NBER WORKING PAPER SERIES

A REVEALED PREFERENCE ANALYSIS OF ASSET PRICING UNDER RECURSIVE UTILITY

Larry G. Epstein
Angelo Melino

Working Paper No. 4524

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November, 1993

We are grateful to the Social Sciences and Humanities Research Council of Canada for financial support and to Phillipe Weil for helpful discussions. This paper is part of NBER's research program in Asset Pricing. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #4524
November 1993

A REVEALED PREFERENCE ANALYSIS OF ASSET PRICING UNDER RECURSIVE UTILITY

ABSTRACT

This paper considers a representative agent model of asset prices based on a recursive utility specification. A constant elasticity of intertemporal substitution is assumed but the risk-preference component of utility is restricted only by qualitative, nonparametric regularity conditions. The principal contribution is to determine the exhaustive implications of this semiparametric recursive utility model for the one-step ahead joint probability distribution for consumption growth and asset returns.

Larry G. Epstein
Department of Economics
University of Toronto
Toronto, Ontario M5S 1A1
CANADA

Angelo Melino
Department of Economics
University of Toronto
Toronto, Ontario M5S 1A1
CANADA

A REVEALED PREFERENCE ANALYSIS OF ASSET PRICING UNDER RECURSIVE UTILITY

Larry G. Epstein, University of Toronto

Angelo Melino*, University of Toronto

August 30, 1993

1 INTRODUCTION

While the consumption-based model of asset pricing (C-CAPM) is predominant in macroeconomics and important in finance, its empirical difficulties when faced with aggregate time series data have been recognized for some time (see, for example, Grossman and Shiller (1981), Hansen and Singleton (1983) and Mehra and Prescott (1985)). In this model, equilibrium relations are derived from the optimizing behaviour of a representative agent whose intertemporal utility function is both state separable (conforms to expected utility theory) and time separable (the von Neumann-Morgenstern index is intertemporally additive). A number of researchers have considered more

*We are grateful to the Social Sciences and Humanities Research Council of Canada for financial support and to Phillipe Weil for helpful discussions.

general utility functions in an attempt to remedy the noted empirical failures. The class of generalizations known as recursive utility (Epstein and Zin(1989)), that relaxes the assumption of state separability, is the focus of this paper.

Recursive intertemporal utility permits a degree of separation between certainty preferences (the elasticity of intertemporal substitution) and risk preferences (the degree of risk aversion). A number of papers have, with mixed results, tested the equilibrium relations implied by various parametric functional forms for each of these components of recursive utility. In this paper, we are concerned with the empirical restrictions implied by the recursive utility model itself, unaccompanied by ad hoc functional form specifications. We argue below that the problem of functional form specification is particularly problematic for risk preferences. Therefore, we adopt the common CES specification for certainty preferences, but we impose only qualitative restrictions, such as monotonicity and risk aversion, on risk preferences. Consequently, our model is nonparametric with respect to risk preferences and semiparametric overall.

Our prinicipal contribution is to determine the exhaustive implications of this semiparametric recursive utility model for the law of motion of the joint stochastic process governing consumption growth and asset returns, that is, for the one-step ahead joint probability distribution for these variables. These implications form the counterpart for our model of the strong axiom of revealed preference in consumer demand theory. Moreover, our arguments are constructive. Therefore, when our revealed preference conditions are satisfied, we can describe the complete class of recursive utility functions that rationalize the data. In particular, we can describe the set of elasticity of intertemporal substitution-risk aversion pairs that are implied by the data with 'minimal' interference from ad hoc functional form specifications.

An alternative direction in which to generalize the standard expected additive utility specification is to weaken the assumption of time separability, in order to model habit formation or the durability of consumption goods (see Constantinides(1990), Heaton(1993) and Gallant and Tauchen(1989), for example). It is a still unresolved empirical question whether time or state nonseparability is more useful for explaining and organizing the observed behavior of consumption and asset returns. (While the studies just cited report results that support their models, Cochrane and Hansen (1992) show that they fail in other dimensions.) From a theoretical perspective, we point out that the relaxation of time separability does not permit a separation between the two conceptually distinct aspects of preference, substitution and risk aversion, while such separation seems important for the proper understanding of numerous issues in macroeconomics. (See Epstein (1992) for further discussion and also for arguments that this separation is achieved via recursive utility at a reasonable 'cost' in terms of axiomatic underpinnings. For instances where such separation has been used to advantage see Epstein(1988), Kandel and Stambaugh (1991), Obstfeld(1992), Hansen and Sargent(1992) and Hansen, Sargent and Tallarini (1993).) This paper's revealed preference characterization provides another theoretical 'argument' in support of the recursive utility route; a comparable theoretical result is not available for the time-nonseparable models.

Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) provide a 'nonparametric' approach to testing dynamic asset pricing models that has been widely used. It is worthwhile, therefore, to clarify the differences from our approach. The above authors consider a very general class of pricing models, those for which pricing relations can be expressed in terms of stochastic discount factors, and derive moment restrictions that must be satisfied by these discount factors if they are to be consistent with a given set of data. These restrictions can then be checked for any specific model of discount factors. In particular, they can be examined for any parametric recursive utility model. However, if the restrictions are violated, then one is left wondering whether it is the particular functional form or the recursive utility model that is to blame. Secondly, even if the restrictions are satisfied, since they are only necessary (but not sufficient) for the pricing relations to hold, the test is inconclusive regarding the validity of the model. In contrast, the revealed preference implications derived here are both **necessary and sufficient** for consistency with some recursive utility model in our semiparametric class. On the other hand, of course, our analysis is not relevant to pricing models outside this recursive utility class.

A further important limitation of our analysis is that it is nonstatistical. The empirical implications of our model take the form of a lack of stochastic dominance in pairwise comparisons of conditional probability distributions for consumption growth and asset returns, while data typically consist of sample realizations for these variables rather than their law of motion. One can employ seminonparametric statistical techniques of Gallant and Tauchen (1989), for example, to estimate the law of motion and then check whether or not the stochastic dominance conditions are satisfied by the estimated law, but there do not appear to be statistical methods available that take sampling error into account. We feel, nevertheless, that there are several ways in which our analysis makes a contribution to the empirical asset pricing literature. Firstly, a set of exhaustive implications, especially when as simple as below, can help to assess the vailidity of the model on informal grounds and clarify dimensions where the model fails. Secondly, there exist statistical techniques for testing stochastic dominance in related but simpler settings, (see McFadden (1989) and Klecan and McFadden (1991), for example), and our analysis may help to motivate suitable extensions of these techniques. Finally, our stochastic dominance conditions are tractable in the context of simple calibration exercises such as conducted by Mehra and Prescott(1985). Indeed, below we apply our revealed preference conditions to a simple general equilibrium model in order to clarify the potential of recursive utility for resolving the 'equity premium puzzle'. In so doing, we also provide perspective on Weil(1989) and Epstein and Zin(1990), who examine the equity premium puzzle in the context of parametric recursive utility models.

Finally, we emphasize that our principal contribution is to the asset pricing literature rather than to the theory of revealed preference, whether at the level of abstract choice theory or more particularly, at the level of choice under risk (see Fishburn(1975) and Border(1992), for example). Our contribution is to recognize the relevance of revealed preference theory for our intertemporal framework and to adapt the arguments in the above literature appropriately. On the other hand, it should be pointed out that the intertemporal framework facilitates empirical applications of revealed preference tests. That is because the repeated choice observations that are called for are available from observations of behavior over time, while they are less readily available in the atemporal frameworks adopted in the papers cited above and in Dybvig and Ross(1982), Varian(1983) or Green and Srivastava(1986), all of which deal with (one-shot) portfolio choice problems. We suspect that it is for this reason that we have not found any empirical investigations of the revealed preference conditions derived in these studies.[1]

We proceed as follows: Section 2 provides an outline of recursive utility and Section 3 contains our main results. An application to the equity

[1]Another difference from our study is that the cited papers deal exclusively with expected utility preferences. For a revealed preference analysis that is more thoughtful about the data requirements for implementation, but in other respects differs substantially from this paper, see Green and Osband (1991). Finally, some papers consider the uniqueness of preferences that rationalize a set of portfolio (inverse) demand functions, assuming existence; see Green, Lau and Polemarchakis (1979) for an atemporal expected utility analysis and Wang (1993) for an intertemporal, parametric recursive utility analysis. Since we want to rationalize a finite set of data, as opposed to functions, uniqueness does not hold in our model.

premium puzzle is provided in Section 4. Formal proofs are collected in an appendix.

2 RECURSIVE UTILITY[2]

Consider an infinitely lived agent who receives utility from consumption of a single good. At time t, current consumption c_t is deterministic but future consumption is uncertain. Thus intertemporal utility is defined over random consumption sequences. We assume that utility is recursive in the sense that the utility U_t derived from consumption at time t and beyond satisfies the recursive relation

$$U_t = W(c_t, \mu(U_{t+1})), \qquad t \geq 0, \tag{1}$$

where: $\mu(U_{t+1})$ represents the certainty equivalent of random future utility U_{t+1} conditional upon period t information; and W is termed an aggregator since it aggregates c_t with a risk-adjusted index of the future.

Before interpreting this structure, it is convenient to define the subclass of recursive utility models that will be considered below. First, we adopt the following CES form for the aggregator:

$$W(c,z) = [(1-\beta)c^\rho + \beta z^\rho]^{1/\rho}, \tag{2}$$

where $0 < \beta < 1$ is a discount factor and $0 \neq \rho \leq 1$. Certainty equivalents are required to satisfy:

$$\mu(b) = b \qquad \text{for all nonegative reals } b, \text{ and} \tag{3}$$

$$\mu(\lambda x) = \lambda\, \mu(x), \qquad \text{for all } \lambda > 0, \tag{4}$$

and for all random variables x in the domain of μ (specified below). The first condition states simply that the certainty equivalent of a prospect that yields b with certainty should equal b. The second condition, as will be confirmed below, imposes constant relative risk aversion. Its only justification is tractability; such linear homogeneity is a feature of all parametric empirical studies of which we are aware. Further regularity conditions on certainty equivalents that are specified below impose forms of monotonicity and risk aversion and thus do not restrict μ to lie in a parametric class.

[2]See Epstein and Zin (1989) and Epstein (1992) for further details regarding recursive utility.

The structure defined above admits a simple interpretation. First, in the case of a certain consumption sequence $(c_0, c_1, ..., c_t, ...)$, we derive the CES form for utility

$$U_t = [(1-\beta)\, \Sigma_{i=0}^{\infty} \beta^i \, c_{i+t}^{\rho}]^{1/\rho}. \tag{5}$$

Therefore, $(1-\rho)^{-1}$ is the elasticity of intertemporal substitution. It follows that only risk attitudes are affected by a change in the specification of μ. The separation permitted in this way between intertemporal substitution and risk aversion represents an important theoretical argument for the study of recursive utility functions. Epstein and Zin (1989) show that μ represents the agent's induced preference ordering over timeless wealth gambles, that is, gambles that are resolved immediately, before any consumption/savings decisions are made. Such gambles are the prospects with which subjects are usually confronted in experimental studies of decision making under risk; therefore, available experimental evidence reflects upon appropriate specifications for μ. Finally, for any fixed deterministic sequence $(c_0, c_1, ..., c_t, ...)$ and for any scalar random rescaling s of future consumption that is measurable with respect to time 1 information, the utility of the random consumption sequence $(c_0, sc_1, ..., sc_t, ...)$ equals $[(1-\beta)c_0^{\rho} + \beta\mu^{\rho}(s)U_1^{\rho}]^{1/\rho}$, where U_1 is defined by (5). Therefore, μ represents the preference ranking of any two such sequences that differ only in the random rescaling, that is,

$$(c_0, s'c_1, ..., s'c_t, ...) \succ (c_0, sc_1, ..., sc_t, ...) \Longleftrightarrow \mu(s') > \mu(s).$$

Combined, the preceeding describes the multiple senses in which it is appropriate to interpret μ as representing the agent's 'risk preferences'.

A number of parametric specializations of (1)-(4) have been adopted in empirical asset pricing models. The most common and the basis for empirical analyses of the C-CAPM has

$$\mu(x) = (Ex^{\rho})^{1/\rho}, \tag{6}$$

implying that

$$U_t = [(1-\beta)\, E_t\, \Sigma_{i=0}^{\infty} \beta^i c_{t+i}^{\rho}]^{1/\rho}. \tag{7}$$

(Here and below E denotes the expected value operator, with the conditioning information set either suppressed in the notation or indicated by a subscript as in (7).) Since the certainty equivalent μ depends on the substitution parameter ρ, both risk aversion and intertemporal substitution are confounded in a single parameter. In terms of empirical performance, we have already cited in the introduction the failures of this model to rationalize aggregate consumption and asset return data.

A natural generalization of the standard model introduces a separate parameter to model risk aversion, that is, μ is defined by

$$\mu(x) = (Ex^{\alpha})^{1/\alpha}, \qquad 0 \neq \alpha \leq 1. \tag{8}$$

While allowing $\alpha \neq \rho$ adds statistically significant explanatory power to the resulting model (see Epstein and Zin(1991) and Bufman and Leiderman(1990)), empirical difficulties remain (see Weil (1989)). Similarly, only limited success has been achieved by adopting specifications for the certainty equivalent that do not conform to expected utility theory (see Epstein and Zin (1990, 1992)).

While the above studies provide some information about the potential explanatory power of the recursive utility model (1), they deal only with a small number of functional forms. As indicated earlier, in this paper we continue to adopt the CES aggregator (2), but we weaken substantially the *a priori* restrictions on risk preferences. Obviously, there is a case to be made for being nonparametric also in the specification of W. However, we feel that functional form specification is more problematic for risk preferences than for W, as we now elaborate.

In contemplating alternatives to (8), it is natural to seek guidance from the experimental evidence regarding individual behavior under risk and the decision theory literature that it has spawned. Due to evidence such as the Allais paradox contradicting the positive accuracy of expected utility theory, a large number of generalizations have been developed. However, no single model clearly stands out in terms of axiomatic underpinnings and consistency with experimental evidence (see Camerer (1989) and Starmer (1992) for evidence regarding some of these models).

From a more practical 'curve-fitting' perspective, Kandel and Stambaugh (1989, 1991) have raised the issue of 'functional form flexibility'. They point out the limited ability of the specification (8) to model "plausible" risk attitudes over a broad range of gambles. Roughly, the risk premia implied for hypothetical small gambles are plausible only if $1-\alpha$ is large, but then risk premia for moderately sized gambles are unrealistically large. This feature is of particular concern if, as advocated by Mehra and Prescott (1985), for example, the success of the empirical model is evaluated in part on the basis of the plausibility of the risk premia for a range of gambles implied by the estimated certainty equivalent. The functional forms employed by Epstein and Zin (1990, 1992) are more flexible in the above respect, but not uniquely so.

Yet another argument for being nonparametric in risk preferences may be derived from Epstein and Wang (1993). They show that some deviations from (6) or (8) are appropriate if the agent does not know the underlying law of motion precisely, but only imprecisely or 'vaguely' in a sense corresponding intuitively to 'Knightian uncertainty'. Since there are a large number of ways in which such uncertainty could be modeled, the class of conceivable functional forms is enormous and a nonparametric approach has obvious advantages.

3 IMPLICATIONS FOR CONSUMPTION AND ASSET RETURNS

Let a representative agent have recursive utility defined by (1)-(4). The agent operates in a standard competitive environment. There are N assets and the i[th] has gross real return $r_{i,t+1}$ over the interval $[t, t+1]$. Write $r_{t+1} = (r_{1,t+1}, ..., r_{N,t+1})$ and denote by M_{t+1} the return to the market portfolio over the same interval.

Epstein and Zin (1989) show that intertemporal optimization implies the following conditions:[3] For all t, and conditional upon the information available at t,

$$\mu(\beta^{1/\rho}(c_{t+1}/c_t)^{(\rho-1)/\rho} M_{t+1}^{1/\rho}) = 1 \qquad \text{and} \tag{9}$$

$$\mu((c_{t+1}/M_{t+1})^{(\rho-1)/\rho} M_{t+1}) = \max\{\mu((c_{t+1}/M_{t+1})^{(\rho-1)/\rho} \omega r_{t+1}) : \omega \in R^N, \Sigma\omega_i = 1, \omega_i \geq \ell)\}, \tag{10}$$

The first condition reflects the optimum trade-off between consumption at t and $t+1$, while the second asserts the optimality of the market portfolio. In the latter, $\ell \leq 0$ defines an exogenous limit on short sales. It is customary to replace the maximization (10) by its associated first-order conditions and to refer to the resulting set of equations as Euler equations. However, at the present level of generality, the objective function may not be differentiable with respect to ω or the first-order conditions may not be sufficient for

[3]These conditions are derived under the assumptions that the stochastic environment is time-homogeneous and that the return to the market portfolio coincides with the return to total wealth.

an optimum. Therefore, we focus on (9)-(10) and refer to them as Euler equations.

Readers for whom these Euler equations are unfamiliar may find some reassurance in verifying that when the common specification (6) is substituted and the first-order conditions for (10) are computed, one obtains the Euler equations familiar from the consumption-CAPM literature, that is,

$$\beta E_t[(c_{t+1}/c_t)^{\rho-1}M_{t+1}] = 1 \qquad \text{and}$$

$$E_t[c_{t+1}^{\rho-1}(r_{i,t+1} - r_{j,t+1})] = 0, \qquad i,j = 1, ..., N.$$

Understanding of the general case of (10) is also aided by noting the factor $(c_{t+1}/M_{t+1})^{(\rho-1)/\rho}$ appearing in the objective function. If this factor is absent (or constant), the optimization problem reduces to an atemporal or one-shot portfolio choice problem such as in the revealed preference analyses cited in the introduction. Its appearance above reflects the fact that portfolio choices at each t are made as part of an intertemporally optimal consumption/savings plan.

If the Euler equations are satisfied, we say that μ **rationalizes** the law of motion for consumption and asset returns, for the given β and ρ. Thus in common with the bulk of the parametric literature, we ignore the information contained in the transversality condition associated with the representative agent's optimization problem; it is testable in any case only given knowledge of the joint conditional distributions for the infinite sequence of future consumption and other variables, while we assume the analyst knows only the one-step ahead conditional distributions. Our objective is to characterize the laws of motion that can be rationalized for some μ. First, however, we narrow further, but in an uncontentious way, the class of admissible certainty equivalents.

We assume that μ assigns the same value to any two random variables that share a common cumulative distribution function (cdf). Therefore, we specify the formal domain of μ as the following set of cdf's:

$$\mathcal{D} \equiv \{cdf's \quad F \ on \ [0,\infty): \quad F \text{ has compact support in } (0,\infty)\}. \qquad (11)$$

Thus all distributions are assumed to be bounded both above and below away from zero. (We assume that the law of motion for consumption and asset returns is restricted accordingly so that the Euler equations (9)-(10) are well defined.) Any F in $\mathcal{D}$ is the cdf F_x for some random variable x and we generally write $\mu(x)$ rather than $\mu(F_x)$. More generally, we identify x and F_x.

Further restrictions on risk preferences are expressed in terms of stochastic dominance. Let $\succeq_1$ and $\succeq_2$ denote the first and second order stochastic dominance relations on $\mathcal{D}$:

$$x \succeq_1 y \quad \text{if } Eu(x) \geq Eu(y) \quad \text{for all } u \text{ satisfying } u' \geq 0; \tag{12}$$

$$x \succeq_2 y \quad \text{if } \; Eu(x) \geq Eu(y) \quad \text{for all } u \text{ satisfying } u' \geq 0,\ u'' \leq 0. \tag{13}$$

These relations and their characterization in terms of cumulative distribution functions are well understood (Hadar and Russell (1969), for example). We employ the following nonstandard strict form of these relations: [4] For $i = 1, 2$,

$$x \succ_i y \qquad if \qquad x \succeq_i \kappa y \quad \text{for some } \kappa > 1. \tag{14}$$

Finally, say that μ **respects** $\succeq_i$ if

$$x \succeq_i y \implies \mu(x) \geq \mu(y) \quad \text{and} \quad x \succ_i y \implies \mu(x) > \mu(y). \tag{15}$$

For obvious reasons, we will be interested in certainty equivalents that are **monotone** (respect $\succeq_1$) or are **monotone and risk averse** (respect $\succeq_2$).

We assume that the analyst observes *both* the law of motion *and* realizations of consumption c_t, and asset returns r_{t+1} and M_{t+1} over the time frame $t = 1, ..., T$. In order to describe the implications of our model for such 'data', define for each $t = 1, ..., T$, the random variable x_t by

$$x_t = \beta^{1/\rho}(c_{t+1}/c_t)^{(\rho-1)/\rho} M_{t+1}^{1/\rho}, \tag{16}$$

and the set of random variables (or associated cdf's) B_t by

$$B_t = \{x_t M_{t+1}^{-1} \omega r_{t+1} : \omega \in R^N, \Sigma\omega_i = 1, \omega_i \geq \ell \text{ for all } i\}. \tag{17}$$

Then we see in referring to (10) that B_t can be thought of as the 'budget set' facing the agent at t, and the Euler equations can be rewritten simply in the form

$$1 = \mu(x_t) \geq \mu(y_t), \qquad \forall y_t \in B_t,\ t = 1, ..., T. \tag{18}$$

[4] A more customary notion of strict dominance simply replaces all weak inequalities in the definitions by strict ones. The distinction between the alternative notions of strict dominance is likely to be empirically insignificant in any statistical framework. Our specification is attractive because it leads to the simple and elegant characterization of Theorem 1.

If we further define $x_0 = 1$ and $B_0 = \{x_0\}$, then we obtain

$$\mu(x_s) \geq \mu(y_t), \qquad \text{for all } t, s = 0, ..., T \quad \text{and all } y_t \in B_t. \tag{19}$$

If we assume, for the moment only, that β and ρ are also known to the analyst, then $\{x_t, B_t\}_0^T$ constitute the 'data' to be rationalized and a rationalization consists of a certainty equivalent μ satisfying (19).

Note that if, as is often assumed in applications such as in Section 4, the law of motion for consumption and asset returns involves only finitely many conditional one-step ahead joint probability distributions for these variables, then it may occur that $(x_t, B_t) = (x_\tau, B_\tau)$ for distinct t, τ between 0 and T. In that case, one of these data points is redundant and may be discarded. It is then convenient to renumber the data so that the subscript t in $\{x_t, B_t\}_0^T$ indexes distinct conditional distributions rather than time.

It is apparent that the following condition is *necessary* for $\{x_t, B_t\}_0^T$ to be rationalizable by some μ respecting $\succeq_i$, $i = 1, 2$:

$$y_t \not\succ_i x_s \quad \text{for all } t, s = 0, ..., T \quad \text{and all } y_t \in B_t. \tag{20}$$

That it is also *sufficient* is the central result of the paper and is established in the following theorem:

Theorem 1 *For each $i = 1, 2$, $0 < \beta < 1$ and $0 \neq \rho \leq 1$, the 'data' $\{x_t, B_t\}_0^T$ can be rationalized by a certainty equivalent μ respecting $\succeq_i$* **if and only if** *(20) is satisfied.*

Moreover, in that case there exist two such certainty equivalents μ_* and μ^*, with $\mu_*(x) \leq \mu^*(x)$ on $\mathcal{D}$, such that: a certainty equivalent μ respecting $\succeq_i$ rationalizes $\{x_t, B_t\}_0^T$ **if and only if**

$$\mu_*(x) \leq \mu(x) \leq \mu^*(x) \qquad \text{on } \mathcal{D}. \tag{21}$$

Thus, given β and ρ, (20) represents the exhaustive implications of the recursive utility model (1)-(2) if risk preferences are required only to satisfy monotonicity and possibly risk aversion, in addition to the constant relative risk aversion property (4). Typically, β and ρ are unknown and the theorem would be applied by searching over (β, ρ) pairs to determine if (20) is satisfied for *some* pair. If such a pair is found, then (21) describes all the certainty equivalents that rationalize the data. In particular, the indicated inequality represents all the information about risk aversion that can be inferred from the data, given β and ρ, in the absence of additional assumptions about

μ. Fortunately, the proof of the theorem is constructive and the upper and lower bounds μ^* and μ_* can in principle be constructed from the data.

The conditions (20) can be expressed alternatively in the form

$$x_s \text{ is } i\text{-efficient in } \cup_0^T B_t, \qquad s = 0, ..., T, \tag{22}$$

where, consistent with common terminology, i-efficiency means undominated (strictly) with respect to $\succeq_i$. This efficiency condition is, naturally, different from others that have been studied in portfolio theory. First, since mean-variance efficiency and efficiency with respect to second degree stochastic dominance are distinct notions (see Hanoch and Levy (1969)), the above is not comparable to the condition that each x_s be mean-variance efficient. Second, (22) does not imply any form of efficiency for the market portfolio, as would an atemporal portfolio analysis. Only in the case $\rho = 1$ and therefore an infinite elasticity of intertemporal substitution, does x_s reduce to βM_{t+1} and does (22) impose efficiency of the market portfolio.

A formal proof of the theorem is provided in the appendix. However, the essence of the argument is easily conveyed informally with the aid of Figure 1, which deals with the case $T = 1$ and where all random variables are defined on a probability space having two equally likely states. Random variables are identified with points in the plane and two such points induce the same cdf if and only if they are mirror images of one another with respect to the certainty line. Therefore, each cdf can be identified with a unique point in the cone below the 45^0 line, where $\succeq_1$ coincides with the usual partial ordering of vectors in the plane. Adopting this identification, it follows that the (open) set S_* shown equals the collection of all points that dominate some x_t in the sense of $\succ_1$ and that the data shown satisfy the conditions (20) for $i = 1$. From (18), any monotone rationalizing μ must be such that ($\star$) all $x_t's$ lie on a single indifference curve that has nonpositive slope and lies on or above every budget set B_t. The boundary of S_* satisfies these requirements and lies everywhere above (weakly) any other curve that satisfies them. Further, by the linear homogeneity (4), the single indifference curve defined by the boundary of S_* defines uniquely an entire indifference map and therefore a certainty equivalent μ_* that both rationalizes the data and serves as a lower (and not upper, as one might guess at first glance) bound in the sense of (21). For the upper bound, note that the boundary of the (open) set S^*, where S^* consists of all points that are $\succ_1$-dominated by some point in $\cup B_t$, also satisfies the criteria ($\star$) and lies below any other curve that does. Next apply linear homogeneity.

Finally, the case of monotone and risk averse certainty equivalents can be illustrated similarly.

To conclude this section, we consider the consequences of adding further restrictions on risk preferences. One possibility is to restrict the degree of risk aversion embodied in μ in conformity with prior information, derived from observations of choice behavior in settings other than the specific one dealt with in Theorem 1. The use of such information is advocated by Mehra and Prescott and defines the meaning for 'plausibility' of the degree of risk aversion as an additional criterion in rationalizing a set of consumption and asset return data. In our nonparametric framework, such information could take the form

$$a_j \leq \mu(z_j) \leq b_j, \qquad j = 1, ..., J, \tag{23}$$

where each z_j, an element of $\mathcal{D}$, represents a timeless wealth gamble for which it is known that the risk premium $Ez_j - \mu(z_j)$ lies between $Ez_j - b_j$ and $Ez_j - a_j$.[5] Since μ is linearly homogeneous, the above inequalities are equivalent to

$$\mu(z_j/b_j) \leq 1, \quad \mu(z_j/a_j) \geq 1, \qquad j = 1, ..., J.$$

Therefore, the consequences of the added restrictions (23) are covered by the following theorem:

Theorem 2 *For each* $i = 1, 2$, $0 < \beta < 1$ *and* $0 \neq \rho \leq 1$, *the 'data'* $\{x_t, B_t\}_0^T$ *can be rationalized by a certainty equivalent* μ *respecting* $\succeq_i$ *and satisfying the set of inequalities*

$$\mu(z_j') \leq 1, \quad \mu(z_j'') \geq 1, \qquad j = 1, ..., J,$$

if and only if *(20) is satisfied* **and** *for all* j, k, t *and* $y_t \in B_t$,

$$z_j' \not\succ_i x_t, \quad y_t \not\succ_i z_j'' \text{ ,and } \quad z_j' \not\succ_i z_k''. \tag{24}$$

The necessity of the conditions (24) is obvious. The proof of sufficiency is similar to that for Theorem 1 and is sketched in the appendix. Note that, by the nature of the bounding certainty equivalents μ_* and μ^* provided

[5] Recall that μ represents the induced ordering over timeless wealth gambles (see Section 2). Note also that if μ is monotone and risk averse, then (23) is necessarily satisfied if $a_j = \inf z_j$ and $b_j = Ez_j$. Thus one would typically want to impose a larger value for a_j and a smaller one for b_j.

by Theorem 1, the first two conditions can be expressed in the simple and intuitive fashion $\mu_*(z_j') \leq 1$ and $\mu^*(z_j'') \geq 1$ for all j.

The decision theory literature suggests alternative added restrictions for μ. To express them, recall first that each x in $\mathcal{D}$ is identified with its cdf. Accordingly, by $\Sigma\lambda_i * y_i$ with $\Sigma\lambda_i = 1$ and $\lambda_i \geq 0$ for all i, we shall mean the corresponding mixture of cdf's. With this notation in mind, say that μ is **quasiconcave** (in cdf's or probabilities) if

$$\mu(x) \geq k \text{ and } \mu(y) \geq k \Longrightarrow \mu(\tfrac{1}{2} * x + \tfrac{1}{2} * y) \geq k.$$

Quasiconvexity is defined by reversing all inequalities. Finally, say that μ satisfies **betweeness** if it is both quasiconcave and quasiconvex, or equivalently, if each indifference set $\{x \in \mathcal{D} : \mu(x) = k\}$ is a convex set with respect to the mixture operation. The betweeness axiom is a weakening of the independence axiom that can accomodate some behavior contradicting the latter, such as that exhibited in the Allais paradox (see Chew (1983, 1989) and Dekel (1986)). In the probability simplex corresponding to gambles having three fixed outcomes, (see Figure 2), betweeness requires that indifference curves be linear while the independence axiom forces them also to be parallel to one another. Note also that whether or not betweeness holds, monotonicity of a certainty equivalent implies that its indifference curves in the probability simplex are upward sloping, while risk aversion implies further that they are everywhere steeper than the constant mean loci.[6]

The next theorem characterizes data that can be rationalized by a certainty equivalent that is monotone and risk averse and also satisfies one of the three properties just described. In each case, the appropriate strengthening of (20) is provided. Let $\Lambda \equiv \{\lambda = (\lambda_0, ..., \lambda_T) \in R^{T+1} : \lambda_t \geq 0$ all t and $\Sigma\lambda_t = 1\}$ and denote by $\succ_{ssd}$ the standard strict second order dominance relation (see footnote 4). Clearly, $x \succ_2 y \Longrightarrow x \succ_{ssd} y$ but the converse is false in general.

Theorem 3 *For each* $0< \beta < 1$ *and* $0 \neq \rho \leq 1$, *the following conditions are* **necessary** *and* **sufficient** *for the data* $\{x_t, B_t\}_0^T$ *to be rationalized by a certainty equivalent respecting* $\succeq_2$ *and satisfying*

[6]See Machina (1987) for more details on probability triangles as a tool for analysing risk preferences.

(a) quasiconcavity, or (b) quasiconvexity, or (c) betweeness:

$$\begin{array}{lll} (a) & y_t \not\succ_2 \Sigma\lambda_s * x_s & \forall t, y_t \in B_t, \lambda \in \Lambda \\ (b) & \Sigma\lambda_s * y_s \not\succ_2 x_t & \forall t, y_s \in B_s, \lambda \in \Lambda \\ (c) & \Sigma\lambda_s * y_s \not\succ_2 (\not\succ_{ssd}) \Sigma\gamma_s * x_s & \forall y_s \in B_s, \lambda \in \Lambda, \gamma \in \Lambda, \end{array} \tag{25}$$

where in (c), the condition involving ($\not\succ_2$) is necessary for the indicated rationalization while the stronger condition involving ($\not\succ_{ssd}$) is sufficient for such a rationalization.

The betweeness class (c) contains the parametric models estimated by Epstein and Zin (1990,91,92) as well as the standard expected utility model estimated by Hansen and Singleton (1983) and many others. Part (c) falls short of providing a characterization in light of the gap between $\not\succ_2$ and $\not\succ_{ssd}$. However, as noted earlier, that gap is likely to be empirically insignificant in a statistical framework. Note that in the context of Figure 1 with y and x denoting vectors in the lower cone, $y \succ_{ssd} x$ if $y \neq x$, $y_2 \geq x_2$ and $y_1 + y_2 \geq x_1 + x_2$, while $y \succ_2 x$ only if the inequalities are strict.

Figure 2 illustrates, in the context of the three-outcome probability simplex, a data set that cannot be rationalized by a certainty equivalent satisfying any of the conditions in Theorem 3, but that is compatible with a monotone and risk averse certainty equivalent as in Theorem 1.

4 AN APPLICATION

We illustrate our analysis by applying it to the equity premium puzzle posed by Mehra and Prescott (1985). Using a Lucas style endowment economy, a 'simple' consumption process calibrated to U.S. data and a representative agent with isoelastic expected utility preferences, Mehra and Prescott found that they could not match historically observed average returns on both equity and Treasury bills. The parametric studies by Weil (1989) and Epstein and Zin (1990) suggest that the generalization to recursive preferences cannot alone resolve the puzzle, while Bonomo and Garcia (1993) report that, in conjunction with a distinction between consumption and dividends and the specification of rich processes for each, recursive preferences can match the first two moments of equity returns and the risk free rate.[7] Here we

[7]Other 'explanations' of observed returns that have been examined in the literature include habit formation and incomplete markets. See references in the three cited papers.

show that without a priori functional form restrictions on risk preferences, recursive utility can contribute substantially to resolving the puzzle even with a simple process for consumption such as considered by Mehra and Prescott. Further, we illustrate some limitations of a parametric approach and thereby provide a useful perspective on the above studies.

We assume that the endowment process for consumption is such that the growth rate $g_{t+1} \equiv c_{t+1}/c_t$ follows a first-order Markov process. The ex-dividend price of 'equity', that is, of the endowment stream, is described by the time-invariant and positive function $p(g_t, c_t)$ of the 'state' variables g_t and c_t. From the homogeneity of preferences, it follows that the price is linearly homogeneous in consumption, that is,

$$p(g,c) = p(g,1)\, c \equiv P(g)\, c. \tag{26}$$

Similarly, the risk free rate, or the return to a one period discount bond, can be expressed in terms of the function $r(g_t)$. Finally, the return to equity or the market portfolio from time t to $t+1$, is given by

$$M_{t+1} \equiv \frac{p(g_{t+1}, c_{t+1}) + c_{t+1}}{p(g_t, c_t)} = \frac{P(g_{t+1}) + 1}{P(g_t)}\, g_{t+1}. \tag{27}$$

The usual approach specifies a preference ordering and uses the associated Euler equations, the given endowment process and (26)-(27), which embody additional implications of the transversality condition and general equilibrium, to solve for the functions P and r. We reverse this procedure and ask whether 'given' functions P and r can be rationalized by some recursive utility function in our class. More precisely, we begin by finding a price-dividend function P such that the joint distribution of consumption growth and market returns can be rationalized for some pair (β, ρ). At this stage, we can impose additional restrictions on the price-dividend function to achieve desired properties for the joint distribution; for example, we can restrict attention to functions that yield a value for the average return to equity that approximates historical values, or we can require the realized return to equity to vary monotonically with the realized consumption growth rate. Given P, β and ρ as above, we proceed to solve for admissable values of the risk free rate function $r(g)$ by checking the appropriate versions of the stochastic dominance conditions (20). At this stage, we exploit the following two implications of Theorem 1: (i) For each state g, the set of admissable values for the risk free rate is an interval; and (ii) any rate in the interval for state g can be chosen, regardless of the rates chosen in other states, that is, 'admissability' is determined for each state separately.[8]

For simplicity and to maintain comparability with previous authors, we restrict attention to processes featuring only two possible rates of growth, a low growth rate g_ℓ and a high rate g_h. In this case the functions P and r are each summarized by a pair of numbers. Because of its analytical simplicity, we first consider the case of i.i.d. consumption growth before allowing some autocorrelation in growth rates. Throughout we consider rationalization by a monotone and risk averse certainty equivalent. Finally, the short sales constraint imposed in (10) is defined implicitly by the requirement that the resulting random variables are everywhere positive.

4.1 I.i.d. consumption growth

Suppose that the consumption growth rate process is i.i.d. Then $P(g_\ell) = P(g_h)$, denoted simply P, so the return on the market portfolio is perfectly correlated with consumption growth and

$$M_{t+1} = K g_{t+1}, \quad K \equiv (P+1)/P \ > 1. \tag{28}$$

Further, the risk free rate is also constant across states at a level denoted r. Simple calculation based on Theorem 1 shows that the values for K consistent with β, ρ and a monotone and risk averse certainty equivalent are those satisfying

$$(\beta K)^{1/\rho} g_\ell \leq 1 \leq (\beta K)^{1/\rho} E(g). \tag{29}$$

Given such a K, one can further rationalize the risk free rate r if and only if

$$K g_\ell \leq r \leq K / (\beta K)^{1/\rho}. \tag{30}$$

Note that the lower bound for the risk free return is the lowest possible return to holding equity. To elaborate, (29) is necessary and sufficient for the data $\{x_t, B_t\}_0^1$ to be rationalizable where following Section 3 and (28), $x_0 = 1$, $B_0 = \{1\}$, x_1 has the distribution of $(\beta K)^{1/\rho} g$ and $B_1 = \{x_1\}$, that is, equity is the only asset. Similarly, (29) and (30) are necessary and sufficient for the rationalizability of $\{x_t, B_t'\}_0^1$, where $B_0' = B_0$ and B_1' is the

[8] More precisely, we exploit: (i) $\forall \omega, t, s$, $\{r : \omega x_t + (1-\omega) x_t M_{t+1}^{-1} r \not\succ_i x_s\}$ is convex for $i = 1, 2$; (ii) if the two data sets $\{x_t, B_t\}$ and $\{x_t, B_t'\}$ are each rationalizable, then so is $\{x_t, B_t \cup B_t'\}$.

budget set (17) corresponding to trading in equity at return $M_{t+1} = Kg_{t+1}$ and in a bond with return r, that is,

$$B_1' = \{y_1(\omega) \equiv (\beta K)^{1/\rho}[\omega r K^{-1} + (1-\omega)g] : \; y_1(\omega) > 0\}.$$

If we match the consumption process to the first two moments of aggregate U.S. data, then $E(g) = 1.018$ and $\sigma(g) = .036$. Assuming that the two growth rates are equally likely yields $g_\ell = .982$ and $g_h = 1.054$. Historically, the average return to equity is about 1.07. With our simple DGP for consumption, we can find many choices for K and hence P that yield this average return and that can be rationalized for some (β, ρ). However, in all cases we must have $r \geq g_\ell E(M)/E(g)$ because of (30), so r must exceed 1.032. Since the historical average for r is about 1.008, we have an example of the equity premium puzzle, or the associated risk free rate puzzle emphasized by Weil (1989). Adding points of support to the consumption growth process can only help if they lead to lower possible growth rates and hence to lower possible rates of return to equity, a suggestion first proposed by Reitz (1989) in an expected utility-based model. The addition of 'low' growth rates may also be playing a role in resolutions of the puzzle, such as Bonomo and Garcia (1993), that employ more complicated DGP's for consumption.

4.2 Autocorrelated consumption growth

The situation is very different if we admit heterogeneity of the conditional distribution for consumption growth. Even with the same points of support for g, the returns to holding equity can be very volatile. If realized payoffs to equity vary monotonically with realized consumption growth rates, then the lowest value of these payoffs is also the infimum of admissable values for the risk free rate in each state. But with volatile equity returns, these lower bounds provide plenty of room for low risk free rates. Indeed, it is not hard to construct economies that match the average equity return and in which the predicted average risk free rate is *lower* than the historical average!

Table 1 reports some examples. We consider two economies. In both economies the consumption endowment process is as above, except that the conditional probability of staying in the same state is assumed to be .45. (The unconditional probability of each state is 1/2, as in the i.i.d. case.) Both economies are populated by a representative agent with monotone and risk averse preferences and with discount factor fixed at $\beta = .99$. The price-dividend ratios in the low and high growth states, denoted P_ℓ and P_h, are determined by specifying the mean and standard deviation of the return to equity. We maintain the same mean equity return but allow the standard deviation to vary across the two economies.

In the first economy, we match both the mean (1.07) and the standard deviation (.165) of equity returns to their historical values as estimated by Mehra and Prescott (1985). There are two pairs of price-dividend ratios that will generate these moments. We restrict attention to the case where the return to equity is procyclical, that is, $(P_\ell + 1)g_l \leq (P_h + 1)g_h$, so that we obtain a unique pair $P_\ell = 23.42$ and $P_h = 27.89$. Then we calculate the set of values for ρ that, in conjunction with the assumed value of .99 for β, allow us to rationalize the joint consumption and equity price process. From Theorem 1, this amounts to checking that there is no second order dominance among the $\{x_t\}$ for a given value of ρ. (In this two state model, $T = 2$, with $t = 1, 2$ corresponding to the low and high growth states respectively.) Numerically, the set of admissible values for ρ appears to be an interval with no lower bound. Because the elasticity of intertemporal substitution (IES) is $(1 - \rho)^{-1}$, all of the admissible values for ρ suggest very little willingness to substitute intertemporally. For each admissible ρ and each state, we can compute (again via Theorem 1, with budget sets expanded to include a risk free asset) the interval of risk free rates that are consistent with the above joint process for consumption and equity prices. It can be shown that the lower bound of each interval does not vary with ρ and equals the lowest possible return to equity conditional on the state. The upper bounds for the risk free rates do vary with ρ. The maximum value of r_ℓ increases with ρ, while the maximum value of r_h decreases.

For each admissible $\hat{\rho}$, any candidate pair $\{\hat{r}_\ell, \hat{r}_h\}$ from the nonparametrically determined intervals indicated in the Table can be rationalized in this economy by a representative agent with utility parameters $\hat{\rho}$, $\beta = .99$ and some monotone and risk averse certainty equivalent. We see, therefore, that this economy produces a "risk free rate puzzle" but of quite a different sort than that described by Weil (1989). Risk free rates are too low and hence the equity premium is too high! Even if we take $\hat{r}_\ell$ and $\hat{r}_h$ to be their maximum admissible values, the average net risk free rate is negative. At $\hat{\rho} = -16$, the average risk free rate is almost 2.5% below its historical value. The average rate can be increased by taking a lower value for ρ, but this relationship is not globally monotonic. For a wide range of values for ρ, we can rationalize an average risk free rate that is about 1% below its historical value and a standard deviation roughly 2/3 higher than its historical value.

What are the attitudes towards risk implied by these various choices for asset returns? Consider the gamble, denoted Z, in which the agent begins with 75 units of wealth and receives ± 25 units with equal probability. For any given admissible $\hat{\rho}$, we can compute, as in the proof of Theorem 1, the upper and lower bounds for the certainty equivalent of this gamble that are implied by the asset return data. The difference between these bounds provides a measure of the information about risk attitudes that can be inferred from the data. In all examples, the lower bound $\mu_*(Z)$ is slightly above 50. With $\hat{\rho} = -16$, the upper bound $\mu^*(Z)$ equals 56.5, while $\mu^*(Z) = 62.7$ if we take $\hat{\rho} = -50$. By way of comparison, the coefficients of relative risk aversion that yield these certainty equivalent values in an expected utility model are approximately 6.5 and 4, respectively.

The bottom half of Table 1 shows the risk free rates implied by two parametric examples. The Kreps-Porteus certainty equivalent is defined by (8); it is applied to the equity premium puzzle by Weil (1989). For a binary gamble with outcomes $Z_1 < Z_2$, the Yaari certainty equivalent, that is applied and described more completely in Epstein and Zin (1990), is given by

$$\mu^Y(Z) = p_1^\gamma Z_1 + (1 - p_1^\gamma) Z_2,$$

where p_1 is the probability of the inferior outcome Z_1 and where γ is a parameter in the unit interval. For the Kreps-Porteus calculations, we choose the parameters (ρ, α) to match the given pair of price-dividend ratios. The parameters (ρ, γ) in the second model are chosen in the same way. Both models imply unique choices for the risk free rates, but for comparison we also report the nonparametric range of values for the risk free rate consistent with the chosen values of ρ.

The Kreps-Porteus parameters produce risk free rates around the midpoint of the admissible intervals. They yield a standard deviation for r about 1.5 times the historical value and an average rate that is about 5% below the historical value. Although the fit is poor, this example is nevertheless interesting. As noted above, the equity premium is too large and the average risk free rate is too low. Further, since the candidate values of ρ and α are roughly equal, it is not surprising that we can get almost the same values for returns using an expected utility specification of preferences (7) with ρ about -35. Previous authors, with the notable exception of Kandel and Stambaugh (1991), have ignored this region of the parameter space because of a priori beliefs about reasonable levels of risk aversion. Indeed, the

Kreps-Porteus functional form assigns an implausibly low certainty equivalent value of 51 to our 'litmus' gamble Z. But this reflects functional form rather than substance. The same predictions about asset returns can be obtained from a variety of certainty equivalents that assign anything from 50.7 to 62.8 to the gamble Z. The Kreps-Porteus functional form yields a value near the bottom of this range.

Turning to the Yaari certainty equivalent, in the i.i.d. case it generates risk free rates at the maximum of the nonparametric range and this property seems to hold in our examples. Epstein and Zin (1990) did not consider such a low value for the IES (they never went below $\rho = -9$), and concluded that they could account for an equity premium of roughly 2%. With our parameter values we match the historical mean return to equity and generate a risk free rate that is on average only 1% below the historical value. On the other hand, although the Yaari functional form matches the first moments well and the standard deviation of equity returns by construction, the standard deviation of the risk free rate is about 2/3 larger than the historical value. Finally, the Yaari functional form assigns a value of 62.3 to the gamble Z, which is very close to the largest possible value (given ρ). Note that in all of the examples, the upper bound $\mu^*(Z)$ is sensitive to short selling constraints. If we rule out borrowing at the risk free rate, the value of $\mu^*(Z)$ rises in each example to about 73.5.

The right half of Table 1 looks at the second economy that has the same consumption endowment process and mean equity return, but less volatility in equity returns. The price-dividend ratios in this economy are again uniquely determined if we require further that equity returns are procyclical. With less volatile equity returns, higher values of ρ are admissible, but the implied IES is still small by conventional standards. For any value of ρ in [-40,-15], we can match the first two moments of the risk free rate exactly to their historical values. The certainty equivalents that do so are slightly more risk averse than those of the first economy in that $\mu^*(Z)$ tends to be about 60 over this range for ρ. The two parametric examples are again calibrated to the economy's price-dividend ratios. Both functional forms provide a good match to the standard deviation of the risk free rate. The Kreps-Porteus functional form yields a mean risk free rate that is still too low, while the Yaari form delivers a mean rate that is too high, but neither discrepancy is large. However, the two parametric examples present vastly different impressions about risk attitudes, assigning certainty equivalent values for the litmus gamble that are at opposite ends of the nonparametric range.

We take several lessons from these examples. First, with recursive preferences, there is no equity premium puzzle, at least involving first moments. With a low IES, we can match the mean risk free rate and return on equity to their historical values. Moreover, this does not require an implausibly large degree of risk aversion. If there is a puzzle, it involves both first and second moments. With equity returns as volatile as their historical values, it is hard to rationalize a nonnegative average net risk free rate, the difficulty being primarily to produce a sufficiently high value for r_h. Further, models that yield an average net risk free rate close to zero also yield risk free rates that are too volatile. We suspect that we could match both the mean and standard deviation of the risk free rate in the first simple economy if we admitted preferences that are risk seeking for some gambles, though we have no reason to believe that the required deviations from global risk aversion would be 'intuitive' or consistent with psychological evidence that people often are risk seeking in the domain of 'losses'.

A second important lesson is the danger of using a parametric model to infer attitudes towards risk from asset return data, as such inference is not robust to the choice of functional form. The same asset return data can be generated by preferences that assign widely different values to a lottery not spanned by the asset returns.

A APPENDIX

Proof of Theorem 1: Necessity of (20) is obvious. We prove sufficiency. Fix i and denote $\succeq_i$ and $\succ_i$ simply by $\succeq$ and $\succ$.

Define

$$S_* \equiv \cup_0^T \{x \in \mathcal{D} : x \succ x_t\} \qquad \text{and}$$

$$\mu_*(x) \equiv \sup\{\lambda > 0 : \lambda^{-1}x \in S_*\}, \quad x \in \mathcal{D}. \tag{31}$$

$0< \mu_*(x) < \infty$: Given the nature of the supports of elements of $\mathcal{D}$, for each x $\exists \lambda, \nu$ such that $\lambda^{-1}x \in S_*$ and $\nu^{-1}x \notin S_*$.

$\mu_*(x_t) = 1$ $\forall t$: $\lambda^{-1}x_t \succ x_t \forall \lambda > 1 \implies \mu_*(x_t) \geq 1$. On the other hand, if $\lambda^{-1}x_t \succ x_s$ for some $\lambda > 1$, then $x_t \succ x_s$, contradicting (20).

$\mu_*(y_t) \leq 1$ $\forall t \forall y_t \in B_t$: $\mu_*(y_t) > 1 \implies \lambda^{-1}y_t \succ x_s$ for some $\lambda > 1 \implies y_t \succ x_s$, contradicting (20).

μ_* satisfies (3) and (4): Obvious.

μ_* **respects** $\succeq$: $x \succ y \implies x \succ \kappa y$ for some $\kappa > 1 \implies \mu_*(x) \geq \kappa\mu_*(y) > \mu_*(y)$. That $x \succeq y \implies \mu_*(x) \geq \mu_*(y)$ is obvious.

It follows that μ_* rationalizes the data and respects $\succeq$.
Similarly, define

$$S^* \equiv \cup_0^T \{x \in \mathcal{D} : \exists t, y_t \in B_t,\, y_t \succ x\} \quad \text{and}$$

$$\mu^*(x) \equiv \inf\{\lambda > 0 : \lambda^{-1} x \in S^*\}, \quad x \in \mathcal{D}. \tag{32}$$

We can prove as above that μ^* rationalizes the data and respects $\succeq$.

Next let μ be as in (21). Then $\mu(x_t) = 1 \,\forall t$ and $\mu(y_t) \leq \mu^*(y_t) \leq 1$ if $y_t \in B_t$. Conversely, let μ rationalize the data and prove (21). Fix x and without loss of generality, let $\mu(x) = 1$. Then $\lambda^{-1}x \in S_*$ for some $\lambda > 1 \implies \exists s, \lambda^{-1}x \succ x_s \implies \mu(x) > \lambda\mu(x_s) = \lambda > 1$, a contradiction. It follows that $\mu_*(x) \leq 1 = \mu(x)$. Similarly, $\lambda^{-1}x \in S^*$ for some $\lambda < 1 \implies \exists s, y_s \in B_s, y_s \succ \lambda^{-1}x \implies \mu(x) < \lambda\mu(y_s) \leq \lambda < 1$, a contradiction. Therefore, $\mu^*(x) \geq 1 = \mu(x)$.

Proof of Theorem 2: Adopt the notation of the previous proof. Define

$$S \equiv S_* \cup \{x \in \mathcal{D} : \exists j,\, x \succ z_j''\}$$

and define μ by the counterpart of (31). The hypotheses imply that $z_j' \notin S$ and $B_t \cap S = \emptyset$ for all j and t. It follows that μ satisfies all the requirements.

Proof of Theorem 3: Adopt the notation of the proof of Theorem 1.
(a): If μ is a quasiconcave rationalization, $\mu(x_s) = 1 \geq \mu(y_t) \,\forall y_t \in B_t \,\forall s \implies \mu(\Sigma\lambda_s * x_s) \geq \mu(y_t) \implies y_t \not\succ \Sigma\lambda_s * x_s$ since μ respects $\succeq$. For the converse, define

$$S^a \equiv \{x \in \mathcal{D} : \exists \lambda, x_s,\; x \succ \Sigma\lambda_s * x_s\},$$

and define the certainty equivalent μ by the analogue of (31). Since S^a is convex (with respect to probability mixtures) and coincides with $\{x : \mu(x) > 1\}$, μ is quasiconcave. The remainder of the proof is as above.

(b): The necessity of the condition in (25) is obvious. For the converse, let

$$S^b \equiv \{x \in \mathcal{D} : \exists \lambda, y_t \in B_t,\, \Sigma\lambda_t * y_t \succ x\}$$

and define μ by the analogue of (32). We claim that S^b is convex: Suppose that $\Sigma\lambda_t * y_t \succ x$ and $\Sigma\lambda'_t * y'_t \succ x'$. Define $\gamma_t = (\lambda_t + \lambda'_t)/2$ and $\delta_t = \lambda_t/(2\gamma_t)$. Since y_t and y'_t are in B_t, we have $y_t = x_t M_{t+1}^{-1}\omega r_{t+1}$ and $y'_t = x_t M_{t+1}^{-1}\omega' r_{t+1} \implies z_t \equiv x_t M_{t+1}^{-1}(\delta_t\omega + (1-\delta_t)\omega')r_{t+1}$ satisfies $z_t \in B_t$ and $z_t \succeq \delta_t * y_t + (1-\delta_t) * y'_t$. (The latter property is true only for $\succeq_2$.) Therefore, $\Sigma\gamma_t * z_t \succeq \Sigma\gamma_t * (\delta_t * y_t + (1-\delta_t) * y'_t) \succ \frac{1}{2} * x + \frac{1}{2} * x' \implies \frac{1}{2} * x + \frac{1}{2} * x'$ lies in S^b. It follows that μ is quasiconvex. The remainder of the proof is as above.

(c): Necessity of the condition involving $\not\succ_2$ is straightforward. It might be expected that one could prove sufficiency of that condition by constructing a separating hyperplane for the sets S^a and S^b. We have already shown that the sets are convex. That they are disjoint is implied by the nondominance hypothesis. However, we have not succeeded in confirming the possibility of such a separation in the present infinite dimensional setting. On the other hand, if the condition involving $\not\succ_{ssd}$ is assumed, then the revealed preference characterizations in Border(1991,1992) can be applied.

It is worth clarifying the reasons that Border's analysis is relevant even though he studies revealed preference implications of expected utility theory, while we are here dealing with the 'nonexpected' utility theory corresponding to the betweeness axiom. Those reasons are: (i) From (18), rationalization of the data requires only the construction of a *single* indifference surface containing all the x_t's; (ii) an expected utility function cannot be distinguished from a betweeness conforming function given knowledge of only a single indifference surface, since in both cases indifference surfaces are convex sets (or straight lines in the probability triangle). That is not to say that expected utility based and betweeness conforming certainty equivalents cannot be distinguished in our framework. That such a distinction *is* possible is apparent once one recalls that linear homogeneity of the certainty equivalent is a maintained assumption in our analysis. Therefore, rationalization by an expected utility certainty equivalent μ requires not only that one of its indifference surfaces be convex, but in fact that μ have the parametric form (8). In contrast, linear homogeneity of μ does not restrict any single indifference surface of μ if only betweeness is assumed.

Now proceed with the proof of sufficiency, assuming $\not\succ_{ssd}$. Define coB_t, the convex hull of B_t in the sense of probability mixtures, and

$$B'_t \equiv \{x \in \mathcal{D} : \exists z_t \in coB_t,\ z_t \succ_{ssd} x\}.$$

Then, by the arguments used in (b) to prove the convexity of S^b, B'_t is convex,

$t = 0, 1, ..., T$. Further define $B'_{t+T} = B'_{t+2T} \equiv$ convex hull of $\{x_0, x_t\}$, $t = 1, ..., T$. We now have $3T+1$ convex choice sets. Consider the choices z_t from B'_t given by $z_t = x_t$, $t = 0, ..., 2T$, and $= x_0, t = 2T+1, ..., 3T$. The $\not\succ_{ssd}$ hypothesis implies the 'nondominance' condition shown by Border to imply that the above choices can be rationalized by an expected utility function; more precisely, there exists a vNM utility index u, continuous and strictly increasing and concave, such that

$$Eu(z_t) \geq Eu(y_t) \ \forall y_t \in B'_t, \ t = 0, ..., 3T.$$

(Border (1992, Theorem 2.4) provides a revealed preference characterization appropriate for first degree dominance. An extension that includes second degree dominance and that is adequate for our purposes is provided in (1991, Section 6).) It follows that

$$u(1) = Eu(x_t) \geq Eu(y_t) \ \forall y_t \in B_t, \ t = 0, ..., T.$$

The set $\{x \in \mathcal{D} : E\,u(x) = u(1)\}$ defines the single indifference surface discussed above. Finally, define the certainty equivalent μ by

$$\mu(x) \equiv \sup\{\lambda > 0 : \ u(\lambda^{-1}x) \geq u(1)\}.$$

Then μ satisfies all the required conditions.

References

[1] Bonomo, M. and R. Garcia (1993), "Disappointment Aversion as a Solution to the Risk-Free Puzzle", U. Montreal, mimeo.

[2] Border, K.C. (1991),"Functional Analytic Tools for Expected Utility Theory", in *Positive Operators, Riesz Spaces and Economics*, C. Aliprantis, K. Border and W. Luxemburg eds., Springer-Verlag.

[3] Border, K.C. (1992), "Revealed Preference, Stochastic Dominance, and the Expected Utility Hypothesis", *Journal of Economic Theory*, 56, 20-42.

[4] Bufman, G. and L. Leiderman (1990), "Consumption and Asset Returns Under Nonexpected Utility", *Economic Letters*, 34, 231-235.

[5] Camerer, C.F. (1989), "An Experimental Test of Several Generalized Utility Theories", *Journal of Risk and Uncertainty*, 2, 61-104.

[6] Chew, S.H. (1983), "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory", *Econometrica*, 51, 1065-92.

[7] Chew, S.H. (1989), "Axiomatic Utility Theories with the Betweeness Property", *Annals of Operations Research*, 19, 273-298.

[8] Cochrane, J.H. and L.P. Hansen (1992), "Asset Pricing Explorations for Macroeconomics", NBER Working Paper #4088.

[9] Constantinides, G.M. (1990), "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, 98, 519-543.

[10] Dekel, E. (1986), "An Axiomatic Characterization of Preferences Under Uncertainty", *Journal of Economic Theory,* 40, 304-318.

[11] Dybvig, P.H. and S.A. Ross (1982), "Portfolio Efficient Sets", *Econometrica*, 50, 1525-1546.

[12] Epstein, L.G. (1988), "Risk Aversion and Asset Prices", *Journal of Monetary Economics*, 22, 179-192.

[13] Epstein, L.G. (1992), "Behavior Under Risk: Recent Developments in Theory and Applications", in *Advances in Economic Theory,* Vol.II, J.J. Laffont ed., Cambridge U. Press.

[14] Epstein, L.G. and T. Wang (1993), "Intertemporal Asset Pricing Under Knightian Uncertainty", U. Toronto Working Paper #9211, revised, to appear in *Econometrica.*

[15] Epstein, L.G. and S. Zin (1989), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, 57, 937-969.

[16] Epstein, L.G. and S. Zin (1990), "First-Order Risk Aversion and the Equity Premium Puzzle", *Journal of Monetary Economics,* 26, 387-407.

[17] Epstein, L.G. and S. Zin (1991), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis", *Journal of Political Economy*, 99, 263-286.

[18] Epstein, L.G. and S. Zin (1992), "The Independence Axiom and Asset Returns", NBER Technical Working Paper #109.

[19] Fishburn, P. (1975), "Separation Theorems and Expected Utilities", *Journal of Economic Theory*, 11, 16-34.

[20] Gallant, A.R. and G. Tauchen (1989), "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications", *Econometrica*, 57, 1091-1020.

[21] Green E. and K. Osband (1991), "A Revealed Preference Theory for Expected Utility", *Review of Economic Studies*, 58, 677-96.

[22] Green, J., L. Lau and H. Polemarchakis (1978), "Identification of the von Neumann-Morgenstern Utility Function from Asset Demands", in *General Equilibrium, Growth and Trade*, J. Green and J. Scheinkman eds., Academic Press.

[23] Green, R. and S. Srivastava (1986), "Expected Utility Maximization and Demand Behavior", *Journal of Economic Theory*, 38, 313-323.

[24] Grossman, S.J. and R.J. Shiller (1981), "The Determinants of the Variability of Stock Market Prices", *American Economic Review*, 71, 222-7.

[25] Hadar, J. and W.R. Russell (1969), "Rules for Ordering Uncertain Prospects", *American Economic Review*, 59, 25-34.

[26] Hanoch, G. and H. Levy (1969), "Efficiency Analysis of Choices Involving Risk", *Review of Economic Studies*, 36, 335-46.

[27] Hansen, L.P. and T.J. Sargent (1992), "Discounted Linear Exponential Quadratic Gaussian Control", mimeo.

[28] Hansen, L.P., T.J. Sargent and T.D.Tallarini Jr. (1993), "Pessimism, Neurosis, and Feelings about Risk in General Equilibrium", mimeo.

[29] Hansen, L.P. and K.J. Singleton (1983), "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns", *Journal of Political Economy*, 91, 249-65.

[30] Hansen, L. P. and R. Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies", *Journal of Political Economy*, 99, 225-262.

[31] Heaton, J. (1993), "The Interaction Between Time-Nonseparable Preferences and Time Aggregation", *Econometrica*, 61, 353-386.

[32] Kandel, S. and R.F. Stambaugh (1989), "Modeling Expected Stock Returns for Long and Short Horizons", mimeo.

[33] Kandel, S. and R.F. Stambaugh (1991), "Asset Returns and Intertemporal Preferences", *Journal of Monetary Economics*, 27, 39-71.

[34] Klecan, L. and D. McFadden (1991), "A Robust Test for Stochastic Dominance", mimeo.

[35] Lucas, R.E. Jr. (1978), "Asset Prices in an Exchange Economy", *Econometrica*, 46, 1429-1445.

[36] Machina, M.J. (1987), "Choice Under Uncertainty: Problems Solved and Unsolved", *Journal of Economic Perspectives*, 1, 121-154.

[37] McFadden, D. (1989), "Testing for Stochastic Dominance", in *Studies in the Economics of Uncertainty*, T. Fomby and T. Seo eds., Springer-Verlag.

[38] Mehra, R. and E. Prescott (1985), "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, 15, 145-161.

[39] Obstfeld, M. (1992), "Evaluating Risky Consumption Paths: The Role of Intertemporal Substitutability", NBER Technical Working Paper #120.

[40] Reitz, T.A. (1988), "The Equity Premium: A Solution", *Journal of Monetary Economics*, 22, 117-133.

[41] Starmer, C. (1992), "Testing New Theories of Choice Under Uncertainty Using the Common Consequence Effect", *Review of Economic Studies*, 59, 813-830.

[42] Varian, H. (1983), "Non-Parametric Tests of Investor Behavior", *Journal of Financial and Quantitative Analysis*, 18, 2269-278.

[43] Wang, S. (1993), "The Local Recoverability of Risk Aversion and Intertemporal Substitution", *Journal of Economic Theory*, 59, 333-363.

[44] Weil, P. (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle", *Journal of Monetary Economics*, 24, 401-22.

[45] Weil, P. (1990), "Nonexpected Utility in Macroeconomics", *Quarterly Journal of Economics*, 105, 29-42.

Fig 1: Illustration of Theorem 1

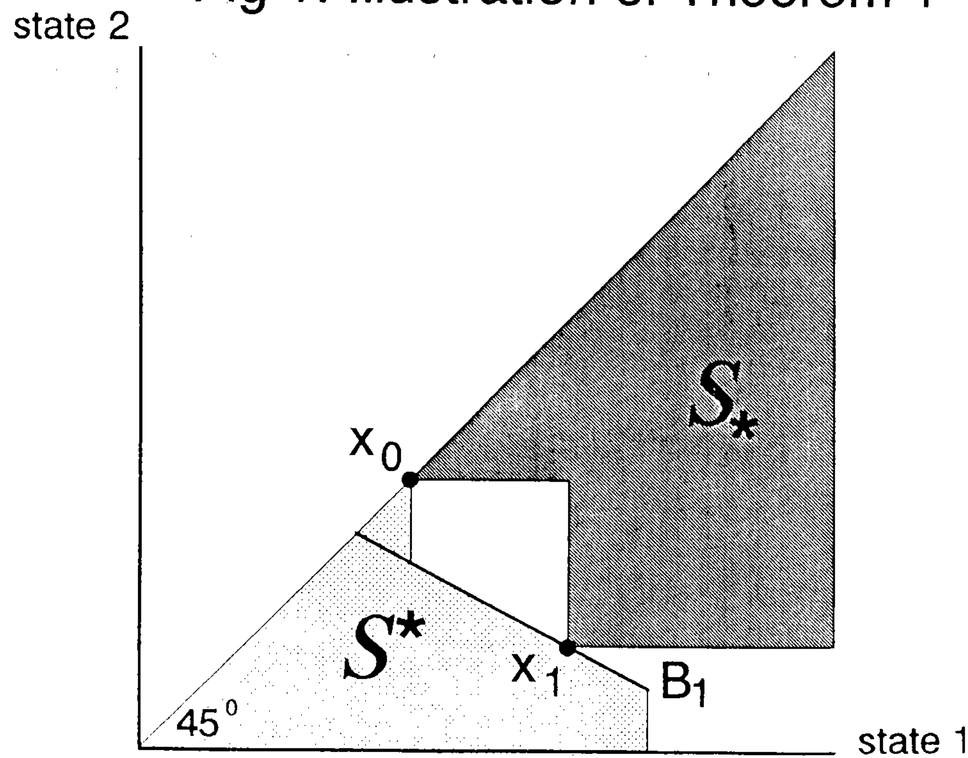
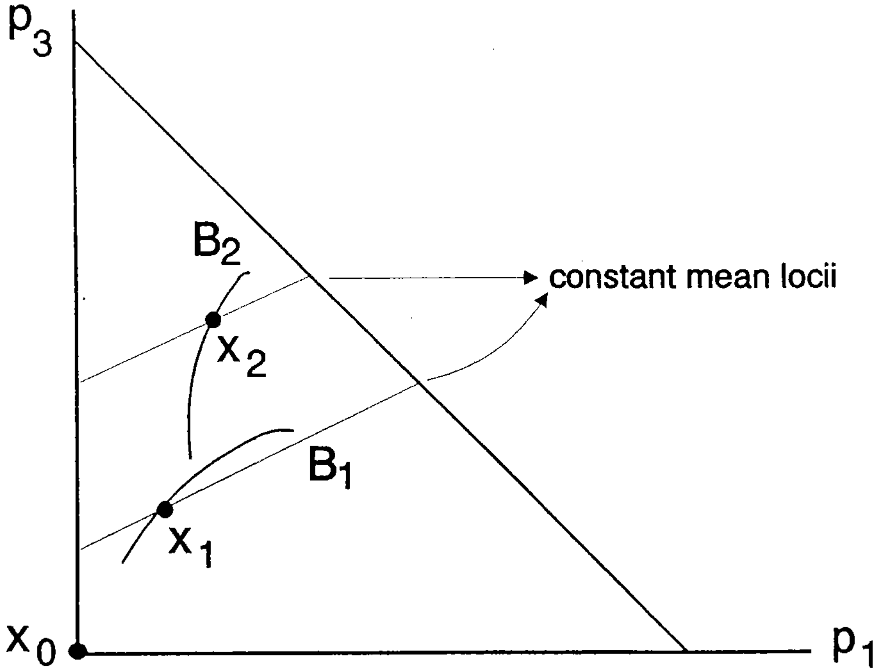


Fig 2: A Rationalizable Data Set



Probability simplex for gambles with outcomes z_1, 1, z_3; $z_1 < 1 < z_3$; and corresponding probabilities p_1, $1-p_1-p_2$, p_3. The data $\{x_t, B_t\}_{t=0}^{2}$ can be rationalized by a monotone and risk averse certainty equivalent that is neither quasiconcave nor quasiconvex.

Table 1. Asset Return Moments

	Economy 1		Economy 2	
	$E(M) = 1.070$ $\sigma(M) = .165$		$E(M) = 1.070$ $\sigma(M) = .115$	
	$P_l = 23.42$ $P_h = 27.89$		$P_l = 20.76$ $P_h = 23.16$	
Nonparametric Examples				
NP Range for ρ	$[-\infty,-15.63]$		$[-\infty,-9.78]$	
	$\hat{\rho} = -16$	$\hat{\rho} = -50$	$\hat{\rho} = -10$	$\hat{\rho} = -40$
NP Range for r_l	[1.024,1.104]	[1.024,1.090]	[1.029,1.094]	[1.029,1.069]
NP Range for r_h	[.860, .865]	[.860, .905]	[.923, .927]	[.923, .961]
$\hat{r}_l$	1.104	1.090	1.089	1.064
$\hat{r}_h$	.865	.905	.927	.952
$E(\hat{r})$	.985	.997	1.008	1.008
$\sigma(\hat{r})$	.119	.093	.081	.056
$[\mu_*(Z),\mu^*(Z)]$	[50.7,56.5]	[50.7,62.7]	[50.8,56.3]	[50.8,59.2]
Parametric Examples				
	Kreps-Porteus	Yaari	Kreps-Porteus	Yaari
	$\hat{\rho}^{KP} = -38.02$ $\hat{\alpha}^{KP} = -36.66$	$\hat{\rho}^{Y} = -41.95$ $\hat{\gamma}^{Y} = .4057$	$\hat{\rho}^{KP} = -23.08$ $\hat{\alpha}^{KP} = -34.44$	$\hat{\rho}^{Y} = -25.50$ $\hat{\gamma}^{Y} = .4243$
NP Range for r_l	[1.024,1.092]	[1.024,1.091]	[1.029,1.082]	[1.029,1.082]
NP Range for r_h	[.860, .903]	[.860, .903]	[.923, .959]	[.923, .959]
$\hat{r}_l$	1.045	1.091	1.048	1.082
$\hat{r}_h$	.873	.903	.934	.959
$E(\hat{r})$	.959	.997	.991	1.020
$\sigma(\hat{r})$	.086	.094	.057	.061
$\mu(Z)$	51.0	62.3	51.0	62.7
$[\mu_*(Z),\mu^*(Z)]$	[50.7,62.8]	[50.7,62.7]	[50.8,59.4]	[50.7,63.2]

M is the return to equity, P is the price-dividend ratio, and r is the risk free rate. NP denotes nonparametrically determined. Z is a fair binary gamble with outcomes 50 and 100. All the calculations assume $\beta = .99$, and the consumption growth process described in Section 4.