WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

On Intertemporal Preferences in Continuous Time: The Case of Certainty*

Ayman Hindy, Chi-fu Huang, and David Kreps

March 1991 WP# 3311

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

On Intertemporal Preferences in Continuous Time:
The Case of Certainty*

Ayman Hindy, Chi-fu Huang, and David Kreps

March 1991 WP# 3311

On Intertemporal Preferences in Continuous Time: The Case of Certainty*

Ayman Hindy[†], Chi-fu Huang[‡], and David Kreps[§]

March 1991

To appear in *Journal of Mathematical Economics*

Abstract

Different topologies on the space of certain consumption patterns in a continuous time setting are discussed. A family of topologies which give an economically reasonable sense of closeness and have an appropriate intertemporal flavor is suggested. The topological duals of our suggested topologies are essentially spaces of Lipschitz continuous functions. Any utility functional whose felicity function at any time t depends explicitly on the consumption at that time and is not linear in it is not continuous in any one of our topologies. A class of utility functionals that are continuous in the suggested topologies and that capture the intuitively appealing notion that consumptions at nearby dates are almost perfect substitutes is provided. We then give necessary and sufficient conditions for a consumption plan to be optimal for this class of utility functionals. We demonstrate our general theory by solving in closed form the optimal consumption problem for a particular utility function. The optimal solution consists of a (possible) initial "gulp" of consumption, or an initial period of no consumption, followed by consumption at the rate that maintains a constant ratio of wealth to an index of past consumption experience.

*This paper merges results reported in two earlier papers, a paper with the same title by Huang and Kreps and a paper titled "Optimal Consumption with Intertemporal Substitution I: The Case of Certainty" by Hindy and Huang. The authors would like to thank Peter Diamond, Darrell Duffie, Oliver Hart, Hua He, John Heaton, Larry Jones, Andreu Mas-Colell, Scott Richard, Tong-sheng Sun, Jean-luc Vila, Thaleia Zariphopoulou, and the anonymous referees for comments. Itzhak Katznelson told us about Orlicz Spaces.

[†]Graduate School of Business, Stanford University, Stanford, CA 94305

[‡]Sloan School of Management, Massachusetts Institutes of Technology, Cambridge, MA 02139

[§]Graduate School of Business, Stanford University, Stanford, CA 94305

1 Introduction and Summary

Consider an economic agent who lives from time 0 to time 1. Suppose there is a single consumption commodity consumable at any time between zero and one. We ask the following questions: First, how might we represent the agent's consumption pattern over his lifetime? Second, if we let X_+ denote the space of possible consumption patterns, what is an appropriate topology on X_+; that is, one which gives an economically reasonable sense of "closeness" with the appropriate intertemporal flavor and which is relatively well-behaved mathematically? Third, what form will equilibrium prices take if they are given by a continuous linear functional? Fourth, what is the optimal consumption-savings behavior of an agent whose preferences are continuous in such a topology?

There are "standard" answers to these questions. Consumption patterns are represented by a function $c: [0,1] \rightarrow \Re_+$, where $c(t)$ represents the *rate* of consumption at time t. Two such consumption patterns $c(t)$ and $\hat{c}(t)$ are "close" if they are close as functions; say, in the sup-norm topology or in some L^p topology. In addition, prices come from L^q, for the appropriate value of q. Preferences are represented by $U(c) = \int_0^1 u(c(t), y(t), t)dt$, where u is the *felicity* function of the individual and $y(t)$ is some functional of the consumption rates.[1] A typical example is $y(t) = \bar{y}(t) + \int_0^1 \theta(t,s)c(s)ds$ for some functions $\bar{y}(t)$ and $\theta(t,\cdot)$. This is a generalized version of the Ryder and Heal [1973] preferences analyzed by Magill [1981].

We are dissatisfied with these standard answers. Basically we feel that consumption at one time should be something of a substitute for consumption at other, near-by times. Consequently, prices for consumption at close-by times should be approximately the same. These, however, are not the properties of the above standard answers. We thus set out to find answers to the questions of the previous paragraph that will conform to these intuitions, and retain, at the same time, mathematical properties important for economic analysis.

Readers familiar with the literature on commodity differentiation, for example, Mas Colell [1975] and Jones [1981, 1984] will at this point rightly think that the answers to our questions are readily available in this literature. Take the commodity space to be the space of signed measures on $[0,1]$ equipped with the topology of weak convergence, or simply the weak topology. A consumption pattern is represented by a measure on $[0,1]$ or by a right-continuous increasing function where the value of this function at time t denotes the cumulative consumption from time 0 to time t. In this representation, continuous injestion as well as gulps are allowed. Then as long as the agents' preferences are continuous in the weak topology and satisfy a strong monotonicity property equilibria exist with prices that are continuous functions of time. Preferences continuous in the weak topology treat nearby consumptions as close substitutes and equilibrium prices are continuous functions of time, all the questions we raised above seem to have been answered!. There is, however, one caveat. A merely continuous function is not, in general, a reasonable representation of

[1] When the felicity function is a function only of c and t, we have the standard time-additive utility functional.

consumption and asset prices over time.

A continuous function can be nowhere differentiable – take, for example, a Brownian motion sample path. In fact, most continuous functions are nowhere differentiable. This is perhaps not so counter-intuitive when time is not the characteristic of the commodity in question. When the commodities are indexed by time, our intuition certainly suggests that prices for consumption should be continuous functions of time. We also think that more should be true. Prices should not fluctuate over time in a nowhere differentiable fashion.[2] In particular, we think that interest rates should exist in a reasonable economic intertemporal model. This condition is equivalent to the requirement that prices for consumptions be *absolutely* continuous functions of time. This additional requirement is what makes the results of the commodity differentiation literature unsatisfactory for our purposes.

We begin in Section 2 with a formulation of the consumption set and consumption space. For the set of feasible consumption patterns, we use the space of positive, increasing, right continuous functions on $[0,1]$, denoted X_+.[3] For given $x \in X_+$ and $t \in [0,1]$, $x(t)$ denotes the *cumulative* consumption from time zero to time t under consumption pattern x. In order to speak of net trades, the space of consumption bundles, or the *commodity space*, is the linear span of X_+. Denoted X, this is the space of functions of bounded variation on $[0,1]$ that are right-continuous.

The space X comes equipped with a number of topologies, any of which could conceivably serve our purposes. In order to investigate their suitability, we construct in Section 3 a "wish list" of properties for a suitable topology. Our wish list comes in two parts. First, we give useful mathematical properties, keeping in mind that, in the end, we will want to do optimization, and conduct equilibrium analysis. Second, and more important, we list desirable economic properties. The discussion here is the heart of this paper in terms of economics; essentially, we are giving our intuition concerning when different consumption patterns should be "close".

Once we have our wish list, we can see how well the standard topologies fare. This is done in Section 4. We quickly look at the topologies induced by the total variation norm and the Skorohod topology and show that these two fail rather substantially. The weak topology has many of the properties we desire, and, indeed as we mentioned above, there is a substantial literature on economies where the commodity space is the space of signed measures, endowed with this topology; see Mas-Colell [1975] and Jones [1981, 1984]. But, as we mentioned above, the weak topology has too large a topological dual, and, as we will see, lacks a certain uniform property that we find appealing.

We turn instead, in Section 5, to a family of norm topologies. For this introduction, we will use one representative member of this family, given by the norm $\|x\| = \int_0^1 |x(t)|dt + |x(1)|$ for $x \in X$; the topology induced by this norm is denoted hereafter by $\mathcal{T}$. We will show

[2] Prices for consumption will in general fluctuate in a nowhere differentiable fashion in a model with uncertainty due to temporal resolution of uncertainty. See, for example, Huang [1985].

[3] We will use weak relations throughout. For example, negative means nonpositive, increasing means nondecreasing, and so forth. If the relation is strict, we will use strictly negative, strictly increasing, and so forth.

that, on X_+, these norm topologies are all topologically equivalent to the weak topology. However, preferences continuous and uniformly proper, a concept due to MasColell [1986] discussed in section 2, in any of the proposed topologies exibit a trade-off property which we find appealing. This property is essantially that agents tolerate delaying or advancing increasingly larger quantities of consumption for decreasingly shorter periods of time. In contrast, preferences continuous and uniformly proper in the weak topology fail to exhibit tolerance for this kind of intertemporal substitution. The norm topologies we introduce have all the properties we find desirable from the point of view of economic intuition and, in addition, they are mathematically very well behaved. For example, preferences continuous in these topologies admit continuous numerical representations.

In Section 7, we turn to the characterization of the topological dual spaces for our topologies. In particular, the dual to $(X, \mathcal{T})$, denoted, hereafter, as X^*, is the space of Lipschitz continuous functions. In this section, we also confirm some desirable mathematical properties of the dual spaces, notably that the duals are vector lattices in the order duals and the order intervals are weakly compact.

Section 8 gives necessary and sufficient conditions for a consumption pattern to be optimal for a class of utility functions continuous in our topologies in a consumption-savings model. An optimal consumption plan in general will be periodic even if the felicity function has an infinite slope at zero. Whether one should consume at a given time depends on the relation between ones wealth and a weighted average of past consumption. A closed form solution for a particular utility function is provided in Section 8.3. Finally, all proofs are in the appendix.

There are three companion papers to this. Mas-Colell and Richard [1991] does the hard work of establishing existence of equilibrium for economics based on our, and other, topologies. Hindy and Huang [1990a, 1990b] generalize and extend our analysis to models where there is uncertainty and temporal resolution of that uncertainty, and they consider the impact of our topology on some of the standard models in the theory of finance.

2 The Consumption Set and Consumption Space

Consider an economic agent who lives from time zero to time one. This agent can consume a consumption commodity at any time during his life span. It is natural to require that the agent consume only positive amounts of the consumption good, and that the cumulative consumption is finite. Thus a life-time consumption pattern can be represented by a positive increasing real-valued function $x : [0,1] \to \Re$, with $x(t)$ denoting the cumulative consumption from time zero to time t. Note that this representation permits consumption at "rates" if $x(t)$ is absolutely continuous and in discrete lumps if $x(t)$ is discontinuous. As x is increasing, it has the following two properties. First, for each $t \in [0,1]$, both unilateral limits $x(t^-) \equiv \lim_{s\uparrow t} x(s)$ and $x(t^+) \equiv \lim_{s\downarrow t} x(s)$ exist and are finite. Second, the only possible kind of discontinuity is a jump.

To avoid complications, we shall further assume that x is *right-continuous:* $\lim_{s\downarrow t} x(s) =$

$x(t)$, for all $t \in [0,1)$. That is, if the agent consumes a nontrivial stock amount of the consumption commodity during the time interval (t^-, t^+), he in fact consumes $x(t^+) - x(t^-)$ at time t.[4] To sum up, we will represent the agent's life-time consumption pattern by a positive increasing right-continuous real-valued function x on $[0,1]$. The set of such functions is denoted by X_+.

Throughout the paper, χ_r for $r \in [0,1]$ will denote the indicator function of $[r,1]$. That is, χ_r represents a unit gulp of consumption at time r. Also, scalars will be denoted by $a, b, \ldots$

Let X denote the linear span of X_+, which is the space of right continuous real-valued functions of bounded variation defined on $[0,1]$. The cone X_+ induces orders $\geq$ and $>$ on X in the usual way: $x_1 \geq x_2$ if $x_1 - x_2 \in X_+$, and $x_1 > x_2$ if $x_1 - x_2 \in X_+ \setminus \{0\}$. We will refer to X as the *commodity space* and to X_+ as the agent's *consumption set*. Note that X is a lattice, in the usual fashion. For $x \in X$, we use x^+ to denote the positive (increasing) part of x. We also use the convention that if $x(0) > 0$, then $x^+(0) = x(0)$. Then we can define $\sup(x,y) = (x-y)^+ + y$, with $\inf(x,y)$ defined similarly.

Our agent will be assumed to have preferences over X_+ that are given mathematically by a complete and transitive binary relation $\succeq$ on X_+. We will assume throughout that these preferences are increasing: $\forall x_1, x_2 \in X_+$, $x_1 + x_2 \succeq x_1$; that they are strictly increasing in the direction of χ_0: $\forall x \in X_+$ and $\forall a > 0$, $x + a\chi_0 \succ x$. The reader may be put off by the special role played by χ_0 here. It would suffice to assume that there is some $x \in X_+ \setminus \{0\}$ in the direction of which preference is always strictly increasing, although condition (c) below and our uniform properness assumption, to be discussed shortly, would need to be modified accordingly. For the analysis of Mas-Colell and Richard [1991], one would wish to use the social endowment in place of χ_0.

Moreover, we will assume that there is given a linear topology T on X; with T^+ denoting the topology T relativized to X_+. Preferences are assumed to be continuous in T^+ and uniformly proper with respect to T. Continuity of preferences in a fixed topology is fairly standard and needs no justification. Uniform properness, on the other hand, is not so standard. This condition, introduced by Mas-Colell [1986], is shown there to be intimately related to the existence of equilibria for economies with an infinite dimensional commodity space. The reader should consult Mas-Colell [1986] for further commentary. Here, we will merely recall his definition:

Definition 1 (Mas Colell) *Preferences $\succeq$ on X_+ are said to be uniformly proper with respect to a linear topology T if there exists a vector $x^* \in X_+$ and a T-open neighborhood V of the origin such that, for every $x \in X_+$, $v \in V$, and scalar $a > 0$, $x - ax^* + av \not\succeq x$.*

In fact, we will assume that preferences are uniformly proper with respect to T for the particular choice of $x^* = \chi_0$.[5]

[4] Here we have used the convention that $x(t) = x(1)\ \forall t \geq 1$ and $x(t) = 0\ \forall t < 0$. We will continue to use this convention throughout, so that the value of any $x \in X$ for arguments greater than one or less than zero is always well defined.

[5] As we mentioned earlier, our use of χ_0 in this special role is purely to ease exposition. Any $x^* \in X_+$

3 A Wish List for T

What properties do we want for the topology T? The list of *desiderata* comes in two parts. First, we want T to be well behaved mathematically. Ideally, T should be a locally convex, metric topology, and X should be complete and separable under this metric. Also, since we wish to invoke uniform properness and undertake optimization for our agents, we will wish for T to be a linear topology. In order to apply the equilibrium existence results of Mas-Colell [1986], one would wish that the lattice operations are continuous with respect to T, or (even) that (X, X_+, T) forms a Banach lattice. But this, as it will turn out, will be too much to hope for, when we come upon economic considerations. Happily, Mas-Colell and Richard [1991] allow us to weaken this desideratum substantially: We will want the order intervals to be weakly compact and the topological dual to be a vector lattice in the order dual.[6]

The second and more substantial part of our wish list concerns the economics of the situation. Our intuition suggests that certain consumption streams ought to be close, and we will look for a topology on X that accomodates this intuition. Consider the following requirements, phrased in terms of properties of $\succeq$.

(a) Two patterns of consumption that have almost equal cumulative consumption at every point in time should be close. Formally, if $\lim_{n\to\infty} \sup_{t\in[0,1]} |x_n(t) - x(t)| = 0$, then $x_n \succeq y$ for all n implies $x \succeq y$, and $y \succeq x_n$ for all n implies $y \succeq x$.[7]

In (a), differences in cumulative consumption between x_n and x must be small at every time t. We wish to weaken this so that shifts in consumption of a fixed size across small amounts of time are also regarded as insignificant. To do so, we begin by introducing a well known metric on X_+, the Prohorov metric. For x and y in X_+, let

$$p(x, y) = \inf\{\epsilon > 0 : x(t+\epsilon) + \epsilon \geq y(t) \geq x(t-\epsilon) - \epsilon\}.$$

The idea of this metric is depicted in Figure 1; the ϵ neighborhood of a consumption path x is any other path y that lives entirely within a "sleeve" around x that is determined by moving ϵ up and to the left and down and to the right (ϵ in each direction) from x at every point $(t, x(t))$. Hence if y is within ϵ of x, then at every time t, the total consumption under x is within ϵ of the total consumption under y at some time no more than ϵ away from t. With this we can pose our second requirement:

(b) If $p(x_n, x) \to 0$ as $n \to \infty$, then $x_n \succeq y$ for all n implies $x \succeq y$, and $y \succeq x_n$ for all n implies $y \succeq x$.[8]

$x^* \neq 0$ could be used instead.

[6] For the definition of a vector lattice and the order dual, see Schaefer (1974).

[7] The astute reader will notice that this rules out the standard model where utility is the integral of felicity of consumption rates, unless felicity is linear. We will demonstrate this explicitly in section 6. The reader will also note that (a) could be said more compactly as: T relativized to X_+, denoted T^+ should be no stronger than the sup norm topology relativized to X_+.

[8] Or: T^+ should be no stronger than the weak topology on X_+ generated by the space of continuous functions on $[0, 1]$.

Our final economic consideration sharpens (b) somewhat. The idea in (b) is that the shift of a given amount of consumption across a small amount of time should not be of great consequence in terms of preference. Suppose that we imagine instead the shift of an increasingly larger amount of consumption over a decreasingly smaller period of time. It may help to give a concrete example: Consider comparing the consumption patterns $a_n\chi_{1/2}$ and $a_n\chi_{(n+1)/2n}$, where a_n is a strictly positive scalar. As n goes to infinity, the shift in time grows small (and so, in particular, $p(a_n\chi_{1/2}, a_n\chi_{(n+1)/2n}) \to 0$), but if a_n grows with n, the amount shifted grows large. We wish to assert that, a *priori*, it should be possible for a_n to go to infinity slowly enough so that the difference in terms of preference becomes negligible uniformly. To be more precise, we know, by the presumed strict monotonicity of $\succ$ in the direction χ_0 that $\chi_0 + a_n\chi_{1/2} \succ a_n\chi_{1/2}$. We wish to assert that the "added utility" from χ_0 eventually compensates for the shift of a_n from 1/2 to $(n+1)/2n$, if a_n goes to infinity sufficiently slowly. For example, we might insist that, if a_n is $o(n)$ (i.e., $a_n/n \to 0$), then the growth in a_n is slow enough. Or we might require that a_n is $o(n^{1/2})$, or $o(\ln n)$, in order to come to the desired conclusion that the shift in consumption through time is small relative to the amount of consumption shifted. But there is some "sufficiently slow" rate of growth in the amount shifted such that, if the shift in time is small, the change in utility is small.

To measure what is meant by a "sufficiently slow" rate of growth, we suppose that we are given a concave, continuous, strictly increasing function $\mu : [0,\infty) \to [0,\infty)$, such that $\mu(0) = 0$ and $\mu(n) \to \infty$ as $n \to \infty$. For examples, think of $\mu(z) = z^{1/p}$ for $p \geq 1$, or $\mu(z) = \ln(z+1)$. By "slow growth" of a sequence of numbers a_n relative to μ is meant $a_n/\mu(n) \to 0$, or $a_n = o(\mu(n))$. Note that linear growth is the fastest "slow growth" that we are willing to contemplate, since μ must be concave. In general, we pose the following condition.

(c) For fixed μ as above, if $\{x_n\}$ and $\{y_n\}$ are two sequences of claims with (i) $\{x_n(1)\}$ and $\{y_n(1)\}$ are both $o(\mu(n))$, and (ii) $p(y_n, x_n) \leq 1/n$, then for all sufficiently large n, $x_n + \chi_0 \succ y_n$.

Note that the condition depends on the fixed μ. It becomes stronger the faster μ grows, so it is strongest when $\mu(z) = z$. We are imagining that μ is given uniformly for all agents in the economy; it enters as part of the data of the economy under consideration. But we do not wish to prespecify what rate μ is "slow enough." We only suppose that some "slow enough" rate can be specified in a particular economic application.[9]

[9] It has been said to us that this condition is not very natural – how would one discover what is the appropriate function μ for a given population of consumers? We suggest the following. First, for a given consumption path x, let $N_\epsilon(x)$ denote the set of $x' \in X_+$ that are within ϵ of x in the Prohorov metric. We suggest that a natural property for consumer preferences is: For each $x \in X_+$, the $\succeq$-worst element of $N_\epsilon(x)$ is x "delayed and lowered by ϵ". That is, in terms of figure 1, the worst consumption pattern in the sleeve drawn around x would be the lower-right boundary of the sleeve. If one admits to this structural property (which, we observe, will hold for the sort of representation we suggest at the end of section 6, if some impatience is added to that representation), then we can ask of a given consumer: Take any path

4 Three Candidate Topologies

Several topologies on the space of functions of bounded variation have been extensively analyzed by mathematicians. In this section, we ask how each of these do in terms of our wish list.

For the space of functions of bounded variation, a mathematically natural norm is *total variation*: For $x \in X$, the total variation of x, denoted by $\mathrm{TV}(x)$, is $\mathrm{TV}(x) = |x(0)| + \int_0^1 |dx(t)|$. (The L^1 norm on consumption rates in the standard models is just the total variation norm on cumulative consumption.) Note well that if we hope to have X be a topological vector lattice with X_+ the positive cone,[10] then we will need to have a topology at least as strong as the total variation norm topology. But while this topology performs well mathematically, aside from lack of separability, it performs miserably in terms of our economic wish list. In the first place, (a) is violated: Consider the consumption patterns x_n given by $x_n(t) = k/n$ for $t \in ((k-1)/n, k/n]$ and x given by $x(t) = t$. According to (a), we should have x_n converging to x, but (of course) we do not, in the total variation norm. In the same way (b) is violated. Total variation, when the underlying time space is a continuum, measures time much too finely to have the sort of economic qualities we desire.

Consider next the Skorohod metric originally introduced on the space of functions which are right-continuous and have finite left limits (RCLL). This metric is defined as follows:[11] Let Λ denote the class of strictly increasing, continuous mappings of $[0,1]$ *onto* itself. If $\lambda \in \Lambda$, then $\lambda(0) = 0$ and $\lambda(a) = 1$. For $x, y \in X$, define $S(x,y)$ to be the infimum of those strictly positive scalars ϵ for which there exists in Λ a λ such that $\sup_{t\in[0,1]} |\lambda(t) - t| \leq \epsilon$ and $\sup_{t\in[0,1]} |x(\lambda(t)) - y(t)| \leq \epsilon$. It can be shown that $S : X \times X \to R_+$ is a metric, and that (X, S) is a separable metric space.

Moreover, this metric does possess some of the economic qualities that we desire. Since two consumption patterns that are close in the sup norm are close under the Skorohod metric, (a) is satisfied. Unhappily, (b) is not satisfied. We note, in this regard, that if we define x_n by $x_n(t) = 0$ for $t < 1/2$, $x_n(t) = 2n(t - 1/2)$ for $t \in [1/2, (n+1)/2n]$, and $x_n(t) = 1$ for $t > (n+1)/2n$, then $p(x_n, \chi_{1/2}) \to 0$ (so (b) should apply), but x_n does not approach $\chi_{1/2}$ in the Skorohod topology. The Skorohod topology is also not very well behaved when it comes to exploiting the linear structure of the space of consumption claims. In particular, it is not a linear topology. Whereas $\chi_{t_n} \to \chi_t$ in this topology if $t_n \to t$, their difference $\chi_{t_n} - \chi_t$ does not approach zero. We still continue our search.

One other topology frequently used on X is the topology of weak convergence denoted by $\mathcal{T}_w$, with $\Rightarrow$ denoting convergence in this topology. Let $C[0,1]$ be the space of continuous

$x \in X_+$ with $x(a) = k$. Suppose you possess this path. What is the most you are willing to have this path "delayed and lowered" by, if in compensation we give you an additional χ_0 up front? Now what path x minimizes this amount, over all paths with $x(a) = k$? (For preferences with representations as at the end of section 6, this optimization problem is tractible.) Finally, how does this amount change as k increases? The answer to this question, of course, gives us μ.

[10] which we would want if we were to rely on the existence result of Mas Colell [1986]

[11] A good reference is Billingsley (1968, p???).

functions on $[0,1]$. Then $x_\alpha \Rightarrow x$ *if and only if* $\int_{0-}^1 f(t)dx_\alpha(t) \to \int_{0-}^1 f(t)dx(t)$ for all $f \in C[0,1]$ [12].[13]

Recall that the weak topology is a linear, Hausdorff, locally convex topology. It is separable (cf. Holmes [1975, exercise 22c]). It is metrizable on X_+ by precisely the Prohorov metric (hence has a countable neighborhood base and is sequential), but it is not metrizable on X.

From the point of view of economic desiderata, the weak topology is well behaved: Preferences continuous in the weak topology on X_+ will obey (a) and (b). Altogether, it is quite an attractive candidate, except for being non-metrizable on all of X. It has the further virtue of having been thoroughly analyzed in the economics literature, e.g., in Mas-Colell [1975] and Jones [1981, 1984].

Preferences that are continuous and uniformly proper in the weak topology need not be well behaved in terms of (c), for any fixed function μ, however. This is so even for linear preferences: Fix some μ, and consider the continuous function $f(t) = 2\sqrt{1/\mu(1/t)} + 1$. (Recall that $\mu(\infty) = \infty$ is assumed.) Define linear preferences by $x \succeq y$ if $\int f dx \geq \int f dy$. That is, preferences are given by evaluating consumption plans at the "prices" given by f. *A fortiori*, $\succeq$ is continuous and uniformly proper in the weak topology, and it is strictly increasing because f is strictly positive.

Consider $x_n = \sqrt{\mu(n)}\chi_0$ and $y_n = \sqrt{\mu(n)}\chi_{1/n}$. Note that x_n and y_n qualify in terms of (c). In particular, $x_n(a)/\mu(n) = 1/\sqrt{\mu(n)} \to 0$. But, for every n,

$$\int f d(x_n + \chi_0) = 1 + \sqrt{\mu(n)} < \int f dy_n = \sqrt{\mu(n)} \cdot (1 + 2\sqrt{1/\mu(n)}).$$

Roughly put, "shadow prices" for consumption at times just after time zero rise very quickly. As a result, shifting an increasing amount of consumption from time zero to time $1/n$ has quite an impact, if that amount increases quickly enough. Condition (c) holds that there is a slow enough increase in the amount shifted so that the shift should become negligible, but the point of (c) is that the slow-enough-increase must be prespecified, and we have picked shadow prices (given by f) to rise around zero quickly enough so that the prespecified slow-enough-increase is still too much.[14]

5 A Family of Norm Topologies

We have discussed three possible topological structures on X. The topology generated by total variation is too strong along the time dimension. The Skorohod topology does better,

[12] Integrals here are defined from $0-$ to 1, so that $f(0)x(0)$ enters as a first term in the integral.

[13] If we norm $C[0,1]$ by the sup norm, then the weak topology is just the weak* topology of X generated by $C[0,1]$.

[14] For any continuous function f, and for preferences defined from f as in the example, we can find a rate function μ that grows slowly enough so that (c) would be satisfied for these preferences and for μ. Thus the order of quantifiers in (c) is all important – μ must be picked first, and then, what we have shown, there are "nice" (i.e., linear) preferences continuous and uniformly proper in the weak topology that violate (c) for this μ.

but it violates (b) and is not linear. The topology of weak convergence is quite suitable, except that preferences continuous and uniformly proper in the weak topology may violate (c). In this section, we will introduce a family of norm topologies which will be free of all these defects.

To do this, the reader must be introduced to the notion of an Orlicz space, for which Musielak [1983] is an accessible reference. To make this paper close to self-contained, we give a very quick introduction here, adequate for, and told in terms of our particular application.

To begin, consider the case where the function $\mu(z) = z^{1/p}$, for some $p \geq 1$. Imagine that we defined the following norm on X:

$$\|x\|_p = \left[\int_0^1 |x(t)|^p dt + |x(1)|^p\right]^{1/p}.$$

Since our functions x are all of bounded variation, it is clear that all these integrals exist. From standard Banach space theory, it is clear that $(X, \|\cdot\|_p)$ is a separable normed topological linear space. It is worth noting, perhaps, that X will not be complete in this norm: Consider, for example, the sequence of right-continuous, bounded variation functions x_n that are indicator functions of $[1/2n, 1/(2n-1)) \cup [1/(2n-2), 1/(2n-3)) \cup \ldots \cup [1/2, 1]$. These are Cauchy and converge to the obvious measurable function, but what they converge to has unbounded variation and is not right-continuous. X_+, however, is complete.

This norm has the right "flavor" for (c) as well, for the case $\mu(z) = z^{1/p}$. Roughly put, $\|\cdot\|_p$ trades off differences in consumption paths in time and in amounts, where differences in amounts are scaled by being raised to the power p. Hence if the difference in two consumption paths is $O(1/n)$ in the time dimension and $o(n^{1/p})$ in the space dimension, then the difference is small.[15]

What, then, do we do if $\mu(z)$ is not of the form $z^{1/p}$? For μ define $\eta = \mu^{-1}$ and the *unit circle* in X, denoted C_η, as

$$C_\eta = \{x \in X : \int_0^1 \eta(|x(t)|)dt + \eta(|x(1)|) = 1\}.$$

The convexity of η ensures that this unit circle is appropriately shaped to be a "gauge" for measuring all other vectors x: Define, for x other than $x \equiv 0$, $\|x\|_\eta$ to be the unique strictly positive scalar a that satisfies $x/a \in C_\eta$. It takes a proof, but because of the continuity of η, $\|x\|_\eta$ is well defined. Moreover, for $\mu(z) = z^{1/p}$, this definition is equivalent to the standard definition of the norm. With this definition:

Proposition 1 *$(X, \|\cdot\|_\eta)$ is a separable normed topological linear space. With the order operations defined in Section 2, order intervals are weakly compact in each of these topologies.*

We omit the proof for the first assertion of this proposition as it is a simple application of the theory of Orlicz spaces; see Musielak [1983, theorem 1.5]. In the appendix, we provide a proof of the second assertion.

[15] See Proposition 4 for an exact statement.

Note that a countable dense set of $(X, \|\cdot\|_\eta)$ is the set of piecewise linear functions that change their values at rational time points and have rational right-hand derivatives. These functions are absolutely continuous. Thus the space of absolutely continuous functions is dense in $(X, \|\cdot\|_\eta)$ – any consumption pattern can be approximated arbitrarily closely in the norm $\|\cdot\|_\eta$ by consumption at rates.

Fixing μ and the corresponding η, the topology induced by $\|\cdot\|_\eta$ is denoted by $\mathcal{T}_\eta$, and the $\mathcal{T}_\eta$ topology relativized to X_+ is denoted by $\mathcal{T}_\eta^+$. For the special case where $\mu(z) = z^{1/p}$, we use notation such as $\mathcal{T}_p$.

The following basic results are established in the appendix.

Proposition 2 *For any μ and corresponding η, $\mathcal{T}_1$ is no stronger than $\mathcal{T}_\eta$, which in turn is weaker than both the Skorohod and total variation topologies.*

The first part of this proposition is fairly intuitive. Since η is convex, it tends to accentuate, more than linearly, large differences in the dimension of space. Note that, in our norm topologies, the lattice operations are most definitely not continuous and $(X, \mathcal{T}_\eta)$ is not a topological vector lattice. For example, $\chi_{1/2} - \chi_{(n+1)/2n} \to 0$, but the corresponding positive parts (or negative parts, or absolute values) do not converge.

We will be looking for preferences on X_+ that are given by a binary relation $\succeq$ that is complete, transitive, continuous in the $\mathcal{T}_\eta^+$ topology, and uniformly proper with respect to $\mathcal{T}_\eta$, for some prespecified η. We will continue to assume that $\succeq$ is increasing as before. The following results establish that we are looking for the "right" thing given our desiderata.

Proposition 3 *Preferences $\succeq$ that are continuous with respect to $\mathcal{T}_\eta^+$ satisfy (a) and (b). Moreover, X_+ is separable under $\mathcal{T}_\eta^+$, and hence every continuous (complete and transitive) preference relation $\succeq$ on X_+ admits a continuous numerical representation.*

Satisfaction of (b) follows as a simple corollary to Proposition 5, which in turn immediately implies the satisfaction of (a). The separability of X_+ follows standard lines and the proof is omitted. The existence of a numerical representation is standard given continuity and separability – see, for example, Fishburn [1970].

Proposition 4 *Fix μ and the corresponding η. Suppose that $\succeq$ is $\mathcal{T}_\eta^+$ continuous, increasing as above, and $\mathcal{T}_\eta$ uniformly proper* for the specific choice $x^* = \chi_0$. *Then $\succeq$ satisfies (c).*

In the appendix, we prove one technical fact: If x_n and y_n are as in (c), then $y_n - x_n \to 0$ in $\mathcal{T}_\eta$. Thus for any neighborhood V of the origin, $y_n - x_n$ is eventually inside V. Apply the definition of uniform properness to $\chi_0 + x_n$, to conclude that, eventually, $\chi_0 + x_n - \chi_0 + y_n - x_n = y_n$ must not stand in the relation $\succeq$ to $\chi_0 + x_n$. That does it.

This leaves open one issue: How do the topologies $\mathcal{T}_\eta^+$ compare with the weak topology on X_+? It is clear from the example of Section 4 and Proposition 4 that uniform properness with respect to $\mathcal{T}_\eta$ is stronger than uniform properness with respect to the weak topology. But:

Proposition 5 *For any μ and η, $\mathcal{T}_\eta^+$ is topologically equivalent to the weak topology on X_+.*

Hence we reaffirm (the obvious): Uniform properness is not just a topological property on X_+.

6 On the Standard Models and the $\mathcal{T}_\eta$ Topologies

The standard models of preferences for consumption through time assume that preferences are defined over consumption in rates and are given by a numerical representation of either the time-additive form

$$U(x) = \int_0^1 u(x'(t), t)dt,$$

or the non-time-additive form

$$U(x) = \int_0^1 u(x'(t), y(t), t)dt, \tag{1}$$

where x' denotes the first derivative of cumulative consumption x, and y is some functional of x'. The time-additive form needs no further explanation. The non-time-additive form includes, for example, a generalization of Ryder and Heal [1973] by Magill [1981], who used

$$y(t) = \bar{y}(t) + \int_0^1 \theta(t, s)x'(s)ds,$$

for some real-valued functions $\bar{y}$ and θ, and a special case of Magill [1981], the so-called habit-formation model, recently analyzed by Constantinides [1990] and Sundaresan [1989] when $\theta(t, s) = e^{-\beta(t-s)}$ for all $s \leq t$ and zero elsewhere.

Preferences so given are not defined for consumption other than at rates (where x is not absolutely continuous), but we might consider the following procedure for extending preferences to general x: Find a sequence of absolutely continuous x_n that approximate x, and then define $U(x) = \lim_{n\to\infty} U(x_n)$, hoping that the limit is well defined and independent of the approximating sequence x_n. We will want preferences to be $\mathcal{T}_\eta^+$ continuous, so the sense of approximation had better be no weaker than $\mathcal{T}_\eta^+$ convergence – and we have already remarked that it is possible to approximate arbitrarily closely, in terms of $\mathcal{T}_\eta^+$, any $x \in X_+$ by absolutely continuous x_n. Hence this program does have some hope of working. It will only work, though, in very special cases:

Proposition 6 *Suppose that preferences on X_+ are represented by the utility function*

$$U(x) = \int_0^1 u(x'(t), y(t), t)dt \tag{2}$$

for some (jointly) continuous function u, for all absolutely continuous consumption paths x, where

$$y(t) = \bar{y}(t) + \int_{t-k_1(t)}^{t+k_2(t)} \theta(t, s)dx(t-s) \tag{3}$$

for some continuous function $\bar{y}$, and nontrivial continuous functions k_1, k_2, and θ. If these preferences are continuous in $\mathcal{T}_\eta^+$, relativized to absolutely continuous consumption paths (for any η), then u must have the form $u(z,y,t) = \alpha(y,t)z + \beta(y,t)$, where α and β are jointly continuous functions of y and t.

The point of this is that when the marginal felicity at every point in time depends explicitly on the consumption rate at that time (a case of great interest in the literature), the utility function is not continuous. In the proof of the above proposition, we constructed a sequence of consumption patterns x_n that converge in $\mathcal{T}_\eta$ to x while $U(x_n) \not\to U(x)$ if $u(x', y, t)$ is not linear in x'. Indeed, x_n converges to x in the sup norm topology, so any treatment of preferences for consumption through time that asked for property (a) would not admit the standard models in the literature. The reason should be clear: Unless marginal felicity at any time is independent of the consumption rate at that time, consuming steadily is not well approximated in the standard models by consumption at quickly varying rates. Yet the cumulative consumption paths of such patterns, and thus the corresponding y, can be made close in the sup norm.

What sort of model might then be used to replace the standard model? Following Jones [1984, p.525], we make the following proposal: The problem with the standard models is that the marginal felicity is too sensitive to current consumption. A more apt model would be one where the felicity at time t is derived only from a functional of the recent past consumption and possibly the nearby future consumption; and the current consumption contributes to the current felicity only through this functional.

Proposition 7 *Let $u(z,t)$ be jointly continuous and continuosly differentiable in its first argument. Then*

$$U(x) = \int_0^1 u(y(t), t)dt$$

for all $x \in X_+$, is continuous in $\mathcal{T}_\eta^+$, where $y(t)$ is defined in (3) for some functions k_1 and k_2 such that either k_1 is a constant or k_1 is differentiable with derivative bounded away from zero, and some differentiable function θ, with bounded derivative.

The form given in the above proposition does indicate the sort of "intertemporal substitution of consumption at near-by times" that motivated this study. One sufficient condition for this form to represent uniform proper preferences is for $\bar{y}(t)$ to be bounded away from zero. Other conditions, which we leave for interested readers, can also be generated,.

7 Duality

For the duration of this section, we assume that some pair μ and η as in Section 5 has been fixed.

Mas-Colell and Richard [1991] tells us that it will be possible to establish existence of equilibrium in economies where preferences are increasing and uniformly proper in our topology if we can establish that the topological dual forms a vector lattice in the order dual. Moreover, equilibrium prices for these equilibria will be drawn from the order dual. Hence our aim in this section is to characterize the topological dual space of $(X, \mathcal{T}_\eta)$, which we denote by X_η^*, and to verify that X_η^* does form a vector lattice in the order dual. More important, we will show that X_η^* is composed of absolutely continuous functions of time. This marks the topological distinction between $\mathcal{T}_\eta$ and the weak topology. Although topologically identical on X_+, on the whole of X, $\mathcal{T}_\eta$ is weaker than the weak topology. As a consequence, X_η^* is a proper subset of the space of continuous functions and does not contain any function that is not absolutely continuous.

Our analysis here will be simple applications of the duality results developed in Musielak [1983, section 13] for Orlicz spaces. However, the reader should note that X is a proper subspace of an Orlicz space, and thus X_η^* is a much smaller space than the topological dual of an Orlicz space.

Before proceeding, a definition is in order.

Definition 2 *η is an N-function if*

$$\lim_{z \downarrow 0} \frac{\eta(z)}{z} = 0 \quad \textit{and} \quad \lim_{z \uparrow \infty} \frac{\eta(z)}{z} = \infty.$$

Putting $\eta^*(s) = \sup\{rs - \eta(r) : r > 0\}$, η^* is the conjugate function of η.

Theorem 1 *Let η be an N function and $\phi : X \to \Re$ a linear functional. Then $\phi \in X_\eta^*$ if and only if there exists an absolutely continuous function f on $[0,1]$ so that*

$$\phi(x) = \int_{0-}^{1} f(t)dx(t) \quad \forall x \in X,$$

with $f' \in L^{\eta^*}$, where

$$L^{\eta^*} \equiv \{y : \int_0^1 \eta^*(\gamma|y(t)|)dt + \eta^*(\gamma|y(t)|) \to 0 \textit{ as } \gamma \to 0^+\},$$

and η^* is the conjugate of η. In addition, in the above representation, ϕ is a positive linear functional if and only f is positive.

Thus, if η is an N function, X_η^* can be identified with absolutely continuous functions whose derivatives y satisfy $||y||_{\eta^*} < \infty$. Since Lipschitz continuous functions have bounded derivatives, X_η^* contains all the Lipschitz continuous functions. Note that examples of N functions include $\eta(z) = z^p$ with $p > 1$. In these cases, X_η^* is composed of absolutely continuous functions whose derivatives lie in L^q with $\frac{1}{p} + \frac{1}{q} = 1$.

Now what if η is asymptotically linear and thus is not an N function? This includes of course the linear case $\eta(z) = |z|$. Arguments identical to those used to prove Theorem 1 show immediately that in the linear case X_η^* is the space of absolutely continuous function with bounded derivatives – exactly the space of Lipschitz continuous functions. For η that is asymptotically linear, one quickly shows that $\mathcal{T}_\eta$ is equivalent to $\mathcal{T}_1$. Thus X_η^* is also the space of Lipschitz continuous function. The following theorem, whose proof is omitted, records this fact.

Theorem 2 *Let η be asymptotically linear in that*

$$\lim_{z\to\infty}\frac{\eta(z)}{z}=\alpha>0,$$

for some scalar α. Then $\phi\in X_\eta^*$ if and only if

$$\phi(x)=\int_{0-}^{1}f(t)dx(t)\quad x\in X$$

for some Lipschitz continuous function f. In addition, ϕ is positive if and only if f is positive.

Given that any η is convex with $\eta(0)=0$ and $\lim_{z\to\infty}\eta(z)=\infty$, if η is not an N function, it must be asymptotically linear. Thus Theorems 1 and 2 have completely characterized X_η^* for all η. Note that if equilibrium prices for consumption are representated by an absolutely continuous function of time, $f(t)$, the interest rate exists for almost all t and is equal to $-f'(t)/f(t)$ at time t.

The following proposition shows that X_η^* is a lattice in the order dual.

Proposition 8 *For any η, the space X_η^* is a lattice in the order dual.*

8 Optimization

In this section we formulate the *optimal consumption problem* for an agent whose preferences are continuous in our topologies. This agent also has a preference for final wealth. Formally, the agent's preferences on $X_+\times\Re_+$ are represented by the functional

$$\int_0^1 u(y(t),t)dt+V(w(1^+)),$$

where

$$y(t)=y(0^-)e^{-\beta t}+\int_{0-}^{t}e^{-\beta(t-s)}dx(t) \tag{4}$$

with $y(0^-)\geq 0$ and $\beta>0$,and where $w(1^+)$ denotes the agent's wealth after consumption at time 1. The function u satisfies the conditions of Proposition 7 and is strictly concave in its first argument, and V continuous, strictly concave, and increasing. Note that y is an exponentially weighted average of past consumption and is a special case of the family of general y's defined in (3). Higher values of β imply higher emphasis on the recent past consumption and less emphasis on the distant past consumption.

The single consumption good is the numeraire and the agent starts at time 0 with wealth $w(0)$. At any time t, he allocates his wealth, $w(t)$, between consumption and investment in a riskless asset with instantaneous rate of return $r(t)$, a continuous function of time. For any consumption pattern x, $\tau_1, \tau_2, \ldots$ denote the times of jumps of x.

We investigate the optimal consumption decision of the agent:

$$\sup_{x \in X_+} \int_0^1 u(y(t), t)dt + V(w(1^+))$$

$$\text{s.t.} \quad \begin{aligned} dw(t) &= w(t)r(t)dt - dx(t) && \text{for } t \in (\tau_i, \tau_{i+1}) \quad \text{for all } i, \\ w(\tau_i^+) &= w(\tau_i) - \Delta x(\tau_i) && \text{for all } \tau_i, \\ w(t) &\geq 0, && \text{for } t \in [0,1], \end{aligned} \tag{5}$$

$$\begin{aligned} dy(t) &= \beta[dx(t) - y(t)dt] && \text{for } t \in (\tau_i, \tau_{i+1}) \quad \text{for all } i, \\ y(\tau_i) &= y(\tau_i^-) + \beta \Delta x(\tau_i) && \text{for all } \tau_i. \end{aligned} \tag{6}$$

Note that $w(t)$ is a left continuous function and (5) is just the intertemporal budget constraint. A consumption pattern $x \in X_+$ that satisfies the budget constraint (5) with the initial wealth $w(0)$ will simply called *budget feasible.*

It is clear that at any time t, given $w(t)$ and $y(t^-)$, the subsequent optimal consumption policy is independent of past decisions. Thus, $(w(t), y(t^-))$ are, using the terminology of dynamic programming, the "state variables" for the above dynamic choice problem. We thus define the "indirect utility function" or the "value function" $J(w(t), y(t^-), t)$ as the supremand of the agent's optimization program starting from t given $(w(t), y(t^-))$ and the dynamics of w and y described in (5) and (6).

8.1 Heuristic Derivation of Necessary Conditions

We now use heuristic arguments to derive a set of necessary conditions for a consumption pattern x to be optimal.[16] These conditions will imply that an optimal consumption pattern can have gulps only at $t = 0$. After $t = 0$, consumption occurs only when $\beta J_y = J_w$.

The reader will see that the following derivations deviate somewhat from the standard approach in dynamic programming. This occurs because consumption other than at rates is feasible, and, more important, the standard Bellman's equation will be linear in the "control" and hence the standard arguments are unapplicable.

Suppose x^* is an optimal consumption pattern, and let the corresponding *state variables* be $(w^*(t), y^*(t^-))$ for $t \in [0,1]$. Assume that the value function J is finite and differentiable in w and y. Let $J_w(w, y, t)$ and $J_y(w, y, t)$ denote the partial derivatives of J with respect to w and y, respectively. Along the optimal solution, J should satisfy the following conditions:

(a) Bellman Optimality Principle: The Bellman optimality principle[17] implies that for all times $0 \leq t \leq \tau \leq 1$

$$J(w^*(t), y^*(t^-), t) = \int_t^\tau u(y^*(s), s)ds + J(w^*(\tau), y^*(\tau^-), \tau). \tag{7}$$

[16] These heuristic arguments can be made formal. The reader is refered to Blaquière [1985] for details.

[17] See, for example, Fleming and Rishel [1975, p.??].

Form this, it is immediate that $J(w^*(t), y^*(t^-), t)$ is a continuous function of time even at times of the jumps of x^*. This is so as the jumps of x^* contribute to the agent's satisfaction only through their contribution to y^* and at the moment of a jump the contribution of the jump to the integral of felicities is zero.

For brevity of notation, we will use $\overline{J}(t)$ to denote $J(w^*(t), y^*(t^-), t)$ and similarly use $\overline{J}_w(t)$ and $\overline{J}_y(t)$ to denote the partial derivatives of $\overline{J}(t)$ with respect to w and y, respectively.

(b) Continuous Marginal Value of w and y: Along the optimal policy, given the interest rates, the marginal value of ϵ units of wealth at time $t \in [0, 1)$ must be equal to the marginal value of $\epsilon e^{\int_t^{t+\Delta t} r(s)\,ds}$ units of wealth at time $t + \Delta t$ for small $\Delta t > 0$ or else some changes in x^* will be called for to change w^* and y^*. That is, we must have

$$\overline{J}_w(t)\epsilon = \overline{J}_w(t + \Delta t)\epsilon e^{\int_t^{t+\Delta t} r(s)ds}. \tag{8}$$

Letting $\Delta t \downarrow 0$, we get $\overline{J}_w(t) = \overline{J}_w(t^+)$. Thus $\overline{J}_w(t)$ must be right-continuous on $[0, 1)$. Along the same line of arguments, we show that $\overline{J}_w(t)$ must be left-continuous on $(0, 1]$ and thus $\overline{J}_w(t)$ is continuous on $[0, 1]$.

Similarly, by the Bellman optimality principle, a marginal value of ϵ increase in $y(t^-)$ at t must be equal to its marginal contribution to the felicity function on $(t, t + \Delta t]$ and to the value function at time $t + \Delta t$. That is,[18]

$$\overline{J}_y(t)\epsilon = \int_t^{t+\Delta t} u_y(y^*(s), s)\epsilon e^{-\beta(s-t)}ds + \overline{J}_y(t + \Delta t)\epsilon e^{-\beta\Delta t}. \tag{9}$$

Letting $\Delta t \to 0$, we conclude that $\overline{J}_y(t)$ must be right-continuous on $[0, 1)$. The left-continuity of $\overline{J}_y(t)$ is established by noting that

$$\overline{J}_y(t - \Delta t)\epsilon = \int_{t-\Delta t}^{t} u_y(y^*(s), s)\epsilon e^{-\beta(s-t+\Delta t)}ds + \overline{J}_y(t)\epsilon e^{-\beta\Delta t} \tag{10}$$

and letting $\Delta t \to 0$.

(c) Dynamics of Marginal Values: By the hypothesis that $r(t)$ is continuous in t, (8) implies, as $\Delta t \to 0$, that

$$\frac{d\overline{J}_w(t)}{dt} = -r(t)\overline{J}_w(t).$$

Next note that if $y^*(t)$ is continuous on some time interval $(t - \Delta t, t + \Delta t)$, then, as $\Delta t \to 0$,

$$\frac{\int_t^{t+\Delta t} u_y(y^*(s), s)e^{-\beta(s-t)}ds}{\Delta t} \to u_y(y^*(t), t)$$

and

$$\frac{\int_{t-\Delta t}^{t} u_y(y^*(s), s)e^{-\beta(s-t+\Delta t)}ds}{\Delta t} \to u_y(y^*(t), t).$$

Hence, in a time interval over which x^* has no "gulps", (9) implies, as $\Delta t \to 0$, that

$$\frac{d\overline{J}_y(t)}{dt} = -u_y(y^*(t), t) + \beta\overline{J}_y(t).$$

[18] We use $u_y(y, t)$ to denote the partial derivative of u with respect to y.

Also note that at $t = 1$, $\overline{J}_w(a) = V'(w(a))$ and $\overline{J}_y(a) = 0$.

(d) Appropriate Times for "Gulps": Applying Bellman's optimality principle in (7) on an interval (τ_i, τ_{i+1}) when our solution prescribes consumption at rates, we get the following Bellman's equation:

$$\max_{x'(t)}\Big\{u(y^*(t),t) + \overline{J}_W[r(t)w^*(t) - x'(t)] + \overline{J}_y\beta[x'(t) - y^*(t)] + \overline{J}_t\Big\} = 0 \quad \text{for all } t \in (\tau_i, \tau_{i+1}). \tag{11}$$

Note that the Bellman equation is *linear* in the consumption rate $x'(t)$. Hence, the consumption rate that satisfies (11) is given by:

$$\begin{aligned} x'(t) &= 0 \quad \text{when} \quad \overline{J}_w(t) - \beta\overline{J}_y(t) > 0\,, \\ x'(t) &\in [0,\infty] \quad \text{when} \quad \overline{J}_w(t) - \beta\overline{J}_y(t) = 0\,, \\ x'(t) &= \infty \quad \text{when} \quad \overline{J}_w(t) - \beta\overline{J}_y(t) < 0\,. \end{aligned}$$

The Bellman equation, therefore, suggests that a "gulp" of consumption might be prescribed at any moment τ when $\overline{J}_w(\tau) - \beta\overline{J}_y(\tau) \leq 0$. Suppose that we prescribe a "gulp" at τ when $\overline{J}_w(\tau) - \beta\overline{J}_y(\tau) < 0$. By continuity, $\overline{J}_w(t) - \beta\overline{J}_y(t)$ will be negative on $(\tau - \epsilon, \tau + \epsilon)$ for some $\epsilon > 0$. Hence our candidate policy should prescribe a jump for all points $t \in (\tau - \epsilon, \tau + \epsilon)$. Since this cannot happen as there can only be a countable number of jumps, our policy might prescribe a "gulp" only at a moment τ, on this particular interval, such that $\overline{J}_w(\tau) - \beta\overline{J}_y(\tau) = 0$.

(e) No "Gulps" after $t = 0$: Now suppose that our policy calls for a "gulp" of size Δ at time $\tau > 0$. Note that by right continuity of x^*, we cannot have two successive jumps at any one moment. This, together with the continuity of the marginal values of w and y, implies that there are two strictly positive numbers ϵ_1, ϵ_2 such that:

$$\begin{aligned} \overline{J}_w(t) - \beta\overline{J}_y(t) &\geq 0 \quad \text{on} \quad (\tau - \epsilon_1, \tau)\,, \\ \overline{J}_w(t) - \beta\overline{J}_y(t) &= 0 \quad \text{when} \quad t = \tau\,, \\ \overline{J}_w(t) - \beta\overline{J}_y(t) &\geq 0 \quad \text{on} \quad (\tau, \tau + \epsilon_2)\,. \end{aligned}$$

Assuming that $\overline{J}_W$ and $\overline{J}_y$ are differentiable functions of time, this implies that $\overline{J}_w(t) - \beta\overline{J}_y(t)$ is decreasing just to the left of τ and increasing just to the right of τ. This condition, however, cannot hold for any size of jump Δ. ¿From the dynamics of $\overline{J}_W$ and $\overline{J}_y$, we get that:

$$\frac{d}{dt}[\overline{J}_w(\tau^+) - \beta\overline{J}_y(\tau^+)] - \frac{d}{dt}[\overline{J}_w(\tau) - \beta\overline{J}_y(\tau)] = \beta\Big[u_y(y^*(\tau) + \beta\Delta, \tau) - u_y(y^*(\tau), \tau)\Big]. \tag{12}$$

But by strict concavity of the felicity function in y, this last quantity is negative. Hence, if $\frac{d}{dt}[\overline{J}_W - \beta\overline{J}_y]$ were negative just before the jump, it would be strictly negative just after the jump, thus violating the condition that $\overline{J}_W - \beta\overline{J}_y$ be increasing just after a "gulp". Note that our analysis is not valid for $t = 0$ and $t = 1$, thus x^* does not have any gulps, except possibly at $t = 0$ and $t = 1$. A gulp at $t = 1$, however, is clearly sub-optimal since the agent cannot derive any consumption satisfaction from it and the final wealth is also reduced.

All the above leads us to conclude that x^* should take the form of a possible "gulp" at $t = 0$, followed by periods of consumption mixed with periods of no consumption. If we should prescribe a jump $\Delta x^*(0)$ at $t = 0$, then the size of the jump should maximize the value function immediately after the jump. That is, the jump size $\Delta x^*(0)$ should solve the following problem

$$\max_{\epsilon} J(w(0) - \epsilon, y(0^-) + \beta\epsilon, 0^+).$$

¿From the first order condition, we then know that $\overline{J}_w(0^+) = \beta\overline{J}_y(0^+)$. Moreover, after the initial possible jump at $t = 0$, the following condition should be satisfied for all $t \in (0, 1]$; at times of no consumption:

$$u(y^*(t), t) + \overline{J}_W r(t) w^*(t) - \overline{J}_y \beta y^*(t) + \overline{J}_t = 0, \tag{13}$$

together with

$$\beta\overline{J}_y(t) - \overline{J}_w(t) < 0\,, \tag{14}$$

and at the times when consumption occurs at rates, we must have

$$[\overline{J}_w(t) - \beta\overline{J}_y(t)]x^{*\prime}(t) = 0. \tag{15}$$

More generally, consumption only occurs when $\overline{J}_w(t) = \beta\overline{J}_y(t)$. During these times the strategy of no consumption is sub-optimal and hence

$$u(y^*(t), t) + \overline{J}_W r(t) w^*(t) - \overline{J}_y \beta y^*(t) + \overline{J}_t \leq 0. \tag{16}$$

The necessary conditions on J can be expressed more compactly by stating that the value function should at all times satisfy the following *differential inequality*:

$$\max\Big\{u(y, t) + \overline{J}_w w r(t) - \overline{J}_y \beta y + \overline{J}_t,\ \beta\overline{J}_y - \overline{J}_w\Big\} = 0\,. \tag{17}$$

Besides this differential inequality, it is clear that J must also satisfy two boundary conditions:

$$\begin{aligned} J(w, y, 1) &= V(w) \quad \text{and} \\ J(0, y, t) &= \textstyle\int_t^1 u(ye^{-\beta(s-t)}, s)\, ds + V(0). \end{aligned} \tag{18}$$

The first of these boundary conditions follows because at $t = 1$ no consumption is optimal. The second is a consequence of the positive wealth constraint – when the agent's wealth reaches zero, he must henceforth refrain from consumption.

We will show in the following section that the differential inequality plus the two boundary conditions are *almost sufficient* for optimality. This deserves some commentary for readers familiar with dynamic programming. We observe that if only consumption rates are allowed, standard results in dynamic programming show that the Bellman equation, together with the appropriate boundary conditions, are not only necessary but also sufficient. This occurs because the optimal consumption rates would be functions of J. Readers unfamiliar with dynamic programming are referred to Fleming and Rishel (1975, pp.??), for example, for a readable account. In our case, however, consumption does not have to be at rates. Hence optimal consumption may not be readily determined once we know J. It turns out that the sufficient conditions for optimality include additional conditions on the optimal policy beyond those imposed on J. These additional conditions are obtained from the preceeding heuristic analysis.

8.2 Sufficiency

We provide two results. First, we show that the solution to (17), with the boundary conditions (18), is an upper bound on the value function J. Second, we give a verification theorem in which we show that if a consumption pattern x^* satisfies certain conditions, then the life-time utility derived from x^* is given by the solution of (17) and (18). Hence J coincides with the upper bound on utility and x^* is optimal.

Proposition 9 *Assume that there exists a differentiable function* $J^*(w,y,t)\colon \Re_+ \times \Re_+ \times [0,1] \to \Re \cup \{-\infty\}$, *concave in* w *and* y, *whose partial derivatives are continuous functions of time, that solves (17) with the boundary conditions (18). Then* $J(w(0),y(0^-),0) \leq J^*(w(0),y(0^-),0)$.

As we mentioned before, the differential inequality and the boundary conditions are almost sufficient in that they produce a function that bounds the value function J from above. The problem is, of course, that a budget feasible consumption pattern that attains J^* may not exist. Note that this is not a problem in the standard models as consumption is at rates and the consumption pattern that produces J^* can be calculated as a function of J^*, and thus $J^* = J$. The following verification theorem gives a set of conditions sufficient for a consumption pattern x^* to attain J^*. In particular, they stipulate that gulps can occur only at $t = 0$ and afterwards, consumption occurs only at times where $J^*_w = \beta J^*_y$.

Theorem 3 *Let* J^* *satisfy the conditions of Proposition 9. Suppose further that there exists a budget feasible consumption pattern* x^*, *which is continuous on* $(0,1]$, *with a possible jump* $\Delta x^*(0)$, *and whose associated state variables are* $w^*(t), y^*(t^-)$, *such that for all* $t \in (0,1]$:

$$u(y^*,t) + J^*_w w^* r(t) - J^*_y \beta y^* + J^*_t \;=\; 0 \quad \textit{and} \tag{19}$$

$$\int_t^{t+\epsilon} [J^*_w(w^*,y^*,t) - \beta J^*_y(w^*,y^*,t)]\, dx^*(t) \;=\; 0 \quad \textit{for all} \quad \epsilon > 0\,, \tag{20}$$

and

$$J^*(w(0),y(0^-),0) = J^*(w(0) - \Delta x^*(0), y(0^-) + \beta \Delta x^*(0), 0). \tag{21}$$

Then $J(w(0),y(0^-),0) = J^*(w(0),y(0^-),0)$ *and* x^* *is the optimal consumption pattern.*

8.3 Example

To demonstrate how to use the verification theorem in constructing an optimal consumption policy, we provide a closed form solution for the optimal consumption problem in an infinite horizon economy with a constant interest rate r. The agent seeks to maximize

$$\int_0^\infty \frac{1}{\alpha} e^{-\delta t} y(t)^\alpha \, dt$$

given $w(0)$ and $y(0^-)$ and where $y(t)$ is given by equation (4). We take $\alpha < 1$ and $\delta > 0$.

This departure from our finite horizon setup in earlier sections is for tractability of calculations. Our topologies generalize in a standard way to this infinite horizon economy by replacing the Lebesgue measure on $[0, \infty)$ with a finite measure, m, that is absolutely continuous with respect to the Lebesgue measure. For example, we can take $m(A) = \int_A e^{-\delta t} dt$ for all Lebesgue measureable sets A. We leave the details to the reader but one can easily show that the agent's preferences for consumption on $[0, \infty)$ specified above satisfy the notion of continuity that we advocate.

The verification theorem for this case, which is simpler than in the case of finite horizon programs, needs a minor modification. Simplicity follows from the observation that the value function is separable in time and in (w, y). With a slight abuse of notation, the value function should have the form $e^{-\delta t} J(w, y)$ at time t for some function J. Then the differential inequality becomes

$$max\left\{\frac{y^\alpha}{\alpha} + J_w wr - J_y \beta y - \delta J,\ \beta J_y - J_w\right\} = 0\,, \tag{22}$$

which is independent of t. The second boundary condition in (18) simplifies to

$$J(0, y) = \int_0^\infty e^{-\delta s} \frac{1}{\alpha} (y e^{-\beta(s-t)})^\alpha \, ds = \frac{y^\alpha}{\alpha(\alpha\beta + \delta)}. \tag{23}$$

The first boundary condition in (18) needs modification as there is no final date and, hence, the agent does not have a preference for final wealth. We again leave the details to the reader, and simply assert that this boundary condition is replaced by

$$\lim_{t \uparrow \infty} e^{-\delta t} J(w(t), y(t^-)) = 0\,, \tag{24}$$

for all budget-feasible policies. Readers familiar with infinite horizon programs would recognize this as a "transversality condition".

We will show that the key feature of the solution is a critical ratio $k^* > 0$ of wealth w to average past consumption y. If the agent starts at time zero with $\frac{w(0)}{y(0^-)}$ strictly less than k^*, then the optimal behavior is to invest all the wealth in the riskless asset and wait while w increases and y declines till the ratio $\frac{w}{y}$ reaches k^*. From then on, the agent consumes at the rate which keeps the ratio $\frac{w}{y}$ equal to k^* forever. If, on the other hand, the agent starts with $\frac{w(0)}{y(0^-)}$ strictly greater than k^*, then the optimal behavior is to take a "gulp" of consumption, reducing w and increasing y, to bring $\frac{w}{y}$ immediately to k^*. Following this gulp, the optimal consumption occurs at the rate that keeps the ratio $\frac{w}{y}$ equal to k^* forever.

To ensure existence of a solution, we make the following assumption. This assumption guarantees, among other things, that the value function will be concave.

Assumption 1 *The parameters of the problem satisfy*

$$\delta + \alpha\beta > 0, \qquad \delta > \alpha r \quad \text{and} \quad (1-\alpha)\beta > \delta - r\,.$$

Now we preceed as follows. First we investigate all the k-ratio consumption policies: the policy of consumption that keeps $\frac{w}{y}$ equal to $k > 0$ forever, after an initial "gulp" if $\frac{w(0)}{y(0^-)} > k$, or after a period of no consumption if $\frac{w(0)}{y(0^-)} < k$. The agent's life-time utility from such a policy, denoted $J^k(w,y)$, satisfies many of the sufficient conditions for optimality:

Proposition 10 *$J^k(w,y)$ is given by:*

$$J^k(w,y) = \begin{cases} \frac{y^\alpha}{\alpha(\delta+\alpha\beta)} + \frac{y^\alpha}{\alpha}\left[\frac{w}{ky}\right]^{\frac{\delta+\alpha\beta}{r+\beta}}\mathcal{A} & \text{if } \frac{w}{y} \le k \\ \frac{1}{\alpha}\left(\frac{y+\beta w}{1+\beta k}\right)^\alpha \mathcal{B} & \text{if } \frac{w}{y} \ge k \end{cases}$$

where $\mathcal{A} = (\delta - \alpha\beta(\frac{rk-1}{\beta k+1}))^{-1} - (\delta+\alpha\beta)^{-1}$ and $\mathcal{B} = (\delta - \alpha\beta(\frac{rk-1}{\beta k+1}))^{-1}$. Furthermore, J^k is continuous, concave, has continuous first derivatives and satisfies

$$\begin{aligned} \frac{y^\alpha}{\alpha} + rwJ^k_w - \beta y J^k_y - \delta J^k = 0 \quad &\textit{if} \quad \frac{w}{y} \le k\,, \\ J^k_w - \beta J^k_y = 0 \quad &\textit{if} \quad \frac{w}{y} \ge k\,, \end{aligned} \tag{25}$$

together with the boundary condition (23) and the transversality condition (24).

However, there is exactly one k so that J^k satisfies the differential inequality of (22). This k-ratio policy is then the optimal policy:

Proposition 11 *Let*

$$k^* = \frac{\frac{r-\delta}{\beta} + (1-\alpha)}{\delta - \alpha r}\,,$$

and note that $k^* > 0$ by Assumption 1. J^{k^*} satisfies the differential inequality (22). It then follows that the k^*-ratio policy is the optimal solution for the agent's problem.

The reader can easily verify that k^* is the unique ratio with the property that the associated J^{k^*} is *twice* continuously differentiable over the whole domain (w,y).

We summarize the solution in the following proposition.

Proposition 12 *The optimal solution is to consume at rate $c(t) = \frac{\delta-\alpha r}{1-\alpha}w(t)$, after an initial gulp of size $\Delta = \frac{1}{(1-\alpha)(r+\beta)}\left[(\delta-\alpha r)w(0) + (\frac{r-\delta}{\beta} + (1-\alpha))y(0^-)\right]$, if $\frac{w(0)}{y(0^-)} > k^*$, or after a period of no consumption of length $t^* = \log\left[\frac{k^* y(0^-)}{w(0)}\right]^{\frac{1}{r+\beta}}$, if $\frac{w(0)}{y(0^-)} < k^*$.*

It is interesting to compare this consumption policy to that adopted by a consumer who starts with the same initial wealth $w(0)$ and the same past consumption experience $y(0^-)$ who seeks to maximize:

$$\int_0^\infty e^{-\delta t}\frac{c(t)^\alpha}{\alpha}\,dt\,,$$

where $c(t)$ is the consumption rate at t and we assume that $\delta > \alpha r$ and $1 - \alpha > 0$. It can be easily verified that his optimal consumption rate is a linear function of his wealth: $c(t) = \frac{\delta - \alpha r}{1-\alpha} w(t)$.

The propensity to consume is the same for agents with time-additive and non-time-additive preferences. However, the quantity of consumption and the trajectory of wealth are different. Consider the case when $\delta = r$. An agent with time-additive preferences will consume $c(t) = rw(t)$, and hence his wealth at any time will be a constant equal to his initial wealth; see figure **3**. A consumer with non-time additive preferences, however, will either start by consuming a "gulp" of size $\frac{rw(0)+y(0^-)}{r+\beta}$, if his wealth is too high relative to $y(0^-)$, and then keep the level of his wealth constant forever, or he will wait for $t^* = \frac{1}{r+\beta}\left[\log \frac{y(0^-)}{rw(0)}\right]$, if his wealth is too low relative to $y(0^-)$, and then start consuming the interest on his wealth forever.

The parameter β, which controls the rate of "decay" of previous consumption, determines the size of the initial "gulp" or the length of the initial waiting period, and hence the level of the subsequent constant wealth. The higher the value of β, the smaller the size of the initial gulp, if the solution calls for a "gulp", and the shorter the initial waiting period, in case the solution entails a waiting period. This is an intuitive result since a higher β implies that past consumption experience has "smaller" effect on current, and hence total, satisfaction. Therefore, when the optimal solution entails an initial waiting period to accumulate wealth, the effect of $y(0^-)$ decays faster with higher β leading to shorter waiting period before the ratio $\frac{w}{y}$ reaches its critical value. Similarly, when the optimal solution calls for an initial "gulp", higher β means that the effect of the "gulp" will decay faster and hence the "gulp" contributes less to future satisfaction. Therefore, agents with higher β choose smaller initial "gulps".

9 References

1. Schaefer, H., *Banach Lattices and Positive Operators*, Springer-Verlag, Berlin, 1974.

2. P. Billingsley (1968), *Convergence of Probability Measures*, John Wiley and Sons, New York.

3. A. Blaquière (1985), "Impulsive Optimal Control with Finite or Infinite Time Horizon", *Journal of Optimization Theory and Applications* 46.

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

5. P. Fishburn (1970), *Utility Theory for Decision Making*, John Wiley and Sons, New York.

6. W. Fleming and R. Rishel (1975), *Deterministic and Stochastic Optimal Control*, Springer-Verlag.

7. A. Hindy and C. Huang (1990a), On Intertemporal Preferences for Uncertain Consumption: A Continuous Time Approach, Sloan School of Management, Massachusetts Institute of Technology.

8. A. Hindy and C. Huang (1990b), Optimal Consumption and Portfolio Rules with Local Substitution, Graduate School of Business, Stanford University.

9. C. Huang, Information Structure and Equilibrium Asset Prices, *Journal of Economic Theory* **35**:33–71, 1985.

10. R. B. Holmes (1975), *Geometric Functional Analysis and its Applications*, Springer-Verlag, Berlin.

11. L. Jones (1981), *Commodity Differentiation in Economic Modelling*, Unpublished Ph.D. Dissertation, University of California, Berkeley.

12. L. Jones (1984), "A Competitive Model of Product Differentiation", *Econometrica* 52, 507-530.

13. M. Magill (1981), "Infinite Horizon Programs", *Econometrica* 49, 679-711.

14. A. Mas Colell (1975), "A Model of Equilibrium with Differentiated Commodities", *Journal of Mathematical Economics* 2, 263-296.

15. A. Mas Colell (1986), "The Price Equilibrium Existence Problem in Topological Vector Lattices", *Econometrica* 54, 1039-1053.

16. A. Mas Colell and S. Richard (1991), "A New Approach to the Existence of Equilibria in Vector Lattices", *Journal of Economic Theory* 53, 1-11.

17. J. Musielak (1983), *Orlicz Spaces and Modular Spaces*, Springer-Verlag, Berlin.

18. H. Ryder and G. Heal (1973), "Optimal Growth with Intertemporally Dependent Preferences", *Review of Economic Studies* 40, 1-31.

19. S. Sundaresan (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth", *Review of Financial Economics* 2, 73-89.

Appendix

A Proofs

Fix a continuous, strictly increasing and convex function η from $[0,\infty)$ to $[0,\infty)$ with $\eta(0)=0$ and $\eta(\infty)=\infty$. Define $I_\eta(x,a)$, for $x \in X$ and $a \in (0,\infty)$, by

$$I_\eta(x,a) = \int_0^1 \eta(|x(t)|/a)dt + \eta(|x(1)|/a).$$

Since functions x are of bounded variation, the integral is certainly well defined for every x and a.

We record a lemma first.

Lemma A *A sequence $\{x_n\}$ in X converges to zero in $\|\cdot\|_\eta$ if and only if, for every scalar $a>0$, $\lim_n I_\eta(x_n,a)=0$.*

PROOF. See Musielak [1983, theorem 1.6]. ∎

PROOF OF PROPOSITION 1. We want to show that order intervals are weakly compact. In fact, we can show sequential compactness in the norm topology itself. The proof follows standard lines, and so we will provide only a sketch.

For $x^* \geq x_*$ let I be the order interval $[x_*, x^*]$. We will assume that $x_* \equiv 0$, this is without loss of generality because all of our topologies are linear. Let Q be any countable dense set of points in $[0,1)$ that excludes all points of discontinuity of x^*.

Now $x \in I$ if and only if x is nondecreasing and $x^* - x$ is nondecreasing. Let $\{x_n\}$ be any sequence out of I. By the usual diagonalization procedure, we can produce a function y such that, along a subsequence (which we will hereafter not bother to mention), $\lim_n x_n(t) = y(t)$ at all arguments $t \in Q$. Since each x_n is nondecreasing, so is y. Define a function x by $x(t) = \lim_{s\downarrow t} y(t)$ for $t \in [0,1)$, where the limit is taken over sequences converging to t from (strictly) above along the set Q and $x(1) = \lim_n x_n(1)$. Since y is nondecreasing, so is x. Moreover, y is right continuous (and has left limits) by construction. Clearly $\lim_n x_n(t) = x(t)$ at all points t that are not points of discontinuity of x – that is, for almost every t. And $\lim_n x_n(1) = x(1)$ by construction. Since all the x_n and x are bounded above by $x^*(1)$, we can apply dominated convergence and Lemma A to conclude that x_n approaches x in $\mathcal{T}_\eta$.

We are done once we show that $x^* \geq x \geq 0$. We have already remarked that x is nondecreasing (i.e., $x \geq 0$). So let us carefully sketch the other half: If $x^*(t) - x(t) < x^*(t') - x(t')$ for $t > t'$, then right continuity of both functions implies that we can find values $q > q'$ from Q such that $x^*(q) - x(q) < x^*(q') - x(q')$. (The case of $t=1$ is left to the reader.) By increasing q' slightly to be some $q'' < q$ if need be (if q' is a point of discontinuity of x, this gives $x^*(q) - y(q) < x^*(q'') - y(q'')$ for $q > q''$. Hence for sufficiently large n, $x^*(q) - x_n(q) < x^*(q'') - x_n(q'')$. Which is a contradiction. ∎

PROOF OF PROPOSITION 2. We note first that for any η, $\mathcal{T}_\eta$ is topologically equivalent to $\mathcal{T}_{a\eta}$ for any scalar a. Hence, without loss of generality, we can assume that $\eta(1)=1$. Note that this implies, by the convexity of η, that $\eta(r) \geq r$ for all $r \geq 1$. Now it is evident that $\|\cdot\|_\eta \leq \|\cdot\|_{\eta'}$ for $\eta \geq \eta'$. Hence, if for fixed η (with $\eta(1)=1$) we let $\eta_1(r) = \max\{|r|, \eta(r)\}$ (which is convex as it is the maximum of two convex functions), we have that $\mathcal{T}$ is no stronger than $\mathcal{T}_{\eta_1}$, for any η. The proof of the first part of Proposition 2 is thus shown if we show that, for any η, $\mathcal{T}_\eta$ is equivalent to $\mathcal{T}_{\eta_1}$. Of course, $\mathcal{T}_\eta$ can be no stronger, so we only need to show that if $x_n \to 0 \quad \mathcal{T}_\eta$, then $x_n \to 0 \quad \mathcal{T}_{\eta_1}$. For this use Lemma A: If $x_n \to 0 \quad \mathcal{T}_\eta$, then for every $a > 0$, $I_\eta(x_n, a) \to 0$. Suppose that for some fixed a, $I_{\eta_1}(x_n, a)$ does not go to zero – look along some subsequence where the limit exceeds $4/M$ for some integer M. Now $I_\eta(x_n, a)$ can be written as the sum of three terms:

$$\int \eta(x_n(t)/a) 1_{\{|x_n(t)|/a>1\}} dt + \int \eta(x_n(t)/a) 1_{\{|x_n(t)|/a\leq 1\}} dt + \eta(x_n(1)/a).$$

Since each term is positive, each must go to zero in n. Pick n large enough so that the first term is less than $1/M$. If the second is to go to zero, since η is strictly positive except at the argument zero, we can find n large enough so that the Lesbesgue measure of the set $\{t \in [0,1] : 1 \geq x_n(t)/a \geq 1/M\} < 1/M$. And for the third term to go to zero, it is necessary that, for n large enough, $x_n(1)/a < 1/M$. Picking n large enough so that all this happen at once, when we expand $I_{\eta_1}(x_n, a)$ in analogous fashion, we will find that (i) the first term in the sum is less than $1/M$, since it is identical with the first term above, (ii) the third term is less than $1/M$, directly from our selection of n, and (iii) the second term is less than $2/M$, when we overestimate $\eta_1(x_n(t)/a)$ in the second term with the function that is $1/M$ if the argument is less than $1/M$ and 1 is the argument is between $1/M$ and 1. Since this is so for all n above some level, we contradict the supposition.

Next we must show that, for any η, $\mathcal{T}_\eta$ is no stronger than the total variatio n and Skorohod topologies. Since total variation is a (linear) norm topology, it suffices to show that if $x_n \to 0$ in total variation, then $x_n \to 0 \quad \mathcal{T}_\eta$. Apply Lemma A – if $x_n \to 0$ in total variation, then the same is true for x_n/a for any scalar $a > 0$. The total variation norm is larger than the sup norm, so the same must be true in the sup norm. And dominated convergence gives the result. As for the Skorohod topology, Billingsley [1968, page 123, problem 2] states the result for the L^1 norm topology, which immediately gives the result for $\mathcal{T}$. For general η, note that if $x_n \to x$ in the Skorohod topology, then for n sufficiently large, $\sup_t |x_n(t)| \leq \sup_t |x(t)| + 1/n$. Fixing any a, then, and evaluating $I_\eta(x_n - x, a)$, we know that the arguments of η will be bounded above by $K/a = 2\sup_t |x(t)|/a + 2/(an)$. Convergence of $I_\eta(x_n - x, a)$ to zero, then, follows immediately if $I_1(x_n - x, a)$ converges to zero, where I_1 means the integral using the function $\eta^\star(t) = t$ in place of η, since we can overestimate $\eta(r)$ with $r\eta(K/a)$. Convergence of $I_1(x_n - x, a)$ to zero follows from the problem in Billingsley, and we are done. ∎

Proof of Proposition 4. We need to show that if x_n and y_n are as in the (2.3), then $x_n - y_n \to 0 \quad \mathcal{T}_\eta^+$. Fixing a scalar $a > 0$, the techniques employed just above show that an upper bound on $I_\eta(x_n - y_n, a)$ is $\frac{1}{n}\eta((3x_n(1) + 2)/a) + \eta((x_n(1) - y_n(1))/a)$. The second term vanishes for any a since $x_n(1) - y_n(1) < p(x_n, y_n) \leq 1/n$. And the first term vanishes for any a: By assumption, $x_n(1)/\mu(n) \to 0$, and thus $(3x_n(1) + 2)/(a\mu(n)) \to 0$. Hence for any ϵ there is N_ϵ such that for all $n > N_\epsilon$, $(3x_n(1) + 2)/a < \epsilon\mu(n)$. Apply η to both sides (η is monotone), and use the shape of η, for $\epsilon < 1$ to conclude that $\eta((3x_n(1) + 2)/a) < \eta(\epsilon\mu(n)) \leq \epsilon\eta(\mu(n)) = \epsilon n$, for all $n > N_\epsilon$. ∎

Proof of Proposition 5. In X_+, the weak topology is described by the Prohorov metric. Accordingly, the topology is sequential, and it suffices to show that a sequence is convergent in one of the two topologies if and only if it is convergent in the other. Suppose first that $x_n \to x$ in $\mathcal{T}_\eta$. Fix $\epsilon > 0$, and, applying Lemma A, pick N sufficiently large so that $I_\eta(x_n - x, \epsilon) < \epsilon$ for $n > N$. We claim that this implies that $p(x_n, x) < \epsilon$. For if not, then there exists some t such that $x_n(t + \epsilon) + \epsilon < x(t)$ or such that $x_n(t - \epsilon) - \epsilon > x(t)$. Suppose that the former is true, and suppose further that $t \in [0, 1 - \epsilon]$. See figure A1. In calculating $I_\eta(x_n - x, \epsilon)$, when we integrate $\eta((x_n(t) - x(t))\epsilon)$ over $[t, t + \epsilon]$, the integrand exceeds $\eta(1) = 1$. (Recall that $\eta(1) = 1$ is wlog.) Hence the integral is at least ϵ, a contradiction. If

$t > 1-\epsilon$, then $x(1) - x_n(1) > \epsilon$, and the last (summed) term in $I_\eta(x_n - x, \epsilon)$ will exceed 1, again a contradiction. The case where, for some t, $x_n(t-\epsilon) - \epsilon > x(t)$ is handled similarly.

Conversely, suppose that $x_n \to x$ in the weak topology. Fix a constant $a > 0$. We will show that $I_\eta(x_n - x, a) \to 0$, so that $x_n \to x$ $\mathcal{T}_\eta^+$ follows by Lemma A. If $p(x_n, x) \to 0$, then it is immediate that $|x_n(1) - x(1)| \to 0$, so that the final term in $I_\eta(x_n, a)$ converges to zero in n. We need, therefore, only be concerned with the first (integral) term. For a given integer m, let $N(m)$ be sufficiently large so that for $n > N(m)$, $p(x_n, x) < 1/m$. Recalling that both x_n and x are nondecreasing, we can bound $\int_0^1 \eta\left(\frac{x_n(t)-x(t)}{a}\right) dt$ above by

$$\frac{1}{m}\sum_{k=1}^{m}\eta\left(\frac{\max\{x(k/m), x_n(k/m)\} - \min\{x((k-1)/m), x_n((k-1)/m)\}}{a}\right).$$

Now since $p(x_n, x) < 1/m$, $\max\{x(t), x_n(t)\} < x(t+1/m) + 1/m$ and $\min\{x(t), x_n(t)\} > x(t-1/m) - 1/m$, so the summation given is bounded above by

$$\frac{1}{m}\sum_{k=1}^{m}\eta\left(\frac{x((k+1)/m) - x((k-2)/m) + 2/m}{a}\right).$$

¿From the shape of η (convex, equal to zero at zero, and increasing), this in turn is bounded above by

$$\frac{1}{m}\eta\left(\sum_{k=1}^{m}\frac{x((k+1)/m) - x((k-2)/m) + 2/m}{a}\right) \le \frac{1}{m}\eta\left(\frac{3x(1)+2}{a}\right).$$

For fixed a, the argument of η is fixed, so this upper bound vanishes with m, proving our assertion. ∎

Proof of Proposition 6. Suppose not. By joint continuity, there exists two scalars r and $\hat{r}$, $y \in \Re_+$ and $t \in [0,1]$ such that $u((r+\hat{r})/2, y, t) \neq u(r, y, t)/2 + u(\hat{r}, y, t)/2$. Without loss of generality, assume that $u((r+\hat{r})/2, y, t) > u(r, y, t)/2 + u(\hat{r}, y, t)/2$. By joint continuity again, there exists $\epsilon > 0$ and some interval I containing t, and some interval Λ containing y such that $u((r+\hat{r})/2, w, s) - \epsilon > u(r, w, s)/2 + u(\hat{r}, w, s)/2$ for all $s \in I$ and $w \in \Lambda$. Consider a sequence of absolutely continuous consumption patterns constructed as follows: Off of I, consume the pattern $m(s)$, to be described shortly, in each x_n. On I, subdivide I into $2n$ equal sized intervals, and consume at rate r on the even subintervals and $\hat{r}$ on the odd. This sequence of consumption patterns converges in the Prohorov metric to the consumption pattern x that has $x'(s) = m(s)$ off I and $x'(s) = (r+\hat{r})/2$ on I. Choose $m(s)$ such that $y(s) \in \Lambda$ for all $s \in [0,1]$. Note that this is always feasible. By Proposition 5, $x_n \to x$ in $\mathcal{T}_\varphi$. Proposition 7 shows that $\lim_{n\to\infty}\int_0^1 u(x', y_n, t)\,dt = \int_0^1 u(x', y, t)\,dt$. But for any y_n, we have:

$$\int_0^1 u(x', y_n, t)\,dt > \int_0^1 u(x_n', y_n, t)\,dt + \epsilon\lambda(I).$$

Taking limits of both sides as $n \to \infty$, we get:

$$\int_0^1 u(x', y, t)\,dt = \lim_{n\to\infty}\int_0^1 u(x', y_n, t)\,dt \ge \lim_{n\to\infty}\int_0^1 u(x_n', y_n, t)\,dt + \epsilon\lambda(I).$$

Hence $U(x) \ge \lim_{n\to\infty} U(x_n) + \epsilon\lambda(I)$, where $\lambda(I)$ is the Lebesgue measure of I. Thus V is not continuous in $\mathcal{T}_\varphi$. Finally, note that if $u(z, y, t) = \alpha(y, t)r + \beta(y, t)$ for $(z, y) \in$

$(0,\infty)\times(0,\infty)$ and if u is jointly continuous, then α and β must be continuous functions. ∎

Proof of Theorem 1. We prove the necessity part of the assertion first. Suppose that ϕ is a $\mathcal{T}_\eta$ continuous linear functional on X. We extend ϕ using the Hahn-Banach theorem to L^η to be a linear functional continuous in $\mathcal{T}_\eta$. Denote this extension still by ϕ. Musielak [1983, theorem 13.17] shows that ϕ can be represented as

$$\phi(z)=\int_0^1 z(t)y(t)dt+z(1)y(1) \tag{26}$$

for some $y\in L^{\eta^*}$.

Now fix $x\in X$. Since x is a right-continuous bounded variation function, we can use integration by parts to rewrite $\phi(x)$ as follows. First, define $f(t)\equiv y(1)+\int_t^1 y(s)ds$ $\forall t\in[0,1]$. Note that f is absolutely continuous with $f'(t)=-y(t)$ and $f(1)=y(1)$. Then integration by parts gives

$$\int_0^1 x(t)y(t)dt+x(1)y(1)=-\int_0^1 x(t)df(t)+x(1)f(1)=\int_{0-}^1 f(t)dx(t).$$

We have thus proved the only if part as $x\in X$ is an arbitrary element of X.

Next we turn to the sufficiency part. First reversing the above integration by parts to show that ϕ can be written as (26) with a $y\in L^{\eta^*}$. Musielak [1983, corollary 13.14] then shows that ϕ is continuous in $\mathcal{T}_\eta$.

The last assertion is obvious. ∎

Proof of Proposition 7. Fix a topology $\mathcal{T}_\varphi$. Let x_n be a sequence of consumption plans that converge to x in $\mathcal{T}_\varphi$. Assume, without loss of generality, that $\sup_n\|X_n\|_\varphi<\infty$. First observe that if $x_n\to x$ in $\mathcal{T}_\varphi$, then $x_n\to x$ in $\mathcal{T}$. Hence,

$$\lim_{n\to\infty}\left[\int_0^1|x_n(t)-x(t)|\,dt+|x_n(1)-x(1)|\right]=0. \tag{27}$$

Now consider $y_n(t)-y(t)=\int_{t-k_1(t)}^{t+k_2(t)}\theta(t,s)\left(dx_n(t-s)-dx(t-s)\right)$. Integrating by parts, we get:

$$y_n(t)-y(t)=-\Big(x_n(k_1(t))-x(k_1(t))\Big)\theta(t,t-k_1(t))-\int_{t-k_1(t)}^{t+k_2(t)}\Big(x_n(t-s)-x(t-s)\Big)\theta_2(t,s)\,ds, \tag{28}$$

where $\theta_2(t,s)$ denotes the partial derivative of $\theta(t,s)$ with respect to its second argument. But since both θ and its derivative are bounded, the left-hand side in (28) is bounded by $M_1|x_n(k_1(t))-x(k_1(t))|+M_2[\int_0^1|x_n(t-s)-x(t-s)|\,ds]$, where M_1 and M_2 are two constants.

¿From this, we can easily conclude that $\lim_{n\to\infty}\left[\int_0^1|y_n(t)-y(t)|\,dt\right]$ is less than or equal to $M_1\lim_{n\to\infty}\left[\int_0^1|x_n(t)-x(t)|\,dt\right]+M_2\lim_{n\to\infty}\left[\int_0^1\int_0^1|x_n(t-s)-x(t-s)|\,ds\,dt\right]$, in the case when $k_1(t)$ is a constant, $\frac{M_1}{c_1}\lim_{n\to\infty}\left[\int_0^1|x_n(t)-x(t)|\,dt\right]+M_2\lim_{n\to\infty}\left[\int_0^1\int_0^1|x_n(t-s)-x(t-s)|\,ds\,dt\right]$, in the case when $0<c_1\le|k_1'(t)|\le c_2$. In both cases, $\lim_{n\to\infty}\left[\int_0^1|y_n(t)-y(t)|\,dt\right]=0$. This follows from (27) and the uniform convergence theorem applied to the inside intergral in the second term of the previous expression.

Now consider the convergence of $u(y_n)$. First, assume that u is Lipschitz continuous in y. It then follows that $|u(y_n) - u(y)| \leq \alpha|y_n - y|$, where α is independent of y. It then follows that $\lim_{n\to\infty}\left[\int_0^1 |u(y_n) - u(y)|\,dt\right] \leq \alpha \lim_{n\to\infty}\left[\int_0^1 |y_n(t) - y(t)|\,dt\right] = 0$. It then follows that $\mathcal{U}(x_n)$ converge to $\mathcal{U}(x)$.

Next consider the case when u is merely continuously differentiable in y. For every integer m, define $u_m(y) = u(\frac{1}{m})\chi_{[0,\frac{1}{m}]} + u(y)\chi_{[\frac{1}{m}\leq y\leq m]} + u(m)\chi_{[m,\infty)}$. For a fixed m, the function u_m is Lipschitz continuous in y, and hence the arguments in the previous paragraph show that $\mathcal{U}_m(x_n)$ converge to $\mathcal{U}_m(x)$, where $\mathcal{U}_m(x)$ is defined exactly as $\mathcal{U}(x)$, except that u is replaced by u_m. Now fix x and the corresponding y. Let $A \equiv \{t \in [0,1]\colon u(y(t),t) \geq 0\}$ and let $A^c \equiv [0,1] \setminus A$. Note that either A or A^c might be of measure zero. It is clear that $u_m(y(t),t)$ converge monotonically to $u(y(t),t)$ on both A and A^c as $m \to \infty$. Applying the monotone convergence theorem, we conclude that $\mathcal{U}_m(x) \equiv \int_A u_m(y(t),t)\,dt + \int_{A^c} u_m(y(t),t)\,dt$ converge to $\mathcal{U}(x)$ as $m \to \infty$. It then follows that:

$$\lim_{n\to\infty} \mathcal{U}(x_n) = \lim_{m,n\to\infty} \mathcal{U}_m(x_n) = \lim_{m\to\infty} \mathcal{U}_m(x) = \mathcal{U}(x).$$

∎

PROOF OF PROPOSITION 8. Theorems 1 and 2 establish that the cone defining the order dual corresponds to nonnegative functions f. Hence in the order dual, the join of f and g (for $f, g \in X_\eta^*$ in the obvious embedding) would be given by $\max\{f,g\}$, and the meet by $\min\{f,g\}$, these being taken pointwise. Now the absolute value of the derivative of $\max\{f,g\}$ for absolutely continuous f and g is (almost everywhere) less than $|f'| + |g'|$. Thus if $f, g \in X_\eta^*$, $\max\{f,g\}$ and $\min\{f,g\}$ are also elements of X_η^*. This completes the proof. ∎

PROOF OF PROPOSITION 9. Assume that the agent adopts a budget feasible consumption pattern x with associated wealth w and y. For convenience of notation, we at times denote $J^*(w(t),y(t^-),t)$, $J_w^*(w(t),y(t^-),t)$, $J_y^*(w(t),y(t^-),t)$, and $J_t^*(w(t),y(t^-),t)$ by $J^*(t)$, $J_w^*(t)$, $J_y^*(t)$, and $J_t^*(t)$, respectively. We have

$$\begin{aligned}&\textstyle\int_0^t u(y(s),s)\,ds + J^*(w(t),y(t^-),t) = \int_0^t u(y(s),s)\,ds + J^*(w(0),y(0^-),0) + \sum_i \left(J^*(\tau_i^+) - J^*(\tau_i)\right)\\ &\textstyle + \sum_i \int_{\tau_i^+}^{\tau_{i+1}^-} \left(J_w^*(s)rw(s)ds - J_w^*(s)dx(s) + J_y^*(s)\beta(dx(s) - y(s)\,ds) + J_s^*(s)ds\right),\end{aligned} \tag{29}$$

where the equality follows from fundamental theorem of calculus and we have set $\tau_0 = 0$. Note that the boundary conditions for J^* ensure that the left-hand expression is the correct expression of utility even for consumption plans that exhaust all the wealth before time 1.

By the hypothesis that J^* is concave in its first two arguments, we know that

$$J^*(\tau_i^+) - J^*(\tau_i) \leq J_w^*(\tau_i)(w(\tau_i^+) - w(\tau_i)) + J_y^*(\tau_i)(y(\tau_i) - y(\tau_i^-)).$$

Substituting this into (29), while noting that $w(\tau_i^+) - w(\tau) = x(\tau_i^-) - x(\tau)$, gives

$$\begin{aligned}&\int_0^1 u(y(s),s)\,ds + J^*(w(t),y(t^-),t)\\ \leq\ & J^*(w(0),y(0^-),0) + \int_0^1 [u(y(s),s) + J_w^*(s)w(s)r(s) - J_y^*(s)\beta y(s) + J_s^*(s)]ds\\ & + \int_{0^-}^1 [\beta J_y^*(s) - J_w^*(s)]dx(s) \leq J^*(w(0),y(0^-),0),\end{aligned}$$

where the second inequality follows from the hypothesis that the integrands are negative and x is increasing. ∎

PROOF OF THEOREM 3. With the hypotheses added in addition to the ones in Proposition 9, arguments identical to those used to prove the proposition prove this theorem. ∎

PROOF OF PROPOSITION 10. Suppose that $\frac{w}{y} > k$. The size of the initial "gulp" required to bring the ratio immediately to k is $\Delta = (w - ky)/(1 + \beta k)$ and we define $J^k(w, y) \equiv J^k(w - \Delta, y + \beta\Delta)$. Concavity of J^k in both w and y follows from assumption 1, and the rest of the proposition can be easily verified by direct computations. For the transversality condition note that for any point $(w(t), y(t^-))$, $J^k(w(t), y(t^-)) \le \gamma y(t^-)^\alpha$, for some constant γ. Noting that for any feasible policy, $y(t^-) \le \beta w(0)e^{rt}$, and hence $e^{-\delta t}J^k(w(t), y(t^-)) \le \gamma\beta e^{(-\delta+\alpha r)t}w(0)$, the transversality condition follows. ∎

PROOF OF PROPOSITION 11. In light of Proposition 10, we only need to prove that

$$\frac{y^\alpha}{\alpha} + J_w^{k^*} rW - J_y^{k^*}\beta y - \delta J^{k^*} \le 0 \quad \text{if} \quad \frac{w}{y} \ge k^* \quad \text{and}$$

$$\beta J_y^{k^*} - J_w^{k^*} \le 0 \quad \text{if} \quad \frac{w}{y} \le k^* .$$

To prove the first inequality, consider any point $\underline{x} \equiv (w, y)$ such that $w \ge k^* y$. Referring to Figure **2**, let $\underline{a}$ be the point on the intersection of the line $w = k^* y$ and the straight line passing through the point $\underline{x}$ with slope $\frac{dw}{dy} = -\frac{1}{\beta}$. In other words, $\underline{a}$ is the point on the boundary $w = k^* y$, to which one would jump if one starts at $\underline{x}$. By construction, $J^{k^*}(\underline{x}) = J^{k^*}(\underline{a})$. Let $f(w, y) = \frac{y^\alpha}{\alpha} + J_w^{k^*} rw - J_y^{k^*}\beta y$, and note that $f(\underline{a}) - \delta J^{k^*}(\underline{a}) = 0$. Applying the fundamental theorem of calculus along the straight line connecting $\underline{a}$ and $\underline{x}$, we obtain $f(\underline{x}) - f(\underline{a}) = \int_{\underline{a}}^{\underline{x}}(f_w\, dw + f_y\, dy)$. Noting that along the line connecting $\underline{a}$ and $\underline{x}$, we have $dy = -\beta dw$, and that $dw > 0$ in the direction from $\underline{a}$ to $\underline{x}$, it then follows that $f_w - \beta f_y \le 0$ is sufficient to conclude that

$$\frac{y^\alpha}{\alpha} + J_w^{k^*} rw - J_y^{k^*}\beta y - \delta J^{k^*} \le 0 \quad \text{if} \quad \frac{w}{y} \ge k^* .$$

Computing $f_w - \beta f_y$ in the region $w \ge k^* y$, the reader can easily verify that $f_w - \beta f_y \le 0$ for $k^* = \frac{\frac{r-\delta}{\beta} + (1-\alpha)}{\delta - \alpha r}$.

To prove the second inequality, consider the function $J_w^{k^*} - \beta J_y^{k^*}$ in the region where $w \le k^* y$. We can write

$$\begin{aligned} J_w^{k^*} - \beta J_y^{k^*} &= y^{\alpha-1} g(\frac{k^* y}{W}) \quad \text{where} \\ g(z) &= \mathcal{A}\Big[\frac{z}{k^*}\Big(\frac{\delta + \alpha\beta}{r + \beta}\Big) + \frac{\beta(\delta - \alpha r)}{r + \beta}\Big] z^{-\frac{\delta+\alpha\beta}{r+\beta}} - \Big[\frac{\beta}{\alpha\beta + \delta}\Big], \end{aligned}$$

where $\mathcal{A}$ is given in proposition 10. Note that $g(1) = 0$, and that $1 - \frac{\delta+\alpha\beta}{r+\beta} > 0$, by assumption 1, hence $g(z) \uparrow \infty$ as $z \uparrow \infty$.

Computing the derivative of g, we get

$$\frac{dg}{dz} = \frac{\mathcal{A}(\delta + \alpha\beta)}{r + \beta} z^{-\frac{\delta+\alpha\beta}{r+\beta}} \left[\Big(\frac{1}{k^*} - \frac{\delta + \alpha\beta}{k^*(r + \beta)}\Big) - \frac{\beta(\delta - \alpha r)}{z(\beta + r)}\right].$$

Substituting k^* in the expression for $\frac{dg}{dz}$, we conclude that $\frac{dg}{dz} = 0$ for $z = 1$, and that $\frac{dg}{dz} > 0$ for $z > 1$. It then follows that $J_w^{k^*} - \beta J_y^{k^*} \geq 0$, for all points (w, y) such that $w \leq k^* y$. ∎