Technical Memorandum Number 724F

Discounting and Risk Neutrality

by

Matthew J. Sobel

Revised September 2006

Department of Operations
Weatherhead School of Management
Case Western Reserve University
330 Peter B Lewis Building
Cleveland, Ohio 44106

DISCOUNTING AND RISK NEUTRALITY*

Matthew J. Sobel**

September 1990; Revision September 15, 2006

Abstract

Let $(\succeq, V)$ be a binary relation in which V is a real vector space of vector-valued discrete-time stochastic processes with **0** as the sequence of zero vectors. If $(\succeq, V)$ satisfies four axioms, then there are unique discount factors such that preferences regarding stochastic processes induce prefences among present value random vectors. Three of the axioms are familiar: weak ordering, continuity, and non-triviality. The fourth axiom, *decomposition*, is $X - Y \succeq (\preceq)$ **0** implies $X \succeq (\preceq) Y$ where **0** is the sequence of zero vectors. Also, if preferences satisfy the four axioms then the following properties are equivalent: the converse of decomposition, the existence of a felicity function on the set of random vectors, and risk neutrality. In this limited sense, discounting implies risk neutrality.

JEL

D81, D99, G12

*In honor of Martin Shubik and in memory of Robert Rosenthal.

**Department of Operations, Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106-7235.

Samuelson (1937) proposed the discounted utility model in a deterministic framework which was subsequently axiomatized by Koopmans (1960), Williams and Nasser (1966), and others. An axiomatic foundation for discounting in a stochastic setting, the subject of this paper, addresses the basic properties of intertemporal preferences which yield utility functions having particular structures. Since the discounted utility model is additively separable, the literature on separability is relevant to this issue (Blackorby, Primont, and Russell (1998)). In spite of this considerable attention, inconsistencies abound in professional practice and in research literatures.

Under a strong form of Koopmans' axioms, there exists an *intraperiod* utility function u and a discount factor $\beta \in (0,1)$ such that, for all deterministic scalar sequences $x = (x_1, x_2, ...)$ and $y = (y_1, y_2, ...)$,

(1) $$x \text{ is weakly preferred to } y \quad \Leftrightarrow \quad \sum_t \beta^{t-1} u(x_t) \geq \sum_t \beta^{t-1} u(y_t)$$

If x and y were replaced by stochastic processes $X = (X_1, X_2, ...)$ and $Y = (Y_1, Y_2, ...)$, then each side of the inequality would be a random variable. Would a combination of Koopmans' axioms and those from utility theory, von Neumann and Morgenstern (1944) for example, imply the existence of an *interperiod* utility function v such that

(2) $$X \text{ is weakly preferred to } Y \quad \Leftrightarrow \quad E\Big(v[\sum_t \beta^{t-1} u(X_t)]\Big) \geq E\Big(v[\sum_t \beta^{t-1} u(Y_t)]\Big)$$

(where E denotes expected value)?

In the axiomatization of preferences among scalar sequences by Williams and Nassar (1966), u is the identity function. So

(3) $$x \text{ is weakly preferred to } y \quad \Leftrightarrow \quad \sum_t \beta^{t-1} x_t \geq \sum_t \beta^{t-1} y_t$$

Would a combination of axioms from Williams and Nassar and utility theory lead to

(4) $$X \text{ is weakly preferred to } Y \quad \Leftrightarrow \quad E[v(\sum_t \beta^{t-1} X_t)] \geq E[v(\sum_t \beta^{t-1} Y_t)]$$

The preference model in numerous research fields is

(5) $$X \text{ is weakly preferred to } Y \quad \Leftrightarrow \quad E[\sum_t \beta^{t-1} u(X_t)] \geq E[\sum_t \beta^{t-1} u(Y_t)]$$

for an intraperiod utility function u. Of course this is (2) with risk neutrality and risk sensitivity, respectively, concerning interperiod utility and intraperiod utility. An elementary ordering that is widely used in professional practice and several research fields is

(6) $$X \text{ is weakly preferred to } Y \quad \Leftrightarrow E(\sum_t \beta^{t-1} X_t) \geq E(\sum_t \beta^{t-1} Y_t)$$

Of course this is (2) with risk-neutral interperiod and intraperiod utility functions, and it is (5) with a risk-neutral intraperiod utility function.

Alternative decision rules are compared via (5) and (6) in diverse areas such as accounting [Moriarity and Allen (1984)], advertising [Monahan (1983)], banking [Shubik and Sobel (1992)], biology [McNamara (1990), Mendelssohn (1979)], capital accumulation [Sethi (1998)], energy [Murphy, Toman and Weiss (1989)], engineering [Yeh (1985)], finance [Altug and Labadie (1994), Cochrane (2000), Duffie (2001), Sharpe (1985), Ho and Lee (1986)], game theory [Shapley (1953)], production [Arrow, Karlin and Scarf (1962)], replacement [Jorgenson, McCall and Radner (1967)], social psychology [Shubik (1970)], and technological change [Balcer and Lippman (1984)].

Can (2), (4), (5), and (6) be reconciled? Do successively stronger assumptions underlie each of (2), (4), and (6) and (2), (5), and (6)? Researchers utilize (5) instead of (6) to investigate the effects of risk sensitivity. So it is not interesting that (2), (4), (5), and (6) are equivalent if the decision maker is risk neutral.

There is a fundamental difficulty with using (5) to investigate the effects of risk sensitivity. When axioms that imply the existence of discount factors (Theorem 1 in §2) are augmented with the existence of an *intra*period utility function, then that function is linear (Theorem 2 in §3). Professor James E. Smith, Duke University, notes that (4) implies that an interperiod utility function is an intraperiod utility function in a single period mdel. So it too must be linear. In this limited sense, discounting implies risk neutrality.

Miyamoto and Wakker (1996) unify much of the literature on multiattribute preference orderings and utility functions that is reviewed by Dyer and Sarin (1979), Farquhar (1977), Fishburn (1978), and Keeney and Raiffa (1976). Wakker (1993) has axioms implying that there are positive numbers β_1, $\beta_2, \ldots$ and a real-valued function $r(\,\cdot\,)$ of random variables such that X is weakly preferred to Y if and only if $U(X) \geq U(Y)$ where $U(X) = \Sigma_t \beta_t r(X_t)$. Let D_a be a degenerate random variable that takes the value a with probability one, and let w be the real-valued function on the reals such that $w(\text{a}) = r(D_a)$. The linearity of r (in probabilities in the sense of (11) in §3) implies $U(X) = E[\Sigma_t \beta_t w(X_t)]$. This paper shows that $w(\,\cdot\,)$ is linear (*i.e.*, risk neutral) under weak sufficient conditions for additive separability. This result diminishes the justification for (5).

Literatures which depend on (5), such as asset pricing and, more generally, theoretical financial economics, are left in an ambiguous position. However, individuals *do* exhibit sensitivity to risk (gambling, insurance, etc.); so this paper casts further doubt on the descriptive validity of conclusions predicated on expected present value. Experimental work challenges expected utility theory and comparisons of deterministic temporal sequences via present values [Frederick, Loewenstein, and O'Donoghue (2003)].

A relatively small body of work provides axiomatic justifications for comparing deterministic sequences via their present values. Koopmans (1960) and Koopmans, Diamond, and Williamson (1964) study orderings of countable sequences of vectors of real numbers. They postulate the existence of an interperiod utility function $v(\,\cdot\,)$ such that deterministic sequence $x = \ < x_t >$ is weakly preferred to $y = \ < y_t >$ if and only if $v(x) \geq v(y)$. They obtain sufficient conditions for the existence of $\beta \in (0,1)$ and an intraperiod utility function $u : \mathbb{R}^M \to \mathbb{R}$ such that $v(x) = \sum_{t=1}^{\infty} \beta^{t-1} u(x_t)$. Koopmans (1972) specifies conditions on a binary relation which imply the existence of $v(\,\cdot\,)$ with the properties assumed in the earlier studies.

Williams and Nassar (1966) obtain sufficient conditions for a weak ordering $\succeq$ on a real vector space of finite sequences $x = \ (x_1,...,x_T)$ of real numbers to have the following property: there exists $\beta \in$ (0,1) such that $x \ \succeq \ y$ if and only if $\sum_{t=1}^{T} \beta^{t-1}(x_t - y_t) \geq 0$.

Lancaster (1963), Fishburn (1970), and others cited by Fishburn and Rubinstein (1982) and Frederick, Loewenstein, and O'Donogue (2003) investigate closely related deterministic models of time preference and impatience. Epstein (1983) studies a stochastic model that relates time preference and risk preference. He specifies necessary and sufficient conditions for a binary relation $\succeq$ on the set of real-valued discrete-time stochastic processes to have the following property. Let $x = (x_1,x_2,...)$ and

$y = (y_1, y_2, ...)$ be degenerate (*i.e.*, constant with probability one) real-valued stochastic processes. There exists $\beta \in (0,1)$ and $u : \mathbb{R} \to \mathbb{R}$ such that $x \succeq y$ if and only if $\sum_{t-1}^{\infty} \beta^{t-1} u(x_t) \geq \sum_{t-1}^{\infty} \beta^{t-1} u(y_t)$. Technical differences aside, this paper presents conditions under which u is affine. The axioms in this paper are briefly compared in §1 with the assumptions made by Williams and Nassar, Koopmans, Epstein, and Miyamoto and Wakker.

Meyer (1976) uses the same rationale that leads from (1) to (2). He develops a cardinal comparison of alternative sample paths which are outcomes in the sample space, say of $x = (x_1, x_2, ...)$ and $y = (y_1, y_2, ...)$, and then, observing that the cardinal measures are random variables, he applies expected utility theory. He implicitly assumes orthogonality of the axioms underlying time preference and risk preference. *The present paper answers a question addressed to the author by Robert Rosenthal (1987) who wondered if the two axiom systems might be related*. They are indeed related.

It is not clear how the previously cited studies and this paper are are related to investigations of consistency in intertemporal choice preference structures. See Kreps and Porteus (1979), Johnson and Donaldson (1985), Machina (1989), and their references.

Section 1 presents notation and discusses axioms. Section 2 has sufficient conditions for discounting, *i.e.*, four axioms that imply that comparisons between vector-valued stochastic processes correspond to comparisons between their present values (which are random vectors). A further axiom yields a vector of familiar geometric discount factors. Three of the axioms are familiar: weak ordering, continuity, and non-triviality. Instead of an independence or Archimedean axiom, the fourth is *decomposition*, namely $X - Y \succeq (\preceq)$ **0** implies $X \succeq (\preceq) Y$ where **0** is the sequence of zero vectors. Section 3 shows that if preferences satisfy the four axioms then the following properties are equivalent: the converse of decomposition, the existence of a felicity function, and risk neutrality. In §4, a combination of results in earlier sections with some in Koopmans (1972) yields sufficient conditions for a scalar discount factor to generate vectors of geometric discount factors. Section 5 briefly examines relationships among alternative continuity axioms and definitions of impatience. Section 6 sketches the implications of §3 for attribute decomposition in random vectors.

1. Notation and Axioms

Fix a probability space (Ω,$\mathcal{F}$,P). Let I and V be respective sets of positive integers and countable stochastic sequences $X = (X_1, X_2, ...)$ defined on (Ω,$\mathcal{F}$,P) with $X_t(\omega) \in \mathbb{R}^M$ for all $(t, \omega) \in I \times \Omega$. Let X_{mt} denote the m-*th* component of X_t. If $X \in V$, $Y \in V$, and $b \in \mathbb{R}$, let $X + Y \in V$ and $bX \in V$ have values at $(t, \omega) \in I \times \Omega$ given by $X_t(\omega) + Y_t(\omega)$ and $bX_t(\omega)$. Let θ be the zero vector in $\mathbb{R}^M$, and let $\mathbf{0} = (\theta,\theta,...) \in V$. So V is a real vector space with zero element $\mathbf{0}$.

Let e_m be the m-*th* unit vector in $\mathbb{R}^M$ and let $e_{mt} \in V$ be the sequence $\mathbf{0}$ of θ vectors except that the t-*th* vector is e_m. For $u = (u_i) \in \mathbb{R}^M$ and $\upsilon = (\upsilon_i) \in \mathbb{R}^M$ let $u \cdot \upsilon$ denote $\sum_{i=1}^{M} u_i \upsilon_i$.

Let $\succeq$ be a binary relation on V, write $X \sim Y$ if $X \succeq Y$ and $Y \succeq X$, and write $X \succ Y$ if $X \succeq Y$ but not $Y \succeq X$. The following notations are written interchangeably: $X \succeq Y$ and $Y \preceq X$, $X \sim Y$ and $Y \sim X$, and $X \succ Y$ and $Y \prec X$. A *weak ordering* (sometimes called a *strong ordering*) is a complete transitive binary relation.

Let S be the set of $\mathbb{R}^M$-valued random vectors. For $C \in S$, let (C,$\mathbf{0}$) $\in V$ denote (C,θ,θ,...). Then ($\succeq$,V) induces the following binary relation $\geqslant$ on S: $A \geqslant B$ if and only if (A,$\mathbf{0}$) $\succeq$ (B,$\mathbf{0}$). Let $W \approx Z$ denote $W \geqslant Z$ and $Z \geqslant W$ and let $W \gg Z$ denote $W \geqslant Z$ but not $Z \geqslant W$. Notations employed interchangeably include $W \geqslant Z$ and $Z \leqslant W$, $W \approx Z$ and $Z \approx W$, and $W \gg Z$ and $Z \ll W$. For $A \in S$, let $E(A) \in \mathbb{R}^M$ denote the vector of expectations of components (when the expectations exist).

Axioms

In the following axioms, $X = (X_1, X_2, ...) \in V$, $Y = (Y_1, Y_2, ...) \in V$, $t \in I$, $m = 1, ..., M$, and $\lambda > 0$.

(A1) *Transitivity and rationality:* $\succeq$ *weakly orders* V

(A2) *Decomposition:* $X - Y \succeq (\preceq) \mathbf{0}$ *implies* $X \succeq (\preceq) Y$

(A3) *Continuity:* $\{\alpha\text{: } \alpha X - Y \succeq (\preceq) \mathbf{0}\}$ is closed

(A4) *More is better:* $e_{mt} \succ \mathbf{0}$

(A5) *Stationarity:* $(X_1,X_2,...) \sim \mathbf{0}$ *implies* $(\theta,X_1,X_2,...) \sim \mathbf{0}$

(A6) *Sooner is better:* $X + \lambda(e_{mt} - e_{m,t+1}) \succ X$

The first four axioms lead to a present value formula. The fifth axiom yields the usual geometric form of discount factors; the sixth axiom implies that discount factors are less than one. A loose interpretation of (A5) is that indifference between receiving a cash flow and not receiving it implies indifference for the one-period delay.

Reasonable alternative versions of the continuity axiom (A3) include the following statements:

(A3′) $X \succ (\prec) \mathbf{0}$ *implies that there exists* $\alpha^* > 0$ *such that* $\alpha X \prec (\succ) X$ *if* $0 < \alpha < \alpha^*$

(A3″) $X \succeq (\preceq) \mathbf{0} \;\Rightarrow\; bX \succeq (\preceq) \mathbf{0} \; \forall \, b \geq 0$

However, (A3′) and (A3″) are equivalent to each other and to the restriction of (A3) in which $Y = \mathbf{0}$ (Proposition 5 in §7).

Decomposition axiom and its converse

Axiom (A2) is particularly important and may be contrasted with its converse:

(A2^c) $X \succeq (\preceq) Y \;\Rightarrow\; X - Y \succeq (\preceq) \mathbf{0}$ $\qquad (X \in V; Y \in V)$

The present value formula in §2 invokes (A2) but not (A2^c).

Proposition 1: *Suppose (A1) is valid.*

(a) (A2) $\Leftrightarrow$ $[W \succeq (\preceq, \sim) \; \mathbf{0} \Rightarrow W + Z \succeq (\preceq, \sim) Z \;\; \forall \, Z \in V]$

(b) (A2^c) $\Leftrightarrow$ $[W + Z \succeq (\preceq, \sim) Z \;\Rightarrow\; W \succeq (\preceq, \sim) \mathbf{0}]$

Proof: For sufficiency in (a), use $X = W + Z$ and $Y = Z$. For necessity in (a), use $W = X - Y$ and $Z = Y$. For (b), let $X = W + Z$ and $Y = Z$. $\square$

Thus, with (A2), W is as good as the status quo only if incrementing *any* Z with W is as good as Z. With (A2^c), W is not as good as the status quo if there is *any* Z which is better than Z augmented by W. The combination of (A2) and (A2^c) is a version of the “independence" assumptions that recur (and whose descriptive validity is questioned) in axiomatic theories of decision making. Also, this combination corresponds to “persistence," “monotonicity," and “consistent choice" in dynamic programming [Denardo (1967), Sobel (1975, 1980), and Blair (1984)]. Two applications of part (a) imply

$$[X - Y \sim \mathbf{0} \;\Rightarrow\; X \sim Y] \;\Leftrightarrow\; [W \sim \mathbf{0} \Rightarrow W + Z \sim Z \;\; \forall \, Z \in V]$$

The counterpart to V in Miyamoto and Wakker (1996) can be construed as a set of sequences of time-indexed marginal probability distributions. Allowing for difference in the models, their axioms include (A1) and (A3) and, instead of (A2), assumptions of outcome monotonicity and the Thomsen condition. Define outcome monotonicity here as $A \succeq (\preceq) B$ and $W \succeq (\preceq) Z$ imply $A + W \succeq (\preceq) B + Z$. The following result shows that (A2) alone is less restrictive than outcome monotonicity because the former is equivalent to (A2) together with (A2^c).

Proposition 2: *If* $(\succeq, V)$ *satisfies (A1) then it satisfies outcome monotonicity if and only if it satisfies both (A2) and (A2^c).*

Proof: *(A2) and (A2^c)* $\Rightarrow$ *outcome monotonicity:*

Using (A2^c), $A \succeq B$ and $W \succeq Z$ imply $A - B \succeq \mathbf{0}$ and $W - Z \succeq \mathbf{0}$. So Lemma 1(b) in §2 implies $A + W - B - Z \succeq \mathbf{0}$ yielding $A + W \succeq B + Z$ by (A2).

Outcome monotonicity $\Rightarrow$ *(A2)*:

Let $A = X - Y, B = \mathbf{0}$, and $W = Z = Y$. Thus, $A \succeq B$ and $W \succeq Z$ are $X - Y \succeq \mathbf{0}$ and $Y \succeq Y$. So outcome monotonicity, $A + W \succeq B + Z$, is $X \succeq Y$.

Outcome monotonicity $\Rightarrow$ *(A2^c)*:

Let $A = X, B = Y$, and $W = Z = -Y$. Thus, $A \succeq B$ and $W \succeq Z$ are $X \succeq Y$ and $-Y \succeq -Y$. So outcome monotonicity is $X - Y \succeq \mathbf{0}$. $\square$

Preferences are *constant risk averse* if $X \succeq Y$ implies $X + \epsilon \succeq Y + \epsilon$ for every $\epsilon \in V$ that is constant with probability one. However, $X \succeq Y$ and (A2^c) imply $X - Y \succeq \mathbf{0} \Rightarrow X + \epsilon \succeq Y + \epsilon$ from (A2). Preferences are *relative risk averse* if $X \succeq Y$ implies $aX \succeq aY$ for all $a > 0$. However, $X \succeq Y$, (A2^c), and Lemma 2(a) in §2 imply $X - Y \succeq \mathbf{0} \Leftrightarrow a(X - Y) \succeq \mathbf{0}$ for all $a \geq 0$. Then (A2) yields $aX \succeq aY$. Since axioms (A1), (A2), (A2^c), and (A3) imply both constant and relative risk aversion, one should anticipate risk neutrality (cf. Theorems 1 and 2 in Miyamoto and Wakker).

Axiom sets

The first four axioms imply that there are discount factor vectors such that a vector-valued stochastic process is weakly preferred to another if and only if there is a corresponding weak preference concerning the present value random vectors. Conversely, axioms (A1) and (A3) are satisfied by any preference relation among stochastic processes that is determined by preference among present value random vectors. Also, (A4) is satisfied if the preference among random vectors has the property that more is preferred to less. However, (A2) is necessarily satisfied only if X and Y are independent. So the key axiom is decomposition, namely (A2).

The following example satisfies (A1), (A3), and (A4) but neither (A2) nor (A2^c). Let $M = T = 1$, let $\gg$ denote preferences among random variables corresponding to variance minus expected value, $\Omega = \{a, b\}$ with probabilities 3/4 for a and 1/4 for b, and $X(a) = 0,\ X(b) = -1,\ Y(a) = -1,$ and $Y(b) = 0$. Then $X - Y \gg \theta$ and $Y \gg X$, so $X - Y \succ \mathbf{0}$ and $Y \succ X$.

A loose interpretation of (A5) is that indifference between receiving a cash flow and not receiving it implies the same indifference for the one-period delay. In §4, the converse of (A5) implies that a scalar discount factor generates a vector-valued discounte factor.

Consider restrictions of the assumptions to finitely long deterministic sequences of scalars. The axioms in Williams and Nasser (1966) correspond to (A1) through (A6) (with (A2^c) and a continuity assumption different from (A3)). The postulates in Koopmans (1972) roughly correspond to (A1) through (A5), (A2^c) and the converse of (A5). It is surprising that “sooner is better" is implied by his postulates (although he does not directly assume the counterpart of (A6)). The preference ordering in Epstein (1983) is defined on the space of probability measures on the Borel field of countable sequences of real numbers. Epstein's assumptions, a restatement of Koopmans' in a stochastic framework, correspond roughly to (A1) through (A5), (A2^c), and the converse of (A5). The continuity assumptions in Williams and Nassar (1966), Koopmans (1972), and Epstein (1983) differ from each other and from (A3).

2. Discounting Theorem

This section owes much to Williams and Nassar (1966). Its main results, Theorem 1 and Corollaries 1 and 2, are stated and then proved with several lemmas.

It is convenient to define $\beta_{m1} = 1$ for each m, $\beta_t = (\beta_{1t},...,\beta_{Mt})$ for each t, and for $X \in V$ to write $\beta_t X_t$ for the random vector with m^{th} component $\beta_{mt} X_{mt}$. Recall that $(\succeq, V)$ induces the binary relation ($\geqslant$,S) on the set of random M-vectors. The following result gives sufficient conditions for preferences in finitely long processes to be consistent with a stochastic discounting formula. The proof would be simplified by replacing the assumption of (A2) with the stronger assumption of outcome monotonicity.

Theorem 1: *Axioms* (A1), (A2), (A3), *and* (A4) *imply the unique existence of* $\beta_{mt} > 0$ (*for each* m *and* t) *such that*

(7) $$X \succeq Y \;\Leftrightarrow\; \sum_{t=1}^{T} \beta_t X_t \geqslant \sum_{t=1}^{T} \beta_t Y_t$$

for all $X = (X_1,...,X_T,\mathbf{0})$ *and* $Y = (Y_1,...,Y_T,\mathbf{0})$ *with* $T \in I$.

The proof of Theorem 1 uses the property that $Y^{(j)} \sim \mathbf{0}$, $j = 1, ..., J$, implies $\sum_{j=1}^{J} Y^{(j)} \sim \mathbf{0}$ (Lemma 1(d)). The following countable analog is one of several ways to extend the theorem:

(8) If $Y^{(j)} \sim \mathbf{0}$ for all j and $\sum_{j=1}^{\infty} Y^{(j)} \in V$ then $\sum_{j=1}^{\infty} Y^{(j)} \sim \mathbf{0}$.

Corollary 1. *Axioms* (A1), (A2), (A3), *and* (A4) *and* (8) *imply*

$$X \succeq Y \;\Leftrightarrow\; \sum_{t=1}^{\infty} \beta_t X_t \geqslant \sum_{t=1}^{\infty} \beta_t Y_t$$

if $\sum_{t=1}^{\infty} \beta_t X_t$ *and* $\sum_{t=1}^{\infty} \beta_t Y_t$ *exist and have finite components with probability one.*

Now the stationarity axiom (A5) yields geometric discount factors.

Corollary 2. *Axioms* (A1), (A2), (A3), (A4), *and* (A5) *imply for each* $m = 1, ..., M$ *there uniquely exists* $\beta_m > 0$ *such that* $\beta_{mt} = (\beta_m)^{t-1}$ *for all* $t \in I$. *If also* (A6) *is valid then* $\beta_m < 1$.

Proof of Theorem 1

Lemma 1. Assuming (A1) and (A2):

(a) $X \succ (\succeq, \sim)\, \mathbf{0} \;\Leftrightarrow\; -X \prec (\preceq, \sim)\, \mathbf{0}$

(b) $X \succeq (\preceq, \sim)\, \mathbf{0}$ *and* $Y \succeq (\preceq, \sim)\, \mathbf{0} \Rightarrow X + Y \succeq (\preceq, \sim)\, \mathbf{0}$

(c) $X \succeq (\preceq)\, \mathbf{0}$ *and* $Y \succ (\prec)\, \mathbf{0} \Rightarrow X + Y \succ (\prec)\, \mathbf{0}$

(d) $X \sim \mathbf{0}$ *and* $Y \sim \mathbf{0} \;\Rightarrow\; X + Y \sim \mathbf{0}$

(e) $X - Y \sim \mathbf{0} \;\Rightarrow\; X \sim Y$

Proof. (a): $-X \preceq \mathbf{0} \Leftrightarrow \mathbf{0} - X \preceq \mathbf{0} \Rightarrow X \succeq \mathbf{0}$ due to (A2). So $X \succeq \mathbf{0} \Leftrightarrow -X \preceq \mathbf{0}$.

Now $X \sim \mathbf{0} \Leftrightarrow X \succeq \mathbf{0}$ and $X \preceq \mathbf{0} \Leftrightarrow -X \preceq \mathbf{0}$ and $-X \preceq \mathbf{0} \Leftrightarrow -X \sim \mathbf{0}$.

Now $X \succ \mathbf{0} \Leftrightarrow X \succeq \mathbf{0}$ and not $X \sim \mathbf{0}$. So not $-X \sim \mathbf{0}$. Also, $X \succeq \mathbf{0} \Leftrightarrow -X \preceq \mathbf{0}$. So $-X \prec \mathbf{0}$.

(b) and (c): The contrapositive of (c) is $X+Y \preceq \mathbf{0}$ implies $X \prec \mathbf{0}$, $Y \preceq \mathbf{0}$, or both. From (A2), $X+Y \preceq \mathbf{0}$ implies $X \preceq -Y$. So (A1) and $X \succeq \mathbf{0}$ imply $-Y \succeq \mathbf{0}$ and (a) yields Y $\preceq \mathbf{0}$. If $Y \succ \mathbf{0}$ then $X \preceq -Y$ and (a) imply $X \prec \mathbf{0}$. Using (c), (b) follows from [$X \succeq \mathbf{0}$ *and* $Y \sim \mathbf{0} \Rightarrow X+Y \succeq \mathbf{0}$] due to Proposition 1(a) and (A1).

(d): Two applications of (b).

(e): Two applications of (A2). $\square$

Lemma 2. Assuming (A1), (A2), *and* (A3):

(a) $X \succeq (\preceq) \mathbf{0} \Leftrightarrow cX \succeq (\preceq) \mathbf{0}$ *for all* $c \geq 0$.

(b) $\mathbf{0} \preceq (\succeq) X$ and $0 \leq c < c+\gamma \Rightarrow cX \preceq (\succeq) (c+\gamma)X$.

(c) $X \sim \mathbf{0} \Rightarrow cX \sim \mathbf{0}$ *for all* $c \in \mathbb{R}$.

(d) $cX \sim \mathbf{0}$ *and* $c \neq 0 \Rightarrow X \sim \mathbf{0}$.

Proof. (a): $\Leftarrow$ is trivial ($c=1$) and $\Rightarrow$ is trivial if $c=0$. For $\Rightarrow$ when $X \succeq \mathbf{0}$ and $c \in I$ (positive integers), Proposition 1(a) implies $X \preceq 2X$ so $\mathbf{0} \preceq 2X$. Inductively, $nX \preceq (n+1)X$ and $\mathbf{0} \preceq (n+1)X$. So $cX \succeq \mathbf{0}$ *for all* $c \in I$. Let $c \in \mathcal{T}$ where $\mathcal{T}$ is the set of positive rational numbers. So $c = m/n$ with $m \in I$ and $n \in I$, and $c = \sum_1^m X/n$. Therefore, $cX \succeq \mathbf{0}$ if $X/n \succeq \mathbf{0}$. If $X/n \prec \mathbf{0}$ then Proposition 1(a) with $W = Z = X/n$ yields $X/n \succeq 2X/n$; so (A1) implies $2X/n \prec \mathbf{0}$. Inductively, $(k+1)X/n \preceq kX/n$ and $kX/n \prec \mathbf{0}$ for $k \in I$. In particular, $X = nX/n \prec \mathbf{0}$. So $X/n \succeq \mathbf{0}$ because $X \succeq \mathbf{0}$. Therefore, $cX \succeq \mathbf{0}$ for all $c \in \mathcal{T}$.

If $c \notin \mathcal{T}$, let $c = n+f$ where $f \in (0,1]$ and $n \in I \cup \{0\}$. Hence, $nX \succeq \mathbf{0}$ and there exists a sequence $f_i \in \mathcal{T}$ with $f_i \to f$ as $i \to \infty$. So $f_i X \succeq \mathbf{0}$ for all i, (A3), and Lemma 1(b) imply $fX \succeq \mathbf{0}$ and $cX = nX + fX \succeq \mathbf{0}$.

(b): $\mathbf{0} \preceq X$, (a), $\mathbf{0} \preceq cX$**,** $\mathbf{0} \preceq \gamma X$**,** and Proposition 1(a).

(c) $X \sim \mathbf{0} \Leftrightarrow [X \succeq \mathbf{0}$ and $X \preceq \mathbf{0}] \Leftrightarrow$ (by Lemma 1(a) and (a)) $cX \succeq \mathbf{0}$ for all $c \in \Re$

(d) Use Lemma 1(a) and replace c with b, in (a) replace X with bX and c with b^{-1}.$\square$

The proof of Lemma 3(a) uses an argument due to Professor James C. Alexander, Department of Mathematics, Case Western Reserve University.

Lemma 3. *Assuming* (A1), (A2), (A3) *and* (A4):

(a) For all $t = 2, 3, ...$ *and* $m = 1, ..., M$ *there uniquely exists* $\beta_{mt} > 0$ *such that*

$$\mathbf{0} \sim \; - e_{m1} + \beta_{mt}^{-1} e_{mt}.$$

(b) Also assuming (A6) implies $\beta_{m,t+1} < \beta_{mt} < 1$.

Proof. (a): *Uniqueness* If $c > 0$ and $c' > 0$ with $-e_{m1} + ce_{mt} \sim \mathbf{0}$ and $-e_{m1} + c'e_{mt} \sim \mathbf{0}$ then $e_{m1} - c'e_{mt} \sim \mathbf{0}$ from Lemma 1(a), so $(c - c')e_{mt} \sim \mathbf{0}$ from Lemma 1(b). Uniqueness follows from (A4) and Lemma 2(d).

Existence Let $A = \{\alpha : \alpha e_{mt} - e_{m1} \succ \mathbf{0}\}$. If $A = \emptyset$ then $e_{mt} - (1/\alpha)e_{m1} \preceq \mathbf{0}$ for all large enough α (Lemma 2(a)), so (A3) implies $e_{mt} \preceq \mathbf{0}$. But $e_{mt} \succ \mathbf{0}$ so $A \neq \emptyset$.

Let $c = inf\{\alpha \in A\}$. If $\alpha \in A$ then $\alpha e_{mt} - e_{m1} \succ \mathbf{0}$ so $\alpha e_{mt} \succeq e_{m1} \succ \mathbf{0}$ ((A2) and (A4)). Now, α is neither zero ((A1)) nor negative (Lemma 2(a) and (A4)). So $c \geq \mathbf{0}$.

From (A3), $ce_{mt} - e_{m1} \succeq \mathbf{0}$. For any $\alpha < c$, $\alpha e_{mt} - e_{m1} \preceq \mathbf{0}$ so $ce_{mt} - e_{m1} \preceq \mathbf{0}$ from (A3). Therefore, $ce_{mt} - e_{m1} \sim \mathbf{0}$. If $c = 0$ then $-e_{m1} \sim \mathbf{0}$ so $e_{m1} \sim \mathbf{0}$ (Lemma 1(a)); but $e_{m1} \succ \mathbf{0}$ ((A4)). So $c > 0$. Let $\beta_{mt} = 1/c$.

(b): (A6) with $\lambda = 1$ and $X = \mathbf{0}$. $\square$

If (A6) were replaced with $X + \lambda(e_{mt} - e_{m,t+1}) \succ X$ for $t \geq \tau$, then Lemma 3(b) would become $\beta_{m,t+1} < \beta_{mt} < 1$ for $t \geq \tau$.

Conclusion of proof of Theorem 1. From Lemma 3, $\mathbf{0} \sim r_{mt} = \; - e_{m1} + \beta_{mt}^{-1} e_{mt}$. So Lemma 2(c) implies $\mathbf{0} \sim \; - \beta_{mt} e_{m1} + e_{mt}$. Therefore,

$$\begin{aligned}\mathbf{0} &\sim \textstyle\sum_{m=1}^{M} \sum_{t=1}^{T} X_{mt}(e_{mt} - \beta_{mt} e_{m1}) \\ &= X - \textstyle\sum_{m=1}^{M} \sum_{t=1}^{T} \beta_{mt} \, X_{mt} e_{m1} \\ &= X - (\textstyle\sum_{m=1}^{M} \sum_{t=1}^{T} \beta_{mt} X_{mt} e_m, \mathbf{0}) = X - (\sum_{t=1}^{T} \beta_t X_t, \mathbf{0}).\end{aligned}$$

Thus Lemma 2(a) yields $X \sim (\sum_{t=1}^{T} \beta_t X_t, \mathbf{0})$ and

$$X \succeq Y \;\Leftrightarrow\; (\textstyle\sum_{t=1}^{T} \beta_t X_t, \mathbf{0}) \;\succeq\; (\sum_{t=1}^{T} \beta_t Y_t, \mathbf{0}) \;\Leftrightarrow\; \sum_{t=1}^{T} \beta_t X_t \geqslant \sum_{t=1}^{T} \beta_t Y_t. \qquad \square$$

Proof of Corollary 2. From the proof of Lemma 3, $r_{mt} = -e_{m1} + \beta_{mt}^{-1} e_{mt} \sim \mathbf{0}$ for each $m = 1, ..., M$ and $t = 2, 3, ...$ From (A5), $\mathbf{0} \sim -e_{m2} + \beta_{mt}^{-1} e_{m,t+1}$ and the theorem implies $\theta \approx g$ where $g =$ (0,...,0, $-\beta_{m2} + \beta_{m,t+1}/\beta_{mt}$,0,...,0). Then $g = \theta$ for each m and t from (A4); so $\beta_{m,t+1} = (\beta_{m2})^t$. $\square$

Let $Y_t = \theta$, $t \in I$. If (A5) is replaced with

$$[(Y_1,...,Y_T,X_1,X_2,...) \sim \mathbf{0} \;\Rightarrow\; (Y_1,...,Y_T,\theta,X_1,X_2,...) \sim \mathbf{0}]$$

then $\beta_{mt} = \beta_{mT}(\beta_m)^{t-T}$ for all $t > T$, where $\beta_m = \beta_{m,T+1}/\beta_{mT}$.

3. Risk Neutrality

Recall that S denotes the set of $\mathbb{R}^M$-valued random vectors. A *felicity function* $u : S \to \mathbb{R}$ is order-preserving and linear, i.e., it satisfies (9) and (10) for all $\lambda \in$ [0,1], $A \in S$, and $B \in S$:

(9) $\quad A \gtrsim B$ *if and only if* $u(A) \geq u(B)$

(10) $\quad u[\lambda A + (1-\lambda)B] \;=\; \lambda u(A) + (1-\lambda)u(B)$

Note that a felicity function u is unique up to a positive affine transformation; that is, if u and w are felicity functions, then there are $b > 0$ and $a \in \mathbb{R}$ such that $w = a + bu$.

The following result axiomatizes the comparison of expected present values of cash flows, namely (6). Let V^* be the subset of V whose present value vectors, using any discount factor β with components $0 \leq \beta_m < 1$, are absolutely bounded with probability one.

Theorem 2: *If* $(\succeq, V)$ *satisfies* (A1), (A2), (A3), *and* (A4), then the following properties are equivalent:

(a) $(\succeq, V)$ *satisfies* (A2^c);

(b) *There exists a felicity function;*

(c) *There exists* $\mathfrak{v} \in \mathbb{R}^M$ *and* $\beta_t \in \mathbb{R}^M$, $t \in I$, *such that*

(12) $\quad \mathrm{X} \succeq \mathrm{Y} \quad \Leftrightarrow \quad \mathfrak{v} \cdot E(\sum_{t=1}^T \beta_t X_t) \;\geq\; \mathfrak{v} \cdot E(\sum_{t=1}^T \beta_t Y_t)$

for all $X \in V^*$, $Y \in V^*$ *with* $X_t = Y_t = \theta$ *for* $t > T$ *for some* $T \in I$.

Proof. Since (a) and (b) are immediate consequences of (c) with the felicity function $u(A) = \mathfrak{v} \cdot E(\beta_1 A)$, the proof shows that (b) implies (c), and that (a) implies (b). Uniqueness of a felicity function up to a positive affine transformation implies that there is no loss of generality from assuming $u(\theta) = 0$.

Lemma 4: *Axioms* (A1), (A2), *and* (A3) *and the existence of a felicity function* u *with* $u(\theta) = 0$ imply:

(a) $u(A) = -u(-A)$, $A \in S$, *and* $v(x) = -v(-x)$, $x \in \mathbb{R}^M$, *with* $v(\theta) = 0$.

(b) $u(A+B) = u(A) + u(B)$, $A, B \in S$, *and* $v(a+b) = v(a) + v(b)$, $a, b \in \mathbb{R}^M$.

Proof: (a): For $A \in S$,

$$A \approx A \;\Rightarrow\; A + (-A) \approx \theta \;\Rightarrow\; (1/2)A + (1/2)(-A) \approx \theta \text{ (Lemma 2(c))}$$

$$\Rightarrow\; 0 = u[(1/2)A + (1/2)(-A)] = (1/2[u(A) + u(-A)] \;\; ((9))$$

$$\Rightarrow\; u(A) = -u(-A).$$

Hence, $v(x) = -v(-x)$ for all $x \in \mathbb{R}^M$ and $v(\theta) = 0$.

(b): From (10),

$$(1/2)A + (1/2)B \;\approx\; (1/2)(A+B) + (1/2)\theta$$

$$\Rightarrow (1/2)u(A) + (1/2)u(B) = u[(1/2)A + (1/2)B] = u[(1/2)(A+B) + (1/2)\theta]$$

$$= \; (1/2)u(A+B) \qquad \square$$

Let $\leq$ partially order S according to first-order stochastic dominance; that is, $A \leq B$ if $P\{A \neq B\} > 0$ and $P\{A \leq z\} \geq P\{B \leq z\}$ for all $z \in \mathbb{R}^M$. Hence, for $x = (x_i) \in \mathbb{R}^M$ and $y = (y_i) \in \mathbb{R}^M$, $x \leq y$ if $x \neq y$ and $x_i \leq y_i$ for $i = 1, ..., M$. Let $S^* \subset S$ be the set of absolutely bounded random vectors: $S^* = \{A \in S \colon P\{|A_m| < K_A, m = 1, ..., M\} = 1 \text{ for some } K_A < \infty\}$.

Lemma 5: *Axioms* (A1), (A2), *and* (A3) *and the existence of an intraperiod utility function* u *with* $u(\theta) = 0$ imply:

(a) $v(x) = \mathfrak{v} \cdot x$, $x \in \mathbb{R}^M$.

(b) $u(A) = \mathfrak{v} \cdot E(A)$ *for all* $A \in S^*$.

Proof: (a): (i) $x \in I^M$: From Lemma 4(b), $v(x) = v(\sum_{m=1}^M x_m e_m) = \sum_{m=1}^M v(x_m e_m) = \mathfrak{v} \cdot x$.

(ii) $x \in \mathcal{T}^M$: For each m, let $x_m = q_m / s_m$ with $q_m, s_m \in I$. So Lemma 4(b) implies $v(x_m e_m) = v(q_m e_m / s_m) = q_m v(e_m / s_m)$. Also, $v(e_m) = v(s_m e_m / s_m)$=$s_m v(e_m / s_m)$. Therefore, $v(e_m / s_m) = v(e_m)/s_m$; so $v(x_m e_m) = q_m v(e_m)/s_m = \upsilon_m x_m$ and $v(x) = \upsilon \cdot x$.

(iii) $y \le x \;\; \Rightarrow \;\; v(y) \le v(x)$: Let $z = x - y$ with $z_m \ge 0$ for each m. Then Lemma 2(a) and $\mathfrak{v}_m > 0$ imply $v(z_m e_m) \ge 0$. So $v(x-y) = v(z) \ge 0$. Therefore, (A2) implies $v(y) \le v(x)$.

(iv) $x \in (0,\infty)^M$: Let x^{i-},$x^{i+} \in \mathcal{T}^M$ with $x^{i-} \le x \le x^{i+}$ for each i=1,2,... with both sequences converging to x. From (b) and (c), $\mathfrak{v} \cdot x^{i-} = v(x^{i-}) \le v(x)v(x^{i+}) = \mathfrak{v} \cdot x^{i+}$. Letting i $\to \infty$ implies $v(x) = \mathfrak{v} \cdot x$.

(v) $x \in \mathbb{R}^M$: If $x_m < 0$, (d) and Lemma 4(a) imply $v(x_m e_m) = -v(-x_m e_m)$ $= -(-x_m)v(e_m) = \mathfrak{v}_m x_m$. If $x_m = 0$ then $v(x_m e_m) = v(\theta) = 0 = \mathfrak{v}_m x_m$.

(b): Part (a) and (11) imply $u(A) = \mathfrak{v} \cdot E(A)$ for all $A \in S'$. For any $A \in S^*$, there is a sequence (A_i,A^i) with A_i, $A^i \in S'$, $A_i \le A \le A^i$ so $u(A_i) \le u(A) \le u(A^i)$ and $E(A_i) \le E(A) \le E(A^i)$, $i \in I$, with $E(A^i) \to E(A)$ and $E(A_i) \to E(A)$. For each i, $\mathfrak{v} \cdot E(A_i) \le U(A) \le \mathfrak{v} \cdot E(A_i)$. Convergence of the expectations to $E(A)$ implies $U(A) = \mathfrak{v} \cdot E(A)$. $\square$

Lemma 6: *Axioms* (A1), (A2), (A2^c), *and* (A3) *imply the existence of a felicity function.*

Proof. From Herstein and Milnor (1953), (9) and (10) are implied by (13), (14), and (15) (for all $A, B, C \in S$) which follow:

(13) $\quad A \approx B$ *implies* $A/2 + C/2 \approx B/2 + C/2$

(14) $\quad \gtrdot$ *weakly orders* S;

(15) $\quad \{\alpha \in [0,1]$: $\alpha A + (1-\alpha)B \gtrdot (\lessdot) C\}$ *is a closed set*

In order to prove (15), let $X = (A - B, \mathbf{0})$ and $Y = (C - B, \mathbf{0})$. Then $\{\alpha \in [0,1]$: $\alpha A + (1-\alpha)B \gtrdot C\} = \{\alpha \in [0,1]$: $\alpha X - Y \succeq \mathbf{0}\}$ = $\{\alpha \in [0,1]$: $\alpha X \succeq Y\}$ which is closed due to (A3). Similarly, $\{\alpha \in [0,1]$: $\alpha A + (1-\alpha)B \lessdot C\}$ is closed.

Axiom (A1) implies (14).

In order to prove (13), use (A2), (A2^c), and parts (d) and (e) of Lemma to obtain

$$
\begin{aligned}
A \gtrdot B \;\; &\Leftrightarrow \;\; (A - B, \mathbf{0}) \succeq \mathbf{0} \;\; \Rightarrow \;\; (1/2)(A-B, \mathbf{0}) \succeq \mathbf{0} \\
&\Rightarrow \;\; (1/2)(A-B, \mathbf{0}) + (1/2)(C, \mathbf{0}) \; \succeq \; (1/2)(C, \mathbf{0}) \\
&\Rightarrow \;\; ((1/2)(A+C), \mathbf{0}) \; \succeq \; ((1/2)(B+C), \mathbf{0}) \\
&\Rightarrow \;\; (1/2)(A+C) \; \gtrdot \; (1/2)(B+C)
\end{aligned}
$$

Similarly, $A \lessdot B \;\; \Leftrightarrow \;\; B \gtrdot A \;\; \Rightarrow \;\; (1/2)(B+C) \gtrdot (1/2)(A+C)$. $\square$

This completes the proof of Theorem 2. □

Extensions of Theorem 2

It is an open question whether there are conditions weaker than (A1) through (A5) and ($A2^c$) that would axiomatize (4) and the optimization of the “expected utility of the present value" of a time stream of rewards. See surveys by White (1988) and Whittle (1990).

The remainder of this section concerns several routes that can be followed to extend Lemma 5(b) beyond S^* (hence Theorem 2 beyond V^*). Let S' be the subset of S whose elements take only finitely many values (with probability one). Recall that θ and e_m denote the zero vector and the m*-th* unit vector in $\mathbb{R}^M$. Suppose that $u(\,\cdot\,)$ is a felicity function with $u(\theta) = 0$, for $x \in \mathbb{R}^M$ let D_x be a random variable with $P\{D_x = x\} = 1$, $v(x) = u(D_x)$, $\mathfrak{v}_m = v(e_m)$, $m = 1, \ldots, M$, and $\mathfrak{v} = (v_m)$. Then (A4) and $v(\theta) = 0$ imply $\mathfrak{v}_m > 0$ for each m. The expected utility formula, a consequence of (10), is

(11) $$u(A) = \sum_i P\{A = x_i\} v(x_i) = E[v(A)], \quad A \in S'$$

First, at the expense of additional assumptions and complexity, versions of (11) are valid for larger collections than S' [Fishburn (1982)]. These versions would expand the set of random vectors that can be approximated in the proof of Theorem 2. Second, additional assumptions would induce continuity of $u(\,\cdot\,)$ in the topology of weak convergence, and that would permit approximating $A \in S - S^*$ with a sequence in S^*. For a third route, let $S^o \subset S$ be the set of random vectors whose components have finite expectations. The next result assumes essentially that $\gg$ is continuous at θ.

Lemma 7: *Suppose that a felicity function* u *satisfies* $u(\theta) = 0$ *and*

(16) *For all sequences* $\langle C_i \rangle$ *in* S^o *with* $P\{C_i = \theta\} \to 1$ *as* $i \to \infty$ *and for all* $D \in S^o$ *with* $D \gg \theta$, *there is a* j *such that* $i \geq j$ *implies* $-D \ll C_i$ *and* $C_i \ll D$

Then (A1), (A2), *and* (A3) *imply* $u(A) = \mathfrak{v} \cdot E(A)$ *for all* $A \in S^o$.

Proof: (a) From Lemma 5(b), for all $\gamma > 0$ there exists $D \in S^*$ with $U(D) = \gamma$. Let C_i be a sequence in S^o with $P\{C_i = \theta\} \to 1$ as $i \to \infty$. So (16) implies that there exists j such that $|U(C_i)| < \gamma$ for all $i \geq j$. Therefore, $U(C_i) \to 0$ as $i \to \infty$.

(b) For $A \in S^o$, let $A^+ \in S^o$ have m^{th} component max $\{A_m, 0\}$, $m = 1, \ldots, M$. Then $A = A^+ - (-A)^+$ and $U(A) = U(A^+) + U[-(-A)^+]$ with each component of A^+ $[-(A)^+]$

bounded below (above) by 0. So it suffices to prove the assertion for those elements in S^{o} whose components are all either bounded above by zero or bounded below by zero.

(c) Let $A \in S^{o}$ with $A \geq \theta$ and $z > 0$. Define $D_z = A$ and $D^z = \theta$ if $A_m \leq z$ for $m = 1, ..., M$; otherwise, $D_z = \theta$ and $D^z = A$. Hence, $A = D_z + D^z$ for each z and $P\{D^z = \theta\} \to 1$ as $z \to \infty$. So part (a) and Lemmas 4(b) and 5(b) imply $U(A) = U(D_z) + U(D^z) = \upsilon \cdot E(D_z) + U(D^z)$ $\to \upsilon \cdot E(A)$ as $z \to \infty$. A similar argument suffices for $A \leq \theta$. □

Corollary 1 and Lemma 7 extend Theorem 2 to stochastic processes with countable time indices.

COROLLARY 3: *Assumptions (8) and (16) imply that Theorem 2 is valid with (12) replaced by (17) if both expectations exist with finite components:*

(17) $\quad \mathrm{X} \succeq \mathrm{Y} \quad \Leftrightarrow \quad \upsilon \cdot E(\sum_{t=1}^{\infty} \beta_t X_t) \ \geq \ \upsilon \cdot E(\sum_{t=1}^{\infty} \beta_t Y_t) \quad X \in V,\ Y \in V$

Corollary 2 augments (A1) through (A4) with (A5) to obtain $\beta_{mt} = (\beta_m)^{t-1}$; including (A6) yields $\beta_m < 1$. So Corollaries 2 and 3 give conditions which imply that preferences are consistent with the maximization of $\mathfrak{v} \cdot E(\sum_{t=1}^{\infty} \beta^{t-1} X_t)$. The algorithm in Feinberg and Schwartz (1994) computes an optimal policy for this criterion in a Markov decision process with $M > 1$ and finitely many states and actions. Furukawa (1980) and Henig (1983) study Pareto optimization of the vector $E(\sum_{t=1}^{\infty} \beta^{t-1} X_t)$ when $\beta_1 = \ \cdot \ \cdot \ \cdot \ = \beta_M$.

4. Scalar Discount Factor

This section combines Corollary 2, Theorem 2, and Koopmans (1972) to specify conditions which imply $\beta_1 = \ \cdot \ \cdot \ \cdot \ = \beta_M$. Let $\mathbb{D} \subset V$ be the sequences that are constant with probability one, and for $x \in \mathbb{D}$ let $\| x \| \ = \ sup_t \ max_m \ |x_{mt}|$. Let $\mathbb{D}_T \subseteq \mathbb{D}$ contain the sequences x for which $x_t = \theta$ if $t > T$. Let $x^{(n)} \in \mathbb{D}$ and $y^{(n)} \in \mathbb{D}$ with $x^{(n)} \succeq y^{(n)}$ for each $n \in \mathcal{I}$ and let $x \in \mathbb{D}$ and $y \in \mathbb{D}$ with $\| x - x^{(n)} \| \ \to 0$ and $\| y - y^{(n)} \| \ \to 0$ as $n \to \infty$. Say that $\succeq$ *is continuous on* $\mathbb{D}$ *with respect to* $\| \cdot \|$ if $x \succeq y$. There are two additional assumptions:

(A3*) $\quad \succeq$ *is continuous on* $\mathbb{D}$ *with respect to* $\| \cdot \|$

(A5^c) $\quad (\theta, X_1, X_2, ...) \sim \mathbf{0}$ *implies* $(X_1, X_2, ...) \sim \mathbf{0}$

Axiom (A5^c) is the converse of (A5).

Proposition 4: *Axioms* (A1) *through* (A5), (A3*), *and* (A5ᶜ) *and the existence of an intraperiod utility function imply* $\beta_1 = \cdot \cdot \cdot = \beta_M$.

Proof: The assumptions satisfy Postulates 1 through 4 in Koopmans (1972). So in $\mathbb{D}_\mathrm{T}$

(18) $$x \sim y \quad \Leftrightarrow \quad \sum_{t=1}^{T} \alpha^{t-1} u(x_t) = \sum_{t=1}^{T} \alpha^{t-1} u(y_t)$$

where $\alpha \in (0,1)$ is unique and $u : \mathbb{R}^M \to \mathbb{R}$ generates a class that is unique up to a positive affine transformation. Let $u(\theta) = 0$ and $x_t = y_t = \theta$ for all $t > 1$; from Theorem 2, $u(w) = \mathfrak{v} \cdot w$ generates a class that is unique up to a positive linear transformation.

Let $X = e_{mt}$ and $Y = \beta_{mt} e_{m1}$; so $X \sim Y$ from Lemma 3 and (A2). From (18), $e_{mt} \sim \beta_{mt} e_{m1}$ if and only if $\alpha^{t-1} \mathfrak{v}_m = \beta_{mt} \upsilon_m$. Thus, $\beta_{mt} = \alpha^{t-1}$. $\square$

5. Continuity Axioms and Impatience

Reasonable alternative versions of the continuity axiom

(A3) $$\{\alpha\text{:}\ \alpha X - Y \succeq (\preceq) \mathbf{0}\} \text{ is closed}$$

include these statements:

(A3′) $$X \succ (\prec) \mathbf{0} \Rightarrow \exists\, \alpha^* > 0 \text{ such that } \alpha X \prec (\succ) X \ \forall\, \alpha \in (0,\alpha^*)$$

(A3″) $$X \succeq (\preceq) \mathbf{0} \Rightarrow bX \succeq (\preceq) \mathbf{0} \ \forall\, b \geq 0$$

Let (A3ᵒ) denote the restriction of (A3) in which $Y = \mathbf{0}$.

Proposition 5. *Assume* (A1), (A2), *and* (A2ᶜ). *Then* (A3ᵒ), (A3′), *and* (A3″) *are equivalent*.

Proof. Let (3ᵒ), (3′), and (3″) denote the respective unions of {(A1),(A2),(A2ᶜ)} with {(A3ᵒ)}, {(A3′)}, and {(A3″)}.

(3ᵒ) $\Rightarrow$ (3″): Lemma 2(a) and its proof.

(3′) $\Rightarrow$ (3″): For $b > 0$, let $\langle b_i \rangle$ be a sequence in $\mathcal{T}$ that converges to b from below. To initiate a contrapositive proof, if $bX \prec \mathbf{0}$ then $-bX \succ \mathbf{0}$ (due to (A2ᶜ)). So $-bX \succ -(b - b_i)X$ for b_i sufficiently near b (due to (A3′)). Therefore, there exists i^* such that $i \geq i^*$ implies $\mathbf{0} \prec -bX + (b - b_i)X = -b_i X$. So $b_i X \prec \mathbf{0}$ contrary to $b_i X \succeq \mathbf{0}$ for all i because $X \succeq \mathbf{0}$ and $b_i > 0$ is rational.

(3″) $\Rightarrow$ (3′): To initiate a contrapositive proof,

$$X \preceq \alpha X \;\Leftrightarrow\; \alpha X - X \succeq \mathbf{0} \;\Leftrightarrow\; ((1-\alpha)(-X) \succeq \mathbf{0}$$

Let $\alpha^* < 1$ and $0 < \alpha < \alpha^*$; so $1-\alpha > 0$. Now (3″) implies $b(1-\alpha)(-X) \succeq \mathbf{0}$ for all $b \geq 0$. If $b = (1-\alpha)^{-1}$ then

$$\mathbf{0} \preceq b(1-\alpha)(-X) \;\Leftrightarrow\; \mathbf{0} \succeq -X \;\Leftrightarrow\; X \preceq \mathbf{0}$$

(3″) $\Rightarrow$ (3°): Let $\alpha_i \to \alpha$ with $\alpha_i X \succeq \mathbf{0}$ for all i. If $\alpha = 0$ then $\alpha X \succeq \mathbf{0}$ ((A1)). If $\alpha > (<)\, 0$ then without loss of generality $\alpha_i > (<)\, 0$ for all i. So (3″) implies $b\alpha_1 X \succeq (\preceq)\, \mathbf{0}$ for all $b \geq 0$. Let $b = \alpha/\alpha_1$ to obtain $\alpha X \succeq \mathbf{0}$. □

Impatience

For any $X \in V$ and M-dimensional random vector Z, let (Z,X) denote the sequence in V with first component Z and t^{th} component X_{t-1} for all $t \geq 2$. Let ${}_2X$ be the sequence in V with t^{th} component X_{t+1} for all t $\in \mathcal{I}$. Versions of impatience that may appear to differ from

(A6) *Sooner is better:* $\qquad X + \lambda(e_{mt} - e_{m,t+1}) \succ X$

include the following statements:

(19) $\quad X \succ Y \;\Leftrightarrow\; X - Y \succ (Z,X) - (Z,Y) \;\forall\; Z$

(20) $\quad X \succ (Z,{}_2X) \;\Leftrightarrow\; (X_1,Z,{}_2X) \succ (Z,X)$

To interpret (19), which is implicit in Koopmans (1960), suppose that X is preferred to Y. Then X is preferred to Y with an "intensity" which is greater than the intensity with which a delay of X is preferred to the same delay of Y. For an interpretation of (20) that corresponds to the formal definition of impatience in Koopmans (1960), suppose that X is preferred to the sequence in which X_1 is replaced by Z. Then $(X_1,Z,X_2,X_3,...)$ is preferred to $(Z,X_1,X_2,X_3,...)$.

Proposition 6. *Axioms* (A1) *through* (A6) *and* (A2^{c}) *imply* (19) *and* (20).

Proof. To prove the claim regarding (19), let $\beta = (\beta_m) \in (0,1)^M$ have components whose existence is asserted in Corollary 2, let β^t be the vector with m^{th} component $(\beta_m)^t$, and let $\mathbf{1}$ be the M-vector of ones. Then

$$X - Y \succ (Z,X) - (Z,Y) \;\Leftrightarrow\; [X - (Z,X)] - [Y - (Z,Y)] \succ \mathbf{0}$$

$$\Leftrightarrow\; \theta \ll \textstyle\sum_{t=1}^{T}(\beta^{t-1} - \beta^t)(X_t - Y_t) = (\mathbf{1} - \beta)\sum_{t=1}^{T}\beta^{t-1}(X_t - Y_t)$$

$\Leftrightarrow\ \ \theta \ \ll\ \sum_{t=1}^{T}\beta^{t-1}(X_t - Y_t)\ \ \Leftrightarrow\ \ X \succ Y.$

For the claim concerning (20),

$$X \succ (Z,{}_2X)\ \ \Leftrightarrow\ \ X_1 \gg Z\ \ \Leftrightarrow\ \ (\mathbf{1}-\beta)X_1 \gg Z$$

$$\Leftrightarrow\ \ X_1 + \beta Z + \textstyle\sum_{t=2}^{T}\beta^t X_t\ \gg\ Z + \beta Z + \sum_{t=2}^{T}\beta^t X_t$$

$$\Leftrightarrow\ \ (X_1, Z, {}_2X) \succ (Z, X) \qquad\qquad \square$$

6. Attribute Decomposition in Random Vectors

Several of the previous results have interpretations for preference orderings of M-dimensional random vectors. This section is confined to the special case of Theorem 2 when $\succeq$ is a binary relation on sequences $(A,\mathbf{0})$ for all $A \in S$. If axiom (A2^c) were invoked, the restriction would be equivalent to fixing $X \in V$ and considering the sequences $(A, X) \in V$ for all $A \in S$. This yields the following version of Theorem 2 in which $\gtrsim$ is regarded as a primitive binary relation on S (rather than a binary relation that is induced by $\succeq$ on V). Recall that S^* denotes the set of absolutely bounded random vectors and $S^o \subset S$ consists of vectors whose components have finite expectations.

Corollary 4: *If* $(\gtrsim, S)$ *is weakly ordered,* $e_m \gg \theta$ $(m = 1, ..., M)$, and

(21) $\qquad A - B \gtrsim (\lesssim)\,\theta$ *implies* $A \gtrsim (\lesssim)\,B \qquad (A \in S, B \in S)$

$\qquad\qquad \{\alpha \in [0,1] : \alpha A \gtrsim (\lesssim)\,B\}$ *is closed* $\qquad (A \in S, B \in S)$

then the following properties are equivalent:

(a) $(\gtrsim, S)$ *satisfies* (A2^c);

(b) *There exists a felicity function;*

(c) *There exists* $\mathfrak{v} \in \mathbb{R}^M$ *with positive components such that*

$$A \gtrsim B \text{ if and only if } \mathfrak{v}\cdot E(A) \geq \mathfrak{v}\cdot E(B) \qquad (A \in S^*, B \in S^*)$$

The assertion above is valid with S^* *replaced by* S^o *if also*

for all sequences $\langle C_i\rangle$ *in* S *with* $P\{C_i = \theta\} \to 1$ as $i \to \infty$ *and for all* $D \in S$ *with* $D \gg \theta$, *there exists* j *such that* $i \geq j$ *implies* $-D \ll C_i$ *and* $C_i \ll D$.

Assumption (21) corresponds to the decomposition axiom (A2) and is a strong form of *utility independence* which is investigated in references cited in §1 and Blackorby, Primont, and Russell (1978).

7. Conclusion

Assumptions that imply discounting are *not* orthogonal to assumptions that yield an intraperiod utility function. Weak axioms that are consistent with discounting in a deterministic environment carry the seeds of risk neutrality in a stochastic environment. The following assumptions imply that preferences among stochastic processes correspond to comparisons of random present value vectors: weak ordering, decomposition, continuity, and more is better. If there exists an intraperiod utility function, the same assumptions imply risk neutrality. An axiomatic justification of expected present value of intraperiod utility (with nonlinear intraperiod utility) seems to require a justification of discounting in a stochastic environment (*i.e.*, (7)) with an assumption that is weaker than decomposition (A2).

Acknowledgements

The author is grateful to Professor James C. Alexander for comments and the argument on which the proof of Lemma 3(a) is based, to Professor Eugene Feinberg and Dr. Charu Sinha for comments, and to an anonymous commentator on an earlier version of the paper who noted that confining Y to $\mathbf{0}$ in (A3) permitted a counterexample to a lemma.

References

Altug, S. and Labadie, P. *Dynamic Choice and Asset Markets,* Academic Press, New York, 1994.

Arrow, Kenneth J., Karlin, Samuel and Scarf, Herbert. *Studies in the Mathematical Theory of Inventory and Production*, Stanford University Press, Stanford, CA, 1962.

Balcer, Yves and Lippman, Steven A. "Technology Expectations and Adoption of Improved Technology." *J. Economic Theory,* 1984, 34, pp. 292-318.

Blackorby, C., Primont, D., and Russell, R. R. *Duality, Separability, and Functional Structure: Theory and Economic Applications*. North-Holland, New York, 1978.

____________________ "Separability: A Survey." *Handbook of Utility Theory, Volume 1: Principles*. Kluwer Academic Publishers, Dordrecht, 1998.

Blair, Charles E. "Axioms and Examples Related to Ordinal Dynamic Programming." *Mathematics of Operations Research*, 1984, 8, pp. 345-347.

Cochrane, John H. (2000). *Asset Pricing*, Princeton University Press, Princeton, NJ.

Denardo, Eric V. "Contraction Mappings in the Theory Underlying Dynamic Programming." *SIAM Review*, 1967, 9, pp. 165-77.

Duffie, Darrell. *Dynamic Asset Pricing Theory,* 3rd ed., Princeton, NJ: Princeton University Press, 2001.

Dyer, James S. and Sarin, Rakesh "Measureable Multiattribute Value Functions." *Operations Research*, 1979, 27, pp. 810-22.

Epstein, Larry G. "Stationary Cardinal Utility and Optimal Growth under Uncertainty." *J. Economic Theory*, 1983, 31, pp. 133-151.

Farquhar, Peter H. "A Survey of Multiattribute Utility Theory and Applications." pp. 59-89 in *Multiple Criteria Decision Making,* Starr, M. K. and Zeleny, M. eds., Amsterdam: North-Holland, 1977.

Feinberg, Eugene and Schwartz, Adam. "Markov Decision Models with Weighted Discounted Criteria," *Mathematics of Operations Research*, 1994, 19, pp. 152-168.

Fishburn, Peter C. *Utility Theory for Decision Making*. Huntington, NY: Robert E. Krieger Publishing Co., 1970.

________________. "A Survey of Multiattribute/Multicriterion Evaluation Theories," pp. 181-224 in *Multiple Criteria Problem Solving*, Zionts, S. ed., Springer-Verlag, Berlin, 1978.

________________. *The Foundations of Expected Utility*. Dordrecht, Holland: D. Reidel, 1982.

________________ and Rubinstein, Ariel. "Time Preference." *International Economic Review*, 1982, 23, 2, pp. 677-94.

Frederick, Shane, Loewenstein, George, and O'Donoghue, Ted. "Time Discounting and Time Preference: A Critical Review." *J. of Economic Literature*, 2003, to appear.

_________________ *The Foundations of Expected Utility*, Dordrecht, Holland: D. Reidel, 1982.

Furukawa, N. "Characterization of Optimal Policies in Vector-Valued Markovian Decision Processes," *Mathematics of Operations Research*, 1980, 5, pp. 271-79.

Henig, Mordecai I. "Vector-Valued Dynamic Programming." *SIAM J. Control & Optimization*, 1983, 21, pp. 490-99.

Herstein, Israel N. and Milnor, John W. "An Axiomatic Approach to Measureable Utility." *Econometrica*, 1953, 21, pp. 291-297.

Ho, T. S. Y. and S-B. Lee, S-B. "Term Structure Movements and Pricing Interest Rate Contingent Claims," *J. Finance*, 1986, XLI, pp. 1011-29.

Johnson, T. H. and Donaldson, John B. "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans," *Econometrica*, 1985, 53, pp. 1451-58.

Jorgenson, Dale W., McCall, John J. and Radner, Roy. *Optimal Replacement Policy*, Amsterdam: North-Holland, 1967.

Keeney, Ralph L. and Raiffa, Howard. *Decisions with Multiple Objectives*, New York: John Wiley, 1976.

Koopmans, Tjallings C. "Stationary Ordinal Utility and Impatience," *Econometrica*, 1960, 28, pp. 287-309.

Koopmans, Tjallings C., Diamond, Peter. A., and Williamson, R. E. “Stationary Utility and Time Perspective," *Econometrica*, 1964, 32, pp. 82-100.

Koopmans, Tjallings C. “Representation of Preference Orderings Over Time." Chap. 4 in *Decisions and Organizations,* McGuire, C. B. and Radner, R. eds., , Amsterdam: North-Holland, 1972.

Kreps, David M. and Porteus, Evan L. “Temporal von Neumann-Morgenstern and Induced Preferences," *J. Economic Theory*, 1979, 20, pp. 81-109.

Lancaster, Kelvin J. “An Axiomatic Theory of Consumer Time Preference." *International Economic Review*, 1963, 4, pp. 221-31.

Machina, Mark J. “Dynamic Consistency and Non-Expected Utility Models of Choice." *J. Economic Literature*, 1989, 27, pp. 1622-68.

McNamara, John M. "The Policy which Maximises Long-Term Survival of an Animal Faced with the Risks of Starvation and Predation." *Advances in Applied Probability*, 1990, 22, pp. 295-308.

Mendelssohn, Roy. “Using Markov Decision Models for Purposes other than Simple Optimization: Analyzing the Consequences of Policy Alternatives on the Management of Salmon Runs." *Fisheries Bulletin*, 1979, 78, pp. 35-50.

Meyer, Richard F. "Preferences Over Time." Chapter 9 in Keeney, Ralph L. and Raiffa, Howard, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, New York: Wiley, 1976.

Miyamoto, J. M. and Wakker, P. "Multiattribute Utility Theory without Expected Utility Foundations." *Operations Research*, 1996, 44, pp. 313-326.

Monahan, George E. P. “Optimal Advertising with Stochastic Demand." *Management Science*, 1983, 29, pp. 106-17.

Moriarity, S. and Allen, C. P. *Cost Accounting*, New York: Harper and Row, 1984.

Murphy, F. H., Toman, M. A.and Weiss, H. J. “A Dynamic Nash Game Model of Oil Market Disruptions and Strategic Stockpiling." *Operations Research*, 1989, 37, pp. 958-71.

Rosenthal, Robert. Personal communication, Stony Brook, NY, 1987.

Samuelson, Paul A. "A Note on Measurement of Utility." *Review of Economic Studies*, 1937, 4, pp. 155-161.

Sethi, Suresh P. "Optimal Consumption-Investment Decisions allowing for Bankruptcy: A Survey," pp. 397-426 in *Worldwide Asset and Liability Modeling*, Ziemba, W. T. and Mulvey, J.M. eds., Cambridge, UK: Cambridge University Press, 1998.

Shapley, Lloyd S. "Stochastic Games." *Proc. National Academy of Sciences, USA* , 1953, 39, pp. 1095-1100.

Sharpe, William F. *Investments*, 3rd ed.,Englewood Cliffs, NJ: Prentice-Hall, 1985.

Shubik, Martin. "Game theory, behavior and the paradox of the prisoner's dilemma: Three solutions." *J. Conflict Resolution*, 1970, 14, pp. 181-93.

Shubik, Martin and Sobel, Matthew J. "On Matching Book: A Problem in Banking and Corporate Finance." *Management Science*, 1992, 38, pp. 827-39.

Shubik, Martin and Thompson, Gerald L. "Games of Economic Survival." *Naval Research Logistics Quart., 1959,* 6, pp. 111-23.

Sobel, Matthew J. “Ordinal Dynamic Programming," *Management Science*, 1975, 21, pp. 967-75.

Sobel, Matthew J. “Ordinal Sequential Games." *Économies et Sociétés*, 1980, XIV, pp. 1571-82.

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

Wakker, P. P. "Additive Representations on Rank-Ordered Sets II. The Topological Approach." *J. Mathematical Economics*, 1993, 22, pp. 1-26.

White, Douglas J. “Mean, Variance, and Probabilistic Criteria in Finite Markov Decision Processes: A Review." *J. Optimization Theory & Applications*, 1988, 56, pp. 1-29.

Whittle, Peter C. *Risk-Sensitive Optimal Control*, Wiley, New York, 1990.

Williams, Albert C. and Nassar, J. I. “Financial Measurement of Capital Investments." *Management Science*, 1966, 12, pp. 851-64.

Yeh, William W-G. “Reservoir Management and Operations Models: A State-of-the-Art Review," *Water Resources Research*, 1985, 21, pp. 1797-1818.