STOCHASTIC DIFFERENTIAL UTILITY

BY DARRELL DUFFIE AND LARRY G. EPSTEIN
APPENDIX C WITH COSTIS SKIADAS[1]

This paper presents a stochastic differential formulation of recursive utility. Sufficient conditions are given for existence, uniqueness, time consistency, monotonicity, continuity, risk aversion, concavity, and other properties. In the setting of Brownian information, recursive and intertemporal expected utility functions are observationally distinguishable. However, one cannot distinguish between a number of non-expected-utility theories of one-shot choice under uncertainty after they are suitably integrated into an intertemporal framework. In a "smooth" Markov setting, the stochastic differential utility model produces a generalization of the Hamilton-Jacobi-Bellman characterization of optimality. A companion paper explores the implications for asset prices.

KEYWORDS: Choice under uncertainty, stochastic control, recursive utility, Bellman equation.

1. INTRODUCTION

THIS PAPER DESCRIBES an extension of the standard additive utility specification in which the utility at time t for a consumption process c is defined by

$$(1) \qquad V_t = E_t \left[\int_{s \geq t} e^{-\beta(s-t)} u(c_s) \, ds \right], \quad t \geq 0,$$

where E_t denotes expectation given information available at time t. The more general utility functions, called *recursive*, exhibit intertemporal consistency and admit Bellman's characterization of optimality. Much of the tractability of (1) is therefore preserved. For example, a companion paper shows that the recursive utility specification implies a model of asset pricing that incorporates Breeden's (1979) consumption-based capital asset pricing model (CCAPM) as a special case.

A special example of the utility model presented in this paper is obtained by making the assumption that information at time t is that generated by Brownian motion. In this case, after a simplifying change of variables, the utility process V for a consumption process c is determined by an "aggregator" function f and is

shown to be the unique solution V to

$$(1') \qquad V_t = E_t \left[\int_{s \geq t} f(c_s, V_s) \, ds \right], \quad t \geq 0,$$

which reduces to the additive model (1) in the case $f(c, v) = u(c) - \beta v$.

A discrete-time analogue of the utility model studied in this paper appears in Epstein and Zin (1989), which builds upon Kreps and Porteus (1978). The former paper pointed out the following advantage of recursive utility, which is present also in continuous time: It is well-known that in the standard specification (1), risk aversion and intertemporal substitutability are both reflected in the curvature of $u(\cdot)$. In contrast, these two aspects of preferences can be disentangled within the more general class of recursive utility functions (in the sense of Proposition 6 of this paper). The possibility of such a separation is useful in clarifying the determinants of asset prices (Epstein (1988) gives a discrete-time demonstration) and presumably for a number of other issues in capital theory and finance. In a continuous-time setting, the implications of recursive utility for asset pricing are simpler than is the case in discrete-time, as is amply demonstrated in Duffie and Epstein (1991).

A continuous-time setting also allows one to more clearly connect the manner in which uncertainty is resolved over time with the ability to deduce axiomatic differences in preferences by observing actions. In order to be more precise, it is necessary to elaborate upon the structure of recursive utility functions. One component of the specification of a recursive utility function is a certainty equivalent functional that is defined on (a subset of) probability distributions on the real line. This functional represents preferences over a restricted set of consumption processes (see Section 4). The certainty equivalent could be specified according to any of a number of generalizations of expected utility theory that have been proposed recently in response to the experimental evidence against the expected utility model. The new theories all deal with static or one-shot choice environments and specify utility functionals for probability distributions on the real line. They may be integrated into our temporal framework via the specification of the certainty equivalent. One might wonder, however, whether such alternative specifications for the certainty equivalent have differing implications for behavior in a continuous-time setting. Indeed, we show that in settings of continuous information resolution (a special case being the information generated by Brownian motion), it is impossible to distinguish, by observing actions alone, whether the underlying preference order is determined by an expected utility certainty equivalent or by one conforming to Machina's (1982) notion of smooth local expected utility. In particular, these two preference models cannot be empirically distinguished using security price data within the standard diffusion models of asset prices. Similarly, we show that suitably "smooth" functionals of the Chew (1983, 1989) and Dekel (1986) axiomatic class are empirically indistinguishable, given continuous information resolution, from expected-utility-based certainty equivalents. It is essential to realize that the latter specification does not imply an intertemporal expected

utility preference ordering. In particular, neither (1) nor (3) below is implied. Rather, one obtains a continuous-time version of a preference structure studied by Kreps and Porteus (1978) in discrete-time, where an expected utility ordering applies to a restricted domain of probability distributions.

As a further incentive for exploring a continuous-time model for recursive preferences, we point to the literature on continuous-time general equilibrium in security markets. In Harrison and Kreps (1979), for example, an equivalence is shown between the existence of an optimal strategy and the existence of an equivalent martingale measure for security prices (under technical restrictions). In Duffie and Huang (1985), for another example, a continuous-time general equilibrium is obtained from an abstract static infinite-dimensional Arrow-Debreu style equilibrium, whose consumption allocation is dynamically implemented with appropriate price processes and security strategies. The apparent generality with regard to preference restrictions allowed in these two examples, as in much of this literature, is somewhat illusory. The optimality criterion for agents' choices is applied at the initial date of the economy, but not at intermediate dates, given the newly available information. With the additively-separable criterion (1), of course, any strategy that is initially optimal is always optimal. This follows from the fact that additively-separable utility functions induce *time-consistent* intertemporal preferences. Time-consistent preferences are defined by the property that, for any two consumption processes c and $\tilde{c}$ and any stopping time τ, if c and $\tilde{c}$ are identical up to time τ, and if the continuation of c is preferred to the continuation of $\tilde{c}$ at time τ, then c is preferred to $\tilde{c}$ at time zero. The recursive preferences examined in this paper are shown to be time-consistent, and thus extend the range of application of continuous-time general equilibrium models. See Johnson and Donaldson (1985) for the link between consistency and the lack of market re-opening typically assumed in the Arrow-Debreu model of contingent commodity markets. They also clarify that consistency is an issue only if one requires, as we do implicitly, that preferences do not depend on unrealized alternatives.

In order to elucidate and further motivate the nature of recursive utility, we conclude this introduction by considering choice among three hypothetical consumption programs, which are informally defined as follows. Consumption during the interval $[0, 1)$ is fixed at the same level for all three programs. In the first program, c^A, a fair coin is tossed at $t = 1$. If the outcome is a head, then consumption is constant at the level l for the entire remaining horizon $[1, T]$. Otherwise, it is constant at $L > l$. For consumption program, c^B, $T - 1$ independent fair coins are flipped, one for each integer time $t \in [1, T)$, and all coin tosses are revealed at time 1. If the tth toss yields a head (tail) then consumption over the interval $[t, t + 1)$ is $l(L)$. Finally, c^C differs from c^B only in that the tth coin toss is not revealed until time t. Given the standard utility specification (1), $c^A \sim c^B \sim c^C$, while neither indifference is imposed by recursive utility. Moreover, introspection suggests that neither indifference is compelling. For example, in comparing c^A and c^B, the former involves perfect correlation between consumption at all times; that is, c_t^A is high if and only if

c_τ^A is high for all t and τ. In contrast, consumption levels in nonoverlapping intervals are serially independent in c^B. Diversification motives suggest that c^B may be strictly preferred to c^A, although the psychic cost of serially fluctuating consumption would work in the opposite direction. The essential difference between c^B and c^C is that, in the former, all uncertainty is resolved at $t = 1$, while in c^C uncertainty is resolved gradually on $[1, T]$. Indifference between c^B and c^C is thus not compelling and arguments, based on the psychic costs and benefits of early resolution, can be proposed for either preference direction. (See Kreps and Porteus (1978) and Chew and Epstein (1990) for details in a discrete-time setting. We have nothing to add in this paper with regard to attitudes towards the timing of the resolution of uncertainty.)

We proceed as follows: Section 2 provides a definition of recursive utility in continuous time. A "morally equivalent" but mathematically more workable definition is presented in Section 3, along with a proof of existence. Section 4 contains some examples and observations regarding the potential usefulness for continuous-time modeling of generalizations of expected utility that have been proposed recently (for example, Machina (1982)). Sufficient conditions for several desirable properties of recursive utility are established in Section 5. In Section 6, an extension of the usual Hamilton-Jacobi-Bellman equation is shown to characterize optimality. Section 7 includes some extensions. Readers interested mainly in the asset pricing implications of stochastic differential utility could begin directly with Duffie and Epstein (1991).

2. RECURSIVE UTILITY IN CONTINUOUS TIME

This section defines recursive utility in a continuous-time stochastic setting that will be the basis for the remainder of the paper. Figure 1 relates our work to existing literature on recursive utility. Consider first a discrete-time framework, for which the seminal paper is Koopmans (1960). In a deterministic setting, Koopmans studied utility functions that satisfy the recursive relation (2),

Setting	Discrete-Time		Continuous-Time	
Certainty	$V_t = W(c_t, V_{t+1})$ $U(\{c_1, c_2, \dots\}) = V_0$ Koopmans (1960) Lucas and Stokey (1984)	(2)	$dV_t = -f(c_t, V_t)\,dt$ $U(\{c_t : t \geq 0\}) = V_0$ Epstein (1987)	(5)
Uncertainty state-separable	$U(\{c_t\}) = E[\sum_{t=0}^\infty u(c_t) e^{-\sum_{s=0}^{t-1} \beta(c_s)}]$ Epstein (1983)	(3)	$U(\{c_t\}) = E(\int_0^\infty u(c_t) e^{-\int_0^t \beta[c(s)]\,ds}\,dt)$ Uzawa (1968)	(6)
Uncertainty state-separable or nonseparable	$V_t = W[c_t, m(V_{t+1})]$ $U(\{c_t\}) = V_0$ Epstein and Zin (1989)	(4)	$dV_t = [-f(c_t, V_t) - A(V_t)\sigma_V^2(t)/2]\,dt$ $\quad\quad + \sigma_V(t)\,dB_t$ $U(\{c_t\}) = V_0$ This paper	(7)

FIGURE 1.—Recursive utility.

where c_t denotes consumption in period t and $V_t = U(c_t, c_{t+1}, \ldots)$. The function W is called an *aggregator* by Lucas and Stokey (1984), since it combines c_t and the utility index V_{t+1} of future consumption in order to compute the current utility V_t.

If Koopmans' axioms are extended in a natural fashion to a model with uncertainty, and if expected utility is assumed, then one obtains the functional form (3). (See Epstein (1983).) The von Neumann-Morgenstern index in (3) is closely related to the functional form first proposed by Uzawa (1968) in a continuous-time deterministic setting. A limitation of (3), and the time-additive special case (in which $\beta(\cdot)$ is a constant), is that risk aversion and intertemporal substitution are indistinguishable within this class. (See Epstein and Zin (1989) and Chew and Epstein (1990, Section 4).)

In order to overcome this inflexibility of the (recursive) intertemporal expected utility model, Epstein and Zin (1989) propose the general recursive structure (4), with the following interpretation: From the perspective of time t, the intertemporal utility V_{t+1} for period $t+1$ and beyond is a random variable. Thus the agent first computes the certainty equivalent $m(\sim V_{t+1} | \mathscr{F}_t)$ of the conditional distribution $\sim V_{t+1} | \mathscr{F}_t$ of V_{t+1} given information $\mathscr{F}_t$ at time t, and then combines the latter with c_t via the aggregator W. This functional structure has two components: the aggregator W, which encodes the intertemporal substitutability of consumption; and the certainty equivalent function m, which encodes risk aversion in the sense described in Epstein and Zin (1989) and also below. In a finite-horizon setting, Kreps and Porteus (1978) study (4) under the assumption that m is an expected-utility-based certainty equivalent such as (8) below.

Here, we develop the continuous-time formulation of Epstein and Zin's class of recursive utility. A parallel exercise for the deterministic case was undertaken in Epstein (1987), in which the continuous-time version of Koopmans' class of preferences is formulated. The key observation in the latter paper is that, for a given program, (2) could be viewed as a difference equation in the utility levels V_t. This observation led naturally to the differential equation (5). Similar intuition applies under uncertainty if we begin with (4). At time t, V_t is known with certainty and thus $m(\sim V_t | \mathscr{F}_t) = V_t$. Relation (4) therefore defines a difference equation in the certainty equivalent utility. The continuous-time analogue, we argue in this paper, is the stochastic differential representation (7), at least in the case of information generated by Brownian motion.

In the remainder of this section we give a definition of continuous-time recursive utility that, in the following section, will be used to justify the stochastic differential expression (7) for utility. We also introduce some notation and terminology to be used later on.

The primitives for uncertainty in our model are:
- a time set $\mathscr{T} = [0, T]$, for some finite T,
- a probability space $(\Omega, \mathscr{F}, P)$, and
- a filtration $\mathbb{F} = \{\mathscr{F}_t: t \in \mathscr{T}\}$ of sub-σ-algebras of $\mathscr{F}$ satisfying the usual conditions (right-continuous, increasing, and augmented).

For notational simplicity alone, we suppose throughout that $\mathscr{F}_0$ is trivial, in that it contains only events for probability zero or 1. For technical definitions, we refer the reader throughout to Protter (1990) or Chung and Williams (1990). An Appendix co-authored with Costis Skiadas treats the infinite time horizon case.

Consumption processes are valued in a closed convex subset $\mathscr{C}$ of some separable Banach lattice, which, for concreteness, the reader can think of as $\mathbb{R}^l$, for l commodities. We use $\|c\|$ to denote the norm of a consumption choice $c \in \mathscr{C}$. The space D of consumption processes is then taken to be the optional[2] $\mathscr{C}$-valued square-integrable processes. (As usual, we equate any two consumption processes that are equal almost everywhere.) The optional restriction implies basically that the consumption rate c_t can depend only on the information available at time t. The square-integrability restriction means that any c in D has finite norm $\|c\|_D = [E(\int_0^T \|c_t\|^2 \, dt)]^{1/2}$.

The utility process for a given consumption process $c \in D$ is defined by a semimartingale[3] V; for any time t, the random variable V_t is treated as the utility for the "continuation" $\{c_s : s \geqslant t\}$ given $\mathscr{F}_t$, as with (1). Of course, V_0 is the utility of the entire process c.

For any interval X of the real line, let $\mathscr{P}(X)$ denote the space of probability measures on X whose mean exists. By a *certainty equivalent* m, we mean a function $m : \mathscr{P}(X) \to \mathbb{R}$ that assigns to a probability measure p (representing the distribution of utility) its certainty equivalent $m(p)$, satisfying:

ASSUMPTION 1: (i) $m(\delta_x) = x$, $x \in X$, where δ_x denotes the Dirac measure. (ii) (*Monotonicity*) $m(p') \geqslant m(p)$ if p' exhibits first order stochastic dominance[4] over p.

One example of such a function—others are provided below—is the expected-utility based specification

$$(8) \qquad m(\sim V) = h^{-1}(E[h(V)]),$$

where V is a real-valued integrable random variable, $\sim V$ denotes its distribution, and the von Neumann-Morgenstern index h is continuous, strictly increasing, and satisfies a growth condition.[5]

Risk aversion for m does not necessarily correspond to risk aversion of intertemporal utility, and thus the former will not generally be assumed for a certainty equivalent. Occasionally, however, we will adopt the following assumption.

---

[2] A $\mathscr{C}$-valued process c is optional if $c : \Omega \times \mathscr{T} \to \mathscr{C}$ is measurable with respect to the σ-algebra on $\Omega \times \mathscr{T}$ generated by the space of right-continuous with left limits (càdlàg) $\mathbb{F}$-adapted processes, and the Borel σ-algebra on $\mathscr{C}$.
[3] A semimartingale is an adapted process that can be written in the form $H + Y$, with Y a local martingale and H a finite variation process.
[4] A distribution p' exhibits first order stochastic dominance over a distribution p, by definition, if the cumulative distribution function of p' is dominated by the cumulative distribution function of p.
[5] That is, for some constant k, $h(x) \leqslant k(1 + |x|)$, implying integrability of $h(V)$.

ASSUMPTION 2: *For all $p \in \mathscr{P}(X)$, $m(p) \leqslant m(\delta_{\bar{p}})$, where $\bar{p} = \int_X x \, dp(x)$ is the mean of p.*

For purposes of motivating our formulation of recursive utility in continuous-time, we will suppose that V is the utility process for a consumption process c. In a discrete-time setting, with sufficient regularity on W, we can apply the implicit function theorem to (4) in order to obtain a representation of the form:

$$(9) \qquad m(\sim V_{t+1} | \mathscr{F}_t) = G(c_t, V_t),$$

for some $G: \mathscr{C} \times \mathbb{R} \to \mathbb{R}$. Subtracting $m(\sim V_t | \mathscr{F}_t) = m(\delta_{V(t)}) = V_t$ from each side of (9) leaves

$$(10) \qquad m(\sim V_{t+1} | \mathscr{F}_t) - m(\sim V_t | \mathscr{F}_t) = G(c_t, V_t) - V_t.$$

Of course, the definition of G depends on the length of a time interval, so we could think in terms of discrete-time approximations to continuous time by writing $G(c, v) \equiv \mathscr{G}(c, v, \Delta t)$, where Δt is the length of a time interval, and then re-expressing (10) as

$$(11) \qquad m(\sim V_{t+\Delta t} | \mathscr{F}_t) - m(\sim V_t | \mathscr{F}_t) = \mathscr{G}(c_t, V_t, \Delta t) - \mathscr{G}(c_t, V_t, 0).$$

Assuming differentiability with respect to Δt, we have

$$(12) \qquad \frac{d}{ds} m(\sim V_{t+s} | \mathscr{F}_t)_{|s=0} = -f(c_t, V_t) \quad \text{a.s.,}$$

where $f(c, v) = -\partial \mathscr{G}(c, v, 0)/\partial \Delta t$.

With this background, it is convenient to take as primitives for our continuous-time model of utility a certainty equivalent m and a measurable function $f: \mathscr{C} \times \mathbb{R} \to \mathbb{R}$. The pair (f, m) is called an aggregator.

Based on the informal derivation of (12), we have a definition: The recursive utility process for a consumption process c under an aggregator (f, m) (if it exists) is the unique integrable semimartingale V satisfying $V_T = 0$ and, for all $t < T$, relation (12).

This definition has some intuitive appeal. Since m satisfies Assumption 1(i), (12) reduces to (5) when c is deterministic. Thus, f determines the degree of intertemporal substitution of consumption and other aspects of "certainty preferences." Of course, just as with the additive special case, f also generates collateral risk attitudes under uncertainty. Given f, however, risk attitudes are finally fixed by the certainty equivalent m, which has no effect on intertemporal substitution. Some degree of separation is therefore achieved.

Our next step is to derive an essentially equivalent definition of the utility process based on smoothness assumptions on the certainty equivalent m, show conditions for existence, and finally derive a number of properties of this utility function. The reverse direction of providing an axiomatic derivation of (12) is not pursued here. Our analysis therefore parallels Lucas and Stokey (1984), which begins with W and (2) and examines the implied utility functions, rather than Koopmans (1960), which axiomatizes (2). In the discrete-time case with

uncertainty, Epstein and Zin (1989) begin with the functional structure (4), while Chew and Epstein (1990) provide an axiomatic basis. A different axiomatic approach is given by Skiadas (1991).

3. A STOCHASTIC DIFFERENTIAL MODEL OF RECURSIVE UTILITY

In order to derive a stochastic differential model of utility based in spirit on the recursive definition (12), we adopt smoothness assumptions on the certainty equivalent function m. Our objective is to reach via informal arguments the model of utility defined by (18) and subsequently by (25).

3.1. *Smoothness Assumptions*

Consider the following smoothness hypothesis which is related to Machina's (1982) local expected utility hypothesis described in Example 4 of Section 4. The Gateaux derivative of a certainty equivalent m at a measure ν in the direction of a measure p, when it exists, is defined by

$$(13) \qquad \nabla m(\nu; p) = \lim_{\alpha \downarrow 0} \frac{m(\nu + \alpha p) - m(\nu)}{\alpha}.$$

(We take p of the form $\nu - \mu$, where μ is a probability measure.) We define a certainty equivalent m to be *smooth at certainty* if, for each x in $\mathbb{R}$, there is some $M(\cdot, x)$: $\mathbb{R} \to \mathbb{R}$ with two continuous derivatives such that, for any probability measure p with compact support, $\nabla m(\delta_x; p)$ exists and

$$(14) \qquad \nabla m(\delta_x; p) = \int M(y, x) \, dp(y).$$

We call M: $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$ the local gradient representation (LGR) of m. Since, for any x and y in $\mathbb{R}$, we have $\nabla m(\delta_x; \delta_y) = y$, it is always the case that the first partial derivative of $M(y, x)$ with respect to y at $y = x$ is $M_1(x, x) = 1$.

For the expected-utility certainty equivalent m in (8), for instance, the LGR is defined by

$$M(v, x) = \frac{h(v)}{h'(x)}.$$

In order to see this, write (8) in the form $h[m(\sim V)] = E[h(V)]$ and note that the LGR of the right side is $h(v)$ and that of the left is $h'(x)M(v, x)$, by a suitable form of the chain rule. The required conditions for M are satisfied if h satisfies the standard assumptions of twice continuous differentiability with a positive first derivative. To see what is excluded by these assumptions, suppose that h has distinct right and left derivatives at some x^*. Fix a random variable ε having zero mean and positive (finite) variance, and for each scale factor $t > 0$,

define the risk premium $\pi(t)$ by

$$h[x^* - \pi(t)] = E[h(x^* + t\varepsilon)].$$

Then $\pi'(0^+) \neq 0$, implying that for small gambles about x^* the risk premium is proportional to the standard deviation of the gamble. Segal and Spivak (1990) refer to this property (for all, not only expected, utility functions) as 'first-order' local risk aversion (or risk loving if $\pi' < 0$), in order to contrast it with the more customary 'second-order' local risk aversion, for which the premium is proportional to the variance. We do not have a characterization of certainty equivalents that are smooth at certainty, but can refer to Allen (1987) for a characterization of smooth certainty equivalents. The examples in Section 4, however, show that this assumption is not overly restrictive and they suggest the conjecture that the essential property of preferences that is excluded by smoothness at certainty is "first-order" risk aversion. Section 7.3 sketches an example characterizing utility without a smooth certainty equivalent in terms of "local time" risk aversion. See Epstein and Zin (1990) and (1991) for some implications of "first-order" risk aversion for asset pricing.

For later reference it is convenient to note the following implication of Assumption 2, risk aversion of a certainty equivalent m with LGR M. Denoting $M_{11}(v, x) \equiv \partial^2 M(v, x)/\partial v^2$, we have

(15) $M_{11}(x, x) \leqslant 0, \quad x \in \mathbb{R},$

based on the same (Jensen Inequality) reasoning used by Machina (1982).

3.2. The Brownian Case

For now, we will restrict ourselves to the standard filtration $\mathbb{F} = \{\mathscr{F}_t\}$ of a standard Brownian motion B in $\mathbb{R}^d$, for some dimension $d \in \mathbb{N}$. That is, $\mathscr{F}_t$ is the σ-algebra generated by the null sets of $\mathscr{F}$ and $\{B_s: 0 \leqslant s \leqslant t\}$. Based on this assumption of "Brownian information," we can deduce the form of the utility process V corresponding to a given consumption process c in D. At the end of the paper, we make some remarks concerning extensions to more general information.

With Brownian information given an aggregator (f, m) with certainty equivalent m that is smooth at certainty, it is natural to conjecture that a utility process V is an Ito process. That is, we conjecture that V has a stochastic differential representation of the form

$$dV_t = \mu_t \, dt + \sigma_t \, dB_t,$$

where μ and σ are progressively measurable processes valued in $\mathbb{R}$ and $\mathbb{R}^d$, respectively. At any time t, under strong technical conditions (that we avoid here since we have no need to rigorously justify the following calculation) we

would have[6]

$$\frac{d}{ds} m(\sim V_{t+s} | \mathcal{F}_t)_{|s=0}$$

$$= \nabla m \left(\delta_{V(t)}; \frac{d}{ds} (\sim V_{t+s} | \mathcal{F}_t)_{|s=0} \right)$$

$$= \lim_{\alpha \downarrow 0} \frac{m \left(\delta_{V(t)} + \alpha (d/ds)(\sim V_{t+s} | \mathcal{F}_t)_{|s=0} \right) - m(\delta_{V(t)})}{\alpha}$$

$$= \lim_{\alpha \downarrow 0} \frac{m \left(\delta_{V(t)} + \alpha \lim_{s \downarrow 0} \left[(\sim V_{t+s} | \mathcal{F}_t) - \delta_{V(t)} \right]/s \right) - m(\delta_{V(t)})}{\alpha}$$

$$= \lim_{\alpha \downarrow 0} \lim_{s \downarrow 0} \frac{m \left(\delta_{V(t)} + \alpha \left[(\sim V_{t+s} | \mathcal{F}_t) - \delta_{V(t)} \right]/s \right) - m(\delta_{V(t)})}{\alpha}$$

(16)

$$= \lim_{s \downarrow 0} \lim_{\alpha \downarrow 0} \frac{m \left(\delta_{V(t)} + \alpha \left[(\sim V_{t+s} | \mathcal{F}_t) - \delta_{V(t)} \right]/s \right) - m(\delta_{V(t)})}{\alpha}$$

$$= \lim_{s \downarrow 0} \frac{1}{s} E \left[M(V_{t+s}, V_t) - M(V_t, V_t) | \mathcal{F}_t \right]$$

$$= \lim_{s \downarrow 0} \frac{1}{s} E \left(\int_t^{t+s} \left[M_1(V_\tau, V_t) \mu_\tau + \frac{1}{2} \sigma_\tau \cdot \sigma_\tau M_{11}(V_\tau, V_t) \right] d\tau \bigg| \mathcal{F}_t \right)$$

$$= \mu_t M_1(V_t, V_t) + \frac{1}{2} M_{11}(V_t, V_t) \sigma_t \cdot \sigma_t \quad \text{almost everywhere}$$

$$= \mu_t + \frac{1}{2} A(V_t) \sigma_t \cdot \sigma_t \quad \text{almost everywhere,}$$

where $A(x) \equiv M_{11}(x, x)$. From (12), this would imply that, almost everywhere on $\Omega \times [0, T]$,

(17) $\mu_t = -f(c_t, V_t) - \frac{1}{2} A(V_t) \sigma_t \cdot \sigma_t$.

Since $\sigma_t \cdot \sigma_t$ is the derivative of the quadratic variation $[V]$ of V, we henceforth write, $d[V]_t/dt = \sigma_t \cdot \sigma_t$. We will also refer to A as the *variance multiplier* of m. Since square-integrability of V implies that, for any t, $E(V_T | \mathcal{F}_t) = V_t + E(\int_t^T - \mu_s \, ds | \mathcal{F}_t)$ and since $V_T = 0$, the above rough calculations suggest that

---

[6] The first equality is the chain rule for differentiation. (See, for example, Luenberger (1969) for sufficient technical conditions, essentially Fréchet differentiability throughout.) The second equality is the definition of ∇m. The third equality is the definition of the derivative of the distribution of V_{t+s} given $\mathcal{F}_t$ with respect to s at $s = 0$. The fourth equality calls for continuity of m. The fifth equality, interchanging limits, calls for uniform convergence assumptions. The sixth equality uses the definition of smoothness of m at certainty, plus the linearity of $\nu \to \nabla m(\delta_{V(t)}; \nu)$, which allows $1/s$ to reappear outside the expectation. The seventh equality is Ito's Lemma plus the fact that $\int M_1(V_\tau, V_t) \sigma_\tau \, dB_\tau$ is a martingale under mild technical restrictions. The eighth equality follows (almost everywhere) by Fubini's Theorem for conditional expectation, as in Ethier and Kurtz (1986). The last equality uses $M_1(x, x) = 1$.

we might characterize V as a solution to the integral equation

$$(18) \qquad V_t = E\left[\int_t^T \left(f(c_s, V_s) + \frac{1}{2}A(V_s)\frac{d}{ds}[V]_s\right)ds \,\bigg|\, \mathcal{F}_t\right], \quad t \in [0, T].$$

Technical conditions under which (18) and (12) could easily be proven equivalent would be much stronger and clumsier than those that we will ultimately use to prove existence and uniqueness of a solution V to (18). For our purposes, the following "morally equivalent" definition of the utility process is therefore proposed.

DEFINITION: *Let (f, m) be an aggregator such that m is smooth at certainty and has a measurable variance multiplier A. Then V is the stochastic differential utility (SDU) process for $c \in D$ under (f, m) if V is the unique square-integrable semimartingale satisfying* (18).

If there is an SDU process V^c for each $c \in D$, then the function $U: D \to \mathbb{R}$ defined by $U(c) = V_0^c$ is the *recursive utility function* generated by (f, m). (Strictly speaking, V_0^c is a random variable taking a particular value, denoted $U(c)$, with probability one.) The existence of a recursive utility function under conditions on (f, m) is established at the end of this section. The infinite-horizon case is treated in an appendix co-authored with Costis Skiadas.

3.3. *Ordinally Equivalent Utility Processes*

Only the ordinal properties of a utility function are of interest, and there may exist many different aggregators generating ordinally equivalent utility functions. In order to explore this, we define a *change of variables* as any $\varphi: \mathbb{R} \to \mathbb{R}$ that is strictly increasing and continuous with $\varphi(0) = 0$. Two utility functions U and $\overline{U}$ are ordinally equivalent if there is a change of variables φ such that $\overline{U} = \varphi \circ U$. Two aggregators (f, m) and $(\bar{f}, \overline{m})$ are then defined to be *ordinally equivalent* if they generate ordinally equivalent recursive utility functions.

Consider two aggregators (f, m) and $(\bar{f}, \overline{m})$ generating utility functions, with m and $\overline{m}$ smooth at certainty. We claim that (f, m) and $(\bar{f}, \overline{m})$ are ordinally equivalent if there is a smooth (C^2) change of variables φ with

$$(21) \qquad f(c, z) = \frac{\bar{f}[c, \varphi(z)]}{\varphi'(z)}, \quad (c, z) \in \mathscr{C} \times \mathbb{R},$$

$$m(\sim \cdot) = \varphi^{-1}(\overline{m}[\sim \varphi(\cdot)]).$$

In order to show this, we use the fact that $\varphi[m(\sim \cdot)] = \overline{m}[\sim \varphi(\cdot)]$ and the properties of the respective LGR's m and $\overline{m}$ to get, for $(x, y) \in \mathbb{R}^2$,

$$(22) \qquad \varphi'(x)M(y, x) = \overline{M}[\varphi(y), \varphi(x)],$$

$$A(x) = \varphi'(x)\overline{A}[\varphi(x)] + \frac{\varphi''(x)}{\varphi'(x)},$$

where A and $\bar{A}$ are the respective variance multipliers. For any consumption process c, let V and $\bar{V}$ denote the utility processes for c under (f, m) and $(\bar{f}, \bar{m})$ respectively. By Ito's Lemma, we then know that $\bar{V}_t = \varphi(V_t)$ since

$$
d\bar{V}_t = \left(\varphi'(V_t) \left[-f(c_t, V_t) - \frac{1}{2} A(V_t) \frac{d}{dt}[V]_t \right] \right.
$$
$$
\left. + \frac{1}{2} \varphi''(V_t) \frac{d}{dt}[V]_t \right) dt + \varphi'(V_t) \sigma_t \, dB_t
$$
$$
= \left[-\bar{f}(c_t, \bar{V}_t) - \frac{1}{2} \bar{A}(\bar{V}_t) \frac{d}{dt}[\bar{V}]_t \right] dt + \bar{\sigma}_t \, dB_t,
$$

where $\bar{\sigma}_t = \varphi'(V_t)\sigma_t$. Relying on the fact that both V and $\bar{V}$ are square-integrable by definition, this shows that $\bar{U} = \varphi \circ U$, so indeed (f, m) and $(\bar{f}, \bar{m})$ are ordinally equivalent.

We extend the ordinal equivalence relation on aggregators by defining (f, m) and $(\bar{f}, \bar{m})$ to be ordinally equivalent if they generate ordinally equivalent utility functions, or (in case the existence of utility functions is unresolved) if there is a C^2 change of variables φ satisfying (21). In what follows, we will identify ordinally equivalent aggregators.

At this point, it is evident why the risk-aversion Assumption 2 was not maintained for certainty equivalents—it is not preserved by the ordinal equivalence relation and hence is not a statement about the underlying intertemporal preference ordering. However, as mentioned near the end of Section 2, an assumption such as risk aversion for m is meaningful for intertemporal preferences if f is fixed. See, for example, Propositions 6 and 7 below and the discussion preceding the latter.

Even for fixed f, distinct certainty equivalents m and $\bar{m}$ can generate ordinally equivalent aggregators (f, m) and $(f, \bar{m})$. Given our specialization here to Brownian information, the drift function μ depends on m only via its variance multiplier A. Thus, we are justified in referring to (f, m) and $(f, \bar{m})$, or m and $\bar{m}$ as being *observationally equivalent* with Brownian information if their respective variance multipliers A and $\bar{A}$ are identical functions. In that case, an individual's choices between consumption processes would be identical given m or $\bar{m}$. Some comments on the implications for "observability" with more general information structures are offered in Section 4.

3.4. *Normalization of Stochastic Differential Utility*

In this subsection we pursue conditions under which, for each consumption process c, there is a unique solution V to the defining equation (18).

In order to reduce the existence issue to a standard fixed point problem, our first step is to design a change of variables φ that eliminates the variance multiplier A from the formulation. That is, we consider the possibility of choosing a φ so that the new variance multiplier $\bar{A}$ defined by (22) is zero. It is

enough that φ satisfies the differential equation

$$(23) \qquad \varphi''(x) = A(x)\varphi'(x).$$

Solutions to (23) are defined by

$$(24) \qquad \varphi(v) = C_2 + C_1 \int_{v_0}^{v} \exp\left[\int_{v_0}^{u} A(x)\, dx\right] du,$$

where v_0 is arbitrary and C_2 and C_1 are constants, with $C_1 > 0$, chosen so that $\varphi(0) = 0$. This change of variables was developed in Duffie and Lions (1990) for the purpose of demonstrating PDE solutions of stochastic differential utility in a Markov setting as a function of the underlying Markov state. (The "upper" limits of the integrals defining φ may in fact be less than v_0, in which case the direction of integration is reversed. If A is continuous, for example, these integrals exist.) If $\bar{V} = \varphi \circ V$ is square-integrable, it follows that V satisfies (18) if and only if

$$(25) \qquad \bar{V}_t = E\left[\int_t^T \bar{f}(c_s, \bar{V}_s)\, ds \,\middle|\, \mathscr{F}_t\right], \quad t \in [0, T].$$

By the above construction, any aggregator (f, m) with a continuous variance multiplier has an ordinally equivalent aggregator $(\bar{f}, \bar{m})$ whose variance multiplier is zero. For our purposes, this has the same effect as having $\bar{m}(\sim V) = E(V)$. We refer to $(\bar{f}, \bar{m})$, or $\bar{f}$ itself, as the normalized version of (f, m). We will henceforth place assumptions directly on a *normalized aggregator*, that is, a measurable function $f: \mathscr{C} \times \mathbb{R} \to \mathbb{R}$.

We remind the reader once again that the reduction to a normalized aggregator $(\bar{f}, \bar{m})$ does not mean that intertemporal utility is risk-neutral or that we have lost the ability to disentangle substitution from risk aversion. The point is simply that the modeler can select which of the many ordinally equivalent aggregators to use depending on the question at hand. The normalized aggregator is advantageous for proof of existence of stochastic differential utility, while the un-normalized aggregator (f, m) is convenient for achieving the desired disentangling by changing m with f fixed, as described in Section 5.6. Such a change in risk aversion is much less readily described in terms of a normalized aggregator. (For instance, consider the effect of a change in α on the normalized and un-normalized aggregators in Example 3 below.)

3.5. *Existence of Stochastic Differential Utility*

In summary, for the case of Brownian information, the utility process can be viewed without essential loss of generality as the solution $\bar{V}$ to an integral equation of the form (25). Once we have posed the problem in this normalized form, we do not actually need to assume Brownian information in order to prove existence and uniqueness of the utility process. Of course, for non-Brownian information, (25) characterizes recursive utility for only a sub-class of aggregators.

Our main result of this section is the following set of sufficient conditions on a normalized aggregator f for the existence of recursive utility:

- f is *Lipschitz in utility*, in the sense that there exists some constant k such that, for all $c \in \mathscr{C}$ and all $(v, w) \in \mathbb{R}^2$, we have $|f(c, v) - f(c, w)| \leqslant k |v - w|$.
- f satisfies a *growth condition in consumption*, in the sense that there are constants k_1 and k_2 such that, for all $c \in \mathscr{C}$, we have $|f(c, 0)| \leqslant k_1 + k_2 \|c\|$.

The conditions allow, for example, $f(c, v) = c^\alpha - \beta v$ (for constants $\alpha \in (0, 1)$ and β), which generates the additively separable specification (1) for $u(c) = c^\alpha$. The growth condition on f and square-integrability condition defining D can both be relaxed. It is actually enough for all of the results of the paper that $\|c_t\|^\alpha$ is integrable for some $\alpha > 1$ and that $f(c, 0) \leqslant k_1 + k_2 \|c\|^\gamma$ for some $\gamma < \alpha$. Our growth and Lipschitz conditions are typical for the existence of solutions to stochastic differential equations, a similar fixed point problem.

THEOREM 1: *Suppose $(\Omega, \mathscr{F}, \mathbb{F}, P)$ is a filtered probability space, where $\mathbb{F} = \{\mathscr{F}_t: t \in [0, T]\}$ satisfies the usual conditions. Let $\bar{f}: \mathscr{C} \times \mathbb{R} \to \mathbb{R}$ be measurable, Lipschitz in utility, and satisfy a growth condition in consumption. Then, for any consumption process $c \in D$, there is a unique square-integrable semimartingale $\bar{V}$ satisfying* (25).

Proof is shown in Appendix A and extended to the infinite horizon case in Appendix C. In Section 5, several properties are established for the recursive utility function $U: D \to \mathbb{R}$ defined by $U(c) \equiv V_0^c$, whose existence is guaranteed by Theorem 1.

4. EXAMPLES

We offer several examples of aggregators and, in some cases, the corresponding stochastic differential utility function. A closed-form expression for the utility function is not generally available.

In order to clarify the examples and differences among them, we describe how they rank the three consumption programs described in the introduction. One can verify that, apart from the special cases of Examples 2 and 3 which coincide with (1), we have:

$$c^A \nsim c^B \sim c^C \qquad \text{for Example 2,}$$

$$c^A \nsim c^B \nsim c^C \qquad \text{for Example 3.}$$

That is, neither $c^A \sim c^B$ nor $c^B \sim c^C$ is implied by stochastic differential utility.

EXAMPLE 1 (Standard Additive Utility): The standard additive expected utility function (1), with the utility process $V_t = E[\int_{s \geqslant t} u(c_s) e^{-\beta(s-t)} ds \mid \mathscr{F}_t]$, corresponds to the aggregator (f, m), where

(26) $\quad f(c, v) = u(c) - \beta v \quad \text{and} \quad m(\sim V) = E(V).$

Suppose $\mathscr{C} = \mathbb{R}_+$ or $\mathscr{C} = \mathbb{R}$. If $u(0) = 0$ and u has the usual properties, then an

ordinally equivalent aggregator $(\bar{f}, \overline{m})$ is given by

$$(27) \qquad \bar{f}(c,v) = \beta \frac{u(c) - u(v)}{u'(v)} \quad \text{and} \quad \overline{m}(\sim V) = u^{-1}(E[u(V)]).$$

This aggregator satisfies (29) below. The corresponding utility process $\bar{V}$ satisfies

$$\bar{V}_t = u^{-1} \circ E\left[\beta \int_{s \geq t} u(c_s) e^{-\beta(s-t)} ds \,\middle|\, \mathscr{F}_t\right].$$

EXAMPLE 2 (Uzawa Utility): Let $f(c,v) = u(c) - \beta(c)v$ and $m(\sim V) = E(V)$. The intertemporal expected utility functional (6) is generated by (f, m). The utility process is

$$V_t = E\left[\int_{s \geq t} u(c_s) e^{-\int_t^s \beta(c_\tau) d\tau} ds \,\middle|\, \mathscr{F}_t\right].$$

EXAMPLE 3 (Example of Kreps-Porteus Utility): Let $\mathscr{C} = \mathbb{R}_+$, $0 \neq \rho \leq 1$, $0 \leq \beta$, $0 \neq \alpha \leq 1$, and define

$$(28) \qquad f(c,v) = \frac{\beta}{\rho} \frac{c^\rho - v^\rho}{v^{\rho-1}} \quad \text{and} \quad m(\sim V) = [E(V^\alpha)]^{1/\alpha}.$$

The function f coincides with that in (27) if $u(c) \equiv c^\rho/\rho$. Thus, for deterministic consumption processes, the utility function is of the additive CES form, with elasticity of intertemporal substitution $(1-\rho)^{-1}$. If $\alpha = \rho$, then (28) and (26) coincide and the utility functions they generate are identical for stochastic consumption processes as well. This is not so, however, if $\alpha \neq \rho$. It can be shown (by applying Ito's Lemma and taking limits as the length of a time interval goes to zero) that the corresponding stochastic differential equation for the utility process V is the continuous-time limit of the homogeneous CES specification examined in discrete time by Epstein and Zin (1989). It is shown there that risk aversion of the intertemporal ordering increases as α falls. A corresponding result for our continuous-time framework is shown in Proposition 6 below. An ordinally equivalent aggregator, for which the variance multiplier is 0, is given by

$$\bar{f}(c,v) = \frac{\beta}{\rho} \frac{c^\rho - (\alpha v)^{\rho/\alpha}}{(\alpha v)^{(\rho/\alpha)-1}} \quad \text{and} \quad \overline{m}(\sim V) = E(V), \quad \alpha \leq 1, \quad \alpha \neq 0.$$

Although $\bar{f}$ does not satisfy the Lipschitz condition for Theorem 1, existence and uniqueness is shown by PDE methods, in a Markov setting, by Duffie and Lions (1990). This ends Example 3.

The certainty equivalents considered so far are all consistent with expected utility theory. A number of generalizations of the expected utility model have been proposed recently as rationalizations of behavior such as that exhibited in

the Allais paradox. In the final two examples, we integrate two of these generalizations into our model in order to examine their implications for choice in a continuous-time setting. The new theories all deal with static or one-shot choice environments, and specify utility or evaluation functionals for probability distributions on the real line. These functionals enter our model as alternative certainty-equivalent components of an aggregator.

It will be convenient to consider real-valued consumption processes ($\mathscr{C} = \mathbb{R}_+$) over an infinite time horizon, considered in Appendix C, and also to adopt a change of variables for which the constant path at level $\bar{c}$ has utility $\bar{c}$. It follows that the utility of any consumption process c is equal to $\bar{c}$ if the constant consumption path at level $\bar{c}$ is indifferent to c. In other words, utility is measured in consumption units. This normalization of utility corresponds to the condition

$$(29) \qquad f(c, c) = 0, \quad c \in \mathbb{R}_+.$$

Of course, (18) defines the manner in which m affects intertemporal utility. Preference or behavioral interpretations of m are more explicit, however, if we temporarily extend beyond the case of Brownian information used in most of our analysis. Such an extension is also useful in understanding the discussion (in the following examples) of the extent to which different certainty equivalents are observationally equivalent. Consider consumption processes for which consumption is constant at c_0 on $[0, \tau)$ and such that, at time τ, consumption jumps to a new random level $\bar{c}$ that persists thereafter. Both τ and the probability distribution $\sim \bar{c}$ are known at $t = 0$, and no further information regarding the latter is obtained during $(0, \tau)$. Suppose, further, that utility is measured in consumption units and is generated by an aggregator (f, m) satisfying (29). Then, by a natural extension of our analysis, for given τ and c_0, the utility of the process described above is an increasing function of $m(\sim \bar{c})$. Thus m defines the preference ranking between suitable pairs of consumption processes.

EXAMPLE 4 (Machina's Extension of Expected Utility): Machina (1982) proposes that the assumption of an expected-utility functional be replaced by the assumption of Fréchet differentiability of the utility functional. Let Ψ be a real-valued function that is Fréchet differentiable in a suitable norm on the set, say dom Ψ, of probability distributions on the real line having compact support. Let ψ, defined by $\psi(x) \equiv \Psi(\delta_x)$, correspond to the restriction of Ψ to dirac measures. If Ψ is increasing in the sense of first order stochastic dominance, then ψ and Ψ have effectively identical ranges, and we can define a certainty equivalent m by

$$(30) \qquad m(\cdot) \equiv \psi^{-1}[\Psi(\cdot)].$$

Then m is Fréchet differentiable; at each measure ν in dom(Ψ), let $\Gamma(\cdot, \nu)$: $\mathbb{R} \to \mathbb{R}$ denote the Riesz representation of its Fréchet derivative at ν, as in Machina (1982). If for any dirac measure $\delta_x, \Gamma_1(\cdot, \delta_x)$ is bounded and continuously differentiable and V is a process whose conditional distribution is a

Fréchet differentiable function of time, then the local gradient representation M of m is defined by $M(v, x) = \Gamma(v, \delta_x)$. (These assumptions for $\Gamma(\cdot, \delta_x)$ constitute additional restrictions on Ψ or m beyond Fréchet differentability. But they appear to be purely technical in nature and, moreover, are implied by the differentiability assumptions that Machina frequently (for instances, pp. 300 and 307) adopts for his "local" utility functions.) The representation Γ of the Fréchet derivative uniquely identifies m since, as Machina shows, the local information provided by Γ can be "pieced together" to make the global comparisons necessary to uncover m. But the same is not true for M, which provides local information about m only near dirac measures. Thus, there exists $\overline{m}$ distinct from m for which M and $\overline{M}$ coincide.

Let f be arbitrary subject to (29), and consider the aggregator (f, m). Since the characterization (18) of the utility process involves m only via its variance multiplier A, we see from the above discussion that, in general, there exist many Fréchet differentiable certainty equivalent functions that are observationally equivalent. For example, let h be any strictly increasing C^2 function satisfying

$$(31) \qquad \frac{h''(x)}{h'(x)} = A(x), \quad x \in \mathbb{R}_+,$$

and let $\overline{m}$ be the corresponding expected-utility certainty equivalent defined by (8). It follows that (f, m) and $(f, \overline{m})$ imply identical choices among consumption processes adapted to Brownian information. In other words, in this environment, Machina's extension of expected utility, suitably adapted and integrated into our intertemporal setting, is *empirically indistinguishable from expected utility*.

The intuition for this finding is clear: Given recursivity, the choice between two consumption processes is determined, via "integration," by the certainty equivalent assigned at each t to the uncertainty to be faced over a "small" interval $[t, t + \varepsilon]$. At t, V_t is known with certainty, while the uncertainty represented by $V_{t+\varepsilon}$ is "small" given Brownian information. Thus, the only properties of $m(\cdot)$ that are relevant are those that are reflected in evaluating small risks *about certainty*. Machina's local analysis and propositions extend expected utility analysis and results in a substantive way precisely because it is assumed that the relevant domain includes small gambles about *arbitrary* initial probability distributions. (Mathematically, therefore, one can identify the entire Fréchet derivative rather than just the local gradient representation M.) Even if jumps in utility are allowed, as when a Poisson component is present, the uncertainty faced at any instant still represents a small risk about certainty. Though a discrete jump is possible, the size of the jump is deterministic and the conditional probability that one will occur in $[t, t + \varepsilon]$ is small (roughly $1 - e^{-\lambda \varepsilon}$, where λ is the intensity of the Poisson process). This explains why, even with mixed Brownian-Poisson information, one cannot uniquely identify m. Such identification can be achieved, however, if the more general jump processes described prior to this example are admitted into the domain. The gambles

defined by those processes are not small in any sense. Section 7 formulates a stochastic differential model of recursive utility in the presence of mixed Brownian-Poisson information. This ends Example 4.

EXAMPLE 5 (Dekel-Chew Betweenness Certainty Equivalents): An axiomatic generalization of expected utility, which is not generally Fréchet differentiable, is described by Dekel (1986) and Chew (1989). We say that $m(\cdot)$ is a *betweenness* certainty equivalent if there exists some continuous $H: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that $H(x, x) \equiv 0$, $H(\cdot, x)$ is increasing and, for all $p \in \mathscr{P}(\mathbb{R})$, $m(p)$ is the unique solution to

$$(32) \qquad \int H[x, m(p)] \, dp(x) = 0.$$

If $H(x, y) = h(x) - h(y)$, then m reduces to the expected utility form $h^{-1}(E[h(\cdot)])$ of (8). Suppose that, for each y, the partial $H_1(\cdot, y)$ is positive and continuously differentiable. Then, by Ito's Lemma, m is smooth near certainty, with LGR

$$(33) \qquad M(\cdot, y) = \frac{H(\cdot, y)}{H_1(y, y)} .$$

In the case of Brownian information, it follows as in the preceding example that betweenness and expected utility certainty equivalents are empirically indistinguishable from one another. But a caveat must be applied to this assertion since the above differentiability assumptions for $H(\cdot, y)$ rule out some economically interesting betweenness functionals, called semi-weighted utility by Chew (1989). Such utility functions exhibit 'first-order' local risk aversion, as explained in Section 3. First-order local risk aversion is also a generic property of rank-dependent utility theory, which is an alternative axiomatic generalization of expected utility due to Quiggin (1982), Yaari (1987), and Segal (1989). The integration of such theories into a continuous-time framework is addressed briefly in Section 7.3.

To conclude this example, we briefly anticipate the calculations in Section 7.1 dealing with mixed Brownian-Poisson information. If (f, m) and (f, m^*) are two aggregators, then the corresponding stochastic differential utility functions coincide if and only if $M = M^*$. But for certainty equivalents in the betweenness class, $M = M^*$ if and only if $m = m^*$. (To see this, let m^* and H^* also satisfy (32) with $M^* = M$. Then (33) implies that, for all x and y, $H(x, y) = a(y)H^*(x, y)$, where $a(y) \equiv H_1(y, y)/H_1^*(y, y) \neq 0$. It is therefore immediate from (32) that $m = m^*$. The converse is evident.) In other words, given f, the betweenness certainty equivalent is uniquely determined by preferences over consumption processes under the information generated by combinations of Brownian and Poisson processes. In particular, betweenness and expected utility certainty equivalents *can* be distinguished from one another on the basis of observed choices in such an environment. This ends Example 5.

Our statements concerning the observational equivalence or inequivalence of various forms of preferences in a continuous-time setting apply, in particular, to any distinctions that can be drawn from empirical studies of asset prices. In this regard, our statements of observational equivalence also go beyond those of Kocherlakota (1987), which apply to asset pricing in static or i.i.d. settings.

5. PROPERTIES OF STOCHASTIC DIFFERENTIAL UTILITY

This section shows that the stochastic differential utility function defined in Section 3 has a range of natural properties, under natural conditions. We will provide sufficient conditions for: (i) continuity; (ii) monotonicity with respect to terminal value V_T; (iii) monotonicity with respect to consumption c; (iv) time consistency; (v) concavity; (vi) comparative risk aversion; (vii) risk aversion; (viii) homotheticity.

Unless otherwise stated, throughout this section $(\Omega, \mathscr{F}, P, \mathbb{F})$ is general; we require only that the filtration $\mathbb{F} = \{\mathscr{F}_t : t \in [0, T]\}$ satisfies the usual conditions, is augmented, and that $\mathscr{F}_0$ is trivial (has only events of probability zero or 1). We define $f : \mathscr{C} \times \mathbb{R} \to \mathbb{R}$ to be *regular* if f is continuous, Lipschitz in utility, and satisfies a growth in condition in consumption. It is a maintained hypothesis of this section that the given normalized aggregator f is regular. This ensures (by Theorem 1) existence of the associated utility function U.

5.1. *Continuity*

PROPOSITION 1 (Continuity): *The stochastic differential utility function* U: $D \to \mathbb{R}$ *is continuous.*

PROOF: Given c and $\tilde{c}$ in D, let $V = V^c$ and $\tilde{V} = V^{\tilde{c}}$ be the associated utility processes. For any t,

$$|V_t - \tilde{V}_t| \leqslant E\left[\int_t^T \left| f(c_s, V_s) - f(\tilde{c}_s, \tilde{V}_s) \right| ds \,\middle|\, \mathscr{F}_t \right]$$

$$\leqslant E\left[\int_t^T \left(\left| f(c_s, V_s) - f(\tilde{c}_s, V_s) \right| \right.\right.$$

$$\left.\left. + \left| f(\tilde{c}_s, V_s) - f(\tilde{c}_s, \tilde{V}_s) \right| \right) ds \,\middle|\, \mathscr{F}_t \right]$$

$$\leqslant E\left[\int_t^T \left(\left| f(c_s, V_s) - f(\tilde{c}_s, V_s) \right| + k \left| V_s - \tilde{V}_s \right| \right) ds \,\middle|\, \mathscr{F}_t \right].$$

The stochastic Gronwall-Bellman Inequality (Appendix B) yields

$$|U(c) - U(\tilde{c})| \leqslant E\left[\int_0^T e^{kt} \left| f(c_t, V_t) - f(\tilde{c}_t, V_t) \right| dt \right].$$

If $\{c^n\}$ is a sequence converging in D to c, then for all n,

$$|U(c^n) - U(c)| \leqslant E\left[\int_0^T e^{kt}|f(c_t^n, V_t) - f(c_t, V_t)| \, dt\right].$$

Since c^n converges to D to c (and therefore, along some subsequence of every subsequence, pointwise almost everywhere on $\Omega \times [0, T]$ with respect to the optional σ-algebra) and f is continuous and satisfies a growth condition in consumption, the dominated convergence theorem (Ethier and Kurtz (1987, p. 492)) and Cauchy-Schwartz inequality imply that the integral on the right-hand side converges with n to zero, proving continuity of U.   Q.E.D.

5.2. *Monotonicity For Terminal Value*

We momentarily extend the definition of stochastic differential utility functions so that there is a terminal reward at some $[0, T]$-valued stopping time τ. The terminal reward is defined by some $\mathscr{F}_\tau$-measurable $Y \in L^1(P)$. Proposition A1 (Appendix A) implies that there is a unique integrable semimartingale $V^{c, Y}$ solving the equation

$$(34) \qquad V_t^{c, Y} = E\left[\int_t^\tau f(c_s, V_s^{c, Y}) \, ds + Y \bigg| \mathscr{F}_t\right], \quad t \in [0, T].$$

PROPOSITION 2 (Monotonicity for Terminal Value): *Let τ be a $[0, T]$-valued stopping time. Suppose $Y \geqslant W$, for $\mathscr{F}_\tau$-measurable Y and W in $L^1(P)$. For any given $c \in D$, let $V = V^{c, Y}$ and $\tilde{V} = V^{c, W}$ be defined by equation (34). Then $V \geqslant \tilde{V}$.*

PROOF: For starters, suppose $\tau = T$. We have

$$V_t - \tilde{V}_t = E\left[\int_t^T f(c_s, V_s) - f(c_s, \tilde{V}_s) + Y - W \bigg| \mathscr{F}_t\right],$$

and since f is Lipschitz in utility,

$$f(c_s, V_s) - f(c_s, \tilde{V}_s) \geqslant -k|V_s - \tilde{V}_s|.$$

The result follows by Appendix Lemma B2. For general τ, we can replace $f(c_s, V_s)$ by $1_{\{s \leqslant \tau\}} f(c_s, V_s)$ throughout the above and get the same answer.
                                                                Q.E.D.

5.3. *Monotonicity For Consumption*

Now we show monotonicity of recursive utility with respect to consumption under the assumption that f is *increasing in consumption*, meaning that, for any $v \in \mathbb{R}$, $f(\cdot, v)$ is an increasing function. We say that f is *strictly increasing in consumption* if, for all v, the function $f(\cdot, v)$ is strictly increasing. Likewise, we say that the stochastic differential utility function $U: D \to \mathbb{R}$ is increasing if $c \geqslant \tilde{c}$ implies that $U(c) \geqslant U(\tilde{c})$, and that U is strictly increasing if $U(c) > U(\tilde{c})$ whenever $c \geqslant \tilde{c}$ and $c \neq \tilde{c}$.

PROPOSITION 3 (Monotonicity in Consumption): *If f is increasing in consumption, then U is increasing. If f is strictly increasing in consumption, then U is strictly increasing.*

PROOF: Pick any c and $\tilde{c}$ in D with $c \geqslant \tilde{c}$. Let V be the utility process for c under f, and $\tilde{V}$ be the utility process for $\tilde{c}$ under f. We have, for all t,

$$V_t - \tilde{V}_t = E\left(\int_t^T \left[f(c_s, V_s) - f(\tilde{c}_s, \tilde{V}_s) \right] ds \,\middle|\, \mathscr{F}_t \right),$$

and

$$f(c_s, V_s) - f(\tilde{c}_s, \tilde{V}_s) = f(c_s, V_s) - f(\tilde{c}_s, V_s) + f(\tilde{c}_s, V_s) - f(\tilde{c}_s, \tilde{V}_s)$$

$$\geqslant f(c_s, V_s) - f(\tilde{c}_s, V_s) - k|V_s - \tilde{V}_s|.$$

The result follows by Lemma B2 (Appendix B). *Q.E.D.*

5.4. *Time Consistency*

An important property of recursive utility functions is that they exhibit intertemporal consistency. Consider the following axiom on a family $\succcurlyeq = \{\succcurlyeq_{\omega,t}: (\omega, t) \in \Omega \times [0, T]\}$ of binary orders on D. We let $c \succcurlyeq_t \tilde{c}$ denote the event $\{\omega \in \Omega: c \succcurlyeq_{\omega,t} \tilde{c}\}$, and say that $\succcurlyeq$ is adapted if $(c \succcurlyeq_t \tilde{c}) \in \mathscr{F}_t$ for all $(c, \tilde{c}) \in D \times D$ and all $t \in [0, T]$. As usual, $\succ$ denotes the corresponding family of strict preference orders. For example, for each c in D, let V^c denote the utility process generated by c under f. Then f generates a family $\succcurlyeq$ of complete transitive preference orders defined by $c \succcurlyeq_t \tilde{c}$ if and only if $V_t^c \geqslant V_t^{\tilde{c}}$.

DEFINITION (Consistency): An adapted family $\succcurlyeq = \{\succcurlyeq_{\omega,t}\}$ of binary orders on D is *consistent* under the following condition. For each stopping time $\tau \in \mathscr{T}$ and each pair c and $\tilde{c}$ of consumption processes in D, if the restrictions of $\tilde{c}$ and c to $[\![0, \tau]\!]$ coincide, then

$$(35) \qquad P\left(c \succcurlyeq_\tau \tilde{c} \right) = 1 \Rightarrow c \succcurlyeq_0 \tilde{c}$$

and

$$P\left(c \succcurlyeq_\tau \tilde{c} \right) = 1 \quad \text{and} \quad P\left(c \succ_\tau \tilde{c} \right) > 0 \Rightarrow c \succ_0 c.$$

Consistency requires that, for c and $\tilde{c}$ as above, an unambiguous preference for c over $\tilde{c}$ at time τ is respected at time 0. By the nature of recursive utility, $V_\tau^c = V_\tau^{\tilde{c}}$ if c and $\tilde{c}$ coincide on $[\![\tau, T]\!]$; that is, past consumption does not affect preferences. Once consistency or recursive preferences is established, it follows that one can apply dynamic programming to optimization problems in such a way that *state variables reflecting past consumption are unnecessary*. See Section 6 for our dynamic programming results and Section 8.2 for an extension allowing an additional state-variable for "habit persistence" in the sense of Constantinides (1988). For additional utility models in which past consumption

plays a role in determining utility, see Heaton (1988), Sundaresan (1989), Hindy and Huang (1989a and b), Detemple and Zapatero (1989).

That recursive utility functions satisfy consistency is a simple consequence of monotonicity with respect to terminal value (Proposition 2). The following theorem of consistency actually applies under any conditions leading to a stochastic differential representation of utility of the form $V_s = E(\int_s^T g(\omega, t, c_t, V_t)\, dt \mid \mathscr{F}_s)$.

PROPOSITION 4 (Consistency): *The family of preference orders generated by f is consistent.*

PROOF: Pick τ, c, and $\tilde{c}$ as in the definition of consistency. Using the notation of Proposition 2, we have $V_0 = V_0^{c,V(\tau)}$ and $\tilde{V}_0 = V_0^{c,\tilde{V}(\tau)}$. The result then follows from Proposition 2. $\hspace{3cm}$ Q.E.D.

5.5. Concavity

The following conditions for concavity provide, along with the previously shown continuity and monotonicity conditions, sufficient regularity to apply certain general equilibrium results. For existence of general equilibrium in D, however, an additional condition such as uniform properness may be required, as shown by Mas-Colell (1986). See, also, Duffie and Skiadas (1990).

PROPOSITION 5 (Concavity): *Suppose f is a concave function. Then the stochastic differential utility function generated by f is concave.*

PROOF: Let c and b be consumption processes and, for any $\alpha \in [0,1]$, let V^α be the utility process for $\alpha c + (1-\alpha)b$, and let V^c and V^b be the utility processes for c and b, respectively.

For $t \in [0,T]$, let $\delta_t = V_t^\alpha - [\alpha V_t^c + (1-\alpha)V_t^b]$. Then

$$\delta_t = E\left[\int_t^T \left(f[\alpha c_s + (1-\alpha)b_s, V_s^\alpha]\right.\right.$$
$$\left.\left. - \alpha f(c_s, V_s^c) - (1-\alpha)f(b_s, V_s^b)\right) ds \,\middle|\, \mathscr{F}_t\right]$$

$$= E\left[\int_t^T \left(f[\alpha c_s + (1-\alpha)b_s, V_s^\alpha]\right.\right.$$
$$\left.\left. - f[\alpha c_s + (1-\alpha)b_s, \alpha V_s^c + (1-\alpha)V_s^b] + g_s\right) ds \,\middle|\, \mathscr{F}_t\right],$$

where

$$g_s = f[\alpha c_s + (1-\alpha)b_s, \alpha V_s^c + (1-\alpha)V_s^b]$$
$$- \alpha f(c_s, V_s^c) - (1-\alpha)f(b_s, V_s^b).$$

The integrand in the last expression dominates $g_s - k|\delta_s|$, and, by concavity of f, $g_s \geqslant 0$. The result follows by Lemma B2 (Appendix B). *Q.E.D.*

5.6. *Comparative Risk Aversion*

In order to justify our claim that the generalization from additively separable expected utility to recursive utility makes possible a degree of separation between intertemporal substitution and risk aversion, we need the following definition:

DEFINITION (Comparative Risk Aversion): Let U^* and U be two stochastic differential utility functions. We say that U^* is *more risk averse than U* if U^* rejects any gamble that is rejected by U, that is, for any $c \in D$ and any deterministic process $\bar{c}$ in D,

$$U(c) \leqslant U(\bar{c}) \Rightarrow U^*(c) \leqslant U^*(\bar{c}).$$

If U^* and U are comparable according to this definition, then they must rank deterministic programs identically. After a change of variables, they must be identical on the set of deterministic consumption processes, from which it follows that they share the same function f as the first component of their aggregators. Thus the following theorem restricts attention to the case $f^* = f$.

PROPOSITION 6 (Comparative Risk Aversion): *Suppose $\mathbb{F}$ is the standard filtration of a standard Brownian motion in some Euclidean space. Let U^* and U be the stochastic differential utility functions generated by aggregators (f, m^*) and (f, m) respectively, where m^* and m have continuous variance multipliers A^* and A. Suppose the respective normalized aggregators are regular. Then U^* is more risk averse than U if $A^* \leqslant A$.*

The reader will note from the proof that $A^* \leqslant A$ is also close to a necessary condition for U^* to be more risk averse than U.

PROOF: Let φ be defined by (24), the change of variables that eliminates A. This normalizes (f, m) to the aggregator $(\bar{f}, \bar{m})$ whose variance multiplier is zero. Since $A^* \leqslant A$, applying the same change of variables φ to the aggregator (f, m^*) results in a new aggregator $(\bar{f}, \bar{m}^*)$, with a variance multiplier $\bar{A}^* \leqslant 0$. (This is checked by a calculation, using (24) and (22).) Let $\bar{U}$ and $\bar{U}^*$ be the recursive utility functions generated by $(\bar{f}, \bar{m})$ and $(\bar{f}, \bar{m}^*)$, respectively.

In order to prove the result, it is enough to show, for an arbitrary $c \in D$, that $\bar{U}(c) \geqslant \bar{U}^*(c)$. To this end, let $c \in D$ be arbitrary. Let V^* be the utility process associated with c under $(\bar{f}, \bar{m}^*)$. (Its existence is implied by Theorem 1 and the assumptions of this proposition, after inverting the change of variables normaliz-

ing $(\bar{f}, \overline{m}^*)$.) The result is the solution V^* to the equation

$$V_t^* = E\left(\int_t^T \left[\bar{f}(c_s, V_s^*) + \frac{1}{2}\bar{A}^*(V^*)_s \frac{d}{ds}[V^*]_s\right] ds \,\Big|\, \mathscr{F}_t\right), \quad t \in [0, T].$$

Letting V be the utility process for c under $\bar{f}$, we also have

$$V_t = E\left[\int_t^T \bar{f}(c_s, V_s) \, ds \,\Big|\, \mathscr{F}_t\right], \quad t \in [0, T].$$

Subtracting,

$$V_t - V_t^* = E\left[\int_t^T \left(-\frac{1}{2}\bar{A}^*(V^*)_s \frac{d}{ds}[V^*]_s\right.\right.$$

$$\left.\left. + \bar{f}(c_s, V_s^*) - \bar{f}(c_s, V_s)\right) ds \,\Big|\, \mathscr{F}_t\right], \quad t \in [0, T].$$

The integrand dominates

$$-\frac{1}{2}\bar{A}^*(V^*)_s \frac{d}{ds}[V^*]_s - k|V_s - V_s^*|,$$

while

$$-\frac{1}{2}\bar{A}^*(V^*)_s \frac{d}{ds}[V^*]_s \geqslant 0.$$

Lemma B2 then implies that $\bar{U}(c) \geqslant \bar{U}^*(c)$, and the result follows.   Q.E.D.

According to Proposition 6, we can increase the degree of risk aversion of the intertemporal preference ordering without affecting "certainty preferences" by keeping the f-component of the aggregator unchanged and changing the certainty equivalent so that the variance multiplier falls. A sufficient condition for $A^* \leqslant A$ is that $m^*(\cdot) \leqslant m(\cdot)$; that is, m^* is smaller and hence the more risk averse certainty equivalent. This can be proved by applying Ito's Lemma to a judiciously chosen consumption process, and extending the argument used to prove (15). If $m(\sim V) = h^{-1}(E[h(V)])$ and $m^*(\sim V) = h^{*-1}(E[h^*(V)])$, then $m^* \leqslant m$ if and only if h^* is more concave than h on the range of the utility process.

5.7. Risk Aversion

We may not only want to compare the degrees of risk aversion of two recursive utility functions, but also to determine whether a given utility function is risk averse in an absolute sense. For any consumption process c, let $E(c)$ denote the deterministic process defined, almost everywhere, by $E(c)_t = E(c_t)$. (The separability assumption on $\mathscr{C}$ and definition of D imply that c_t is

integrable for almost every t.) A utility function U is *risk averse* if, for all c in D,

$$(36) \qquad U(c) \leqslant U[E(c)].$$

Suppose, momentarily, that $\mathbb{F}$ is the standard filtration of a standard Brownian motion. If (f, m) is an aggregator whose certainty equivalent is risk averse in the sense of Assumption 2, then its variance multiplier A is nonpositive, as shown by (15). This implies, subject to the technical regularity conditions of Proposition 6, that the utility function generated by (f, m) is more risk averse than that generated by (f, m^*), where m^* is merely the expected value function. The following proposition therefore shows, again subject to the technical restrictions of Proposition 6, that (f, m) is risk averse if m is risk averse and $f(\cdot, v)$ is concave for all v. The proposition does not in fact assume Brownian information, but begins with the given regular normalized aggregator f.

PROPOSITION 7 (Risk Aversion): *Suppose, for all $v \in \mathbb{R}$, that $f(\cdot, v): \mathscr{C} \to \mathbb{R}$ is concave. Then the recursive utility function generated by f is risk averse.*

PROOF: Let c be arbitrary in D, and let $\bar{c} = E(c)$. Let V be the utility process for c under f and let $\bar{V}$ be the utility process for $\bar{c}$ under f. We have

$$\bar{V}_t - V_t = E\left[\int_t^T \left[f(\bar{c}_s, \bar{V}_s) - f(c_s, V_s) \right] ds \,\middle|\, \mathscr{F}_t \right].$$

Using Fubini's Theorem for conditional expectations,

$$\bar{V}_t - V_t = E\left[\int_t^T \left(E\left[f(\bar{c}_s, \bar{V}_s) - f(c_s, \bar{V}_s) \middle| \mathscr{F}_t \right] \right.\right.$$
$$\left.\left. + f(c_s, \bar{V}_s) - f(c_s, V_s) \right) ds \,\middle|\, \mathscr{F}_t \right].$$

The integrand dominates

$$E\left[f(\bar{c}_s, \bar{V}_s) - f(c_s, \bar{V}_s) \middle| \mathscr{F}_t \right] - k|\bar{V}_s - V_s|,$$

while, by Jensen's Inequality for conditional expectations,

$$E\left[f(\bar{c}_s, \bar{V}_s) - f(c_s, \bar{V}_s) \middle| \mathscr{F}_t \right] \geqslant 0.$$

The result follows by Lemma B2. $\qquad$ Q.E.D.

5.8. *Homotheticity*

A utility function U is *homothetic* if, for any consumption processes c and c' and any scalar $\lambda > 0$,

$$(37) \qquad U(\lambda c') \geqslant U(\lambda c) \Leftrightarrow U(c') \geqslant U(c).$$

This leads to the following definition for aggregators.

DEFINITION (Homotheticity): Suppose $\mathscr{C}$ is a cone and the aggregator (f, m) generates a recursive utility function. Then (f, m) is *homothetic* if, for any consumption processes c and $\tilde{c}$ and any scalar $\lambda > 0$,

$$(38) \qquad V^{\lambda \tilde{c}} \geqslant V^{\lambda c} \Leftrightarrow V^{\tilde{c}} \geqslant V^{c}.$$

Throughout this subsection, we restrict ourselves to positive-valued utility processes. The function $f: \mathscr{C} \times [0, \infty) \to [0, \infty)$, where $\mathscr{C}$ is a cone, is *linearly homogeneous of degree one* (a definition) if, for any scalar $\lambda > 0$ and any $(c, v) \in \mathscr{C} \times (0, \infty)$, we have $f(\lambda c, \lambda v) = \lambda f(c, v)$. A variance multiplier A is *linearly homogeneous of degree* -1 if, for some k, $A(v) = k/v$ for all $v > 0$. Homothetic aggregators can be characterized as follows.

PROPOSITION 8 (Homotheticity): *An aggregator (f, m) that generates a recursive utility function U is homothetic if there exists an aggregator (f^*, m^*) generating an ordinally equivalent utility function such that f^* is linearly homogeneous of degree one and the variance multiplier A^* of m^* is linearly homogeneous of degree* -1.

PROOF: The linear homogeneity conditions on f^* and A^* imply that the utility process map $c \mapsto V^c$ generated by (f^*, m^*) is linearly homogeneous of degree one, which yields homotheticity.   *Q.E.D.*

We can sketch out an argument that the linear homogeneity conditions (on an ordinally equivalent aggregator) are also necessary for homotheticity. For necessity, we assume that (f, m) is homothetic, which implies that there is an ordinally equivalent aggregator (f^*, m^*) such that the resulting utility process map $c \mapsto V^c$ is linearly homogeneous of degree one; that is, for any scalar $\lambda > 0$ and consumption process c, $V^{\lambda c} = \lambda V^c$. For deterministic c in D, we have

$$(39) \qquad \frac{d}{dt} V_t^c = -f^*(c_t, V_t^c).$$

Applying the linear homogeneity of $c \mapsto V^c$ and (39) to λc produces the linear homogeneity of f. For general $c \in D$ and scalar $\lambda > 0$, if we let $\{\mu_t\}$ denote the drift process of V^c, then the homogeneity of $c \mapsto V^c$ implies that the drift of $V^{\lambda c}$ is $\lambda \mu_t$. Since $f^*(\lambda c_t, V_t^{\lambda c}) = f^*(\lambda c_t, \lambda V_t^c) = \lambda f^*(c_t, V_t^c)$, we have

$$(40) \qquad \lambda \mu_t - \lambda f^*(c_t, V_t^c) = \frac{1}{2} A^*(V_t^{\lambda c}) \frac{d}{dt} [V^{\lambda c}]_t = \lambda \frac{1}{2} A^*(V_t^c) \frac{d}{dt} [V^c]_t.$$

Since the quadratic variation of $V^{\lambda c} = \lambda V^c$ is merely λ^2 multiplied by the quadratic variation of V, we then have

$$(41) \qquad A^*(V_t^{\lambda c}) = A^*(\lambda V_t^c) = \frac{A^*(V_t^c)}{\lambda}.$$

This implies the desired homogeneity of A^*, at least in the range of the utility processes.

6. THE BELLMAN EQUATION

This section shows that optimality for a continuous-time recursive utility function is characterized by Bellman's equation, which, in this setting, is an extended version of the Hamilton-Jacobi-Bellman equation.

The setting chosen is a finite time horizon, a controlled state process $\{X_t\}$, and a regular normalized aggregator f. Let $\Gamma: \mathbb{R}^n \times \mathcal{T} \to 2^c$ define the admissible set of controls, in the sense that c_t must be chosen from the set $\Gamma(x, t)$ at time t when the current state X_t is $x \in \mathbb{R}^n$. For a given control $c \in D$, the state process $\{X_t\}$ exists if it uniquely solves the equation

$$(42) \qquad dX_t = b(X_t, t, c_t)\, dt + a(X_t, t, c_t)\, dB_t,$$

where X_0 is given and $b: \mathbb{R}^n \times \mathcal{T} \times \mathscr{C} \to \mathbb{R}^n$ and $a: \mathbb{R}^n \times \mathcal{T} \times \mathscr{C} \to \mathbb{R}^{n \times d}$ are measurable.

A process c in D is an *admissible control* if: (i) there is an integrable state process $\{X_t\}$ solving (42); (ii) for all t, $c_t \in \Gamma(X_t, t)$. Let D^Γ denote the set of admissible control processes, let U denote the recursive utility function on D generated by f, and consider the *control problem*

$$(43) \qquad \sup_{c \in D^\Gamma} U(c).$$

An admissible control c^* is optimal if $U(c^*) = \sup_{c \in D^\Gamma} U(c)$.

We will show a version of the sufficiency of the Bellman equation for optimality. First, for any $\bar{c} \in \mathscr{C}$ and any $J \in C^{2,1}(\mathbb{R}^n \times \mathcal{T})$, let

$$(44) \qquad \mathscr{D}^{\bar{c}} J(x, t) = J_t(x, t) + J_x(x, t) b(x, t, \bar{c})$$
$$+ \tfrac{1}{2} \operatorname{tr}\left[a(x, t, \bar{c}) a(x, t, \bar{c})^\top J_{xx}(x, t) \right].$$

We define $J \in C^{2,1}(\mathbb{R}^n \times \mathcal{T})$ to be the *value function* of the control problem, when there exists an optimal control c, if $V_t^c = J(X_t^c, t)$, $t \in [0, T]$, where V^c is the utility process for c and X^c is the state process determined by c. By our consistency result, Proposition 4, if c and $\tilde{c}$ are both optimal controls, then $V^c = V^{\tilde{c}}$, so we may refer to V^c as the value process.

Despite the strength of the technical conditions in the following characterization of the Bellman equation as a sufficient condition for optimality, the basic nature of the result is quite natural.

PROPOSITION 9 (Bellman's Equation): *Suppose* $J \in C^{2,1}(\mathbb{R}^n \times \mathcal{T})$ *has a bounded derivative* J_x *and that, for all* $(x, t) \in \mathbb{R}^n \times \mathcal{T}$,

$$(45) \qquad \sup_{\bar{c} \in \Gamma(x, t)} \mathscr{D}^{\bar{c}} J(x, t) + f[\bar{c}, J(x, t)] = 0,$$

with boundary condition $J(x, T) = 0$, $x \in \mathbb{R}^n$. *Suppose, moreover, there exists a measurable function* $C: \mathbb{R}^n \times \mathcal{T} \to \mathscr{C}$ *such that, for all* $(x, t) \in \mathbb{R}^n \times \mathcal{T}$,

$$(46) \qquad C(x, t) \in \operatorname*{arg\,max}_{\bar{c} \in \Gamma(x, t)} \mathscr{D}^{\bar{c}} J(x, t) + f[\bar{c}, J(x, t)].$$

Finally, suppose that X^* is the unique square-integrable process solving the stochastic differential equation

$$(47) \qquad dX_t^* = b[X_t^*, t, C(X^*, t)] \, dt + a[X_t^*, t, C(X_t^*, t)] \, dB_t, \qquad X_0^* = X_0.$$

Let c^* be the process defined by $c_t^* = C(X_t^*, t)$, $t \in [0, T]$. If c^* is a square-integrable process, then J is the value function of the control problem and c^* is an optimal control.

REMARK: The bound on J_x can be relaxed; we need only that $\int J_x(X_t, t) a(X_t, t, c_t) \, dB_t$ be a martingale. A key technical issue is the strong (pathwise unique) existence of a solution X^* to (47). It is enough that $x \mapsto b(x, t, C(x, t))$ and $x \mapsto a(x, t, C(x, t))$ satisfy a uniform Lipschitz condition, locally in x, and a uniform growth condition. This is difficult to verify since there is little advance information about the feedback control C. The same issues arise in the traditional additive setting; for one result of this type, see Fleming and Rishel (1975, p. 197). Alternatively, one may look for a "weak" solution to (47), as in Krylov (1980, p. 87). Likewise, it is difficult to know in advance the differentiability of J. On this, see, for example, Krylov (1980, Chapter 4), Lions (1981), and Crandall and Lions (1983).

PROOF: Let c be any admissible control, and let V be the utility process for c under f. Let X^c denote the solution of the state process equation (42) for control c. By Ito's Lemma, since J_x is bounded and $J(X_T^c, T) = 0$,

$$(48) \qquad J(X_t^c, t) = E\left[\int_t^T -\mathscr{D}^{c(s)} J(X_s^c, s) \, ds \,\middle|\, \mathscr{F}_t\right], \quad t \in [0, T].$$

We also have

$$V_t = E\left[\int_t^T f(c_s, V_s) \, ds \,\middle|\, \mathscr{F}_t\right], \quad t \in [0, T].$$

The Bellman equation (45) implies that, for all t,

$$J(X_t^c, t) - V_t = E\left[\int_t^T [-\mathscr{D}^{c(s)} J(X_s^c, s) - f(c_s, V_s)] \, ds \,\middle|\, \mathscr{F}_t\right],$$

where

$$-\mathscr{D}^{c(s)} J(X_s^c, s) - f(c_s, V_s) = -\mathscr{D}^{c(s)} J(X_s^c, s) - f[c_s, J(X_s^c, s)]$$

$$+ f[c_s, J(X_s^c)] - f(c_s, V_s)$$

$$\geq -\mathscr{D}^{c(s)} J(X_s^c, s) - f[c_s, J(X_s^c, s)]$$

$$- k|J(X_s^c, s) - V_s|.$$

Since $-\mathscr{D}^{c(s)}J(X_s^c, s) - f[c_s, J(X_s^c, s)]$ is nonnegative, Appendix Lemma B2 implies that $U(c) = V_0^c \leqslant J(x_0, 0)$, where x_0 is the starting value of the controlled state process. If we take the particular case of $c = c^*$, (46) implies that $U(c^*) = J(x_0, 0)$. Since c and x_0 are arbitrary, c^* is therefore optimal.   *Q.E.D.*

It may often be convenient to begin with an un-normalized aggregator (f, m), for example, for the purpose of comparative risk aversion analysis (Proposition 6). Thus we note the following generalization of (45):

$$(49) \qquad \sup_{\bar{c} \in \Gamma(x,t)} \mathscr{D}^{\bar{c}}J(x, t) + f[\bar{c}, J(x, t)]$$

$$+ \tfrac{1}{2}A[J(x, t)] \| J_x(x, t) a(x, \bar{c}, t) \|^2 = 0.$$

Provided the normalized aggregator is regular, this extension of the Bellman equation (45) follows from Ito's Lemma and (22).

Trudinger (1983) has a limited result on the existence of solutions to the Bellman equation in a related setting. Duffie and Lions (1990) remark how the theory of viscosity solutions due to Crandall and Lions (1983) extends naturally to the case of stochastic differential utility.

7. CONCLUDING REMARKS

This section briefly discusses some extensions of the model.

7.1. *Mixed Poisson-Brownian Information*

Suppose that $\mathbb{F}$ is the filtration generated by a standard Brownian motion B in $\mathbb{R}^d$ and an independent martingale compensated point process N with predictable jump intensity λ, and unit size jumps. (For example, N could be a compensated Poisson process, meaning that λ is a constant.) In this case we conjecture that a utility process V is of the form

$$dV_t = \mu_t \, dt + \sigma_t \, dB_t + G_t \, dN_t,$$

for some μ, σ, and G. Following by analogy the arguments in the pure Brownian case,

$$(50) \qquad \mu_t = -f(c_t, V_t) - \tfrac{1}{2}A(V_t)\sigma_t^2 - \lambda_t[M(V_{t-} + G_t, V_{t-}) - M(V_{t-}, V_{t-})].$$

We leave further analysis of recursive utility in this more general information environment for separate research, which should help elucidate the potential importance of static non-expected-utility theories for continuous-time modeling. (See Examples 4 and 5 of Section 4.)

7.2. *Habit Formation*

For a model combining the features of recursive utility with those of "habit formation," in the sense of Constantinides (1990), we could define V to be the

utility process for $c \in D$ if V is the unique solution to

$$(51) \qquad V_t = E\left[\int_t^T f(c_s, z_s, V_s)\, ds \,\bigg|\, \mathscr{F}_t\right],$$

where z_t is a "weighted average" of $\{c_s: 0 \leqslant s \leqslant t\}$, say $z_t = \int_0^t g(t-s)c_s\, ds$, for some continuous function $g: \mathbb{R}_+ \to \mathbb{R}$ such as $t \mapsto e^{-\beta t}$. See also Sundaresan (1989), Detemple and Zapatero (1989), as well as Heaton (1988).

Alternatively, we could connect with the results of Hindy and Huang (1989a, b) by defining V as the unique solution to

$$(52) \qquad V_t = E\left[\int_t^T f(z_s, V_s)\, ds \,\bigg|\, \mathscr{F}_t\right],$$

where $z_t = \int_0^t g(t-s)\, dC_s$, and where C is a cumulative-to-date total consumption process drawn from the set of increasing, nonnegative, right-continuous adapted processes satisfying

$$\|C\| = \left(E\left[\int_0^T C(t)^2\, dt + C_T^2\right]\right)^{1/2} < \infty.$$

In this case, the utility function automatically inherits the continuity properties with respect to intertemporal substitution as well as the duality results that are studied by Hindy and Huang (1989a, b). This example is further extended in Duffie and Skiadas (1990).

For either (51) or (52), existence and uniqueness of V follow immediately from Appendix Proposition A1, provided only that f is measurable, satisfies a growth condition with respect to the consumption-related variables (that is, with respect to (c, z) for (51), or with respect to z for (52)), and satisfies a Lipschitz condition with respect to utility. The Bellman equation corresponding to (51) is the obvious extension of (45) obtained by adjoining the "habit" process $\{z_t\}$ to the state description.

7.3. Utility with a Kinked Certainty Equivalent

In principle, it is not necessary to have a smooth certainty equivalent $m(\cdot)$ in order to characterize a recursive utility process $\{V_t = V_0 + \int_0^t \mu_s\, ds + \int_0^t \sigma_s\, dB_s\}$. For example, an extension of Ito's Lemma for convex (and not necessarily C^2) functions gives a definition of the utility process, one that replaces the quadratic "risk penalty" $\frac{1}{2}A(V_t)\|\sigma_t\|^2$ in (17) with a natural "local time risk penalty."

Take for instance the Dekel-Chew betweenness certainty equivalent m defined by (32) with the special "kinked" case of $H(x, y) = \varphi(x/y)$, where $\varphi(x) = x - 1$ for $x \leqslant 1$ and $\varphi(x) = \gamma(x - 1)$ for $x > 1$, where $\gamma \in (0, 1)$. The motivation for examining this sort of extension may be apparent from the discussion in Example 5 of Section 4 of "first-order" local risk aversion. Putting aside existence and uniqueness questions, the following calculations characterize the utility process V for a consumption process c under (f, m), for arbitrary well-behaved f.

We suppose that $V_t = K_t + \int_0^t \sigma_s \, dB_s$ for some finite variation process K and progressively measurable σ with $E(\int_0^T \|\sigma_t\|^2 \, dt) < \infty$. We let $X_t = V_t/Z_t$, $t \geqslant 0$, where $Z_t = m(\sim V_t)$. By Ito's Lemma for convex functions (Karatzas and Shreve (1989)),

$$(53) \qquad \varphi(X_t) = \varphi(X_0) + \int_0^t D^- \varphi(X_s) \, dX_s + \int_{-\infty}^{\infty} \Lambda_t^X(a) \, d\nu_\varphi(a),$$

where $D^- \varphi$ is the left-derivative of φ, $\Lambda_t^X(x)$ is the local time of X at x up to time t, and ν_φ is the measure on $\mathbb{R}$ defined by $\nu_\varphi([a, b)) = D^- \varphi(b) - D^- \varphi(a)$. In our case, ν_φ is a dirac measure of mass $\gamma - 1$ at 1. The definition $E[\varphi(V_t/Z_t)] = 0$ of Z_t and the mild assumption that Z is differentiable in t imply, taking expectations through (53), that

$$0 = 0 + E\left(\int_0^t D^- \varphi(X_s) \left[\frac{1}{Z_s} \, dK_s - \frac{V_s \dot{Z}_s}{Z_s^2} \, ds \right] \right) + E\left[(\gamma - 1) \Lambda_t^X(1) \right].$$

More generally, for each t let $X_s^t = V_s/Z_s^t$, where $Z_s^t = m(\sim V_s | \mathscr{F}_t)$, $s \geqslant t$, and let l^t be the local time process of X^t. Reinterpreting the last displayed equation at time t, conditional on $\mathscr{F}_t$, leaves

$$\frac{1}{Z_t^t} \, dK_t - \frac{V_t \dot{Z}_t^t}{Z_t^t} \, dt + (\gamma - 1) \, dL_t = 0,$$

where $dL_t = dl_t^t$. Since $Z_t^t = m(\delta_{V(t)} | \mathscr{F}_t) = V_t$, we would have $dK_t = \dot{Z}_t^t \, dt + (\gamma - 1) V_t \, dL_t$. We put aside the issue of the existence of a process L whose increments are defined by "sewing together" the increments[7] of l^t at time t, t by t, and continue on a purely formal basis. With $f(c_t, V_t) = -\dot{Z}_t^t$ from (12), we have

$$(54) \qquad dK_t = (1 - \gamma) V_t \, dL_t - f(c_t, V_t) \, dt,$$

which replaces the "risk penalty" $\frac{1}{2} A(V_t) \|\sigma_t\|^2 \, dt$ in (17) with the "local-time risk penalty" $(\gamma - 1) V_t \, dL_t$. We have loosely characterized the utility process V as the unique solution to

$$(55) \qquad V_t = E\left(\int_t^T (\gamma - 1) V_s \, dL_s + f(c_s, V_s) \, ds \,\middle|\, \mathscr{F}_t \right), \quad t \in [0, T].$$

The measurement of risk by local time can also be viewed in terms of risk measured by "local mean absolute deviation" with the aid of the Tanaka-Meyer formula (Karatzas and Shreve (1989, p. 220)). This has been a purely formal analysis, without rigorous justification.

---

[7] One approach would be to construct L from $\{l^t : t \geqslant 0\}$ by taking a sequence $\{t_n^i\}$ of time grids with mesh size $\sup_n |t_n^i - t_{n-1}^i|$ going to zero with i. Letting L^i be piecewise constant as defined by the jumps $\Delta L^i(t_{n+1}^i) = l^{t_n^i}(t_{n+1}^i) - l^{t_n^i}(t_n^i)$, the objective would be weak convergence of $\{L^i\}$ to a uniquely defined process L. This construction is well beyond our objectives here, especially as this is merely one part of the utility existence problem.

7.4. *Other Extensions*

We could allow almost arbitrary state and time dependence of a generalized aggregator f: $\mathscr{C} \times \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$, as well as a terminal reward. For existence and uniqueness of the utility process, we need only check the integrability and Lipschitz conditions of Appendix Proposition A1. Under natural assumptions, the properties of such extended recursive utility functions, even with the addition of "habit persistence," are easily studied with the aid of Appendix B, Lemma B2.

Finally, for the infinite time horizon case, see Appendix C, co-authored with Costis Skiadas.

Graduate School of Business, Stanford University, Stanford, CA 94305-5015, U.S.A.

and

Department of Economics, University of Toronto, 150 St. George St., Toronto, Canada M5S 1A1

APPENDIX A: FIXED POINTS OF RECURSIVE INTEGRAL EQUATIONS

Recall that $(\Omega, \mathscr{F}, \mathbb{F}, P)$ is our (abstract) filtered probability space, where $\mathbb{F} = \{\mathscr{F}_t: t \in [0, T]\}$ satisfies the usual conditions. Let $\mathscr{C}$ be a closed convex subset of a Banach lattice (which is usually $\mathbb{R}^l$, representing bundles of l commodities) with norm $\|\cdot\|$, and let D denote the space of $\mathscr{C}$-valued optional square-integrable processes. Let $\|\cdot\|_D$ be defined on D by

$$\|c\|_D = \left[E \left(\int_0^T \|c_t\|^2 \, dt \right) \right]^{1/2}, \quad c \in D.$$

For each given $c \in D$, we will establish the existence and uniqueness of a square-integrable semimartingale V^c satisfying the equation

(a1) $$V_t^c = E \left[\int_t^T f(c_s, V_s^c) \, ds \,\Big|\, \mathscr{F}_t \right], \quad t \in [0, T],$$

provided f: $\mathscr{C} \times \mathbb{R} \to \mathbb{R}$ is measurable, uniformly Lipschitz in its second (utility) argument, and satisfies a growth condition in its first (consumption) argument. (The Lipschitz condition can be weakened.)

First, for $1 \leqslant p \leqslant \infty$ and each semimartingale X, let

$$\|X\|_{\mathscr{S}^p} = \left\| \sup_t |X_t| \right\|_{L^p}.$$

Let $\mathscr{S}^p$ denote the space of semimartingales (always taken to be càdlàg) that are finite in $\|\cdot\|_{\mathscr{S}^p}$-norm. This space is shown in Protter (1990) to have a convenient relationship with the usual space $\mathscr{H}^p$ of p-integrable semimartingales, especially for the purpose of establishing the existence and uniqueness of stochastic differential equations driven by semimartingales. (See Protter (1990, Theorem V.2 and Lemma V.1).)

Let g: $[0, T] \times \Omega \times \mathbb{R} \to \mathbb{R}$ be measurable and such that, for each t and v, $g_t(v) \equiv g(t, \cdot, v)$: $\Omega \to \mathbb{R}$ is $\mathscr{F}_t$-measurable. We usually have in mind $g_t(v) = f(c_t, v)$ for some $c \in D$. We will use the

regularity conditions:

    (i) g is k-lipschitz: The scalar k is such that, for any $t \in [0, T]$ and $(w, v) \in \mathbb{R}^2$,

$$|g_t(v) - g_t(w)| \leqslant k|v - w|.$$

    (ii) g is p-integrable: The scalar $p \in [1, \infty]$ is such that $\int_0^T |g_t(0)| \, dt$ is in L^p.

    LEMMA A1: *Suppose there is some $p \in [1, \infty)$ and constant k such that g is k-lipschitz and p-integrable. Then, for any $\mathscr{F}_T$-measurable Y in L^p and any $V \in \mathscr{S}^p$, the process $F(V)$ defined by*

$$(\text{a2}) \qquad F(V)_t = E\left[\int_t^T g_s(V_s) \, ds + Y \, \middle| \, \mathscr{F}_t \right], \quad t \in [0, T],$$

is also in $\mathscr{S}^p$.

    PROOF: By the Lipschitz assumption,

$$\sup_t |F(V)_t| \leqslant \sup_t E\left[\int_t^T |g_s(V_s)| \, ds + |Y| \, \middle| \, \mathscr{F}_t \right]$$

$$\leqslant \sup_t E\left[\int_t^T (|g_s(0)| + k|V_s|) \, ds + |Y| \, \middle| \, \mathscr{F}_t \right]$$

$$\leqslant kT \sup_t |V_t| + \sup_t E\left[\int_t^T |g_s(0)| \, ds + |Y| \, \middle| \, \mathscr{F}_t \right].$$

The L^p norm of the left-hand side is finite because the L^p norm of the first term on the right is finite by assumption, and because the L^p norm of the second term on the right is finite by Theorem V.2 of Protter (1990) and the definition of the $\mathscr{H}^p$-norm (Dellacherie and Meyer (1982)).   Q.E.D.

    The main result of this appendix is the following fixed point theorem for recursive integral equations. The method of proof is similar to that for stochastic differential equations, as in Chung and Williams (1990).

    PROPOSITION A1: *Suppose g is k-lipschitz for some k and p-integrable for some $p \in (1, \infty)$. Then F: $\mathscr{S}^p \to \mathscr{S}^p$ defined by (a2) has a unique fixed point (that is, a unique V in $\mathscr{S}^p$ such that $F(V) = V$).*

    PROOF: For any U and V in $\mathscr{S}^p$,

$$|F(U)_t - F(V)_t| \leqslant E\left[\int_t^T |g_s(U_s) - g_s(V_s)| \, ds \, \middle| \, \mathscr{F}_t \right]$$

$$\leqslant k E\left[\int_t^T |U_s - V_s| \, ds \, \middle| \, \mathscr{F}_t \right]$$

$$\leqslant k(T - t) E\left[\sup_{0 \leqslant s \leqslant T} |U_s - V_s| \, \middle| \, \mathscr{F}_t \right].$$

Let[8] $Z \equiv \sup_{0 \leqslant s \leqslant T} |U_s - V_s|$ and $Z_t \equiv E(Z \mid \mathscr{F}_t)$. Since $|F(U)_t - F(V)_t| \leqslant k(T - t)Z_t$, we have, using Fubini's Theorem for conditional expectation,

$$|F^{(2)}(U)_t - F^{(2)}(V)_t| \leqslant E\left[\int_t^T k |F(U)_s - F(V)_s| \, ds \, \middle| \, \mathscr{F}_t \right]$$

$$\leqslant k^2 E\left[\int_t^T (T - s) Z_s \, ds \, \middle| \, \mathscr{F}_t \right]$$

$$= k^2 \int_t^T (T - s) E(Z_s \mid \mathscr{F}_t) \, ds$$

$$= \frac{[k(T - t)]^2}{2!} Z_t.$$

---

[8] We could also take $Z = \int_t^T |U_s - V_s| \, ds$.

In general,

$$\left| F^{(n)}(U)_t - F^{(n)}(V)_t \right| \leqslant \frac{[k(T-t)]^n}{n!} Z_t,$$

$$\left\| \sup_t \left| F^{(n)}(U)_t - F^{(n)}(V)_t \right| \right\|_{L^p} \leqslant \frac{(kT)^n}{n!} \left\| \sup_t Z_t \right\|_{L^p}.$$

But, Doob's inequality gives (for $p > 1$ and $(1/p) + (1/q) = 1$)

$$\left\| \sup_t Z_t \right\|_{L^p} \leqslant q \sup_t \| Z_t \|_{L^p}$$

and

$$\| Z_t \|_{L^p}^p = E\left(E(Z \mid \mathscr{F}_t)^p \right) \leqslant E\left[E(Z^p \mid \mathscr{F}_t) \right] = E(Z^p).$$

Therefore

$$\| Z_t \|_{L^p} \leqslant \| Z \|_{L^p} = \| U - V \|_{\mathscr{S}^p}.$$

We can conclude that

$$\left\| F^{(n)}(U) - F^{(n)}(V) \right\|_{\mathscr{S}^p} \leqslant \frac{q(kT)^n}{n!} \| U - V \|_{\mathscr{S}^p}.$$

So, for n large enough, $F^{(n)}$ is a contraction and hence has a unique fixed point V. Since $F^n(V) = V$, we have $F^n(FV) = (FV)$, and by uniqueness, $V = F(V)$. Finally, F has a unique fixed point since $F^{(n)}$ does.   Q.E.D.

The assumption that f is Lipschitz in utility and satisfies a growth condition in consumption is enough to meet the sufficient conditions of Proposition A1 for $g_t(v) = f(c_t, v)$ and $p = 2$, which proves Theorem 1.

Recently and independently, Pardoux and Peng (1990) have given a related result which is more general than Proposition A1 in the special case of Brownian filtrations. Their method of proof is quite different and relies on the martingale representation property of Brownian filtrations.

APPENDIX B: Extensions of Gronwall's Inequality

Except for homotheticity, each of the properties in Section 5, as well as Proposition 9 (Bellman's equation) is shown by appealing to stochastic extensions of Gronwall's inequality. The first is an extension of what Fleming and Rishel (1975) call the Gronwall-Bellman inequality.

THE GRONWALL-BELLMAN INEQUALITY: *Suppose $h: [0, T] \to \mathbb{R}$ is continuous, α is a constant, and $g: [0, T] \to \mathbb{R}$ is integrable. If $h_t \leqslant \int_t^T (g_s + \alpha h_s) \, ds + h_T$ for all $t \in [0, T]$ then, for any t, $h_t \leqslant e^{\alpha(T-t)} h_T + \int_t^T e^{\alpha(s-t)} g_s \, ds$. If $h_t \geqslant \int_t^T (g_s + \alpha h_s) \, ds + h_T$, then $h_t \geqslant e^{\alpha(T-t)} h_T + \int_t^T e^{\alpha(s-t)} g_s \, ds$.*

PROOF: For the first part of the result, let $H_t = \int_t^T (g_s + \alpha H_s) \, ds + h_T$. We know that $H_t = e^{\alpha(T-t)} h_T + \int_t^T e^{\alpha(s-t)} g_s \, ds$. Now let $p_t = h_t - H_t$, implying $p_T = 0$. Let $G_t = \int_t^T p_s \, ds$. Then $\dot{G}_t = -p_t \geqslant -\alpha G_t$, which implies that $d/dt(e^{\alpha t} G_t) \geqslant 0$, which implies that $G_t \leqslant 0$. This implies that $p_t \leqslant 0$ for all t, for otherwise there is an interval $[t_0, t_1] \subset [0, T]$ such that $p(t_1) = 0$ and $p_t > 0$ on (t_0, t_1). We could then apply the previous arguments on $[0, t_1]$, however, and deduce that $\int_t^{t_1} p_s \, ds \leqslant 0$ for $t \leqslant t_1$, which would be a contradiction. Thus $p_t \leqslant 0$ for all t, and we are done. A like argument shows the opposite inequality.   Q.E.D.

COROLLARY B1 (The Stochastic Gronwall-Bellman Inequality): *Let $(\Omega, \mathscr{F}, \mathbb{F}, P)$ be a filtered probability space whose filtration $\mathbb{F} = \{\mathscr{F}_t : t \in [0, T]\}$ satisfies the usual conditions. Suppose $\{Y_s\}$ and $\{X_s\}$ are optional integrable processes and α is a constant. Suppose, for all t, that $s \mapsto E(Y_s \mid \mathscr{F}_t)$ is*

continuous almost surely. If, for all t, $Y_t \leqslant E(\int_t^T (X_s + \alpha Y_s) \, ds \mid \mathscr{F}_t) + Y_T$, *then, for all* t,

$$Y_t \leqslant e^{\alpha(T-t)} E(Y_T \mid \mathscr{F}_t) + E\left[\int_t^T e^{\alpha(s-t)} X_s \, ds \,\middle|\, \mathscr{F}_t\right] \qquad a.s.$$

Alternatively, if, for all t, $Y_t \geqslant E(\int_t^T (X_s + \alpha Y_s) \, ds \mid \mathscr{F}_t) + Y_T$, *then, for all* t,

$$Y_t \geqslant e^{\alpha(T-t)} E(Y_T \mid \mathscr{F}_t) + E\left[\int_t^T e^{\alpha(s-t)} X_s \, ds \,\middle|\, \mathscr{F}_t\right] \qquad a.s.$$

PROOF: For each t and $s \in [t, T]$, let $h_{t,s} = E(Y_s \mid \mathscr{F}_t)$ and $g_{t,s} = E(X_s \mid \mathscr{F}_t)$. We can then apply Fubini's Theorem for conditional expectations, as in Proposition 4.6, page 74, of Ethier and Kurtz (1987), and the law of iterated expectations, followed by the Gronwall-Bellman inequality applied to the interval $[t, T]$, in order to deduce the result. $\hfill$ Q.E.D.

The stochastic Gronwall-Bellman inequality was the basis for most of the proofs in the first draft of the paper. Costis Skiadas then provided the authors with the following consequence of Gronwall's inequality, which has allowed us to simplify the structure of most of the proofs in Section 5.

LEMMA B2: *Let* $(\Omega, \mathscr{F}, \mathbb{F}, P)$ *be a filtered probability space whose filtration* $\mathbb{F} = \{\mathscr{F}_t: t \in [0, T]\}$ *satisfies the usual conditions. Suppose* $\{X_s\}$ *and* $\{Y_s\}$ *are integrable optional processes,* α *is a constant, and* $\{G_s\}$ *is a measurable process. Suppose, for all* t, *that* $s \mapsto Y_s$ *is right continuous and* $s \mapsto E(Y_s \mid \mathscr{F}_t)$ *is continuous almost surely. If* $Y_T \geqslant 0$ *a.s. and, for all* t, $G_t \geqslant -\alpha |Y_t|$ *a.s. and* $Y_t = E[\int_t^T G_s \, ds + Y_T \mid \mathscr{F}_t]$ *a.s. then, for all* t, $Y_t \geqslant 0$ *a.s.*

PROOF: Fix $t \in [0, T]$ and define the stopping time $\tau = \inf\{s \geqslant t: Y_s \geqslant 0\}$. Since $Y_T \geqslant 0$ a.s., $P\{t \leqslant \tau \leqslant T\} = 1$. Using Doob's Optional Sampling Theorem, it is easy to show that, for all $u \in [t, T]$, we have $Y_u = E[\int_u^\tau G_s \, ds + Y_\tau \mid \mathscr{F}_u]$ whenever $\tau > u$, written as

$$Y_u 1_{\{\tau > u\}} = E\left[\int_u^\tau G_s 1_{\{\tau > u\}} \, ds + Y_\tau 1_{\{\tau > u\}} \,\middle|\, \mathscr{F}_u\right] \qquad a.s.$$

By right continuity of Y, $Y_\tau \geqslant 0$ a.s. Also, for given s and on $\{\omega : t \leqslant s < \tau(\omega)\}$, we have

$$G_s \geqslant -\alpha |Y_s| = \alpha Y_s \quad a.s.$$

Therefore, for all $u \in [t, T]$,

$$Y_u 1_{\{\tau > u\}} \geqslant E\left[\int_u^\tau \alpha Y_s 1_{\{\tau > u\}} \, ds \,\middle|\, \mathscr{F}_u\right] = E\left[\int_u^T \alpha Y_s 1_{\{\tau > s\}} \, ds \,\middle|\, \mathscr{F}_u\right].$$

By the stochastic Gronwall-Bellman inequality, applied on $[t, T]$, $Y_t 1_{\{\tau > t\}} \geqslant 0$ a.s. Since $Y_t 1_{\{\tau = t\}} \geqslant 0$ a.s. by the definition of τ, it follows that $Y_t \geqslant 0$ a.s. $\hfill$ Q.E.D.

APPENDIX C: THE INFINITE HORIZON CASE
WITH COSTIS SKIADAS

This appendix[9] addresses the definition, existence, and properties of an infinite horizon stochastic differential utility function. The issue of existence is analogous to that of stability of a nonlinear feedback system. Stability is guaranteed by imposing a "uniform sector condition" on the "feedback function." The basic properties of the finite horizon recursive utility presented in the body of the paper generalize directly to the infinite horizon case. For the finite time horizon T, given a consumption process c, and under regularity conditions on f, the main body of the paper shows the existence of a unique integrable semimartingale, V^T, the recursive utility process corresponding to

---

[9] A version of this appendix appeared originally as a short paper by Duffie, Epstein, and Skiadas.

c, defined by

(c1) $V_t^T = E\left[\int_t^T f(c_s, V_s^T)\, ds \,\Big|\, \mathscr{F}_t\right]$ a.s., $t \in [0, T]$.

A utility function, U^T, is then defined by letting $U^T(c) = V_0^T$. In this appendix it is shown that, under regularity conditions on f and c, the finite horizon recursive utility process corresponding to c converges to an integrable semimartingale as the horizon length goes to infinity:

(c2) $V_t = \lim_{T \to \infty} V_t^T$ a.s.

Furthermore, V_t satisfies, for all t and $T \geqslant t$,

(c3) $V_t = E\left[\int_t^T f(c_s, V_s)\, ds + V_T \,\Big|\, \mathscr{F}_t\right]$ a.s.

It is also shown that V is the unique integrable semimartingale satisfying (c3) and a transversality condition of the form:

(c4) $\lim_{t \to \infty} e^{-\nu t} E(|V_t|) = 0$,

for a suitable constant ν. We call V the infinite horizon recursive utility process corresponding to c. An infinite horizon stochastic differential utility function, U, is then defined by letting $U(c) = V_0$, and is shown to possess all of the elementary properties of U^T discussed in the body of the paper. Duffie and Lions (1990) show existence of infinite-horizon stochastic differential utility by partial differential equation techniques in a Markov diffusion setting, admitting some weakening of the conditions below on f.

We now proceed with the formal details. For simplicity we will omit the a.s. (almost sure) qualification whenever it obviously applies. The basic primitive of the model is a filtered probability space $(\Omega, \mathscr{F}, \mathbb{F}, P)$, where the filtration $\mathbb{F} = \{\mathscr{F}_t : t \geqslant 0\}$ is assumed to satisfy the usual conditions. It is also assumed that $\mathscr{F}_0$ is trivial, that is, it only contains events of probability one or zero. Consumption processes are valued in a closed convex subset, $\mathscr{C}$, of some separable Banach lattice. The reader may choose to think of $\mathscr{C}$ as a finite dimensional Euclidean space or its positive cone. For fixed $\nu \in \mathbb{R}$, $\mathscr{D}_\nu$ is defined to be the space of all optional, $\mathscr{C}$-valued processes, $c = (c_t; t \geqslant 0)$, such that $E(\int_0^\infty e^{-\nu t} \|c_t\|^2\, dt) < \infty$, and is equipped with the norm $|c|_{\mathscr{D}_\nu} = [E(\int_0^\infty e^{-\nu t}\|c_t\|^2\, dt)]^{1/2}$. Finally, for any horizon length $T < \infty$, $\mathscr{D}[0, T]$ is defined to be the space of all optional, $\mathscr{C}$-valued processes, c, such that $E(\int_0^T \|c_t\|^2\, dt) < \infty$ and $c_t = 0$ for all $t \geqslant T$. Notice that $\mathscr{D}[0, T] \subset \mathscr{D}_\nu$ for all choices of $T < \infty$ and ν. Section 3 defined a recursive utility U^T over $\mathscr{D}[0, T]$. Here, the definition is extended to all of $\mathscr{D}_\nu$ for an appropriate ν.

Assumptions

The issue of existence of the infinite horizon recursive utility has the flavor of that of stability of nonlinear feedback systems. We can view f as a nonlinear feedback function. In control theory literature, f is often required to satisfy a "sector condition." (See for example Vidyasagar (1978).) Here we will employ a more stringent "uniform sector condition," which can also be viewed as a generalized Lipschitz condition.

In the sequel $f: \mathscr{C} \times \mathbb{R} \to \mathbb{R}$ will be taken to be a measurable function satisfying the following regularity conditions:

ASSUMPTION 1: *f satisfies a growth condition in consumption. That is, for some constants* k_1 *and* k_2, *and all* $c \in \mathscr{C}$, $|f(c, 0)| \leqslant k_1 + k_2\|c\|$.

ASSUMPTION 2: *f satisfies the uniform sector condition in utility: For some constants* ν *and* k *with* $\nu < k$, *and all* $c \in \mathscr{C}$,

$$-k \leqslant \frac{f(c, \bar{v}) - f(c, v)}{\bar{v} - v} \leqslant -\nu, \qquad \bar{v} \neq v, \qquad \bar{v}, v \in \mathbb{R}.$$

For $\nu = -k < 0$, Assumption 2 is equivalent to a Lipschitz condition. The notation reflects the fact that, in general, $\nu \geqslant 0$. This is not, however, a condition required by our theory.

Existence and Uniqueness

Section 3 showed that if f satisfies Assumptions 1 and 2, with $\nu = -k$, then for any $c \in \mathscr{D}[0,T]$ there is a unique integrable semimartingale[10] V^T that satisfies the recursive relation $(c1)$ for $t \in [0,T]$ and such that $V_t^T = 0$ for $t \geqslant T$. The finite horizon recursive utility, $U^T: \mathscr{D}[0,T] \to \mathbb{R}$, is then defined by letting $U^T(c) = V_0^T$. The following theorem allows us to extend this definition.

THEOREM C1: *Under Assumptions 1 and 2 and for all $c \in \mathscr{D}_\nu$, the limit in $(c2)$ exists and satisfies $(c3)$. Furthermore, $(c2)$ defines the unique integrable semimartingale V satisfying $(c3)$ and the transversality condition $(c4)$.*

Assumption 2 and the requirement that $c \in \mathscr{D}_\nu$ play antagonistic roles. We want to find the largest possible ν for which Assumption 2 holds, in order to allow the largest possible class of consumption processes implied by the theorem.

The infinite horizon recursive utility, $U: \mathscr{D}_\nu \to \mathbb{R}$, can now be defined by setting $U(c) = V_0$. Notice that, for $0 \in \mathscr{C}$ and $f(0,0) = 0$, U^T is the restriction of U to $\mathscr{D}[0,T]$. Before proving the theorem we give two lemmas, proved at the end of this appendix, which will be useful in subsequent proofs. Until the end of the proof of Theorem 1, we fix some $c \in \mathscr{D}_\nu$.

LEMMA C1: *Suppose V is an integrable semimartingale satisfying $(c3)$ for some $T > 0$. Then, for all stopping times τ_1, τ_2 bounded by T,*

$$V_{\tau_1} = E\left[\int_{\tau_1}^{\tau_2} f(c_s, V_s)\, ds + V_{\tau_2} \,\bigg|\, \mathscr{F}_{\tau_1}\right], \qquad on \quad \{\tau_1 < \tau_2\}.$$

LEMMA C2: *There exist constants K_1 and K_2 such that for all $t \leqslant T$,*

$$\left|V_t^T\right| \leqslant K_1 + K_2 e^{\nu t} E\left[\int_t^T \|c_s\| e^{-\nu s}\, ds \,\bigg|\, \mathscr{F}_t\right].$$

PROOF OF THEOREM 1: *Existence:* We show that the limit in $(c1)$ exists by proving that, for each time $t \geqslant 0$, the sequence $\{V_t^n: n \geqslant 0\}$ is Cauchy, almost everywhere on Ω. Suppose $n, m \geqslant N$, fix t, and define the stopping times: $\tau = \inf\{s: s \geqslant t, V_s^n \leqslant V_s^m\}$ and $\tau_N = \tau \wedge N$. Lemma C1 and Assumption 2 imply, for all $u \in [t,T]$,

$$(V_u^n - V_u^m) 1_{\{\tau > u\}} = E\left[\int_u^{\tau_N}\left(f(c_s, V_s^n) - f(c_s, V_s^m)\right) 1_{\{\tau > u\}}\, ds \,\bigg|\, \mathscr{F}_u\right]$$

$$+ E\left[\left(V_{\tau_N}^n - V_{\tau_N}^m\right) 1_{\{\tau > u\}} \,\big|\, \mathscr{F}_u\right]$$

$$\leqslant E\left[\int_u^{\tau_N} -\nu(V_s^n - V_s^m) 1_{\{\tau > u\}}\, ds + (V_N^n - V_N^m) 1_{\{\tau > N\}} \,\bigg|\, \mathscr{F}_u\right]$$

$$= E\left[\int_u^N -\nu(V_s^n - V_s^m) 1_{\{\tau > s\}}\, ds + (V_N^n - V_N^m) 1_{\{\tau > N\}} \,\bigg|\, \mathscr{F}_u\right].$$

By the stochastic Gronwall-Bellman inequality,

$$(V_t^n - V_t^m)^+ \leqslant e^{-\nu(N-t)} E\left[(V_N^n - V_N^m) 1_{\{\tau > N\}} \,\big|\, \mathscr{F}_t\right]$$

$$\leqslant e^{-\nu(N-t)} E\left[|V_N^n| + |V_N^m| \,\big|\, \mathscr{F}_t\right].$$

Because of symmetry, $|V_t^n - V_t^m| \leqslant 2 e^{-\nu(N-t)} E[|V_N^n| + |V_N^m| \,|\, \mathscr{F}_t]$. Combining this result with Lemma C2 proves our assertion.

To show that V satisfies $(c3)$, we start with $V_t^n = E[\int_t^T f(c_s, V_s^n)\, ds + V_T^n \,|\, \mathscr{F}_t]$. Noting that $|f(c_s, V_s^n)| \leqslant |f(c_s, V_s^n) - f(c_s, 0)| + |f(c_s, 0)| \leqslant \max(k, -\nu)|V_s^n| + k_1 + k_2\|c_s\|$ and using domi-

---

[10] We identify semimartingales that are versions of each other.

nated convergence and Lemma 2, we can let $n \to \infty$ to derive (c3).

$\quad$ *Uniqueness:* Suppose $\hat{V}$ and V both satisfy (c3) and (c4). Arguing exactly as above, we find $|\hat{V}_t - V_t| \leqslant 2e^{-\nu(N-t)}E[|\hat{V}_N| + |V_N| \mid \mathscr{F}_t]$. Taking expectations on both sides and letting $N \to \infty$, we find $\hat{V}_t = V_t$. This completes the proof of Theorem 1. $\hspace{3cm}$ Q.E.D.

We conclude this section by discussing the representation:

$$(\text{c5}) \qquad V_t = E\left[\int_t^\infty f(c_s, V_s)\, ds \,\middle|\, \mathscr{F}_t\right].$$

Tempting as it is, such a representation is not always valid.

EXAMPLE: Let $f(c, v) \equiv 1 - v$. Then, clearly, Assumptions 1 and 2 are satisfied. For any c, the corresponding utilities are $V_t^T = 1 - e^{(t-T)}$ and $V_t = 1$. While (c2) and (c3) are satisfied, (c5) fails. One might suggest that (c5) be modified to

$$V_t = E\left[\int_t^\infty f(c_s, V_s)\, ds \,\middle|\, \mathscr{F}_t\right] + V_\infty,$$

for some random variable V_∞ (identically equal to 1 in our case). But then consider $\bar{f}(c, v) \equiv c - (1 + c^{-1})v$ and $c_t = 1 + t$. Again Assumptions 1 and 2 are satisfied while the corresponding utilities are $\bar{V}_t^T = (1 + t)V_t^T$ and $\bar{V}_t = (1 + t)V_t$, with V^T and V as above. Now (c5) fails even more seriously. However, we have the following positive result.

PROPOSITION: *Suppose $\nu \leqslant 0$, Assumptions 1 and 2 are satisfied, and $c \in \mathscr{D}_\nu$. Then (c3) and (c5) are equivalent.*

PROOF: Clearly, (c5) implies (c3). For the converse, we wish to let $T \to \infty$ and then apply the dominated convergence theorem as in the proof of the existence part of the above theorem. The reader can verify that Lemma 2 and Fubini's theorem yield the required integrability condition. It is here that the condition $\nu \leqslant 0$ is crucial. $\hspace{3cm}$ Q.E.D.

Properties

We conclude by briefly reviewing some properties of the infinite horizon stochastic differential utility. For this section, f is taken to satisfy Assumptions 1 and 2. Most of the basic properties are direct consequences of their finite horizon counterparts and equation (c2). Thus
- U is (strictly) increasing whenever f is (strictly) increasing in consumption.
- U is concave whenever f is concave.
- U is time consistent in the sense of Section 5.

The discussion on risk aversion and homotheticity in Section 5 also extends in an obvious manner to the infinite horizon case. A less obvious result is continuity relative to the norm $|\cdot|_{\mathscr{D}_\nu}$ on $\mathscr{D}_\nu$. Proof of the following is left to the end of the Appendix.

THEOREM C2 (Continuity): *The utility function $U: \mathscr{D}_\nu \to \mathbb{R}$ is continuous provided f is continuous.*

Finally, the discussion of the Bellman equation in Section 6 also generalizes to an infinite horizon. The only change is the terminal value condition for the value function which is now replaced by a transversality condition as in equation (c4). The reader who has read the proofs of this paper will have no trouble in applying the same approach in modifying the finite horizon argument.

$\quad$ An alternative approach to Bellman's equation for solving optimization problems in the presence of convexity is generalized Kuhn-Tucker theory. This requires the computation of the gradient of the utility function. Work in this direction is reported in Duffie and Skiadas (1990).

Remaining Proofs

PROOF OF LEMMA C1: Equation (c3) implies that

$$\int_0^t f(c_s, V_s)\, ds + V_t = E\left[\int_0^T f(c_s, V_s)\, ds + V_T \,\middle|\, \mathscr{F}_t\right],$$

which is a martingale. Doob's Optional Stopping Theorem then allows us to replace t by τ_1 or τ_2 in

the above equation and hence in (c3). Therefore,

$$V_{\tau_1} = E\left[\int_{\tau_1}^T f(c_s, V_s)\, ds + V_T \middle| \mathcal{F}_{\tau_1}\right],$$

and using the fact that $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$,

$$E\left[V_{\tau_2} \middle| \mathcal{F}_{\tau_1}\right] = E\left[\int_{\tau_2}^T f(c_s, V_s)\, ds + V_T \middle| \mathcal{F}_{\tau_1}\right].$$

Subtracting the last equation from the second to last, the result follows. Q.E.D.

PROOF OF LEMMA C2: Fix t and define the stopping times:

$$\tau^+ = \inf\left\{s\colon s \geq t,\, V_s^T \leq 0\right\} \quad \text{and} \quad \tau^- = \inf\left\{s\colon s \geq t,\, V_s^T \geq 0\right\},$$

with the convention that $\inf \varnothing = \infty$. Let τ represent either τ^+ or τ^-. Since $\tau \leq T$, Lemma C1 implies that, for all $u \in [t, T]$,

$$V_u^T 1_{\{\tau > u\}} = E\left[\left(\int_u^\tau f\left(c_s, V_s^T\right) ds + V_\tau^T\right) 1_{\{\tau > u\}} \middle| \mathcal{F}_u\right],$$

and since $V_{\tau^+}^T 1_{\{\tau^+ > u\}} \leq 0$,

$$V_u^T 1_{\{\tau^+ > u\}} \leq E\left[\int_u^T f\left(c_s, V_s^T\right) 1_{\{\tau^+ > s\}}\, ds \middle| \mathcal{F}_u\right].$$

The sector condition on f then implies that, for all $u \in [t, T]$,

$$V_u^T 1_{\{\tau^+ > u\}} \leq E\left[\int_u^T \left(|f(c_s, 0)| - \nu V_s^T\right) 1_{\{\tau^+ > s\}}\, ds \middle| \mathcal{F}_u\right].$$

Similarly, for all $u \in [t, T]$,

$$-V_u^T 1_{\{\tau^- > u\}} \leq E\left[\int_u^T \left(|f(c_s, 0)| - \nu\left(-V_s^T\right)\right) 1_{\{\tau^- > s\}}\, ds \middle| \mathcal{F}_u\right].$$

The stochastic version of the Gronwall-Bellman inequality (stated in Appendix B) gives

$$V_t^T 1_{\{\tau^+ > t\}} \leq E\left[\int_t^T e^{-\nu(s-t)} |f(c_s, 0)| 1_{\{\tau^+ > s\}}\, ds \middle| \mathcal{F}_t\right]$$

and

$$-V_t^T 1_{\{\tau^- > t\}} \leq E\left[\int_t^T e^{-\nu(s-t)} |f(c_s, 0)| 1_{\{\tau^- > s\}}\, ds \middle| \mathcal{F}_t\right].$$

Adding the last two inequalities and noting that $\tau^+ = \tau^- = 0 \Rightarrow V_t^T = 0$, we conclude that

$$|V_t^T| \leq E\left[\int_t^T e^{-\nu(s-t)} |f(c_s, 0)|\, ds \middle| \mathcal{F}_t\right].$$

The result follows by the growth condition on f. Q.E.D.

THEOREM C2 (Continuity): *The utility function* $U\colon \mathscr{D}_\nu \to \mathbb{R}$ *is continuous provided* f *is continuous.*

PROOF: Suppose $c, \hat{c} \in \mathscr{D}_\nu$ and let V and $\hat{V}$ be the respective associated utility processes. Fix t and define the stopping times: $\tau = \inf\{s\colon s \geq t,\, V_s \leq \hat{V}_s\}$ and $\tau_N = \tau \wedge N$. Lemma C1 and Assumption

2 imply that, for all $u \in [t, T]$,

$$
\begin{aligned}
\left(V_u - \hat{V}_u\right) 1_{\{\tau > u\}} &= E\left[\int_u^{\tau_N}\left(f(c_s, V_s) - f(\hat{c}_s, \hat{V}_s)\right) 1_{\{\tau > u\}}\, ds + \left(V_{\tau_N} - \hat{V}_{\tau_N}\right) 1_{\{\tau > u\}}\,\Big|\,\mathscr{F}_u\right] \\
&\leqslant E\left[\int_u^N\left(f(c_s, V_s) - f(\hat{c}_s, V_s) + f(\hat{c}_s, V_s) - f(\hat{c}_s, \hat{V}_s)\right) 1_{\{\tau > s\}}\, ds\,\Big|\,\mathscr{F}_u\right] \\
&\quad + E\left[\left(V_N - \hat{V}_N\right) 1_{\{\tau > N\}}\,\Big|\,\mathscr{F}_u\right] \\
&\leqslant E\left[\int_u^N\left(|f(c_s, V_s) - f(\hat{c}_s, V_s)| - \nu\left(V_s - \hat{V}_s\right) 1_{\{\tau > s\}}\right) ds\,\Big|\,\mathscr{F}_u\right] \\
&\quad + E\left[\left(V_N - \hat{V}_N\right) 1_{\{\tau > N\}}\,\Big|\,\mathscr{F}_u\right].
\end{aligned}
$$

By the stochastic Gronwall-Bellman inequality, applied on $[t, T]$,

$$
\begin{aligned}
\left(V_t - \hat{V}_t\right)^+ &\leqslant e^{-\nu(N-t)} E\left[|V_N| + |\hat{V}_N|\,\big|\,\mathscr{F}_t\right] \\
&\quad + E\left[\int_t^N e^{-\nu(N-s)}|f(c_s, V_s) - f(\hat{c}_s, V_s)|\, ds\,\Big|\,\mathscr{F}_t\right].
\end{aligned}
$$

Therefore, by symmetry,

(c6) $\qquad \begin{aligned}[t] |V_t - \hat{V}_t| &\leqslant 2e^{-\nu(N-t)} E\left[|V_N| + |\hat{V}_N|\,\big|\,\mathscr{F}_t\right] \\ &\quad + 2E\left[\int_t^N e^{-\nu(N-s)}|f(c_s, V_s) - f(\hat{c}_s, V_s)|\, ds\,\Big|\,\mathscr{F}_t\right]. \end{aligned}$

Suppose now that $c^n \to c$ in $\mathscr{D}_\nu$. By Jensen's inequality, this implies that $\|c^n - c\|^2$ converges[11] to zero in $L^1(\Omega \times [0, \infty), \mathscr{F} \otimes \mathscr{B}[0, \infty), \mu)$, where μ is the product measure defined by $\mu(A \times B) = E(\int_B e^{-\nu t} 1_A\, dt)$. It follows that, given any subsequence $(c^{n_k};\ k = 1, 2, \ldots)$, there is a further subsequence $(c^{n'_k};\ k = 1, 2, \ldots)$ that converges to c, μ-a.e. as $k \to \infty$. By continuity of f, $|f(c_t, V_t) - f(c_t^{n'_k}, V_t)| \to 0$ as $k \to \infty$, μ-a.e. Noting that

$$
\begin{aligned}
|f(c_t, V_t) - f(c_t^{n'_k}, V_t)| &\leqslant |f(c_t, V_t) - f(c_t, 0)| + |f(c_t, 0)| + |f(c_t^{n'_k}, 0)| \\
&\quad + |f(c_t^{n'_k}, V_t) - f(c_t^{n'_k}, 0)| \\
&\leqslant 2\max(k, -\nu)|V_t| + 2k_1 + k_2\left(\|c_t\| + \|c_t^{n'_k}\|\right),
\end{aligned}
$$

and using Lemma C2, the reader can verify that the dominated convergence theorem allows us to conclude that for every $N > 0$,

$$
E\left[\int_0^N e^{-\nu(N-s)}|f(c_s, V_s) - f(c_s^{n'_k}, V_s)|\, ds\right] \to 0 \quad \text{as} \quad k \to \infty.
$$

Let $V^{n'_k}$ be the utility process corresponding to $c^{n'_k}$. Then given $\varepsilon > 0$, the above result and (c6), applied at $t = 0$, imply that there exists K such that for all $k > K$,

$$
|V_0 - V_0^{n'_k}| \leqslant 2e^{-\nu N} E\left[|V_N| + |V_N^{n'_k}|\right] + \varepsilon.
$$

Letting $N \to \infty$ and using (c4), it follows that $|V_0 - V_0^{n'_k}| \leqslant \varepsilon$ for all $k > K$.

We have shown that every subsequence of $\{U(c^n) - U(c)\}$ has a further subsequence converging to zero. Therefore, the original sequence also converges to zero.   *Q.E.D.*

---

[11] Recall that an optional process is necessarily progressively measurable.

REFERENCES

ALLEN, B. (1987): "Smooth Preferences and the Approximate Expected Utility Hypothesis," *Journal of Economic Theory*, 41, 340–355.

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

CHEW, S. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51, 1065–1092.

——— (1989): "Axiomatic Utility Theories with the Betweenness Property," *Annals of Operations Research*, 19, 273–298.

CHEW, S., AND L. EPSTEIN (1990): "Recursive Utility under Uncertainty," in *Equilibrium Theory with an Infinite Number of Commodities*, ed. by A. Khan and N. Yannelis. New York: Springer Verlag.

CHUNG, K.-L., AND R. WILLIAMS (1990): *An Introduction to Stochastic Integration* (Second Edition). Boston: Birkhäuser.

CLARK, J. M. C. (1971): "The Representation of Functionals of Brownian Motion by Stochastic Integrals," *Annals of Mathematical Statistics*, 41, 1282–1295; 42, 1778.

CONSTANTINIDES, G. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519–543.

CRANDALL, M., AND P.-L. LIONS (1983). "Viscosity Solutions of Hamilton-Jacobi Equations," *Transactions of the American Mathematics Society*, 277, 1–42.

DEKEL, E. (1986): "An Axiomatic Characterization of Preferences Under Uncertainty," *Journal of Economic Theory*, 40, 304–318.

DELLACHERIE, C., AND P.-A. MEYER (1982): *Probabilities and Potential B: Theory of Martingales*. New York: North-Holland.

DETEMPLE, J., AND F. ZAPATERO (1989): "Optimal Consumption-Portfolio Policies with Habit Formation," School of Business, Columbia University.

DUFFIE, D., AND L. EPSTEIN (1991): "Asset Pricing with Stochastic Differential Utility," Graduate School of Business, Stanford University.

DUFFIE, D., AND C.-F. HUANG (1985): "Implementing Arrow-Debreu Equilibria by Continuous Trading of Few Long-Lived Securities," *Econometrica*, 53, 1337–1356.

DUFFIE, D., AND P.-L. LIONS (1990): "PDE Solutions of Stochastic Differential Utility," Research Paper 1095, Graduate School of Business, Stanford University; forthcoming in the *Journal of Mathematical Economics*.

DUFFIE, D., AND C. SKIADAS (1990): "Continuous-Time Security Pricing: A Utility Gradient Approach," Research Paper 1096, Graduate School of Business, Stanford University.

EPSTEIN, L. (1983): "Stationary Cardinal Utility and Optimal Growth under Uncertainty," *Journal of Economic Theory*, 31, 133–152.

——— (1987): "The Global Stability of Efficient Intertemporal Allocations," *Econometrica*, 55, 329–358.

——— (1988): "Risk Aversion and Asset Prices," *Journal of Monetary Economics*, 22, 179–192.

EPSTEIN, L., AND S. ZIN (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

——— (1990): "'First-Order' Risk Aversion and the Equity Premium Puzzle," *Journal of Monetary Economics*, 26, 387–407.

——— (1991): "The Independence Axiom and Asset Returns," Unpublished, Department of Economics, University of Toronto.

ETHIER, S., AND T. KURTZ (1986): *Markov Processes: Characterization and Convergence*. New York: Wiley.

FLEMING, W., AND R. RISHEL (1975): *Deterministic and Stochastic Optimal Control*. New York: Springer-Verlag.

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

HEATON, J. (1988): "The Interaction Between Time-nonseparable Preferences and Time Aggregation," Unpublished, Department of Economics, University of Chicago.

HINDY, A., AND C.-F. HUANG (1989a): "On Intertemporal Preferences with a Continuous Time Dimension II: The Case of Uncertainty," Working Paper 2105-89, Alfred P. Sloan School of Management, Massachusetts Institute of Technology.

——— (1989b): "Optimal Consumption with Intertemporal Substitution II: The Case of Uncer-

tainty," Working Paper Nov. 89, Alfred P. Sloan School of Management, Massachusetts of Technology.

HUANG, C.-F., AND D. KREPS (1987): "On Intertemporal Preferences with a Continuous Time Dimension I: The Case of Certainty," Unpublished, Massachusetts Institute of Technology.

JOHNSEN, T., AND J. DONALDSON (1985): "The Structure of Intertemporal Preferences Under Uncertainty and Time Consistent Plans," *Econometrica*, 53, 1451–1463.

KARATZAS, I., AND S. SHREVE (1988): *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

KOOPMANS, T. (1960): "Stationary Ordinary Utility and Impatience," *Econometrica*, 28, 287–309.

KOCHERLAKOTA, N. (1987): "State Nonseparability: Theory and Empirical Implications," PhD Thesis, Department of Economics, University of Chicago.

KREPS, D., AND E. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

KRYLOV, N. (1980): *Controlled Diffusion Processes*. New York: Springer-Verlag.

LIONS, P.-L. (1981): "Control of Diffusion Processes in $\mathbb{R}^n$," *Communications in Pure and Applied Mathematics*, 34, 121–147.

LUCAS, R., AND N. STOKEY (1984): "Optimal Growth with Many Consumers," *Journal of Economic Theory*, 32, 139–171.

LUENBERGER, D. (1969): *Optimization by Vector Space Methods*. New York: Wiley.

MACHINA, M. (1982): "'Expected Utility' Analysis without the Independence Axiom," *Econometrica*, 50, 277–323.

MAS-COLELL, A. (1986): "The Price Equilibrium Existence Problem in Topological Vector Lattices," *Econometrica*, 54, 1039–1054.

PARDOUX, E., AND S.-G. PENG (1990): "Adapted Solutions of a Backward Stochastic Differential Equation," *Systems and Control Letters*, 14, 55–61.

PROTTER, P. (1990): *Stochastic Integral and Differential Equations*. New York: Springer-Verlag.

QUIGGIN, J. (1982): "Anticipated Utility Theory," *Journal of 'Economic Behavior and Organization*, 3, 323–343.

SEGAL, U. (1989): "Axiomatic Representation of Expected Utility with Rank-Dependent Probabilities," *Annals of Operations Research*, 19, 359–373.

SEGAL, U., AND A. SPIVAK (1990): "'First-Order' versus 'Second-Order' Risk Aversion," *Journal of Economic Theory*, 51, 111–125.

SKIADAS, C. (1991): "Time-Consistent Choice and Preferences for Information," Unpublished, Department of Operations Research, Stanford University.

SUNDARESAN, S. (1989): "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Economics*, 2, 73–89.

SVENSSON, L. E. O. (1989): "Portfolio Choice with Non-Expected Utility in Continuous Time," *Economic Letters*, 30, 313–317.

TRUDINGER, N. (1983): "Fully Nonlinear Uniformly Elliptic Equations Under Natural Structure Conditions," *Transactions of the American Mathematical Society*, 278, 751–769.

UZAWA, H. (1968): "Time Preference, The Consumption Function and Optimum Asset Holdings," in *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, ed. by J. Wolfe. Chicago: Aldine.

VIDYASAGAR, M. (1978): *Nonlinear Systems Analysis*. Englewood Cliffs, New Jersey: Prentice-Hall.

YAARI, M. (1987): "The Dual Theory of Choice under Risk: Risk Aversion without Diminishing Marginal Utility," *Econometrica*, 55, 95–115.