3-, 6-, 9-, and 12-month discount bonds for 1964:7–1986:12. Nominal returns were converted to real returns using the implicit price deflator on nondurables and services. These bond price data (with the most recent periods excluded) have been used by Fama (1984), Dunn and Singleton (1986), and Stambaugh (1988), among others, to investigate time variation in risk premia and particular models of bond prices. In figure 6 we report the resulting regions S and S^+ . The region S^+ is shaded, and the lower boundary of the region S is given by the dashed line below S^+ . The resulting standard deviation bounds for m are quite striking. For means of m in the vicinity of one, the bound on the standard deviation is near one. Given the magnitude of these bounds, it is not surprising that restriction 2 has an important incremental contribution vis-à-vis restriction 1.

The bounds reported in figure 6 appear to us to pose quite a challenge to a large class of asset valuation models. For instance, the quarterly counterpart to the $\theta = 0$ curve in figure 5 ranges from (1, 0) to (.90, .08) as γ ranges from zero to -14 . Volatility bounds of a similar magnitude were also obtained using monthly data on 1-month holding period returns for Treasury bills with maturities from 1 to 6 months. These bounds are directly comparable to the three curves plotted in figure 5. Since these latter bounds apply to IMRSs measured over a shorter time period (1 month instead of 3 months), they are even more startling. However, short-term Treasury bills are often held to maturity, and trading of these Treasury bills takes place in secondary markets except for the 3-, 6-, and 12-month bills. The bid-ask spreads for the short-term bills can be quite substantial (see Stambaugh 1988; Knez, Litterman, and Scheinkman 1989), so the

¹⁴ In fig. 1, one of the two moment conditions, $E(m\mathbf{x} - \mathbf{q}) = 0$, is satisfied by construction. The other condition was tested using the method suggested in Hansen and Singleton (1982): the $\chi^2(1)$ statistic is 1.40 with probability value .24. Similarly, for fig. 5, four of the eight moment conditions are satisfied by construction. The $\chi^2(4)$ statistic for the four remaining conditions is 4.88 with a probability value of .30.

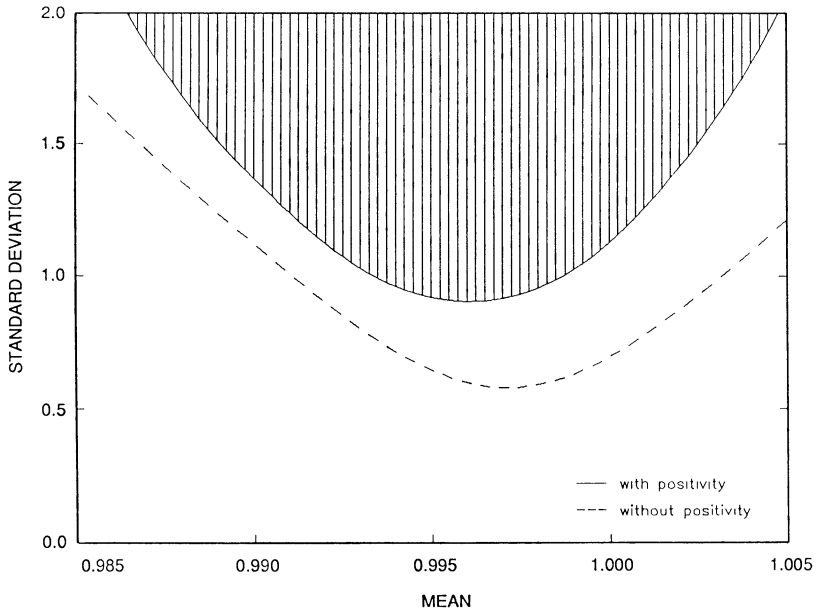


FIG. 6.—IMRS frontier computed using quarterly holding period returns

prices used in our calculations may be less reliable. These concerns should be less problematic for the results in figure 6 since they were computed using price data from the more richly traded 3-, 6-, 9-, and 12-month Treasury bills.

As emphasized by Knez et al., short-term Treasury bills often may be held to maturity as cash substitutes for particular transactions. As such, these bills may generate important liquidity services that are not measured appropriately by the implied ex post real returns. Hence the measured real returns may understate the value of the assets to the holders of the securities. Recall from Section IIIC that large standard deviation bounds for m occur when the slope of the mean–standard deviation frontier for R is steep. For the Treasury bill data, this means that the reason that the volatility bounds on m are large (as reflected in the region S) is that the expected short-term gain associated with holding longer-term bills is large relative to the increase in the standard deviation. Abstracting from the liquidity services of the short-term bills may distort the magnitude of the resulting volatility bounds on m .

Refinements of real asset pricing models to incorporate money, such as the cash-in-advance models of Svensson (1985), Lucas and Stokey (1987), and Townsend (1987), are designed to accommodate the rate of return dominance between one-period bonds and money.

However, in their current form they are not well suited to differentiate among short-term Treasury bills with different maturity dates. Although the link between measured *real* IMRSs and security market data may be confounded in these models, there is an alternative notion of the *indirect* IMRS for money that reflects the fact that the cash-in-advance constraint may not always be binding. Hence for an appropriate interpretation of m , these monetary models are compatible with restrictions 1 and 2 as long as money is not included among the vector of assets used to generate S^+ .

VI. Conclusions and Extensions

In this paper we have characterized the implications of security market data for means and standard deviations of IMRSs. This exercise is important in evaluating alternative models of dynamic economies because IMRSs are the channels by which the attributes of these models impinge on asset prices. Abstracting from the restriction that IMRSs are positive, we established the connection between volatility bounds on IMRSs and mean–standard deviation frontiers for asset payoffs. Thus we showed how diagnostics commonly used in empirical finance can be translated into information about IMRSs. We also showed how to extract sharper volatility bounds by taking account of the fact that IMRSs should be positive. These sharper bounds exploit more fully the absence of arbitrage opportunities in the underlying economic environment than, say, linear factor representations of asset prices.

There are three important directions in which the ideas in this paper can be developed further. An earlier version of this paper has already provoked some work along these three lines.

i) In this paper we focused exclusively on deriving implications for IMRSs expressed in terms of population moments of asset payoffs and prices. In practice, these attributes of asset market data will not be known a priori, but can be approximated only by using time-series averages in place of population moments. This introduces sampling error into the analysis. A major drawback of the discussion in Section V is that it abstracted from the presence of approximation error introduced by using sample averages from historical time series in place of population moments. Hansen and Jagannathan (1990) show how to use large sample theory both to assess whether there is sufficient statistical evidence to reject that the bounds are degenerate (equal to zero) and to assess the magnitude of the approximation errors.

ii) The restrictions on IMRSs derived in this paper all pertain to first and second moments. More generally, it would be desirable to

characterize the admissible family of distributions for IMRSs given asset market data. An additional step toward such a characterization is taken by Snow (1990), who shows how to extend the analysis in this paper to obtain bounds on other moments of the IMRSs.

iii) The diagnostics derived in this paper can be applied to any intertemporal asset pricing model for which moments of m can be computed. While the calculations in Section V were performed by first constructing hypothetical time series on m , such a construction is not necessary. All that is really essential is the ability to compute the moments of m implied by the model. As an alternative to constructing a time series on m , these moments can be deduced from the equilibrium stochastic law of motion for the model (see, e.g., Heaton 1990). Therefore, calculations like those illustrated in Section V can be performed for an extensive array of intertemporal asset pricing models including models that take account of measurement errors in consumption, seasonality, and aggregation-over-time biases.

Appendix A

Let L^2 be the Hilbert space of all random variables with finite second moments that are Borel measurable functions of \mathbf{x} . Let P be a closed linear subspace of L^2 and π be a continuous linear functional on P . In contrast to the analysis in the text, we allow the space P to be infinite-dimensional. Define $R \equiv \{r \in P: \pi(r) = 1\}$, $R^+ \equiv \{r^+: r \in R\}$, and $Z \equiv \{z \in P: \pi(z) = 0\}$. Throughout our analysis we assume that R is not empty. Let C denote the closure (in L^2) of R^+ . In this Appendix we establish several results that support conclusions in Section IV.

Consider the following two minimum norm problems. The first problem is

$$\delta \equiv \inf_{r \in R} \|r^+\|^2. \tag{P1}$$

A closely related minimum norm problem is

$$\eta \equiv \inf_{y \in C} \|y\|^2. \tag{P2}$$

This second problem has the advantage that the inf is attained.

There are two additional problems that are closely related to (P1) and (P2). The first one is an orthogonality problem:

$$\text{Find } \hat{y} \in C \text{ such that } E(\hat{y}z) = 0 \text{ for all } z \in Z. \tag{P3}$$

As in standard minimum norm problems on Hilbert spaces, it is often the case that (P3) has the same solution as (P1) and (P2). The focal point of our analysis is the following problem:

$$\text{Find } y^* \in L^2 \text{ such that } y^* \geq 0 \text{ and } \|y^*\|^2 = \frac{1}{\delta} \text{ and } \pi(p) = E(y^*p) \text{ for all } p \in P. \tag{P4}$$

We now investigate the relation among these four problems. First we establish the connection between (P1) and (P2).

LEMMA A1. There is a unique \bar{y} in C such that $\|\bar{y}\|^2 = \delta = \eta$.

Proof. Let $\{r_j\}$ be a sequence in R such that $\{\|(r_j)^+\|^2\}$ converges to δ . Then for any positive integers j and k ,

$$\begin{aligned} \|(r_j)^+ - (r_k)^+\|^2 &= -\|(r_j)^+ + (r_k)^+\|^2 + 2\|(r_j)^+\|^2 + 2\|(r_k)^+\|^2 \\ &\leq -\|(r_j + r_k)^+\|^2 + 2\|(r_j)^+\|^2 + 2\|(r_k)^+\|^2 \end{aligned}$$

because $\|(r_j)^+ + (r_k)^+\|^2 \geq \|(r_j + r_k)^+\|^2$. Since r_j and r_k are both in R , $(r_j/2) + (r_k/2)$ is also in R . Consequently, $\|[(r_j/2) + (r_k/2)]^+\|^2 \geq \delta$, and

$$\begin{aligned} \|(r_j)^+ - (r_k)^+\|^2 &\leq -4 \left\| \left(\frac{r_j}{2} + \frac{r_k}{2} \right)^+ \right\|^2 + 2\|(r_j)^+\|^2 + 2\|(r_k)^+\|^2 \quad (A1) \\ &\leq -4\delta + 2\|(r_j)^+\|^2 + 2\|(r_k)^+\|^2. \end{aligned}$$

If we take limits as $j, k \rightarrow \infty$, it follows that $\{(r_j)^+\}$ is Cauchy and hence converges to some \bar{y} in C . Therefore, $\{\|(r_j)^+\|^2\}$ converges to $\|\bar{y}\|^2 = \delta$.

Since C is the closure of R^+ , for any y in C there is a sequence $\{(r_j)^+\|^2\}$ that converges to $\|y\|^2$. Therefore, $\eta = \delta$.

Finally, let \hat{y} be any member of R^+ for which $\|\hat{y}\|^2 = \delta$, and let $\{\hat{r}_j\}$ be a sequence in R such that $\{(\hat{r}_j)^+\}$ converges to \hat{y} . Analogous to (A1),

$$\|(r_j)^+ - (\hat{r}_j)^+\|^2 \leq -4\delta + 2\|(r_j)^+\|^2 + 2\|(\hat{r}_j)^+\|^2.$$

Since $\{\|(r_j)^+\|^2\}$ and $\{\|(\hat{r}_j)^+\|^2\}$ both converge to δ , $\{(r_j)^+\}$ and $\{(\hat{r}_j)^+\}$ have the same limit points. Therefore, \hat{y} and \bar{y} are equal (with probability one). Q.E.D.

Next we establish the connection between (P2) and (P3).

LEMMA A2. A solution \bar{y} to (P2) is also a solution to (P3).

Proof. To prove this result we use the following inequality:

$$[(r + cz)^+]^2 \leq (r^+ + cz)^2. \quad (A2)$$

To see that it holds, first suppose that $r + cz \leq 0$. In this case the left side of (A2) is zero while the right is greater than or equal to zero. Second, suppose that $r + cz \geq 0$. Then $0 \leq r + cz \leq r^+ + cz$, which also implies (A2).

Let $\{r_j\}$ be a sequence in R such that $\{(r_j)^+\}$ converges to \bar{y} , and let z be any member of Z distinct from zero. Then

$$\liminf_{j \rightarrow \infty} \|(r_j + cz)^+\|^2 \leq \|\bar{y} + cz\|^2. \quad (A3)$$

The right side of (A3) is minimized by $\bar{c} = -E(\bar{y}z)/E(z^2)$. In order for \bar{y} to be the solution to (P2), it must be that $\bar{c} = 0$ or, equivalently, that $E\bar{y}z = 0$. Q.E.D.

Lemma A2 has the following partial converse.

LEMMA A3. If $\hat{r}^+ \in R^+$ is the solution to (P3) and the solution to (P2) is in R^+ , then \hat{r}^+ is the solution to (P2).

Proof. Let \bar{r}^+ denote the solution to problem (P2). It follows from lemma A2 that $E[(\bar{r}^+ - \hat{r}^+)(\bar{r} - \hat{r})] = 0$. Also, $(\bar{r}^+)^+ \hat{r} \leq (\bar{r}^+)^+(\hat{r}^+)^+$ and $\bar{r}(\hat{r}^+)^+ \leq (\bar{r}^+)^+(\hat{r}^+)^+$. Hence

$$0 = E[(\bar{r}^+ - \hat{r}^+)(\bar{r} - \hat{r})] \geq E[(\bar{r}^+ - \hat{r}^+)^2] \geq 0.$$

Therefore, $\bar{r}^+ = \hat{r}^+$ (with probability one). Q.E.D.

We now use the Hahn-Banach theorem and the Riesz representation theorem to establish the existence of a solution to (P4).

LEMMA A4. If $\delta > 0$, (P4) has a solution.

Proof. The first half of this proof follows closely the proof of lemma 1 in Kreps (1981). Since $\delta > 0$,

$$\pi(p) \leq \left(\frac{1}{\delta}\right)^{1/2} \|p^+\| \quad \text{for all } p \in P. \tag{A4}$$

Among other things, inequality (A4) implies that $\pi(p) \geq 0$ whenever $p \geq 0$ because π is linear and $(-p)^+$ is zero. The right side of (A4), $(1/\delta)\|(\cdot)^+\|$, is a particular version of the sublinear function used by Kreps in applying the Hahn-Banach theorem to extend P to a larger space, say L^2 . (The analogues to the spaces P and L^2 are much more general in Kreps's analysis.) Let Π denote such an extension. Then Π satisfies the counterpart to (A4):

$$\Pi(y) \leq \left(\frac{1}{\delta}\right)^{1/2} \|y^+\| \quad \text{for all } y \in L^2. \tag{A5}$$

Clearly Π is continuous and $\Pi(y) \geq 0$ whenever $y \geq 0$. It follows from the Riesz representation that there exists a $y^* \in L^2$ such that

$$\Pi(y) = E(y^*y) \quad \text{for all } y \in L^2. \tag{A6}$$

It remains to show that $y^* \geq 0$ and $\|y^*\| = (1/\delta)^{1/2}$. Consider any $r \in R$ and note that $\Pi(r^+) \geq \Pi(r) = 1$. Since (A6) is satisfied, it follows from the Cauchy-Schwarz inequality that $\|y^*\| \|r^+\| \geq E(y^*r^+) \geq 1$. Consequently,

$$\|y^*\| \delta^{1/2} = \|y^*\| \inf_{r \in R} \|r^+\|^2 \geq 1,$$

or, equivalently, $\|y^*\| \geq (1/\delta)^{1/2}$. Relations (A5) and (A6) imply

$$\|y^*\|^2 = \Pi(y^*) \leq \left(\frac{1}{\delta}\right)^{1/2} \|(y^*)^+\| \leq \left(\frac{1}{\delta}\right)^{1/2} \|y^*\|.$$

Therefore, $\|y^*\| = \|(y^*)^+\| = (1/\delta)^{1/2}$. Q.E.D.

For our next set of results we find it convenient to restrict (P, π) to satisfy the no-arbitrage condition:

For any $p \in P$ such that $p \geq 0$ and $\|p\| > 0$, $\pi(p) > 0$.

LEMMA A5. If (P, π) satisfies the no-arbitrage condition, then $\delta > 0$ and the solution to (P2) is in R^+ .

Proof. Let $\{r_j\}$ be a sequence in R such that $\{(r_j)^+\}$ converges to \bar{y} , where \bar{y} is a solution to (P2). Our goal is to show that there exists a convergent subsequence of $\{r_j\}$ with limit payoff \bar{r} . Given this convergence, we then argue that $(\bar{r})^+ = \bar{y}$.

First we show that $\{\|r_j\|^2\}$ is bounded. Suppose to the contrary that $\{\|r_j\|^2\}$ is unbounded. Without loss of generality, we may assume that this sequence is increasing (otherwise we could extract a subsequence that is increasing and unbounded). Form $p_j \equiv r_j/\|r_j\|$. Since the closed unit ball of L^2 is weakly compact, $\{p_j\}$ has a subsequence that converges weakly to a random variable \bar{p} in L^2 . The weak limit of a sequence in P must be orthogonal to the orthogonal complement of P . In other words, \bar{p} must be in P . Furthermore, $\pi(\bar{p}) = 0$ because $\pi(p_j) = 1/\|r_j\|$ and $\{\|r_j\|\}$ is increasing and unbounded. The sequence $\{(p_j)^+\}$ converges strongly to zero because $\|(p_j)^+\| = \|(r_j)^+\|/\|r_j\|$ and $\{(r_j)^+\}$ converges strongly to \bar{y} . Form the orthogonal decomposition $p_j = (p_j)^+ - (-p_j)^+$.

Then $\{\|(-p_j)^+\|\}$ converges to one and a subsequence of $\{(-p_j)^+\}$ converges weakly to $-\hat{p}$. It follows from Clark (1990, lemma 2) that $-\hat{p} \geq 0$ and $\|-\hat{p}\| > 0$. This finding violates the no-arbitrage condition because $\pi(-\hat{p}) = 0$. Therefore, $\{\|r_j\|\}$ must be bounded.

Since $\{\|r_j\|\}$ is bounded, $\{r_j\}$ has a subsequence that converges weakly to a limit point \bar{r} in R . Consequently, $\{(-r_j)^+\}$ has a subsequence that converges weakly to $\bar{y} - \bar{r} \geq 0$. Note that

$$0 = E[(r_j)^+(-r_j)^+] = E\{(-r_j)^+[(r_j)^+ - \bar{y}]\} + E\{(-r_j)^+\bar{y}\}. \tag{A7}$$

It follows from the Cauchy-Schwarz inequality that

$$\begin{aligned} |E\{(-r_j)^+[(r_j)^+ - \bar{y}]\}| &\leq \|(-r_j)^+\| \|(r_j)^+ - \bar{y}\| \\ &\leq \|r_j\| \|(r_j)^+ - \bar{y}\|. \end{aligned} \tag{A8}$$

The right side of (A8) converges to zero because $\{\|r_j\|\}$ is bounded and $\{(r_j)^+\}$ converges strongly to \bar{y} . If we take limits of the right side of (A7) along the weakly convergent subsequence, it follows that $\bar{y} - \bar{r}$ is orthogonal to \bar{y} . Since both of these random variables are nonnegative and $\bar{r} = \bar{y} + (\bar{r} - \bar{y})$, $(\bar{r})^+ = \bar{y}$.

To verify that $\delta > 0$, suppose to the contrary that $\delta = 0$. In this case, $-\bar{r} \geq 0$, implying a violation of the no-arbitrage condition because $\pi(-\bar{r}) = -1$ and π is linear. Q.E.D.

In light of lemmas A1, A2, A3, and A5, when (P, π) satisfies the no-arbitrage condition, the solutions to (P1)–(P3) coincide and are in R^+ . As is shown in Section IV, in this case a solution to (P4) is given by $y^* = \bar{y}/\|\bar{y}\|^2$, where \bar{y} is the solution to (P1)–(P3).

Consider now the special case in which 1 is not in P . As in Section IV, let $P^a \equiv P + \{1\}$ and define the arbitrage bounds

$$\begin{aligned} \bar{\pi}(1) &\equiv \inf\{\pi(p) : p \geq 1\}, \\ \underline{\pi}(1) &\equiv \sup\{\pi(p) : p \leq 1\}. \end{aligned}$$

Extend π from P to P^a by assigning v to 1, and let π_v denote the resulting extension.

LEMMA A6. Suppose that (P, π) satisfies the no-arbitrage condition. Then (P^a, π_v) satisfies the no-arbitrage condition if and only if $v \in (\underline{\pi}(1), \bar{\pi}(1))$.

Proof. Part of this result is an implication of theorem 4 in Kreps (1981). For completeness we include a simple proof.

Suppose that $v \in (\underline{\pi}(1), \bar{\pi}(1))$. Let $p + w \geq 0$ for some $p \in P$ and some $w \neq 0$. If $w > 0$, then $p/(-w) \leq 1$ and $\pi[p/(-w)] \leq \underline{\pi}(1) < v$. Hence $\pi(p) + vw > 0$. A similar argument applies to the case in which $w < 0$.

Next suppose that (P^a, π_v) satisfies the no-arbitrage condition. Then clearly $v \in [\underline{\pi}(1), \bar{\pi}(1)]$. Suppose that $v = \underline{\pi}(1)$. Consider a sequence $\{p_j\}$ such that $p_j \leq 1$, $\pi(p_j) \leq v$, and $\{\pi(p_j)\}$ converges to v . Since 1 is not in P and P is closed, the lim inf of the sequence $\{\|(1 - p_j)\|\}$ is strictly positive. Form the sequence $\{(1 - p_j)/\|(1 - p_j)\|\}$. Following the logic of the proof of lemma A5, this sequence has a weakly convergent subsequence with a nonnegative limit point of the form $1 - p_0$ for some $p_0 \in P$, where $v - \pi(p_0) = 0$ and $\|1 - p_0\| > 0$. However, this violates the no-arbitrage condition. A similar argument applies when $v = \bar{\pi}(1)$. Therefore, v must be in the open interval $(\underline{\pi}(1), \bar{\pi}(1))$. Q.E.D.

Appendix B

In this Appendix we describe in more detail the series used to perform the calculations underlying each of the figures.

Figures 1 and 4.—For a description of the stock, bond, and consumption data, see table 1 of Campbell and Shiller (1988) under the heading Cowles/S&P 500, 1871–1986.

Figures 2 and 3.—Monthly observations for 1959:4–1986:12 on 1-month holding period returns on 1-, 2-, 3-, 4-, 5-, and 6-month Treasury bills were constructed using bond prices from the Fama term structure yield file of data tapes from the Center for Research in Security Prices (CRSP) at the University of Chicago. Nominal returns were converted to real returns using the implicit price deflator for consumption of nondurables and services from the Personal Consumption Expenditure data tape of the National Income and Product Accounts.

Figure 5.—Monthly observations on the 1-month return on Treasury bills and on the 1-month value-weighted return on the New York Stock Exchange were taken from the CRSP data tapes. Nominal returns were converted to real returns using the implicit price deflator for the consumption of nondurables and services. Monthly observations on eight series of asset payoffs were constructed using these two returns. The first two payoffs are the two original returns. The prices of these payoffs are one by construction. The second two payoffs were formed by multiplying the two returns by the one-period lagged value of the real Treasury bill return. The prices of these two payoffs are equal to the one-period lag of the real Treasury bill return. The third two payoffs were formed by multiplying the original two returns by the one-period lagged value of the real value-weighted return. The prices of these two payoffs are equal to the one-period lag of the real value-weighted return. Finally, the last two payoffs are the original two payoffs multiplied by the ratio of per capita real consumption in the two previous time periods. The prices of the last two payoffs are both equal to the lagged consumption ratio. The time period is 1959:3–1986:12.

The consumption series was taken from the Personal Consumption Expenditure data tape of the National Income and Product Accounts, and the total population series from the Citibase data tape.

Figure 6.—The bond prices were taken from the Fama term structure yield file of the CRSP data tapes. Four monthly time series of 3-month holding period returns were constructed from the monthly price data on 3-, 6-, 9-, and 12-month discount bonds. Nominal returns were converted to real returns using the monthly implicit deflator for consumption of nondurables and services described previously. The time period is 1964:7–1986:12.

References

- Arditti, Fred D., and John, Kose. "Spanning the State Space with Options." *J. Financial and Quantitative Analysis* 15 (March 1980): 1–9.
- Breeden, Douglas T. "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities." *J. Financial Econ.* 7 (September 1979): 265–96.
- Breeden, Douglas T., and Litzenberger, Robert H. "Prices of State-Contingent Claims Implicit in Option Prices." *J. Bus.* 51 (October 1978): 621–51.

- Breen, William; Glosten, Lawrence R.; and Jagannathan, Ravi. "Economic Significance of Predictable Variations in Stock Index Returns." *J. Finance* 44 (December 1989): 1177-89.
- Brown, David P., and Gibbons, Michael R. "A Simple Econometric Approach for Utility-based Asset Pricing Models." *J. Finance* 40 (June 1985): 359-81.
- Campbell, John Y., and Shiller, Robert J. "Dividend-Price Ratios and Expectations of Future Dividends and Discount Factors." *Rev. Financial Studies* 1 (Fall 1988): 195-228.
- Chamberlain, Gary. "Funds, Factors, and Diversification in Arbitrage Pricing Models." *Econometrica* 51 (September 1983): 1305-23.
- Chamberlain, Gary, and Rothschild, Michael. "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets." *Econometrica* 51 (September 1983): 1281-1304.
- Clark, Stephen A. "The Valuation Problem in Arbitrage Price Theory." *J. Math. Econ.* 19 (1990), in press.
- Connor, Gregory. "A Unified Beta Pricing Theory." *J. Econ. Theory* 34 (October 1984): 13-31.
- Connor, Gregory, and Korajczyk, Robert A. "Risk and Return in an Equilibrium APT: Application of a New Test Methodology." *J. Financial Econ.* 21 (September 1988): 255-89.
- Constantinides, George M. "Habit Formation: A Resolution of the Equity Premium Puzzle." *J.P.E.* 98 (June 1990): 519-43.
- Cox, John C.; Ross, Stephen A.; and Rubinstein, Mark. "Option Pricing: A Simplified Approach." *J. Financial Econ.* 7 (September 1979): 229-63.
- Dunn, Kenneth B., and Singleton, Kenneth J. "Modeling the Term Structure of Interest Rates under Non-separable Utility and Durability of Goods." *J. Financial Econ.* 17 (July 1986): 27-55.
- Dybvig, Philip H., and Ingersoll, Jonathan E., Jr. "Mean-Variance Theory in Complete Markets." *J. Bus.* 55 (April 1982): 233-51.
- Eichenbaum, Martin S., and Hansen, Lars Peter. "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data." *J. Bus. and Econ. Statis.* 8 (January 1990): 53-69.
- Eichenbaum, Martin S.; Hansen, Lars Peter; and Richard, Scott F. "Aggregation, Durable Goods and Nonseparable Preferences in an Equilibrium Asset Pricing Model." Working Paper no. 87-9. Chicago: Nat. Opinion Res. Center, Program Quantitative Analysis, 1987.
- Eichenbaum, Martin S.; Hansen, Lars Peter; and Singleton, Kenneth J. "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice under Uncertainty." *Q.J.E.* 103 (February 1988): 51-78.
- Epstein, Larry G., and Zin, Stanley E. "Substitution, Risk Aversion, and Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis." *J.P.E.*, this issue.
- Fama, Eugene F. "The Information in the Term Structure." *J. Financial Econ.* 13 (December 1984): 509-28.
- Gallant, Ronald; Hansen, Lars Peter; and Tauchen, George. "Using Conditional Moments of Asset Payoffs to Infer the Volatility of Intertemporal Marginal Rates of Substitution." *J. Econometrics* 45 (July/August 1990): 141-79.
- Gallant, Ronald, and Tauchen, George. "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications." *Econometrica* 57 (September 1989): 1091-1120.
- Green, R. C., and Jarrow, R. A. "Spanning and Completeness in Markets with Contingent Claims." *J. Econ. Theory* 41 (February 1987): 202-10.

- Hansen, Lars Peter. "Consumption, Asset Markets, and Macroeconomic Fluctuation: A Comment." *Carnegie-Rochester Conf. Ser. Public Policy* 17 (Autumn 1982): 239–50.
- Hansen, Lars Peter, and Jagannathan, Ravi. "Detection of Specification Errors in Stochastic Discount Factor Models." Manuscript. Chicago: Univ. Chicago, 1990.
- Hansen, Lars Peter, and Richard, Scott F. "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica* 55 (May 1987): 587–613.
- Hansen, Lars Peter, and Singleton, Kenneth J. "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models." *Econometrica* 50 (September 1982): 1269–86.
- Harrison, J. Michael, and Kreps, David M. "Martingales and Arbitrage in Multiperiod Securities Markets." *J. Econ. Theory* 20 (June 1979): 381–408.
- Heaton, John. "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications." Manuscript. Cambridge: Massachusetts Inst. Tech., 1990.
- Huberman, Gur, and Kandel, Shmuel. "Mean-Variance Spanning." *J. Finance* 42 (September 1987): 873–88.
- Knez, Peter; Litterman, Robert; and Scheinkman, José A. "Explorations into Factors Explaining Money Market Returns." Manuscript. New York: Goldman, Sachs, Financial Strategies Group, 1989.
- Kreps, David M. "Arbitrage and Equilibrium in Economies with Infinitely Many Commodities." *J. Math. Econ.* 8 (March 1981): 15–35.
- Lehmann, Bruce N., and Modest, David M. "The Empirical Foundations of the Arbitrage Pricing Theory." *J. Financial Econ.* 21 (September 1988): 213–54.
- LeRoy, Stephen F. "Risk Aversion and the Martingale Property of Stock Prices." *Internat. Econ. Rev.* 14 (June 1973): 436–46.
- Lucas, Robert E., Jr. "Asset Prices in an Exchange Economy." *Econometrica* 46 (November 1978): 1429–45.
- Lucas, Robert E., Jr., and Stokey, Nancy L. "Money and Interest in a Cash-in-Advance Economy." *Econometrica* 55 (May 1987): 491–514.
- Mehra, Rajnish, and Prescott, Edward C. "The Equity Premium: A Puzzle." *J. Monetary Econ.* 15 (March 1985): 145–61.
- Merton, Robert C. "Theory of Rational Option Pricing." *Bell J. Econ. and Management Sci.* 4 (Spring 1973): 141–83.
- Novales, Alfonso. "Solving Nonlinear Rational Expectations Models: A Stochastic Equilibrium Model of Interest Rates." *Econometrica* 58 (January 1990): 93–111.
- Reitz, Thomas A. "The Equity Risk Premium: A Solution." *J. Monetary Econ.* 22 (July 1988): 117–31.
- Ross, Stephen A. "The Arbitrage Theory of Capital Asset Pricing." *J. Econ. Theory* 13 (December 1976): 341–60. (a)
- . "Options and Efficiency." *Q.J.E.* 90 (February 1976): 75–89. (b)
- . "A Simple Approach to the Valuation of Risky Streams." *J. Bus.* 51 (July 1978): 453–75.
- Rubinstein, Mark. "An Aggregation Theorem for Securities Markets." *J. Financial Econ.* 1 (September 1974): 225–44.
- . "The Valuation of Uncertain Income Streams and the Pricing of Options." *Bell J. Econ.* 7 (Autumn 1976): 407–25.
- Shanken, Jay. "Multivariate Proxies and Asset Pricing Relations: Living with the Roll Critique." *J. Financial Econ.* 18 (March 1987): 91–110.

- Shiller, Robert J. "Consumption, Asset Markets and Macroeconomic Fluctuations." *Carnegie-Rochester Conf. Ser. Public Policy* 17 (Autumn 1982): 203–38.
- Singleton, Kenneth J. "Specification and Estimation of Intertemporal Asset Pricing Models." In *Handbook of Monetary Economics*, edited by Benjamin Friedman and Frank H. Hahn. Amsterdam: North-Holland, 1990.
- Snow, Karl N. "Diagnosing Asset Pricing Models Using the Distribution of Asset Returns." Manuscript. Chicago: Univ. Chicago, 1990.
- Stambaugh, Robert F. "The Information in Forward Rates: Implications for Models of the Term Structure." *J. Financial Econ.* 21 (May 1988): 41–70.
- Sundaresan, Mahadevan. "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth." *Rev. Financial Studies* 2 (1989): 73–89.
- Svensson, Lars E. O. "Money and Asset Prices in a Cash-in-Advance Economy." *J.P.E.* 93 (October 1985): 919–44.
- Townsend, Robert M. "Asset-Return Anomalies in a Monetary Economy." *J. Econ. Theory* 41 (April 1987): 219–47.
- Wilson, Robert B. "The Theory of Syndicates." *Econometrica* 36 (January 1968): 119–32.