Theoretical Economics 2 (2007), 71–113 1555-7561/20070071

Delay aversion

JEAN-PIERRE BENOÎT
Economics Area, London Business School

EFE A. OK
Department of Economics, New York University

We address the following question: When can one person properly be said to be more delay averse than another? In reply, several (nested) comparison methods are developed. These methods yield a theory of delay aversion which parallels that of risk aversion. The applied strength of this theory is demonstrated in a variety of dynamic economic settings, including the classical optimal growth and tree cutting problems, repeated games, and bargaining. Both time-consistent and time-inconsistent scenarios are considered.

KEYWORDS. Delay aversion, impatience, consumption smoothing, time consistency.

JEL CLASSIFICATION. D11, D90.

"The present always gets its rights." (Eugen von Böhm-Bawerk, The Positive Theory of Capital, 1891)

1. INTRODUCTION

Most, if not all, of an individual's decisions have consequences through time, making it imperative for analysts to have a clear understanding of agents' attitudes towards delay, and a framework for discussing these attitudes. Unsurprisingly, then, the study of time preference has a rich history, going back at least as far as Rae (1834) and Böhm-Bawerk (1891). In recent years, this topic has received an exceptional amount of attention from economists, much of it questioning the canonical exponential discounting model.[1] Despite this large body of work, one straightforward question seems to have gone almost entirely unaddressed in the economics literature: How does one compare the attitudes of two different agents towards time delay?

Jean-Pierre Benoît: jpbenoit@london.edu
Efe A. Ok: efe.ok@nyu.edu
We thank Dirk Bergemann, Juan Dubra, Larry Epstein, Ed Green, Faruk Gul, Jim Jordan, Boyan Jovanovic, Tom Sargent, Martin Schneider, Ennio Stacchetti, and an anonymous referee for various discussions about the content of this work. We are grateful also for the many comments of participants in seminars at Columbia, NYU, Penn State, Rutgers, the University of Texas at Austin, and Washington University in St. Louis.

[1] See Frederick et al. (2002) for a recent survey on time preferences.

Drawing a parallel to the modern theory of choice under risk and uncertainty is telling. By the end of the 1960s, a powerful theory of comparative risk aversion was well-established within the expected utility paradigm, thanks to the seminal contributions of Arrow (1964) and Pratt (1964), among others. Comparative risk aversion is now a textbook topic with a plethora of economic applications. Moreover, it has been generalized to the context of various non-expected utility models. By stark contrast, time preference theory at large lacks methods for comparing the attitudes of individuals towards delay, even within the entirely standard setting of exponential discounting.[2]

That there is such a lacuna may not be apparent at first glance. After all, at least within the exponential discounting model, there is reason to view the discount factor of an individual as a natural index of his "impatience." This may tempt one to view the problem of making impatience comparisons across individuals as one that is readily settled by comparing the discount factors of the involved parties. There are at least three problems with this position, however.

First, at a fundamental level, a comparison of attitudes towards delay should not be tied to a particular representation of preferences. Indeed, it should be possible to reject the exponential discounting model and still make comparisons about the relative impatience of two or more decision makers.

Second, even if one accepts the exponential discounting model, using discount factors to make comparisons is not always meaningful. Consider an environment in which the choice objects are dated monetary outcomes. (A prototypical example of such an environment is provided by sequential bargaining, where the game ends with each player receiving a one-time payment.) As is usual, denote the dated outcome in which x dollars are received in period t as (x,t). Take an individual whose preferences over dated outcomes are represented by the intertemporal utility function $(x,t)\mapsto \alpha^t u(x)$, where u is the agent's instantaneous utility function and $0<\alpha<1$ is his discount factor. Now choose any $0<\beta<1$. One can show that the very *same* preferences, which are represented by $(x,t)\mapsto \alpha^t u(x)$, can also be represented by $(x,t)\mapsto \beta^t v(x)$, for some instantaneous utility function v.[3] Hence, the choice of the discount factor used to represent an individual's preferences is entirely arbitrary in this environment; it follows that, here, discount factors cannot possibly form a meaningful basis for comparing attitudes towards delay.

Third, even in contexts where the exponential discounting model is appropriate, and discount factors are uniquely determined by preferences, making a comparison based solely on these discount factors is questionable. Consider two infinitely lived agents, 1 and 2, who evaluate an arbitrary consumption path $(x_0,x_1,\ldots)$ as $\sum_{t=0}^{\infty}\delta_1^t u_1(x_t)$ and $\sum_{t=0}^{\infty}\delta_2^t u_2(x_t)$, respectively. Here, unlike in the example considered above, the discount

[2]Starting with Koopmans (1960) and Koopmans et al. (1964), there has been much work on the formalization of the notion of "impatience" and the link between this concept and the continuity of intertemporal utility functions. (See Epstein 1987 for a related survey.) While there is still some interest in this matter today (cf. Marinacci 1998), these works are not helpful for comparing the "impatience" of two decision makers. To the best of our knowledge, the only article that studies this issue is a nice, albeit largely unknown, note by Horowitz (1992), about which more in Section 3.

[3]The proof follows upon setting $v:=u^{\ln\beta/\ln\alpha}$. (See Theorem 3 of Fishburn and Rubinstein 1982.)

factor δ_i is uniquely determined by the ith person's intertemporal preferences, and the instantaneous utility function u_i is uniquely determined up to a positive affine transformation. Nevertheless, it would still not be reasonable to compare the relative delay aversion of these individuals by looking *only* at their discount factors. To see the flaw with such an approach, consider the special case in which $u_1(x) = x$, $u_2(x) = \sqrt{x}$, $x \geq 0$, and $1 > \delta_1 > \delta_2 > 0$. Suppose that each person must decide how to allocate a fixed total wealth over time. Clearly, the first person maximizes by consuming the entire amount in the first period, while the second person spreads her wealth through time, since $u_2'(0) = \infty$. Thus, from an observational point of view, the first person exhibits a much stronger bias towards the present, although the second person seems to be the more present-oriented based upon a comparison of discount factors alone.

The culprit behind the second and third points is clear. The instantaneous utility function of an agent, along with his discount factor, plays an essential role in shaping his attitude toward time delay. (For instance, in the last example, the square root function reveals a desire for consumption smoothing.) As a result, even within the exponential discounting framework, comparing the aversion of two decision makers towards delay solely on the basis of their discount factors is troublesome. In the literature, this difficulty has commonly been circumvented by the practice of making impatience comparisons across decision makers only when these agents *share* a single instantaneous utility function.[4] As a basis for formal comparative static exercises, this practice is, as we shall see, sound. Beyond such exercises, however, it is severely wanting. Consider, for instance, the following question: To what extent can differences in the time series wealth profiles of two countries be explained by differences in the delay aversion of the citizens of the countries? Suppose, to address the question, each country is modeled by a representative agent. While this standard simplification may be acceptable, there would seem to be little justification for further assuming that the two representative agents have the same instantaneous utility functions. Indeed, once it is admitted that the representatives of the two countries may differ, as it must be to address the question, it is entirely arbitrary to restrict this difference to one of discount factors. Granted, one might wish to argue that differences in discount factors capture differences in delay aversion provided that instantaneous utility functions do not differ too much, so that positing identical utilities may be an an acceptable simplification. However, merely to formulate such an argument an independent notion of relative delay aversion is first needed. Moreover, the idea that one needs to keep the instantaneous utility functions of two individuals constant to compare their delay attitudes runs into obvious difficulties the moment one departs from the separable time preference model.

The objective of this paper is to develop rigorous techniques for comparing the aversion of decision makers towards time delay. The discussion above, and the parallel we

[4]For instance, studies that aim to compare the impatience of the representative members of distinct socioeconomic classes are often forced to postulate the homogeneity of static preferences for agents across the classes under consideration (cf. Lawrance 1991). Similarly, in dynamic economic analysis, the implications of one party being more impatient than another is explored almost exclusively by varying the discount factors while holding the instantaneous utility functions constant. (Section 4 contains several examples of this nature.) Finally, to our knowledge, all experimental studies that estimate personal discount rates work under the assumption that the subjects have the same utility function for money.

seek to risk theory, suggests at least three constraints in this endeavor. First, the proposed techniques should not, at least at the level of their primitive definitions, depend on the way in which intertemporal preferences are modeled. Second, they should be "easy" to apply, at least within specialized models (such as that of separable intertemporal preferences). Third, these methods should be useful in dynamic economic analysis. That is, they should allow for rigorous comparative statics exercises in actual economic models.[5] With these considerations in mind, we organize this paper in three major parts, each one of which contains subdivisions that address these three concerns.

In Section 3 we introduce a simple (partial) method of comparing two agents' eagerness to enhance earlier consumption at the expense of future consumption, without subscribing to a particular model of intertemporal preferences. Roughly speaking, we qualify individual A as more *delay averse* than individual B if whenever B prefers receiving an increase in consumption at an earlier date to receiving an increase at a later date, A does too, *ceteris paribus*. We go on to obtain several characterizations of the induced ordering for the special case of separable intertemporal utilities. These characterizations yield further insight into the structure and appeal of our delay aversion ordering. In particular, we find that in the exponential discounting model, if one agent is more delay averse than another, then her discount factor *must* be lower than that of the latter, but not conversely. These characterizations also show how to transform a given separable intertemporal utility function into a more delay averse one.

The approach to relative delay aversion taken in Section 3 is "bottom line" in nature, in that it inquires into an individual's desire for enhanced early consumption without distinguishing to what degree this desire reflects a specific bias towards the present and to what degree it is instead a reaction to any unevenness in the individual's underlying endowment stream. Note in this regard that the same individual could want to borrow if his endowment stream was increasing and lend if it was decreasing. This approach to delay aversion can be contrasted with an inquiry into what may be thought of as "pure time bias," namely, a psychological preference for early gratification, which may or may not overwhelm other considerations. In Section 4, we refine our approach to provide an ordering that aims at capturing a pure time preference motive instead. We term this (more complete) ordering an *impatience* ordering, and provide various characterizations of it in the case of separable intertemporal preferences.

Specific economic applications, which demonstrate the applicability of the relative delay aversion and impatience notions, and their various characterizations, are interspersed throughout the paper. In particular, in the context of the one-sector optimal growth problem with exponential discounting, we show that the optimal capital stock of a country can never fall strictly below that of a more delay averse country. Becker (1983)

[5]Recall that, in risk theory, the basic definitions of risk aversion and related concepts do not depend on how one's preferences over risky prospects are represented. Moreover, within specific models (such as the expected utility model), these abstract definitions yield characterizations (via the Arrow–Pratt coefficients, Jensen's inequality, etc.) that accentuate their applicability substantially. Finally, there are numerous economic applications (e.g. the models of demand for insurance and portfolio diversification) that mesh extremely well with these definitions and their characterizations.

obtains a similar result by varying the discount factors of agents with *identical* instantaneous utility functions. It is not clear from Becker's analysis if this result stems from differing degrees of *delay aversion* of the countries or from their relative *impatience*. Our approach clarifies that a pure time motive alone cannot account for this result—one needs the full strength of comparative delay aversion.

We also apply our delay aversion theory to dynamic games. First, for infinitely repeated games, we provide conditions under which the Nash equilibrium of a game remains an equilibrium as players become less delay averse. A particular novelty of this application is that the players are not assumed to be exponential, or even separable, utility maximizers. Second, in Rubinstein bargaining, we revisit a well-known result of Roth (1985) that shows that an agent's equilibrium share decreases as his utility function becomes more concave. While Roth interprets this to mean that becoming more risk averse harms an agent, this interpretation has long been considered to be dubious, as there is no risk in the model. Our results clarify the nature of Roth's observation by showing formally that it is properly understood as a result about delay aversion, not risk aversion.

The above applications concern agents who either have time-consistent preferences or can commit to their plans. The delay aversion theory we develop is applicable in the absence of these assumptions as well. To demonstrate, we study the classical Wicksellian tree-cutting problem for (naive and sophisticated) individuals with time-inconsistent preferences who cannot commit to their plans. While naive agents behave according to intuition, we find, somewhat paradoxically, that a sophisticated person may cut his tree later as he becomes more delay averse. Notably, however, this reversal disappears when preferences are *present-biased*, as in the case of hyperbolic and quasi-hyperbolic discounting models.

2. Preliminaries

For expositional simplicity we develop our formalism only for infinitely-lived agents and *bounded* streams of outcomes—the entire analysis adapts in a straightforward manner to the case of finitely lived agents. Accordingly, we take an intertemporal choice item to be a real sequence $(x_0, x_1, \ldots)$ with $x_t \geq 0$ for all t, and $\sup\{x_t : t \in \mathbb{Z}_+\} < \infty$. Of course, we think of x_t as the level of consumption at time t. We denote the set of all such real sequences as $\mathscr{X}$, and endow $\mathscr{X}$ with the product topology.

The generic members of $\mathscr{X}$ are denoted as $\mathbf{x}, \mathbf{y}$, etc.; we adopt the convention of denoting the t-th term of $\mathbf{x}$ as x_t, so that $\mathbf{x} \equiv (x_0, x_1, \ldots)$. By $(a, \mathbf{x}_{-t})$, we denote the sequence $\mathbf{y}$, where $y_t = a$ and $y_m = x_m$ for all $m \neq t$. Similarly, for any $k \in \mathbb{N}$ and distinct positive integers $t_1, \ldots, t_k$, the expression $(a_{t_1}, \ldots, a_{t_k}, \mathbf{x}_{-(t_1,\ldots,t_k)})$ stands for the sequence $\mathbf{y}$ where $y_{t_i} = a_{t_i}$, $i = 1, \ldots, k$, and $y_m = x_m$ for all $m \in \mathbb{Z}_+ \setminus \{t_1, \ldots, t_k\}$.

We work with strictly increasing preferences over consumption streams, that is, preferences for which more consumption is preferred to less at any period. Moreover, we consider only those preference relations that can be represented by a utility function U on $\mathscr{X}$ such that the restriction $U|_{[0,a]^\infty}$ is continuous for any $0 \leq a < \infty$. Such maps are referred to as *cube-continuous* in what follows.[6] For the case of finitely-lived agents, cube-continuity reduces to the standard notion of continuity.

DEFINITION 1. A cube-continuous and strictly increasing map $U : \mathscr{X} \to \mathbb{R}$ such that $U(0,0,\ldots)=0$ is called an *intertemporal utility function*.[7] We denote the set of all intertemporal utility functions by $\mathfrak{U}$.

The most prevalent types of intertemporal utility functions in economic analysis posit the separability of the evaluation of time and outcomes, presume an additively separable form, and consider stationary instantaneous utility functions. In this paper we refer to such members of $\mathfrak{U}$ succinctly as *separable*. To define this subclass formally, let us call a function $\boldsymbol{\delta} : \mathbb{Z}_+ \to (0,1]$ a *discount function* if $\boldsymbol{\delta}$ is strictly decreasing, $\boldsymbol{\delta}(0)=1$, and $\boldsymbol{\delta}(0)+\boldsymbol{\delta}(1)+\cdots<\infty$.[8] We refer to a continuous and strictly increasing map $u : \mathbb{R}_+ \to \mathbb{R}$ as an *instantaneous utility function* provided that $u(0)=0$ and $u(\infty)=\infty$.[9] The class of all discount functions is denoted by $\mathscr{D}$ and the class of all instantaneous utility functions by $\mathscr{U}$. At times we work with differentiable members of $\mathscr{U}$. The following class receives particular attention:

$$\mathscr{V} := \{u \in \mathscr{U} : \ u \text{ is continuously differentiable and } u'|_{(0,\infty)} > 0\}.$$

Note that any member u of $\mathscr{V}$ has an inverse u^{-1}, which is continuously differentiable with a finite positive derivative on $(0,\infty)$.[10]

We call a map $U : \mathscr{X} \to \mathbb{R}$ a *separable intertemporal utility function* if there exists $(u,\boldsymbol{\delta}) \in \mathscr{U} \times \mathscr{D}$ such that

$$U(\mathbf{x}) = \sum_{t=0}^{\infty} \boldsymbol{\delta}(t) u(x_t) \quad \text{for all } \mathbf{x} \in \mathscr{X}. \tag{1}$$

We denote the class of all separable intertemporal utility functions by $\mathfrak{U}_{\text{sep}}$. The following elementary result justifies our terminology by establishing that every separable intertemporal utility function is indeed an intertemporal utility function. In particular, every separable intertemporal utility function is cube-continuous.

[6] Put differently, U is cube-continuous if and only if the restriction of U to any Hilbert cube in $\mathscr{X}$ is continuous. (Recall that a compact subset $[a,b]^\infty$ of $\mathscr{X}$ is called a *Hilbert cube* in $\mathscr{X}$ for any real numbers a and b with $0 \le a < b$.) Hence, the term "cube-continuous." Given that $\mathscr{X}$ is endowed with the product topology, the *continuity* of a real map on $\mathscr{X}$ is much more demanding than its cube-continuity. For example, the map $f : \mathscr{X} \to \mathbb{R}$ defined by $f(\mathbf{x}) := \sum_{t=0}^{\infty} \delta^t x_t^\sigma$ is not continuous for any choice of $0<\delta,\sigma<1$, but it is cube-continuous for any such choice of δ and σ.

[7] The requirement $U(0,0,\ldots)=0$ is merely a normalization that simplifies the exposition.

[8] The convergence of the series $\sum_{t=0}^{\infty} \boldsymbol{\delta}(t)$ ensures that the map $x \mapsto \sum_{t=0}^{\infty} \boldsymbol{\delta}(t) u(x_t)$ is real-valued on $\mathscr{X}$ whenever u is a continuous function on $\mathbb{R}_+$. The assumption that $\boldsymbol{\delta}$ is strictly decreasing is standard in the literature; it amounts to saying that people dislike time delay in general. We adopt this formulation here to conform with the literature, but note that the main results of this section remain valid in the absence of this assumption.

[9] The assumption $u(\infty)=\infty$ considerably simplifies the subsequent analysis, but is not essential for it. In particular, the "if" parts of all of our characterization theorems remain valid without this assumption.

[10] u^{-1} is also right-differentiable at 0, but its right-derivative at 0 may belong to $\{0,\infty\}$.

LEMMA 1. $\mathfrak{U}_{sep} \subseteq \mathfrak{U}$.

The proof of this result is in the Appendix, together with the proofs of other results not given in the text.

In what follows, by the tuple $(u, \boldsymbol{\delta})$ in $\mathscr{U} \times \mathscr{D}$, we mean the separable intertemporal utility function induced by u and $\boldsymbol{\delta}$ by way of (1). We use the notation $(u, \boldsymbol{\delta})$ and U interchangeably when (1) holds.

Perhaps the most important subclass of $\mathfrak{U}_{\text{sep}}$ is the one consisting of exponential intertemporal utility functions. Formally, an *exponential intertemporal utility function* is defined as a map $U : \mathscr{X} \to \mathbb{R}$ with

$$U(\mathbf{x}) = \sum_{t=0}^{\infty} \delta^t u(x_t) \quad \text{for all } \mathbf{x} \in \mathscr{X}, \tag{2}$$

where $u \in \mathscr{U}$ and $0 < \delta < 1$. In this case, δ is called a *discount factor*. The class of all exponential intertemporal utility functions is denoted by $\mathfrak{U}_{\text{exp}}$. Obviously, $\mathfrak{U}_{\text{exp}} \subseteq \mathfrak{U}_{\text{sep}}$. Again, by the tuple (u, δ) in $\mathscr{U} \times (0, 1)$, we mean the exponential intertemporal utility function induced by u and δ through (2), and hence, with a slight abuse of terminology, we refer to any such pair (u, δ) as an "exponential intertemporal utility function."

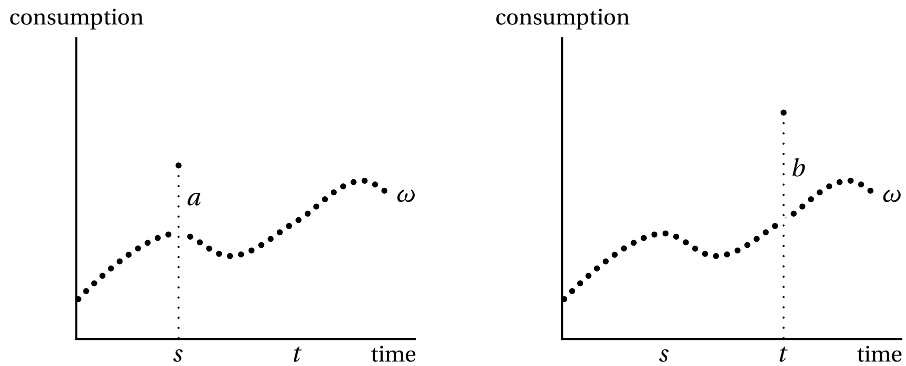
In recent years, a large amount of experimental data has been gathered that questions the exponential discounting model in particular, and the stationarity of time preferences in general. This has led economists to give serious consideration to certain generalizations of the exponential discounting model.[11] Most of these generalizations, such as the quasi-hyperbolic and hyperbolic discounting models, still carry the form of a separable intertemporal utility function (viewed as representing the *commitment* preferences of the individuals), and are thus captured by our results that pertain to $\mathfrak{U}_{\text{sep}}$. It is, however, fair to say that the exponential discounting model remains the most widely used framework in dynamic economic analysis, and hence we emphasize the exact nature of our subsequent results for this specific model.

3. COMPARATIVE DELAY AVERSION

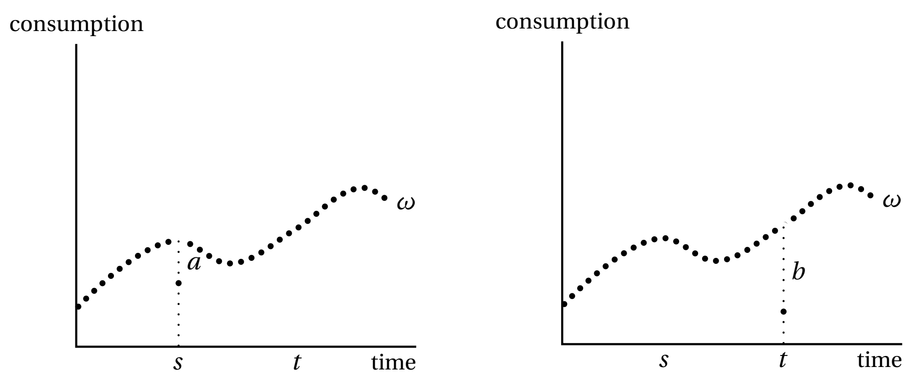
3.1 *Main definition*

Consider an individual, A, facing a fixed endowment stream $\omega \in \mathscr{X}$. Suppose that he has won a prize that allows him to choose between an additional consumption of 100 in period s and 120 in period $s+1$. Without knowing his preferences and endowment, we cannot, of course, predict which of the two options he prefers, but suppose that, as it happens, he favors the additional 100 in period s. Thus, we understand that A does not consider it to be worth waiting an extra period beyond s in order to receive the larger amount of 120. Now consider an individual, B, in *identical* circumstances, but, who is known to dislike delay more than A does. Naturally, we expect that she too considers the

[11] A very good survey of the recent developments in time preference theory is provided by Frederick et al. (2002).



a. If an agent prefers the consumption path on the left, so does a more delay averse agent.



b. If an agent prefers the consumption path on the right, so does a more delay averse agent.

FIGURE 1.

larger amount of 120 to be insufficient compensation for waiting an extra period, and, instead, prefers receiving 100 in period s.

The fact that the above thought experiment was couched in terms of payments made *to* the individuals is not essential. Precisely the same reasoning would apply in terms of payments made *by* them. Suppose that individual A prefers paying b dollars at time t to paying a dollars at an earlier date s. Individual B, who faces the same objective circumstances as individual A, but is more delay averse, should also prefer making the later payment to the earlier one.

Of course, there is nothing special in a particular choice of endowment stream, payments, and time periods. These considerations prompt the following definition, which is illustrated in Figure 1.

DEFINITION 2. An intertemporal utility function $V \in \mathfrak{U}$ is *more delay averse than* $U \in \mathfrak{U}$ if, for any given $s, t \in \mathbb{Z}_+$ with $s < t$ and any $\boldsymbol{\omega} \in \mathscr{X}$,

$$U(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\omega_t + b, \boldsymbol{\omega}_{-t}) \quad \text{implies} \quad V(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\omega_t + b, \boldsymbol{\omega}_{-t}) \tag{3}$$

for all $a, b \geq 0$, and

$$U(\omega_t - b, \boldsymbol{\omega}_{-t}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\omega_s - a, \boldsymbol{\omega}_{-s}) \quad \text{implies} \quad V(\omega_t - b, \boldsymbol{\omega}_{-t}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\omega_s - a, \boldsymbol{\omega}_{-s}) \tag{4}$$

for all $\omega_s \geq a \geq 0$ and $\omega_t \geq b \geq 0$. In turn, we define the binary relation $\precsim$ on $\mathfrak{U}$ as $U \precsim V$ if and only if V is more delay averse than U. The asymmetric part of this relation is denoted as $\prec$, that is, $U \prec V$ if and only if $U \precsim V$ but not $V \precsim U$. When $U \prec V$, we say that V is *strictly more delay averse than* U.[12]

The first condition of this definition says that if agent B is more delay averse than agent A, then B prefers receiving an earlier payment to a (possibly different) later payment whenever A does.[13] The second condition says that the relatively delay averse B prefers making a later payment whenever A does.

A few remarks on the mathematical structure of the binary relation $\precsim$ are in order. This relation is a vector preorder on $\mathfrak{U}$; that is, $\precsim$ is a reflexive and transitive binary relation on $\mathfrak{U}$ such that $U \precsim V$ if and only if $\lambda U + W \precsim \lambda V + W$ for all $\lambda > 0$ and $U, V, W \in \mathfrak{U}$.[14] Moreover, this ordering is continuous in the sense that if $U \precsim V$ holds and W is close enough to U uniformly, then $W \precsim V$. The following proposition summarizes these facts.

PROPOSITION 1. *The binary relation $\precsim$ is a vector preorder on $\mathfrak{U}$. Moreover, if (U_n) and (V_n) are any two sequences in $\mathfrak{U}$ with $U_n \precsim V_n$ for each n, and U and V are intertemporal utility functions with $U_n \to U$ uniformly and $V_n \to V$ uniformly, then $U \precsim V$.*

It is not difficult to see that our delay aversion ordering $\precsim$ is not complete. In Section 3.3 we encounter a few interesting intertemporal utility functions that cannot be compared on the basis of this preorder.

3.2 *Single-crossing and investments*

To the best of our knowledge, the only other preorder introduced in the literature to compare the attitudes of two individuals towards time delay is that of Horowitz (1992). While Horowitz formulates his ordering in a continuous time framework, it is easy to adapt it to the present discrete time setting.[15] First we need to introduce the auxiliary concept of single-crossing streams.

[12]We should note that, from a formal point of view, Definition 2 contains a redundancy. Given that we are working with *all* endowment streams ω, the second part of this definition (i.e. the part that concerns (4)) is implied by its first part, and vice versa. (See Lemma 3 in the Appendix.) We maintain this redundancy because these two parts of the definition are conceptually distinct. Moreover, in the subsequent sections we work with particular subsets of endowment streams and this redundancy disappears.

[13]Naturally, we say that agent B is more delay averse than agent A if B's intertemporal utility function is more delay averse than A's.

[14]Throughout this paper, we refer to any reflexive and transitive binary relation as a *preorder* and *ordering* interchangeably.

[15]Horowitz's choice to model time continuously proves to be unfortunate in a number of respects. Most notably, within the standard exponential utility discounting model it results in a vastly incomplete ordering, which applies *only* when the decision makers have the same instantaneous utility functions. As will become clear shortly, this finding contrasts markedly with our present results.

DEFINITION 3. For any $\mathbf{x}, \mathbf{y} \in \mathscr{X}$, we say that $\mathbf{y}$ *single-crosses* $\mathbf{x}$ *from above* if there exists $t^* \in \mathbb{N}$ such that $y_m \geq x_m$ for all $m = 0, \ldots, t^* - 1$, and $y_m \leq x_m$ for all $m = t^*, t^* + 1, \ldots$.

Horowitz's idea is that if one individual favors a consumption stream that single-crosses another from above, then so should a more delay averse person. This is the content of the following definition.

DEFINITION 4. An intertemporal utility function V is *single-crossing more delay averse than* $U \in \mathfrak{U}$ if, for any $\mathbf{x}, \mathbf{y} \in \mathscr{X}$,

$$U(\mathbf{y}) \left\{ \begin{matrix} \geq \\ > \end{matrix} \right\} U(\mathbf{x}) \quad \text{implies} \quad V(\mathbf{y}) \left\{ \begin{matrix} \geq \\ > \end{matrix} \right\} V(\mathbf{x})$$

whenever $\mathbf{y}$ single-crosses $\mathbf{x}$ from above.

A comparison of $\precsim$ and the single-crossing ordering readily reveals that the latter implies the former. Indeed, our ordering is defined in terms of particularly simple single-crossing consumption streams while Horowitz's ordering uses *all* single-crossing streams. This suggests that the latter ordering might be significantly more demanding than $\precsim$. In fact, however, these two orderings are equivalent.

THEOREM 1. *For any intertemporal utility functions U and V, V is more delay averse than U if and only if V is single crossing more delay averse than U.*

The simplicity of the comparisons involved in the definition of our ordering $\precsim$ is an obvious advantage—one that we exploit in deriving many of the results of the subsequent sections. Theorem 1 shows that this simplicity comes at no conceptual cost.

An analogy may be helpful here. In risk theory, the notion of mean preserving spreads is used to get a basic handle on ranking lotteries on the basis of their riskiness. While the simplicity of this method is appealing, its usefulness seems limited, for in practice it is unlikely that one would deal with two lotteries one of which is derived from the other by means of a single mean preserving spread. In this regard, the second order stochastic dominance ordering seems superior. A celebrated result, however, tells us that these two methods are formally equivalent. Theorem 1 has the same flavor. It shows that the simpler (and apparently less applicable) ordering $\precsim$ is equivalent to the more complicated (but apparently more applicable) ordering of Horowitz.

There are, of course, other ways of thinking about the notion of relative delay aversion. Notably, one may wish to base this notion on the comparative investment behavior of decision makers. Since an investment is a form of delayed gratification, relatively delay averse people should undertake relatively few investments, and conversely. In fact, this point of view is completely consistent with that of the delay aversion ordering $\precsim$.

Consider a person with initial endowment stream $\boldsymbol{\omega}$ who has an investment opportunity that costs $a \leq \omega_s$ units of consumption in period s and yields returns $x_i \geq 0$ in ensuing periods. If he undertakes the investment, he ends up with the stream

$$\boldsymbol{\omega}' := (\omega_0, \ldots, \omega_{s-1}, \omega_s - a, \omega_{s+1} + x_1, \omega_{s+2} + x_2, \ldots).$$

In concert with intuition, if this person prefers $\boldsymbol{\omega}$ to $\boldsymbol{\omega}'$ (that is, chooses not to undertake the associated investment), then any more delay averse person does too. This fact follows immediately from Theorem 1, since $\boldsymbol{\omega}$ single crosses $\boldsymbol{\omega}'$ from above. The converse is also true:

Suppose that, whenever person A prefers $\boldsymbol{\omega} \in \mathscr{X}$ to an $\boldsymbol{\omega}'$ obtained from $\boldsymbol{\omega}$ through an investment, as above, then person B does as well. Then B is more delay averse than A.[16]

3.3 *Comparative delay aversion with separability*

An especially important class of intertemporal utility functions is that of separable intertemporal utility functions. In this section we provide several characterizations of our "more delay averse than" ordering $\precsim$ for these utility functions. The following is the main result of the paper.

THEOREM 2. *For any two separable intertemporal utility functions* $(u,\boldsymbol{\alpha})$ *and* $(v,\boldsymbol{\beta})$*, the following two statements are equivalent.*

(a) $(v,\boldsymbol{\beta})$ *is more delay averse than* $(u,\boldsymbol{\alpha})$.

(b) There exists a map $h:\mathbb{R}_+ \to \mathbb{R}_+$ *such that* $v = h\circ u$ *and*

$$h\left(x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right) \geq h(x)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(h(y+z)-h(z)) \tag{5}$$

for all $s,t\in\mathbb{Z}_+$ *with* $s<t$ *and* $x,y,z\geq 0$.

Moreover, if u *and* v *belong to* $\mathscr{V}$*, then either of the above statements is equivalent to either of the following statements.*

(c) There exists a continuously differentiable map $h:\mathbb{R}_+\to\mathbb{R}_+$ *such that* $v=h\circ u$ *and*

$$\inf\{h'(x):x>0\} \geq \left(\frac{\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)}{\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)}\right)\sup\{h'(x):x>0\} \quad \textit{whenever } s<t.$$

(d) $\dfrac{\boldsymbol{\beta}(s)}{\boldsymbol{\beta}(t)}\dfrac{v'(x)}{v'(y)} \geq \dfrac{\boldsymbol{\alpha}(s)}{\boldsymbol{\alpha}(t)}\dfrac{u'(x)}{u'(y)}$ *for all* $s,t\in\mathbb{Z}_+$ *with* $s<t$ *and* $x,y\geq 0$.[17]

A basic result of risk theory states that a given von Neumann-Morgenstern utility function is at least as risk averse as another if and only if the former is a concave transformation of the latter. This observation enables one to generate more risk averse utility functions from a given von Neumann-Morgenstern utility function and leads to useful characterizations of the "more risk averse than" ordering in the case of differentiable utility functions (via the Arrow–Pratt coefficients).

Theorem 2 provides analogous results for the preorder $\precsim$. Part (b) tells us that $(v,\boldsymbol{\beta})$ is more delay averse than $(u,\boldsymbol{\alpha})$ if and only if v is a particular transformation of u. This

---

[16] This claim follows from the fact that person B prefers $(\omega_t+b,\boldsymbol{\omega}_{-t})$ to $(\omega_s+a,\boldsymbol{\omega}_{-s})$, $s<t$, whenever A does, since $(\omega_t+b,\boldsymbol{\omega}_{-t})$ can be viewed as an investment relative to the endowment stream $(\omega_s+a,\boldsymbol{\omega}_{-s})$.

[17] We note that the ratios $\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)$ and $\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)$ can be replaced by $\boldsymbol{\beta}(t+1)/\boldsymbol{\beta}(t)$ and $\boldsymbol{\alpha}(t+1)/\boldsymbol{\alpha}(t)$, $t=0,1,\ldots$, in the statements of (b), (c) and (d). The same holds for Corollaries 1 and 2 below.

transformation is captured by the functional inequality (5) which, of course, incorporates the influence of the discount functions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. This inequality is extremely useful. In particular, it allows us to obtain (d) by means of a straightforward application of the Inverse Function Theorem.

In turn, the statement in (d) of Theorem 2 is easily interpretable. Think of x as consumption in period s and y as consumption in period $t > s$. The statement then says that $(v,\boldsymbol{\beta}) \precsim (u,\boldsymbol{\alpha})$ if and only if $(v,\boldsymbol{\beta})$ has a larger marginal rate of intertemporal substitution of (the earlier) s-th period consumption for (the later) t-th period consumption, regardless of the levels of consumption at periods s and t.[18]

While the instantaneous utility function and the discount function of an agent both contribute to the determination of his attitude towards delay, the following corollaries point to a greater contribution on the part of the discount function.

COROLLARY 1. *For any separable intertemporal utility functions* $(u,\boldsymbol{\alpha})$ *and* $(v,\boldsymbol{\beta})$,

$$(u,\boldsymbol{\alpha}) \precsim (v,\boldsymbol{\beta}) \quad \textit{implies} \quad \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} \textit{ whenever } s < t.$$

In particular, $(u,\boldsymbol{\alpha}) \precsim (v,\boldsymbol{\beta})$ *implies* $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$.

As we argue in the Introduction, and as Examples 1 and 2 below confirm, discount factors are not sufficient for making delay aversion comparisons for exponential utility maximizers. On the other hand, as Corollary 1 shows, they are necessary for such comparisons. More precisely, if agent A is more delay averse than agent B, then A's discount factor is lower than B's. It follows that, while the instantaneous utility function can undo the effect of the discount factor, it cannot reverse it: If agent A has a lower discount factor than agent B, then either A is more delay averse than B, or the two agents cannot be ranked.

The next corollary shows that the common comparative static exercise of lowering an agent's discount factor while holding his instantaneous utility function constant amounts to rendering the agent more delay averse. At the same time, for preferences that are separable but not exponential, lowering an agent's discount function everywhere is not sufficient to render him more delay averse. In that case, the relative discount functions $\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)$ and $\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)$ must also be considered. The corollary also establishes that holding an agent's discount factor constant while changing his instantaneous utility function (in a nontrivial manner), results in a noncomparable agent.

COROLLARY 2. *For any separable intertemporal utility functions* $(u,\boldsymbol{\alpha})$ *and* $(v,\boldsymbol{\beta})$,

$$(u,\boldsymbol{\alpha}) \precsim (u,\boldsymbol{\beta}) \quad \textit{if and only if} \quad \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} \textit{ whenever } s < t.$$

[18] For Fisher (1930), a person is delay averse—he uses the term "impatient"—in an absolute sense, if his marginal rate of intertemporal substitution is *always* greater than one. Part (d) of Theorem 2 shows that, in an obvious way, our notion of relative delay aversion is the logical extension of Fisher's definition to comparisons.

Moreover,

$$(u,\boldsymbol{\alpha}) \precsim (v,\boldsymbol{\alpha}) \quad \textit{if and only if} \quad u = \theta v \textit{ for some } \theta > 0.$$

The next corollary is a simplification—and a trivial consequence—of Theorem 2 for the case of *exponential* intertemporal utility functions.[19]

COROLLARY 3. *Let (u,α) and (v,β) be exponential intertemporal utility functions with $u,v \in \mathscr{V}$. The following statements are equivalent.*

(a) (v,β) is more delay averse than (u,α).

(b) There exists a continuously differentiable map $h:\mathbb{R}_+ \to \mathbb{R}_+$ such that $v = h \circ u$ and

$$\inf\{h'(x): x>0\} \geq \frac{\beta}{\alpha} \sup\{h'(x): x>0\}. \tag{6}$$

(c) $\dfrac{v'(x)}{\beta v'(y)} \geq \dfrac{u'(x)}{\alpha u'(y)}$ for all $x,y>0$.

We conclude with three simple exponential discounting examples, illustrating the applicability of the above results.

EXAMPLE 1. Let $u,v \in \mathscr{V}$ be instantaneous utility functions such that $u'(0+) = \infty$ and $v'(0+) < \infty$. Regardless of the values of the discount factors α and β, (u,α) and (v,β) *cannot* be ranked on the basis of $\precsim$. This follows immediately from the equivalence of the statements (a) and (c) in Corollary 3.

More generally, if u and v in $\mathscr{V}$ are such that the right derivative of the function $h := v \circ u^{-1}$ at 0 (or at any other point in $\mathbb{R}_+$) belongs to $\{0,\infty\}$, then (u,α) and (v,β) cannot be ranked by $\precsim$. ◊

EXAMPLE 2. For any $0 \leq \sigma < 1$, define the instantaneous isoelastic utility function $u_\sigma \in \mathscr{U}$ by

$$u_\sigma(x) := \frac{x^{1-\sigma}}{1-\sigma},$$

and let $\mathscr{U}_0 := \{u_\sigma : 0 \leq \sigma < 1\}$. This class is widely used in intertemporal macroeconomic models. Let us take two exponential intertemporal utility functions (u_{σ_1},α) and (u_{σ_2},β) in $\mathscr{U}_0 \times (0,1)$. When can these intertemporal utility functions be ranked by $\precsim$? The answer is *only* when the agents have *identical* instantaneous utility functions.[20] More precisely,

[19] The substance of this corollary is investigated by Horowitz (1992) in a continuous time framework. As noted earlier, in that context one ends up with the rather limited result that two agents are comparable if and only if their instantaneous utilities are positive affine transformations of each other.

[20] In practice, however, it may be possible to order this class more completely. Note that the definition of delay aversion implicitly allows for any possible consumption stream. Suppose, however, that there is reason to believe that agents' consumptions in any single period always lie within certain bounds, say for all time periods s, $x_s \in [x_*, x^*]$. Then it makes sense to restrict (3) of Definition 2, to a, b, and $\boldsymbol{\omega}$ such that $x_* \leq \omega_s + c \leq x^*$ for all s, where $c = a$, b, and similarly for (4). With this restricted definition, it is possible to compare some members of this class with different instantaneous utilities.

$(u_{\sigma_1}, \alpha) \precsim (u_{\sigma_2}, \beta)$ if and only if $\sigma_1 = \sigma_2$ and $\alpha \geq \beta$. The "if" part follows from Corollary 2. To prove the "only if" part, observe that the right derivative of $h := u_{\sigma_1} \circ u_{\sigma_2}^{-1}$ at 0 belongs to $\{0, \infty\}$ unless $\sigma_1 = \sigma_2$, and apply the last observation made in Example 1. ◊

These examples provide instances of the incompleteness of the delay aversion ordering $\precsim$. The next example provides an instance in which $\precsim$ applies in a nontrivial manner.

EXAMPLE 3. Take any exponential intertemporal utility function (u, α). We wish to find $(v, \beta) \in \mathscr{U} \times (0, 1)$ such that $(u, \alpha) \prec (v, \beta)$. This can be done for an arbitrarily chosen $\beta \in (0, \alpha)$. For instance, from Corollary 3 we have $(u, \alpha) \prec (h \circ u, \beta)$ for any continuous differentiable concave function $h : \mathbb{R}_+ \to \mathbb{R}_+$with $h'(\infty) \geq (\beta/\alpha)h'(0+)$. ◊

3.4 *Applications*

3.4.1 *Optimal growth theory* Consider two countries (planners, etc.), each with an initial capital stock $k_0 > 0$, and each with access to a twice differentiable production technology $f : \mathbb{R}_+ \to \mathbb{R}_+$, where $f(0) = 0$, $f' > 0$, and $f'' \leq 0$. In every period t, each country must decide how to divide its capital stock k_t between consumption c_t and investment i_t, where $k_t = f(i_{t-1})$, $t = 1, 2, \ldots$. (There is no capital depreciation.) The preferences of Country 1 over consumption paths is represented by the exponential intertemporal utility function (u, α), and those of Country 2 by (v, β). We assume that u is twice differentiable and that $u' > 0$, $u'' < 0$, and $u'(0+) = \infty$, and similarly for v. The optimization problem of Country 1 is to choose nonnegative sequences $(c_0, c_1, \ldots)$ and $(i_0, i_1, \ldots)$ in order to

$$\text{maximize} \sum_{t=0}^{\infty} \alpha^t u(c_t) \quad \text{subject to } c_0 + i_0 = k_0 \text{ and } c_t + i_t = f(i_{t-1}),\ t = 1, 2, \ldots.$$

The problem of Country 2 is analogous.

Let the optimal consumption and investment paths of Country $\mathbf{j} = 1, 2$ be denoted by $(c_{0,\mathbf{j}}, c_{1,\mathbf{j}}, \ldots)$ and $(i_{0,\mathbf{j}}, i_{1,\mathbf{j}}, \ldots)$, respectively. The optimal capital accumulation path of Country $\mathbf{j}$ is denoted by $(k_{0,\mathbf{j}}, k_{1,\mathbf{j}}, \ldots)$, $\mathbf{j} = 1, 2$.

Becker (1983) shows that if the two countries have the *same* instantaneous utility function, but, say, Country 2 has a smaller discount factor, then Country 2's optimal capital stock is always larger than Country 1's. As we know from Corollary 2, Becker's comparative static makes Country 2 more delay averse than Country 1, but it also does more than this (for instance, it holds any consumption smoothing motive constant). This raises the following question: Is his finding on differences in the optimal capital stock paths a finding (just) about differences in delay aversion, or does it depend upon other factors as well? Put differently, would the conclusion of Becker's analysis be different if Country 2 were assumed to be more delay averse than Country 1, without further assuming that the two countries had identical instantaneous utilities?

To answer this question, recall that the optimal consumption and investment paths of Countries 1 and 2 are determined by the following Ramsey–Euler equations:

$$u'(c_{t,\mathbf{1}}) = \alpha u'(c_{t+1,\mathbf{1}}) f'(i_{t,\mathbf{1}}) \quad \text{and} \quad v'(c_{t,\mathbf{2}}) = \beta v'(c_{t+1,\mathbf{2}}) f'(i_{t,\mathbf{2}}), \quad t = 0, 1, \ldots.$$

Since $(u,\alpha) \precsim (v,\beta)$, these equations and Corollary 3 yield

$$\frac{\beta}{\alpha}\frac{u'(c_{t,\mathbf{1}})f'(i_{t,\mathbf{2}})}{u'(c_{t+1,\mathbf{1}})f'(i_{t,\mathbf{1}})} = \frac{v'(c_{t,\mathbf{2}})}{v'(c_{t+1,\mathbf{2}})} \geq \frac{\beta}{\alpha}\frac{u'(c_{t,\mathbf{2}})}{u'(c_{t+1,\mathbf{2}})}, \quad t=0,1,\dots$$

so that

$$\frac{u'(c_{t,\mathbf{1}})f'(i_{t,\mathbf{2}})}{u'(c_{t,\mathbf{2}})f'(i_{t,\mathbf{1}})} \geq \frac{u'(c_{t+1,\mathbf{1}})}{u'(c_{t+1,\mathbf{2}})}, \quad t=0,1,\dots. \tag{7}$$

Now suppose the optimal capital stock of Country 1 falls strictly below that of Country 2 at some period, and let $T \in \mathbb{N}$ be the first period at which this happens. That is, $k_{t,\mathbf{1}} \geq k_{t,\mathbf{2}}$ for all $t=0,\dots,T-1$ and $k_{T,\mathbf{1}} < k_{T,\mathbf{2}}$. Clearly, $i_{T-1,\mathbf{1}} < i_{T-1,\mathbf{2}}$ and $c_{T-1,\mathbf{1}} > c_{T-1,\mathbf{2}}$. Since $f'' \leq 0$ and $u'' < 0$, we have $[u'(c_{T-1,\mathbf{1}})f'(i_{T-1,\mathbf{2}})]/[u'(c_{T-1,\mathbf{2}})f'(i_{T-1,\mathbf{1}})] < 1$ and, by (7), also $u'(c_{T,\mathbf{1}})/u'(c_{T,\mathbf{2}}) < 1$. Thus $c_{T,\mathbf{1}} > c_{T,\mathbf{2}}$. Since $k_{T,\mathbf{1}} < k_{T,\mathbf{2}}$, we then also have $i_{T,\mathbf{1}} < i_{T,\mathbf{2}}$, and repeating the previous argument yields $c_{T+1,\mathbf{1}} > c_{T+1,\mathbf{2}}$. Continuing this way inductively, we find that $(c_{T,\mathbf{1}}, c_{T+1,\mathbf{1}},\dots) > (c_{T,\mathbf{2}}, c_{T+1,\mathbf{2}},\dots)$. This contradicts the optimality of the consumption path $(c_{0,\mathbf{2}}, c_{1,\mathbf{2}},\dots)$ for Country 2, since, given that $k_{T,\mathbf{1}} < k_{T,\mathbf{2}}$, it is feasible for Country 2 to consume $(c_{T,\mathbf{1}}, c_{T+1,\mathbf{1}},\dots)$ instead of $(c_{T,\mathbf{2}}, c_{T+1,\mathbf{2}},\dots)$.

It follows that $k_{T,\mathbf{1}} < k_{T,\mathbf{2}}$ cannot hold for *any* $T \in \mathbb{N}$. Thus, Becker's result is indeed about relative delay aversion. More precisely, *in the one-sector optimal growth model, the optimal capital stock of a country can never fall strictly below that of a more delay averse country.*[21]

3.4.2 *Repeated games* When a single-shot game is repeated, new equilibrium possibilities arise as players trade off present and future gains. In some general sense, one expects that the more people value the future, the greater is the equilibrium set. In this section we address this issue.

Although it is standard to model the players in repeated games as exponential utility maximizers, we do not impose that restriction here. Rather, we allow players to use arbitrary intertemporal utility functions in their evaluations of payoff streams. With such a level of generality, it might seem difficult to derive a revealing result. Nevertheless, the notion of relative delay easily yields a positive finding.

Let $n \in \{2,3,\dots\}$. Consider an arbitrary (single-shot) game $\mathscr{G} := (N, \{A_i, p_i\}_{i\in N})$, where $N := \{1,\dots,n\}$ is the set of players, A_i is the action space of player i, $A := A_1 \times \cdots \times A_n$ is the outcome space, and $p_i : A \to \mathbb{R}$ is the function that maps each outcome to a *monetary* payoff. For each $i \in N$, let m_i denote a pure strategy profile that minmaxes player i in the game $\mathscr{G}$. For any $a \in A$ and $i \in N$, let $\mathsf{B}_i(a)$ denote a best response of player i to $a \in A$. We assume that the sets of minmax strategies and the sets of best responses are always nonempty.

Let $(\mathscr{G}, \{V_i\}_{i\in N})$ be the infinitely repeated game in which the stage game is $\mathscr{G}$ and $V_i \in \mathfrak{U}$ is the intertemporal utility function that player $i \in N$ uses to evaluate her monetary payoff streams. Let S_i stand for the set of all pure strategies of player $i \in N$. (We restrict

[21]This result extends readily to the case of countries with preferences that are separable but not exponential (with virtually the identical argument). Since the optimal plans, as viewed from period zero, may then not be time consistent, this extension is best interpreted as concerning countries that can *commit* to their plans.

ourselves to pure strategies as we do not consider individuals' attitudes towards risk.) Formally, for any $i \in N$, a strategy $s_i \in S_i$ is a sequence $(s_i^0, s_i^1, \ldots)$ where $s_i^0 \in A_i$ and $s_i^t : A^t \to A_i$ for each $t \in \mathbb{N}$. We let $S := S_1 \times \cdots \times S_n$, and for any $s \in S$, write s^t for $(s_1^t, \ldots, s_n^t)$, which is a map from A^t into A. Any strategy profile $s \in S$ inductively defines an outcome path $(a_0(s), a_1(s), \ldots)$ for the repeated game $(\mathcal{G}, \{V_i\}_{i\in N})$ as follows: $a_0(s) := s^0$ and $a_t(s) := s^t(a_0(s), \ldots, a_{t-1}(s))$, $t = 1, 2 \ldots$. A *Nash equilibrium* of this game is a strategy profile $s \in S$ such that

$$V_i(p_i(a_0(s), p_i(a_1(s)), \ldots)) \geq V_i(p_i(a_0(s_i, s_{-i}), p_i(a_1(s_i, s_{-i})), \ldots))$$

for all $s_i \in S_i$ and $i \in N$.[22] The corresponding outcome path is a called a *Nash equilibrium outcome path*.

The simple intuition that the equilibrium payoff set is larger with less delay averse players is not correct in general. Indeed, Sorin (1986) shows that an equilibrium payoff stream need not remain one, even in the simple case of exponential utility maximizers whose discount factors increase while their utility functions remain constant. However, our next proposition shows that the intuition is correct, even outside the exponential discounting model, provided that each player gets at least his minmax payoff in every period along the equilibrium path.[23]

PROPOSITION 2. *Suppose that $(a_0, a_1, \ldots)$ is a Nash equilibrium outcome path of the repeated game $(\mathcal{G}, \{V_i\}_{i\in N})$. For all $i \in N$, let $U_i \in \mathfrak{U}$ be less delay averse than V_i, and for each i with $U_i \neq V_i$ suppose that*

$$p_i(a_t) \geq p_i(m_i) \quad t = 0, 1, \ldots.$$

Then $(a_0, a_1, \ldots)$ is also a Nash equilibrium outcome path of $(\mathcal{G}, \{U_i\}_{i\in N})$.

To see this, note first that since an intertemporal utility function is increasing in single period payoffs, the most efficient "threat" against a potential deviator is to minmax him for the remainder of the game. That is, the path $(a_0, a_1, \ldots)$ is an equilibrium of $(\mathcal{G}, \{V_i\}_{i\in N})$ if and only if for each player $i \in N$,

$$V_i(p_i(a_0), p_i(a_1), \ldots) \geq V_i(p_i(\mathsf{B}_i(a_0)), p_i(m_i), p_i(m_i), \ldots)$$

and

$$V_i(p_i(a_0), p_i(a_1), \ldots) \geq V_i(p_i(a_0), \ldots, p_i(a_{k-1}), p_i(\mathsf{B}_i(a_k)), p_i(m_i), p_i(m_i), \ldots)$$

[22] If each V_i is a time-consistent utility function, then this is just the standard definition of an equilibrium. Otherwise, it presupposes that players commit to their strategies in period 0. The reader who is perturbed by this latter case is free to restrict his or her attention to (possibly non-separable) time consistent utility functions.

[23] For exponential utility maximizers, it follows from Theorem 3 of Abreu et al. (1990) that if players have access to a public randomization device, then any equilibrium payoff of an infinitely repeated game remains an equilibrium payoff as the players' discount factors increase, holding their instantaneous utility functions constant.

for all $k \in \mathbb{N}$. For each $i \in N$ with $U_i \neq V_i$, since $p_i(a_t) \geq p_i(m_i)$ for all t, the path $(p_i(a_0), p_i(a_1), \ldots, p_i(a_{k-1}), p_i(\mathsf{B}_i(a_k)), p_i(m_i), p_i(m_i), \ldots)$ single-crosses $(p_i(a_0), p_i(a_1), \ldots)$ from above. It follows immediately from Theorem 1 that $U_i \precsim V_i$ implies

$$U_i(p_i(a_0), p_i(a_1), \ldots) \geq U_i(p_i(\mathsf{B}_i(a_0)), p_i(m_i), p_i(m_i), \ldots)$$

and

$$U_i(p_i(a_0), p_i(a_1), \ldots) \geq U_i(p_i(a_0), \ldots, p_i(a_{s-1}), p_i(\mathsf{B}_i(a_k)), p_i(m_i), p_i(m_i), \ldots)$$

for each $i \in N$ and $k \in \mathbb{N}$. Thus $(a_0, a_1, \ldots)$ is an equilibrium path for $(\mathscr{G}, \{U_i\}_{i \in N})$, as claimed.[24]

3.4.3 *Tree-cutting with time-inconsistent preferences* The applications we have considered so far either use time-consistent preferences, or presume that a commitment technology is available to the decision makers. In this section we consider agents with time-inconsistent preferences who do not have the ability to commit to a plan. It is well-known that such agents may exhibit counter-intuitive behavior (O'Donoghue and Rabin 1999). We too find such anomalies. We show, however, that if agents are "present-biased" these anomalies disappear, at least in the context of the so-called *Wicksell tree-cutting problem.*

Consider a "tree" of initial size x_0, whose growth is described by a strictly increasing (and bounded) production function. An agent with initial endowment stream $\boldsymbol{\omega}$ and intertemporal utility function $U \in \mathscr{U}$ must decide when to cut down the tree and reap its benefits. Formally, a *tree-cutting problem* can be described by a list $(\boldsymbol{\omega}, \mathbf{x}) \in \mathscr{X}^2$ such that $0 < x_0 \leq x_1 \leq \cdots$. The set of potential outcomes of the problem $(\boldsymbol{\omega}, \mathbf{x})$ is

$$\mathscr{X}(\boldsymbol{\omega}, \mathbf{x}) := \{(\omega_\tau + x_\tau, \boldsymbol{\omega}_{-\tau}) : \tau = 0, 1, \ldots\},$$

where τ represents the period in which the tree is cut.[25]

Given the sequential nature of the tree-cutting problem, at any period t in which the tree has not yet been cut the agent in effect faces a fresh tree-cutting problem $((\omega_t, \omega_{t+1}, \ldots), (x_t, x_{t+1}, \ldots))$. In keeping with the recent literature, we assume that for each of these t-period problems, the agent views himself as a different player—his t-*self*—with utility function U_t defined on $\mathscr{X}(\boldsymbol{\omega}, \mathbf{x})$ by

$$U_t(\omega_\tau + x_\tau, \boldsymbol{\omega}_{-\tau}) := \begin{cases} U(\omega_t + x_t, \omega_{t+1}, \ldots) & \tau = t \\ U(\omega_t, \ldots, \omega_{\tau-1}, \omega_\tau + x_\tau, \omega_{\tau+1}, \ldots) & \text{if } \tau > t. \end{cases} \tag{8}$$

[24] Proposition 2 readily implies that, if repeating the symmetric Pareto Optimal outcome in a Prisoner's Dilemma is part of a *subgame perfect* equilibrium, it remains so as the players become less delay averse.

[25] A slightly more involved version of this model would allow for the tree to "bear fruit" in every period—for example, a bond is described by such a tree. We work here with fruitless trees only to simplify the subsequent analysis. All of our conclusions (summarized in Proposition 3 below) remain valid for the more general version of the tree-cutting problem.

In a manner of speaking, whenever a period t is reached the agent treats it as if it is period zero.[26] We consider three (standard) ways in which the agent might solve a tree-cutting problem.

Commitment Agent A commitment agent is presumed to have access to a commitment technology, in the sense that he makes an optimal plan in period 0, and adheres to it.

Naive Agent A naive agent makes an initial optimal plan in period 0, but then reoptimizes in every subsequent period (using the utility function defined in (8)). At any given period $t \geq 0$, the agent optimizes assuming, possibly incorrectly, that this plan *will be followed* in the subsequent periods.

Sophisticated Agent Like the naive agent, a sophisticated agent solves an optimization problem in every period. However, a sophisticated agent makes his optimal plans in every period *taking into account* his reoptimization behavior. In effect, the decision maker perceives his problem as an extensive-form game played noncooperatively between his t-selves, and solves for the subgame perfect equilibrium of this game.[27]

We now ask the following question:

> Take an arbitrary tree-cutting problem $(\boldsymbol{\omega}, \mathbf{x})$ and two agents V and U of the same type (be it commitment, naive, or sophisticated). If agent V is more delay averse than agent U, does V necessarily cut the tree (weakly) sooner than U?

The answer to this question depends upon the type of agents involved. For a commitment agent, the answer is, obviously, yes. Indeed, suppose a commitment U cuts down the tree in period $t \in \mathbb{Z}_+$. This implies that $U(\omega_t + x_t, \boldsymbol{\omega}_{-t}) > U(\omega_T + x_T, \boldsymbol{\omega}_{-T})$ for any $T \in \{t+1, t+2, \ldots\}$.[28] Since V is more delay averse than U, we also have $V(\omega_t + x_t, \boldsymbol{\omega}_{-t}) > V(\omega_T + x_T, \boldsymbol{\omega}_{-T})$ for any $T \in \{t+1, t+2, \ldots\}$, so that V cuts the tree no later than period t. A similar argument applies to naive agents. Specifically, if a naive U cuts the tree in period t, then $U(\omega_t + x_t, \omega_{t+1}, \ldots) > U(\omega_t, \ldots, \omega_{T-1}, \omega_T + x_T, \omega_{T+1}, \ldots)$ for any $T \in \{t+1, t+2, \ldots\}$. Since V is more delay averse than U, $V(\omega_t + x_t, \omega_{t+1}, \ldots) > V(\omega_t, \ldots, \omega_{T-1}, \omega_T + x_T, \omega_{T+1}, \ldots)$ for any $T \in \{t+1, t+2, \ldots\}$, so that again V cuts the tree no later than U does.

The most interesting case is that of a sophisticated agent. As the next example shows, a sophisticated agent may cut a tree *later* as he becomes more delay averse.

[26] An implicit assumption here is that, for the t-self of the agent, consumption in periods before t does not matter. This assumption is satisfied by any separable utility function U.

[27] The basic idea behind the sophisticated approach was introduced by Strotz (1956) and Phelps and Pollak (1968). The game-theoretical modeling of time consistent decision making (with time-inconsistent preferences) was then developed by several authors (cf. Peleg and Yaari 1973, Goldman 1980, Kocherlakota 1996, and Asheim 1997). This approach, which coincides with the naive approach for time consistent preferences, is adopted routinely in applications of the hyperbolic discounting model (cf. Laibson 1997, O'Donoghue and Rabin 1999, Carillo and Mariotti 2000).

[28] As a simplification, we assume away the possibility of indifferences in this discussion.

EXAMPLE 4. Consider the tree cutting problem $((0,0,\ldots),(10,21,33,33,\ldots))$, and an agent with separable preferences $(u,\boldsymbol{\alpha})$, where u is the identity function, and $\boldsymbol{\alpha}(0):=1$, $\boldsymbol{\alpha}(1):=\frac{2}{3}$ and $\boldsymbol{\alpha}(t):=(\frac{1}{5})^{t-1}$, $t=2,3,\ldots$. Clearly, the only issue here is whether the agent will cut the tree in period 0, 1, or 2.

Suppose the agent is sophisticated. Presented with an uncut tree, his 1-self would not cut it, preferring that it be cut in period 2, since $21<\frac{2}{3}(33)$. Anticipating this, his 0-self cuts the tree immediately, since $10>\frac{1}{5}(33)$.

Now suppose the agent's preferences change to $(u,\boldsymbol{\beta})$, where $\boldsymbol{\beta}(0):=1$, $\boldsymbol{\beta}(1):=\frac{1}{2}$ and $\boldsymbol{\beta}(t):=(\frac{1}{10})^{t-1}$, $t=2,3,\ldots$. Corollary 2 readily yields that $(u,\boldsymbol{\beta})$ is strictly more delay averse than $(u,\boldsymbol{\alpha})$. With the preferences $(u,\boldsymbol{\beta})$, the 1-self would choose to cut the tree, if given the opportunity, since $21>\frac{1}{2}(33)$. Given this, the 0-self does *not* cut down the tree, but waits for the tree to be cut in period 1, since $10<\frac{1}{2}(21)$. ◊

Thus we find that a sophisticated agent may act in a less delay averse manner as he becomes more delay averse. While somewhat paradoxical, this reversal is arguably not terribly surprising, given the game-theoretic nature of a sophisticated agent's decision problem. Less paradoxical, yet more surprising, at least a priori, is that such a reversal is precluded if the agent's preferences exhibit a present bias, as in the next definition.[29]

DEFINITION 5. An intertemporal utility function $U\in\mathscr{U}$ is said to exhibit *present bias* if, for any $\boldsymbol{\omega}\in\mathscr{X}$, $a,b\geq 0$ and $s,t\in\mathbb{Z}_{++}$ with $t>s>1$,

$$U(\omega_0,\ldots,\omega_{s-1},\omega_s+a,\omega_{s+1},\ldots)\begin{Bmatrix}\geq\\>\end{Bmatrix}U(\omega_0,\ldots,\omega_{t-1},\omega_t+b,\omega_{t+1},\ldots).$$

implies

$$U(\omega_1,\ldots,\omega_{s-1},\omega_s+a,\omega_{s+1},\ldots)\begin{Bmatrix}\geq\\>\end{Bmatrix}U(\omega_1,\ldots,\omega_{t-1},\omega_t+b,\omega_{t+1},\ldots).$$

Loosely speaking, the closer a period is to the present, the more important it is to a present-biased agent. To see this, suppose an agent is present-biased and that he initially prefers receiving a payment in some period s to receiving a possibly different payment at a later time. Then, in every subsequent period up to, and including, period s, he continues to favor receiving the earlier s-period payment.[30] It is easy to show that a separable intertemporal utility function $(u,\boldsymbol{\delta})$ displays present bias if and only if $\boldsymbol{\delta}(t)/\boldsymbol{\delta}(s)\geq\boldsymbol{\delta}(t-1)/\boldsymbol{\delta}(s-1)$ for all s, $t\in\mathbb{Z}_{++}$ with $t>s$. Present bias is a property that is shared by all hyperbolic discounting models, and which has received support from psychological research. (See Frederick et al. 2002 for a large set of references.)

[29]Although we were not able to find the type of formalization presented in the definition in the existing literature, the basic idea is commonly known. (The only formal definition we know is that of Prelec (2004). While consistent with our definition, Prelec's formalization is confined to a very specialized time preference context.) To our knowledge, the term "present bias" was first coined by O'Donoghue and Rabin (1999).

[30]This fact follows inductively from Definition 5.

The preferences $(u,\boldsymbol{\alpha})$ in Example 4 are not present-biased. It turns out that this is not happenstance. As the next proposition indicates, a sophisticated and present-biased agent always cuts a tree sooner if he becomes more delay averse. To simplify the statement of the proposition, we consider only problems $(\boldsymbol{\omega},\mathbf{x})$ that are generic in the sense that, for a given intertemporal utility function U, for any $t\in\mathbb{Z}_+$,

$$U_t(\omega_\tau+x_\tau,\boldsymbol{\omega}_{-\tau})\neq U_t(\omega_{\tau'}+x_{\tau'},\boldsymbol{\omega}_{-\tau'})\quad\text{for all }\tau,\tau'\in\{t,t+1,\ldots\}.$$

We refer to any such tree-cutting problem as *generic for* U.

PROPOSITION 3. *Let $U,V\in\mathscr{U}$ and $(\boldsymbol{\omega},\mathbf{x})$ be a tree-cutting problem that is generic for U, and suppose $U\precsim V$. Then the optimal cutting time for agent U is no later than the optimal cutting time for agent V (provided these optima exist), if both agents can commit or if both agents are naive. If the agents are sophisticated, this need not be the case. However, if agent U is present-biased, then agent U's optimal cutting time is no later than agent V's (provided these optima exist), whether they both can commit, are both naive, or are both sophisticated.*

Given our earlier discussion, we need to prove only the final assertion of this proposition. We use the following auxiliary observation for this purpose.

LEMMA 2. *Let $U\in\mathscr{U}$ and $(\boldsymbol{\omega},\mathbf{x})$ be a tree-cutting problem that is generic for U. Suppose that a sophisticated U optimally cuts the tree at time $t\in\mathbb{Z}_+$, while, conditional on reaching period $t+1$ without cutting the tree, he would have cut the tree at time T. If U exhibits present bias, then*

$$U_t(\omega_t+x_t,\boldsymbol{\omega}_{-t})>U_t(\omega_\tau+x_\tau,\boldsymbol{\omega}_{-\tau})\quad\textit{for all }\tau=t+1,\ldots,T.$$

Before proving this fact, let us see how it allows us to settle the final assertion of Proposition 3. Take any $U\in\mathscr{U}$ that has present bias, and fix a tree-cutting problem $(\boldsymbol{\omega},\mathbf{x})$ that is generic for U. Now suppose that the agent U is sophisticated, and chooses to cut the tree at time $t\in\mathbb{Z}_+$. Let $T_1>t$ be the time that this agent would choose to cut the tree, conditional on reaching period $t+1$ without cutting it. By Lemma 2, $U_t(\omega_t+x_t,\boldsymbol{\omega}_{-t})>U_t(\omega_\tau+x_\tau,\boldsymbol{\omega}_{-\tau})$ for all $\tau=t+1,\ldots,T_1$. Since $U\precsim V$, we also have $V_t(\omega_t+x_t,\boldsymbol{\omega}_{-t})>V_t(\omega_\tau+x_\tau,\boldsymbol{\omega}_{-\tau})$ for all $\tau=t+1,\ldots,T_1$, and hence V cannot choose to cut the tree in $\{t+1,\ldots,T_1\}$. Now let $T_2>T_1$ be the time that the agent U would choose to cut the tree, conditional on reaching period T_1+1 without cutting it. Then, by applying Lemma 2 to the tree-cutting problem $((\omega_{T_1},\omega_{T_1+1},\ldots),((x_{T_1},x_{T_1+1},\ldots))$ and reasoning analogously, we find that the optimal cutting time of the agent V cannot be in $\{T_1+1,\ldots,T_2\}$ either. The final assertion of Proposition 3 obtains inductively.

We conclude by proving Lemma 2. Begin by noting that

$$U(\omega_t+x_t,\omega_{t+1},\ldots)>U(\omega_t,\ldots,\omega_{T-1},\omega_T+x_T,\omega_{T+1},\ldots),\tag{9}$$

otherwise cutting the tree at time t would not have been optimal for the t-self of the agent. Thus, if $T=t+1$, we are done. Assume then that $T>t+1$, and take any

$\tau \in \{t+1,\ldots,T-1\}$. To derive a contradiction, let us suppose that $U_t(\omega_t + x_t, \boldsymbol{\omega}_{-t}) \leq U_t(\omega_\tau + x_\tau, \boldsymbol{\omega}_{-\tau})$. Combining this inequality with (9) yields

$$U(\omega_t,\ldots,\omega_{\tau-1},\omega_\tau + x_\tau,\omega_{\tau+1},\ldots) > U(\omega_t,\ldots,\omega_{T-1},\omega_T + x_T,\omega_{T+1},\ldots).$$

Since U has present bias, we then have

$$U(\omega_\tau + x_\tau,\omega_{\tau+1},\ldots) > U(\omega_\tau,\ldots,\omega_{T-1},\omega_T + x_T,\omega_{T+1},\ldots).$$

But this implies that the τ-self of the agent would prefer to cut the tree in period τ instead of waiting for period T, contradicting the hypothesis that, conditional on reaching period $t+1$ without cutting it, the agent would cut the tree in period T.

4. Comparative impatience

4.1 *Main definition*

Consider an agent whose endowment stream is $(0,10,10,\ldots)$, and who receives a lump sum award of 10 which he may add to his consumption in any single period. Suppose he chooses to consume the 10 in period zero. In so doing, he is certainly exhibiting some aversion towards delay, but how much does this choice really tell us about his attitude towards time? Arguably, very little. After all, his decision to consume the 10 immediately may stem more from a reaction to the uneven endowment stream than a taste for early gratification. Indeed, it would hardly seem surprising, or inconsistent, if the same person informed us that he would have consumed the additional 10 in period one, rather than period zero, had his endowment stream been $(10,0,10,10,\ldots)$ instead of $(0,10,10,\ldots)$.

The same ambiguity arises when making comparative statements. Consider agents A and B, both with the endowment stream $(10,0,10,\ldots)$. Suppose that agent A chooses to consume an additional 10 immediately, while agent B waits one period. Then agent A is certainly acting in a more delay averse manner than agent B, but does that mean that A has the greater bias towards the present per se? Not necessarily. It may well be that the two agents are equally present-oriented, but that agent B has a stronger reaction to the uneven endowments.

The definition of relative delay aversion (purposely) makes no attempt to disentangle the various motives that go into an agent's allocation decisions. Rather, it blends them to yield a very strong notion: One agent is more delay averse than another if and only if his behavior is *always* more biased towards the present. While this universal requirement is somewhat demanding, it is significantly less stringent than the common practice of holding an agent's instantaneous utility function fixed while varying his discount factor (in the case of the exponential discounting model). Furthermore, it is in the spirit of many prior intertemporal analyses. Of the earlier major thinkers about time preferences, Böhm-Bawerk (1891) and Fisher (1930) were particularly clear about the attitudes of an individual towards time delay being a consequence of two effects: a reaction to an uneven endowment stream and a preference for early consumption.[31] In line with this point of view, the notion of delay aversion commingles these two effects.

At the same time, many authors, including Friedman (1972), Olson and Bailey (1981), and Stigler (1966), wish to concentrate on agents' pure time preferences.[32] Our methodology also affords a means for isolating pure time biases. Recall that the delay aversion ordering $\precsim$ is defined through simple choice problems relative to *all* endowment streams, however uneven these streams may be. It is due to this fact that it picks up influences besides pure time considerations. The latter effect can be isolated by comparing the preferences of agents only when their endowment streams are neutral with respect to time, as in the following definition. (See Figure 2.) We reserve the term *impatience* for this pure time notion.

DEFINITION 6. Let U and V be two intertemporal utility functions. We say that V is *more impatient than U* if (3) and (4) hold for all endowment streams $\boldsymbol{\omega} \in \mathscr{X}$ such that $\omega_0 = \omega_1 = \cdots$. If V is more impatient than U but not conversely, we say that V is *strictly more impatient than U*.

Note that by its very definition, this impatience ordering is a refinement of our delay aversion ordering.[33] Like $\precsim$, this ordering is a preorder on $\mathscr{U}$ that is continuous (relative to the topology of uniform convergence).

4.2 *Comparative impatience with separability*

We next examine the comparative impatience ordering in the context of separable intertemporal utility functions. We first state the analogue of the first part of Theorem 2 for this ordering.

THEOREM 3. *Let $(u, \boldsymbol{\alpha})$ and $(v, \boldsymbol{\beta})$ be separable intertemporal utility functions. Then $(v, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$ if and only if there exists a map $h : \mathbb{R}_+ \to \mathbb{R}_+$ such that $v = h \circ u$ and*

$$h\left(\left(1 - \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}\right)x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right) \geq \left(1 - \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}\right)h(x) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}h(y) \tag{10}$$

and

$$h\left(\left(1 - \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}\right)y + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}x\right) \leq \left(1 - \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}\right)h(y) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}h(x) \tag{11}$$

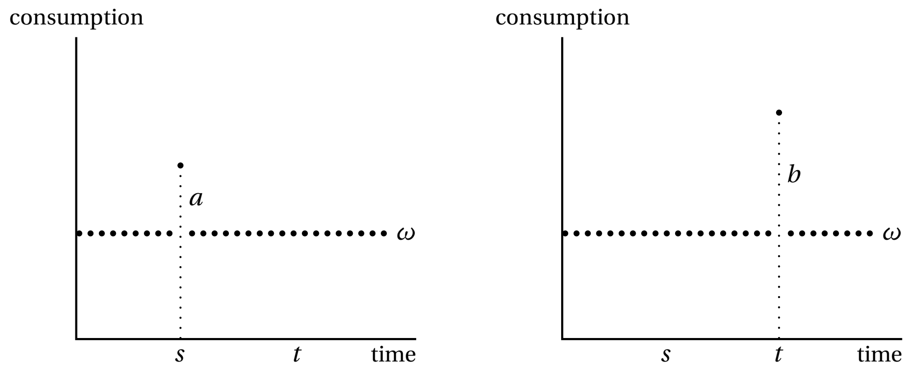
for all $s, t \in \mathbb{Z}_+$ with $s < t$ and $0 \leq x \leq y$.

---

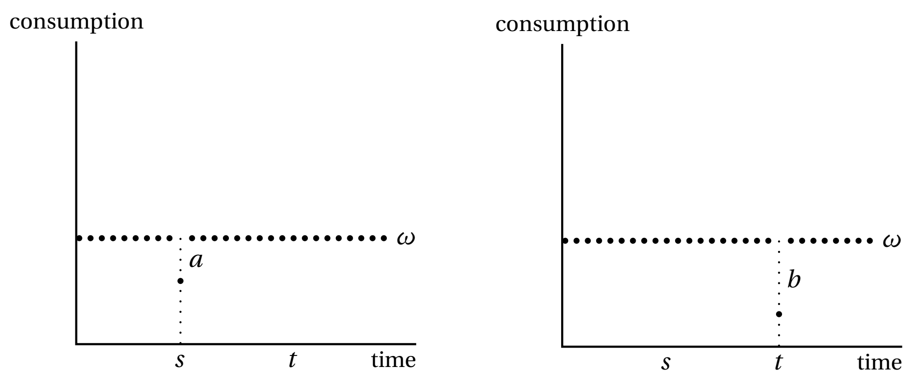
[31] In the words of Böhm-Bawerk (1891, p. 275), "First, the individual may be badly off in the present. In that case the pressing wants of the moment will, by themselves, absorb the small stock of present goods, and, on the ground of this bad provision in the present, these goods will obtain a high value and a preference over future goods. ... Or, second, the individual may be equally well provided as regards both present and future, but may have less forethought. This case leads to a similar result."

[32] To this end, working with exponential utilities, these authors define an individual's "absolute impatience" by considering his marginal rate of substitution between earlier and later consumption along *constant* endowment streams.

[33] We note that (3) may hold for all *constant* endowment streams, while (4) does not. Thus, in contrast to Definition 2, there is no redundancy in Definition 6.



a. If an agent prefers the consumption path on the left, so does a more impatient agent.



b. If an agent prefers the consumption path on the right, so does a more impatient agent.

FIGURE 2.

Technically speaking, functional inequalities (10) and (11) are much less binding than (5), because they depend on two variables (x and y) while (5) depends on three variables (x, y, and z).[34] Notice that since $\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)$ is a number between 0 and 1, (10) is a functional inequality that resembles the definition of concavity. Indeed, if $\boldsymbol{\alpha} \geq \boldsymbol{\beta}$, this inequality is trivially satisfied by any concave and increasing function on $\mathbb{R}_+$. On the other hand, (11) is a functional inequality more in line with convexity. Consequently, these two functional inequalities act as checks and balances, and tell us that if $(v, \boldsymbol{\beta})$ is to be more impatient than $(u, \boldsymbol{\alpha})$, then v cannot be a "too concave" or "too convex" transformation of u, where the permissible amount of concavity and convexity (or, more generally, the variation in slope) depends on the discount functions $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The use of Theorem 3 is similar to that of Theorem 2. In particular, the exact analogues of the corollaries we deduced from Theorem 2 in the previous subsection can be obtained from Theorem 3 for our impatience ordering.

[34] Both (10) and (11) (when stated for all $0 \leq x \leq y$) are special cases of (5) (when stated for all $x, y, z \geq 0$), as can be seen by a suitable change of variables.

COROLLARY 4. *Let $(u, \boldsymbol{\alpha})$ and $(v, \boldsymbol{\beta})$ be separable intertemporal utility functions. If $(v, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$, then $\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s) \geq \boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)$ whenever $s < t$. Moreover, $(u, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$ if and only if $(u, \boldsymbol{\alpha}) \precsim (u, \boldsymbol{\beta})$, and $(v, \boldsymbol{\alpha})$ is more impatient than $(u, \boldsymbol{\alpha})$ if and only if $u = \theta v$ for some $\theta > 0$.*

As in the case of delay aversion, the discount and utility functions contribute asymmetrically to the determination of the impatience of an individual. Once again, no two individuals with the same discount function but cardinally non-equivalent instantaneous utilities can be ranked, this time according to their relative impatience. On the other hand, two individuals with the same instantaneous preferences may be ranked, in which case the impatience and delay aversion orderings coincide.[35]

In the differentiable case, Theorem 2 provides easy necessary and sufficient conditions for checking if a given separable intertemporal utility function is more delay averse than another. Unfortunately, we have been unable to derive a similar characterization for our impatience ordering. Nevertheless, the following result reports an easy-to-check sufficient condition for the exponential discounting model, and also specializes Theorem 2 to this setting.

COROLLARY 5. *Let (u, α) and (v, β) be exponential intertemporal utility functions such that $u, v \in \mathscr{V}$. Then (v, β) is more impatient than (u, α) if and only if there exists a map $h : \mathbb{R}_+ \to \mathbb{R}_+$ such that $v = h \circ u$,*

$$h((1-\alpha)x + \alpha y) \geq (1-\beta)h(x) + \beta h(y) \tag{12}$$

and

$$h((1-\alpha)y + \alpha x) \leq (1-\beta)h(y) + \beta h(x) \tag{13}$$

for all $0 \leq x \leq y$. Moreover, if $u, v \in \mathscr{V}$ and $h := v \circ u^{-1}$ satisfies

$$\max\left\{\frac{\alpha}{\beta}, \frac{1-\beta}{1-\alpha}\right\} h'(y) \geq h'(x) \geq \min\left\{\frac{\beta}{\alpha}, \frac{1-\alpha}{1-\beta}\right\} h'(y), \tag{14}$$

for all $0 \leq x \leq y$, then (v, β) is at least as impatient as (u, α).

We now give two applications of this corollary. The first application supplies a transformation h that yields more impatient utility functions. The second application provides examples of impatience ranked utility functions that were earlier found to be non-comparable by our delay aversion ordering. This example also shows that the sufficient conditions of Corollary 5 are not necessary.

[35]It is natural that $\precsim$ and the "more impatient than" orderings coincide in comparing the separable intertemporal utility functions $(u, \boldsymbol{\alpha})$ and $(u, \boldsymbol{\beta})$. Loosely speaking, the consumption smoothing motive is identical in any two such utility functions, so the difference in attitudes towards time delay is due *only* to the differences in impatience.

EXAMPLE 5. Take any $k > 0$ and $\theta > 0$, and define $h_\theta : \mathbb{R}_+ \to \mathbb{R}_+$ by $h(x) := kx^\theta$. We claim: (1) For any $\theta \geq 1$, h_θ satisfies (12) and (13) for all $0 \leq x \leq y$ if and only if $\alpha^\theta \geq \beta$, and (2) for any $1 \geq \theta > 0$, h_θ satisfies (12) and (13) for all $0 \leq x \leq y$ if and only if $1-\beta \geq (1-\alpha)^\theta$.

We prove only the first claim, the proof of the second one being analogous. Fix any $\theta \geq 1$. Since h_θ is a strictly increasing convex function, it readily satisfies (13) for all $y \geq x \geq 0$. Thus, all we need to show is that (12) holds for all $0 \leq x \leq y$ if and only if $\alpha^\theta \geq \beta$. For any $x \geq 0$, let us define the map $f_x : [x, \infty) \to \mathbb{R}$ by

$$f_x(y) := ((1-\alpha)x + \alpha y)^\theta - (1-\beta)x^\theta - \beta y^\theta.$$

For any $0 \leq x < y$, we have

$$\frac{d}{dy} f_x(y) = \alpha\theta((1-\alpha)x + \alpha y)^{\theta-1} - \beta\theta y^{\theta-1} \geq \alpha\theta(\alpha y)^{\theta-1} - \beta\theta y^{\theta-1} = \theta y^{\theta-1}(\alpha^\theta - \beta),$$

so it follows that if $\alpha^\theta \geq \beta$ then f_x is an increasing function for $x \geq 0$. Since $f_x(x) = 0$, we have $f_x([x,\infty)) \geq 0$ for all $x \geq 0$ if $\alpha^\theta \geq \beta$. Conversely, if $\alpha^\theta < \beta$, then $(d/dy) f_0(y) < 0$ for small enough $y > 0$. Then, $f_0(y) < 0$ for small enough $y > 0$; that is, h_θ fails to satisfy (12) for $x = 0$ and small $y > 0$. ◊

EXAMPLE 2 CONTINUED. Consider the class $\mathscr{U}_0 \times (0,1)$ of exponential intertemporal utility functions considered in Example 2. We saw earlier that no two members of this class with distinct instantaneous utility functions can be ranked in terms of delay aversion. In contrast, any two such members can be ranked in terms of impatience for appropriate discount factors. Put precisely, if $0 < \sigma_1 \leq \sigma_2 < 1$, then

$$(u_{\sigma_2}, \beta) \text{ is more impatient than } (u_{\sigma_1}, \alpha) \quad \text{if and only if} \quad (1-\beta)^{\frac{1}{1-\sigma_2}} \geq (1-\alpha)^{\frac{1}{1-\sigma_1}},$$

while if $0 < \sigma_2 \leq \sigma_1 < 1$, then

$$(u_{\sigma_2}, \beta) \text{ is more impatient than } (u_{\sigma_1}, \alpha) \quad \text{if and only if} \quad \alpha^{\frac{1}{1-\sigma_1}} \geq \beta^{\frac{1}{1-\sigma_2}}.$$

To prove this, let $\theta := (1-\sigma_2)/(1-\sigma_1)$, and define $h : \mathbb{R}_+ \to \mathbb{R}_+$ by $h(x) := [(1-\sigma_1)^\theta/(1-\sigma_2)]x^\theta$. Clearly, $u_{\sigma_2} = h \circ u_{\sigma_1}$, so by Theorem 4, (u_{σ_2}, β) is more impatient than (u_{σ_1}, α) if and only if h satisfies (12) and (13) for all $0 \leq x \leq y$. It follows from the results of Example 5 that h does indeed satisfy these two inequalities. ◊

4.3 *Applications*

4.3.1 *Optimal growth theory revisited* In Section 3.4.1 we saw that in the optimal solution of the one-sector optimal growth model, a more delay averse country invests less than the other country at every period. A natural question is whether this finding remains true under the weaker assumption that Country 2 is more *impatient* than Country 1. The answer is no. In fact, a country may invest strictly less than a more impatient country at *every* period. To illustrate this, let us simplify the optimal growth problem of

Section 3.4.1 by setting $k_0 = 1$ and $f(x) = x$ for all $x \geq 0$. The resulting model is that of the standard *cake-eating* problem (the size of the cake being 1).[36]

Assume that $u := u_{\sigma_1}$ and $v := u_{\sigma_2}$ for some $0 < \sigma_1, \sigma_2 < 1$, where $u_\sigma(x) := x^{1-\sigma}/(1-\sigma)$ as in Example 2 of Section 3.3. With this specification, the optimal capital accumulation path of Country $\mathbf{j} = 1, 2$ is found as $(\gamma_{\mathbf{j}}, \gamma_{\mathbf{j}}^2, \gamma_{\mathbf{j}}^3, \ldots)$, where $\gamma_{\mathbf{1}} = \alpha^{1/\sigma_1}$ and $\gamma_{\mathbf{2}} = \beta^{1/\sigma_2}$. Therefore, the optimal capital accumulation path of Country 2 is everywhere above that of Country 1 if and only if $\alpha^{1/\sigma_1} < \beta^{1/\sigma_2}$. Notice that this implies $\sigma_2 > \sigma_1$. Now recall from the continuation of Example 2 above that when $\sigma_2 > \sigma_1$, Country 2 is more impatient than Country 1 if and only if $(1-\beta)^{1/(1-\sigma_2)} \geq (1-\alpha)^{1/(1-\sigma_1)}$. Therefore, for any specification of $0 < \beta < \alpha < 1$ and $0 < \sigma_1 < \sigma_2 < 1$ such that

$$\alpha^{\frac{1}{\sigma_1}} < \beta^{\frac{1}{\sigma_2}} \quad \text{and} \quad (1-\beta)^{\frac{1}{1-\sigma_2}} \geq (1-\alpha)^{\frac{1}{1-\sigma_1}}, \tag{15}$$

the optimal capital accumulation path of Country 2 is everywhere above that of Country 1, even though Country 2 is the more impatient of the two.[37]

4.3.2 *Repeated games revisited* The previous subsection shows that making changes in peoples' impatience, rather than their delay aversion, may result in radically different conclusions. In contrast, in this section we present a setting in which changes in impatience and delay aversion have similar effects. The following proposition shows that with exponential utility maximizers, any stationary equilibrium path of an infinitely repeated game remains an equilibrium path as players become less *impatient*, even if they do not become less delay averse.[38] We adopt the notation introduced in Section 3.4.2.

Proposition 4. *If $(a, a, \ldots)$ is a Nash equilibrium outcome path of the repeated game $(\mathscr{G}, \{(v_i, \beta_i)\}_{i \in N})$, then $(a, a, \ldots)$ is also a Nash equilibrium outcome path of the repeated game $(\mathscr{G}, \{(u_i, \alpha_i)\}_{i \in N})$, where $(u_i, \alpha_i) \in \mathfrak{U}_{exp}$ is more patient than $(v_i, \beta_i) \in \mathfrak{U}_{exp}$ for each $i \in N$.*

To see this, note that $(a, a, \ldots)$ is an equilibrium path for $(\mathscr{G}, \{(v_i, \beta_i)\}_{i \in N})$ if and only if, for each $i \in N$,

$$\frac{1}{1-\beta_i} v_i(p_i(a)) \geq v_i(p_i(\mathsf{B}_i(a))) + \frac{\beta_i}{1-\beta_i} v_i(p_i(m_i)),$$

that is,

$$v_i(p_i(a)) \geq (1-\beta_i) v_i(p_i(\mathsf{B}_i(a))) + \beta_i v_i(p_i(m_i)). \tag{16}$$

Now suppose that (u_i, α_i) is more patient than (v_i, β_i), and let $h_i := v_i \circ u_i^{-1}$ for each $i \in N$. Then (16) and the functional inequality (13) of Corollary 5 yield

$$\begin{aligned} h_i(u_i(p_i(a))) &\geq (1-\beta_i) h_i(u_i(p_i(\mathsf{B}_i(a)))) + \beta_i h_i(u_i(p_i(m_i))) \\ &\geq h_i((1-\alpha_i) u_i(p_i(\mathsf{B}_i(a))) + \alpha_i u_i(p_i(m_i))) \end{aligned}$$

[36]It may be worth noting that in the special cake eating context, one can easily show that *the optimal consumption path of a country single crosses that of a less delay averse country from above.* (This fact need not hold for an arbitrary optimal growth problem.)

[37]The inequalities in (15) are compatible. For instance, they are satisfied for $\sigma_1 = \frac{1}{4}$, $\sigma_2 = \alpha = \frac{1}{2}$, and any $\beta \in (0.25, 0.35)$.

[38]For less delay averse players, the conclusion is an immediate corollary of Proposition 2.

for each $i \in N$. Since h is increasing,

$$u_i(p_i(a)) \geq (1-\alpha_i)u_i(p_i(\mathsf{B}_i(a))) + \alpha_i u_i(p_i(m_i)) \quad \text{for all } i \in N.$$

Thus, $(a, a, \ldots)$ is an equilibrium path for $(\mathscr{G}, \{(u_i, \alpha_i)\}_{i\in N})$, as was sought.[39]

5. Comparative cryonic delay aversion

5.1 *Main definition*

In intertemporal decision theory, there are two major frameworks. The first of these is the one we have worked within so far, namely, that of infinite (or finite) consumption paths. The second of these takes as its basic choice alternatives dated outcomes that specify the receival date and amount of a given consumption item.[40] This setup corresponds to that of a large number of intertemporal choice experiments in which subjects are asked to compare receiving two sums of money at two different time periods, and are presumed not to consider the intervening periods. It is also used for modeling bargaining games and preemptive investment scenarios, where disagreement (or non-investment) periods are abstracted away from, with incomes in those periods taken to be zero. Put differently, these models maintain that the agents solve their associated problems as if they were "frozen" during intervening periods, making what we term *cryonic* comparisons.

In this alternative framework an agent makes comparisons among consumption paths that are necessarily of the form $(a, \mathbf{0}_{-s})$, where $a \geq 0$, $s \in \mathbb{Z}_+$, and $\mathbf{0} := (0, 0, \ldots)$. Consequently, within this framework it makes sense to apply the comparison methods considered above only with respect to such consumption paths. This prompts the following modification of our delay aversion ordering.

DEFINITION 7. Let U and V be two intertemporal utility functions. We say that V is *cryonically more delay averse than U* if for any given $s, t \in \mathbb{Z}_+$ with $s < t$,

$$U(a, \mathbf{0}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(b, \mathbf{0}_{-t}) \quad \text{implies} \quad V(a, \mathbf{0}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(b, \mathbf{0}_{-t})$$

for all $a, b \geq 0$. In this case we write $U \precsim^{\mathbf{0}} V$.

Note that this definition simply asks that condition (3) hold only when the initial endowment stream is identically zero.[41] Thus, while still partial, the "cryonically more delay averse than" ordering $\precsim^{\mathbf{0}}$ is a significant refinement of the "more impatient than" ordering (which is itself a refinement of $\precsim$). We should note that, within the context of

[39] It is clear from the above proof that Proposition 4 remains valid if "Nash equilibrium" is replaced by "trigger-strategy equilibrium" in its statement. (A trigger-strategy equilibrium is a subgame perfect equilibrium in which deviations from the equilibrium path trigger the repeated play of a single-shot (pure strategy) equilibrium until the end of the game.)

[40] See, for instance, Lancaster (1963), Fishburn and Rubinstein (1982), Rubinstein (2003), and Prelec (2004).

[41] Condition (4) holds vacuously here, since no payments can be made from a starting point of 0.

bargaining theory, a stationary version of this ordering is used by Osborne and Rubinstein (1994).

5.2 *Comparative cryonic delay aversion with separability*

Evidently, the "cryonically more delay averse than" ordering $\precsim^{\mathbf{0}}$ is substantially more complete than our delay aversion and impatience orderings. The following attests to this. (Compare with Theorems 2 and 3.)

THEOREM 4. *For any separable intertemporal utility functions $(u,\boldsymbol{\alpha})$ and $(v,\boldsymbol{\beta})$, we have $(u,\boldsymbol{\alpha})\precsim^{\mathbf{0}}(v,\boldsymbol{\beta})$ if and only if there exists a map $h:\mathbb{R}_+\to\mathbb{R}_+$ such that $v=h\circ u$ and*

$$h\left(\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}x\right)\geq\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}h(x)\quad\text{for all } s,t\in\mathbb{Z}_+ \text{ with } s<t \text{ and } x\geq 0. \tag{17}$$

For any exponential intertemporal utility functions (u,α) and (v,β), we have $(u,\alpha)\precsim^{\mathbf{0}}(v,\beta)$ if and only if there exists a map $h:\mathbb{R}_+\to\mathbb{R}_+$ such that $v=h\circ u$ and

$$h(\alpha x)\geq\beta h(x)\quad\text{for all } x\geq 0.$$

The functional inequality (17) is a special case of the functional inequalities (5) and (10), pointing to the fact that the preorder $\precsim^{\mathbf{0}}$ behaves quite differently from our two previous orderings.

One striking difference is that the analogue of Corollary 1 is false here—it is possible that the exponential intertemporal utility function (v,β) is cryonically more delay averse than (u,α) *even* if $\beta>\alpha$.[42] Another important difference is that it may be possible to rank two separable intertemporal utility functions that have the same discount function. To identify exactly when this occurs, we need to recall the following definition from the theory of functional inequalities: A function $f:\mathbb{R}_+\to\mathbb{R}_+$ is said to be *star-shaped* if $f(\lambda x)\leq\lambda f(x)$ for all $x\geq 0$ and $\lambda\in[0,1]$. One can show that f is star-shaped if and only if $f(0)\leq 0$ and $x\mapsto f(x)/x$ is an increasing map on $\mathbb{R}_{++}$. Thus, if f is convex and $f(0)\leq 0$, then f is star-shaped (but not conversely). Recall that an instantaneous utility function $v\in\mathscr{U}$ is said to be *less convex than* $u\in\mathscr{U}$ if $v=h\circ u$, where $-h$ is some convex function. By analogy, we say that v is *less star-shaped than* u if $v=h\circ u$, where $-h$ is a star-shaped function.

COROLLARY 6. *For any instantaneous utility functions $u,v\in\mathscr{U}$, we have $(u,\boldsymbol{\delta})\precsim^{\mathbf{0}}(v,\boldsymbol{\delta})$ for all $\boldsymbol{\delta}\in\mathscr{D}$ if and only if v is less star-shaped than u. In particular, for any $u,v\in\mathscr{V}$*

$$(u,\boldsymbol{\delta})\precsim^{\mathbf{0}}(v,\boldsymbol{\delta})\text{ for all }\boldsymbol{\delta}\in\mathscr{D}\quad\text{if and only if}\quad\frac{u'(x)}{u(x)}\geq\frac{v'(x)}{v(x)}\text{ for all } x>0. \tag{18}$$

The following is almost an immediate consequence of the previous result.

[42] *Example.* Let $\frac{1}{4}<\alpha<\frac{1}{2}$, and define $u(x):=x$ and $v(x):=\sqrt{x}$. Clearly, $h:=v\circ u^{-1}=v$ while $\sqrt{\alpha x}\geq\frac{1}{2}\sqrt{x}$ for all $x\geq 0$. It follows from Theorem 4 that $(u,\alpha)\prec^{\mathbf{0}}(v,\frac{1}{2})$, even though $\alpha<\frac{1}{2}$.

COROLLARY 7. *For any exponential intertemporal utility functions* (u,α) *and* (v,β)*, we have* $(u,\alpha) \precsim^{\mathbf{0}} (v,\beta)$ *whenever* $\alpha \geq \beta$ *and* v *is less star-shaped than* u*. In particular,* $(u,\delta) \precsim^{\mathbf{0}} (v,\delta)$ *whenever* v *is a concave transformation of* u*.*[43]

5.3 *Application: Bargaining theory*

Roth (1985) argues that in Rubinstein's bargaining model, a player's equilibrium share decreases as he becomes more risk averse. This result is generally regarded in the literature as somewhat difficult to interpret, given that Rubinstein bargaining does not involve any risk.[44] In this section, we argue that Roth's result is in fact properly understood as a result about delay aversion, not risk aversion.

Consider the standard complete-information alternating-offers bargaining game where the size of the pie is 1. The utility function of the first mover, player 1, is a concave function $u \in \mathscr{U}$, and his discount factor is $\alpha \in [0,1]$. The utility function of the second mover, player 2, is also a concave function $w \in \mathscr{U}$, and her discount factor is $\delta \in [0,1]$. Under this specification, the game has a unique subgame perfect equilibrium. Let $(x, 1-x)$ be the equilibrium offer of player 1, and $(1-y, y)$ that of player 2. The values of x and y are determined as the unique solution of the following nonlinear equation system in $[0,1]$:

$$\alpha u(x) = u(1-y) \quad \text{and} \quad \delta w(y) = w(1-x). \tag{19}$$

Since player 1 is the first mover, the realized equilibrium allocation is $(x, 1-x)$.

Now replace player 1 with a player whose utility function is $v \in \mathscr{U}$ and discount factor is $\beta \in [0,1]$. Suppose that $(u,\alpha) \precsim^{\mathbf{0}} (v,\beta)$, that is, this new player is cryonically more delay averse than the original player 1. Let us assume that the resulting bargaining game has a unique subgame perfect equilibrium—a sufficient condition for this is that v be concave. We denote the equilibrium offer of the (new) player 1 by $(x', 1-x')$, and that of player 2 by $(1-y', y')$. The values of x' and y' are determined as the unique solution of the following nonlinear equation system in $[0,1]$:

$$\beta v(x') = v(1-y') \quad \text{and} \quad \delta w(y') = w(1-x') \tag{20}$$

The realized equilibrium allocation is $(x', 1-x')$.

Since the main force behind the equilibrium outcomes in the Rubinstein bargaining game is the attitudes of the players towards time delay, a natural conjecture is that the cryonically more delay averse agent (v,α) should perform less successfully than the agent (u,α), that is, $x \geq x'$. That this is indeed true follows from a general result (Proposition 126.1) of Osborne and Rubinstein (1994). Here we provide an alternative proof using Theorem 4.

[43] The converse of Corollary 7 is false. For instance, define $u, v \in \mathscr{U}$ by $u(x) := \sqrt{x}$ and $v(x) := x$. Clearly, $h(x) = x^2$ for all $x \geq 0$, where $h := v \circ u^{-1}$. Thus, $h(\alpha x) \geq \beta h(x)$ holds for all $x \geq 0$ if and only if $\alpha \geq \sqrt{\beta}$. Hence, $(u,\alpha) \prec^{\mathbf{0}} (v,\beta)$ may hold even when v is *more* convex than v.

[44] Roth himself recognizes this difficulty and claims that the game should be viewed as having "strategic risk." However, the concept of strategic risk is ill-defined, and the connection between this concept and the concavity of a player's static utility function is, at best, ambiguous.

Observe first that (19) and (20) yield

$$x = 1 - w^{-1}(\delta w(y)) \quad \text{and} \quad x' = 1 - w^{-1}(\delta w(y')).$$

Letting $A := y - w^{-1}(\delta w(y))$ and $A' := y' - w^{-1}(\delta w(y'))$, we may write $1 - y = x - A$ and $1 - y' = x' - A'$. Now, towards deriving a contradiction, assume that $x' > x$, that is, $y > y'$. Since w is concave, a routine argument shows that the map $a \mapsto a - w^{-1}(\delta w(a))$ is increasing on $\mathbb{R}_+$. Thus, $y > y'$ implies $A \geq A'$. Since u is strictly increasing and concave,

$$\begin{aligned} u(x'-A') - \alpha u(x') &\geq u(x'-A) - \alpha u(x') \\ &\geq \alpha(u(x'-A) - u(x')) + (1-\alpha)u(x'-A) \\ &\geq \alpha(u(x-A) - u(x)) + (1-\alpha)u(x'-A) \\ &> \alpha(u(x-A) - u(x)) + (1-\alpha)u(x-A) \\ &= u(x-A) - \alpha u(x) \\ &= 0 \end{aligned}$$

where the last equality follows from (19). Thus, $u(1-y') = u(x'-A') > \alpha u(x')$, so by (20) we have

$$\beta h(u(x')) = \beta v(x') = v(1-y') = h(u(1-y')) > h(\alpha u(x')),$$

where $h := v \circ u^{-1}$. Letting $z := u(x')$, we see that $\beta h(z) > h(\alpha z)$, which contradicts (v, β) being cryonically more delay averse than (u, α), in view of Theorem 4. Conclusion: *In the Rubinstein bargaining model, a bargainer's share decreases as he becomes cryonically more delay averse.*

Now let us revisit Roth's result on increasing risk aversion. When Roth performs his comparative static, he takes a concave transformation of one player's instantaneous utility function, holding the player's discount factor constant. Presumably, the discount factor is held constant in order to fix the player's attitude towards time. However, fixing the discount factor does not accomplish the task. Rather, as Corollary 7 shows, a concave transformation of the instantaneous utility function holding the discount factor constant, makes a player cryonically more delay averse. Hence, Roth has actually established a special case of the above result; a non-concave but star-shaped transformation would have yielded him the same conclusion.[45]

6. Concluding comments

This paper begins with the observation that discount factors do not provide a proper basis for a definition of comparative delay aversion. In response, we propose the nested notions of comparative delay aversion, impatience, and cryonic delay aversion. We show, by means of various characterization theorems and economic applications, that these notions are tractable, and that the theory that surrounds them parallels in some major ways the classical theory of risk aversion.

[45] We note that Osborne and Rubinstein's general analysis does not yield any insight into Roth's result; Corollary 7 is essential in this regard.

Needless to say, various avenues of related research have been left unexplored. While we provide characterizations of our delay aversion and impatience orderings for separable intertemporal preferences, the characterization of these orderings for other interesting preference classes, such as recursive preferences, remains to be done. Perhaps more importantly, our entire analysis is confined to decision problems in the absence of risk. A natural avenue of further research is the extension of the delay aversion theory introduced here to environments in which the consumption streams are stochastic. The basic definitions of relative delay aversion, impatience, and cryonic delay aversion are all applicable to intertemporal preferences over stochastic streams, so the present work provides a good starting point for such a study.[46]

7. Appendix: Proofs

7.1 *Proof of Lemma 1*

The nontrivial part of the argument is to establish the cube-continuity of the map defined by (1) for any given $(u,\boldsymbol{\delta})\in\mathscr{U}\times\mathscr{D}$. To this end, fix an arbitrary $a>0$, and define $f:=U|_{[0,a]^\infty}$. We wish to show that f is continuous (in the product topology). Take any $\mathbf{x}\in[0,a]^\infty$ and $\varepsilon>0$. Since $\sum_{t=0}^\infty\boldsymbol{\delta}(t)<\infty$, there is a $T\in\mathbb{N}$ such that

$$\sum_{t=T+1}^{\infty}\boldsymbol{\delta}(t)<\frac{\varepsilon}{2u(a)}.$$

Moreover, since u is continuous, there is a neighborhood O of $\mathbf{x}$ such that O is open in the product topology and

$$|u(x_t)-u(y_t)|<\frac{\varepsilon}{2}\sum_{t=0}^{T}\boldsymbol{\delta}(t)\quad\text{for all }\mathbf{y}\in O\text{ and }t=1,\dots,T.$$

A straightforward application of the triangle inequality yields $|f(\mathbf{x})-f(\mathbf{y})|<\varepsilon$ for any $\mathbf{y}\in O$. We thus conclude that f is continuous. Since $a>0$ is arbitrary in this discussion, it follows that U is cube-continuous.

7.2 *Proofs for Section 3*

We begin with the following preliminary result, which facilitates some of the subsequent arguments.

Lemma 3. *For any $U,V\in\mathfrak{U}$, the following statements are equivalent.*

(a) $U\precsim V$.

[46]Disentangling the effects of delay aversion and risk aversion may be an interesting issue in this regard. Similar issues arise in separating the consumption smoothing and risk aversion motives in intertemporal choice theory with risk (cf. Epstein and Zin 1989).

(b) For any given $\mathbf{x} \in \mathscr{X}$ *and* $s, t \in \mathbb{Z}_+$ *with* $s < t$,

$$U(x_s + a, x_t - b, \mathbf{x}_{-(s,t)}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\mathbf{x}) \quad \textit{implies} \quad V(x_s + a, x_t - b, \mathbf{x}_{-(s,t)}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\mathbf{x})$$

for all $a, b \geq 0$ *with* $x_t \geq b$.

(c) For any given $\boldsymbol{\omega} \in \mathscr{X}$ *and* $s, t \in \mathbb{Z}_+$ *with* $s < t$,

$$U(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\omega_t + b, \boldsymbol{\omega}_{-t}) \quad \textit{implies} \quad V(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\omega_t + b, \boldsymbol{\omega}_{-t})$$

for all $a, b \geq 0$.

Proof. That (a) implies (c) is obvious. To prove (c) implies (b), take any $\mathbf{x} \in \mathscr{X}$, $s, t \in \mathbb{Z}_+$ with $s < t$, and fix any $a, b \geq 0$ with $x_t \geq b$. Define $\boldsymbol{\omega} := (x_t - b, x_{-t})$, and notice that $(\omega_s + a, \boldsymbol{\omega}_{-s}) = (x_s + a, x_t - b, \mathbf{x}_{-(s,t)})$ and $(\omega_t + b, \boldsymbol{\omega}_{-t}) = \mathbf{x}$. That (c) implies (b) is thus evident from this change of variables. It remains to prove that (b) implies (a). To this end, take any $\boldsymbol{\omega} \in \mathscr{X}$, $s, t \in \mathbb{Z}_+$ with $s < t$, fix any $a, b \geq 0$, and assume first that $U(\omega_s + a, \boldsymbol{\omega}_{-s}) \geq (>) U(\omega_t + b, \boldsymbol{\omega}_{-t})$. Define $\mathbf{x} := (\omega_t + b, \boldsymbol{\omega}_{-t})$, and notice that $(x_s + a, x_t - b, \mathbf{x}_{-(s,t)}) = (\omega_s + a, \boldsymbol{\omega}_{-s})$. It thus follows from (b) that $V(\omega_s + a, \boldsymbol{\omega}_{-s}) \geq (>) V(\omega_t + b, \boldsymbol{\omega}_{-t})$. On the other hand, suppose that $\omega_s \geq a \geq 0$ and $\omega_t \geq b \geq 0$, and $U(\omega_t - b, \boldsymbol{\omega}_{-t}) \geq (>) U(\omega_s - a, \boldsymbol{\omega}_{-s})$. Defining $\mathbf{x} := (\omega_s - a, \boldsymbol{\omega}_{-s})$ and applying (b) we find $V(\omega_t - b, \boldsymbol{\omega}_{-t}) \geq (>) V(\omega_s - a, \boldsymbol{\omega}_{-s})$. Thus $U \precsim V$. □

Proof of Theorem 1. We need to prove only the "only if" part of the assertion. Take any $U, V \in \mathfrak{U}$ such that $U \precsim V$. For any $t^* \in \mathbb{N}$, define $T(t^*) := \{(\mathbf{x}, \mathbf{y}) \in \mathscr{X}^2 : U(\mathbf{y}) \geq U(\mathbf{x})$, $\mathbf{y}$ single crosses $\mathbf{x}$ from above and $|\{m \in \mathbb{Z}_+ : y_m > x_m\}| \leq t^*\}$. We wish to show that

$$V(\mathbf{y}) \geq V(\mathbf{x}) \quad \text{for all } (\mathbf{x}, \mathbf{y}) \in T(t^*), \; t^* = 1, 2, \ldots$$

(The case $U(\mathbf{y}) > U(\mathbf{x})$ implies $V(\mathbf{y}) > V(\mathbf{x})$ for all $(\mathbf{x}, \mathbf{y}) \in T(1) \cup T(2) \cup \ldots$ is analogous.) The proof is by induction on t^*.

Take any $(\mathbf{x}, \mathbf{y}) \in \mathscr{X}^2$ and define $\sigma := \sup\{\max\{x_i, y_i\} : i = 1, 2, \ldots\}$. Since U and V are cube-continuous, $U|_{[0,\sigma]^\infty}$ and $V|_{[0,\sigma]^\infty}$ are continuous functions. We use this fact below.

Assume first that $(\mathbf{x}, \mathbf{y}) \in T(1)$. If $\mathbf{x} = \mathbf{y}$, there is nothing to prove, so let $\mathbf{x} \neq \mathbf{y}$. Then we have $|\{m \in \mathbb{Z}_+ : y_m > x_m\}| = 1$. Without loss of generality, we assume $y_0 > x_0$. Let $S := \{m \in \mathbb{N} : y_i < x_i\}$, and to focus on the nontrivial case, suppose that S is an infinite set. We define $s_1 := \min S$ and $s_m := \min S \setminus \{s_1, \ldots, s_{m-1}\}$, $m = 2, 3, \ldots$. By monotonicity of U, we have $U(x_0 + (y_0 - x_0), y_{s_1}, \mathbf{x}_{-(0,s_1)}) > U(\mathbf{y}) \geq U(\mathbf{x})$. Therefore, by continuity of U and the Intermediate Value Theorem, there exists $\xi_1 \in (0, y_0 - x_0)$ such that

$$U(x_0 + \xi_1, y_{s_1}, \mathbf{x}_{-(0,s_1)}) = U(\mathbf{x}) \leq U(\mathbf{y}).$$

By Lemma 3, then, we have

$$V(x_0 + \xi_1, y_{s_1}, \mathbf{x}_{-(0,s_1)}) \geq V(\mathbf{x}).$$

Similarly, there exists $\xi_2 > 0$ such that

$$U(x_0 + \xi_1 + \xi_2, y_{s_1}, y_{s_2}, \mathbf{x}_{-(0,s_1,s_2)}) = U(x_0 + \xi_1, y_{s_1}, \mathbf{x}_{-(0,s_1)}) \leq U(\mathbf{y}),$$

and hence Lemma 3 yields

$$V(x_0 + \xi_1 + \xi_2, y_{s_1}, y_{s_2}, \mathbf{x}_{-(0,s_1,s_2)}) \geq V(x_0 + \xi_1, y_{s_1}, \mathbf{x}_{-(0,s_1)}) \geq V(\mathbf{x}).$$

Proceeding inductively, we obtain a sequence (ξ_m) of positive numbers such that

$$U(\mathbf{z}(m)) \leq U(\mathbf{y}) \quad \text{and} \quad V(\mathbf{z}(m)) \geq V(\mathbf{x}),$$

where

$$\mathbf{z}(m) := (x_0 + \sum_{i=1}^{m} \xi_i, y_{s_1}, \ldots, y_{s_m}, \mathbf{x}_{-(0,s_1,\ldots,s_m)}), \quad m = 1, 2, \ldots.$$

Since $U(\mathbf{z}(m)) \leq U(\mathbf{y})$ for each m, the monotonicity of U implies $z_0(m) \leq y_0$, $m = 1, 2, \ldots$. Being an increasing sequence, then, $(z_0(m))$ must converge to some $a \in (0, y_0]$ as $m \to \infty$. Consequently, for any $\varepsilon > 0$, continuity of $V|_{[0,\sigma]^\infty}$ guarantees the existence of $M_1 > 0$ such that

$$|V(\mathbf{z}(m)) - V(a, \mathbf{z}(m)_{-0})| < \frac{\varepsilon}{2} \quad \text{for all } m \geq M_1.$$

On the other hand, notice that $(a, \mathbf{z}(m)_{-0}) \to (a, \mathbf{y}_{-0})$ as $m \to \infty$ (in the product topology). Thus, since $V|_{[0,\sigma]^\infty}$ is continuous, there exists $M_2 \in \mathbb{N}$ such that

$$|V(a, \mathbf{z}(m)_{-0}) - V(a, \mathbf{y}_{-0})| < \frac{\varepsilon}{2} \quad \text{for all } m \geq M_2.$$

Therefore, we find

$$|V(\mathbf{z}(m)) - V(a, \mathbf{y}_{-0})| \leq |V(\mathbf{z}(m)) - V(a, \mathbf{z}(m)_{-0})| + |V(a, \mathbf{z}(m)_{-0}) - V(a, \mathbf{y}_{-0})| < \varepsilon$$

for all $m \geq \max\{M_1, M_2\}$. Then, since $V(\mathbf{z}(m)) \geq V(\mathbf{x})$ for all m, we have $V(a, \mathbf{y}_{-0}) > V(\mathbf{x}) - \varepsilon$. Since $\varepsilon > 0$ is arbitrary here, we may conclude that $V(a, \mathbf{y}_{-0}) \geq V(\mathbf{x})$. But V is increasing and $y_0 \geq a$, so $V(\mathbf{y}) \geq V(a, \mathbf{y}_{-0})$, which yields $V(\mathbf{y}) \geq V(\mathbf{x})$, as sought.

Now, as the induction hypothesis, assume that there exists $k \in \mathbb{N}$ such that $V(\mathbf{y}) \geq V(\mathbf{x})$ holds for all $(\mathbf{x}, \mathbf{y}) \in T(k)$. Take any $(\mathbf{x}, \mathbf{y}) \in T(k+1)$. If $\mathbf{x} = \mathbf{y}$, there is nothing to prove, so let $\mathbf{x} \neq \mathbf{y}$. Then we have $\{m \in \mathbb{N} : y_m > x_m\} \neq \emptyset$. Without loss of generality, assume $y_0 > x_0$. Since $(\mathbf{x}, \mathbf{y}) \in T(k+1)$, $U(\mathbf{y}) \geq U(\mathbf{x})$ and there exists $M \in \mathbb{N}$ such that $y_m \geq x_m$ for all $m = 0, \ldots, M-1$, and $y_m \leq x_m$ for all $m \geq M$. If $y_m = x_m$ for each $m \geq M$, then $V(\mathbf{y}) \geq V(\mathbf{x})$ holds by monotonicity of V, so we assume that $y_m < x_m$ for some $m \geq M$. In that case, by using the monotonicity and continuity of $U|_{[0,\sigma]^\infty}$, we can find $\mathbf{w} \in \mathscr{X}$ such that $U(\mathbf{w}) = U(\mathbf{x})$ and

$$x_0 < w_0 \leq y_0, \quad x_m \leq w_m \leq y_m, m = 1, \ldots, M-1, \quad \text{and} \quad w_m = y_m, m = M, M+1, \ldots$$

Notice that, by the monotonicity of U, we have

$$U(w_0, \mathbf{x}_{-0}) > U(\mathbf{x}) = U(\mathbf{w}) \geq U(w_0, x_1, \ldots, x_{M-1}, \mathbf{w}_{-(0,\ldots,M-1)}).$$

Therefore, by the continuity of $U|_{[0,\sigma]^\infty}$, there exists $\mathbf{z}\in\mathscr{X}$ such that $U(\mathbf{z})=U(\mathbf{x})$, and

$$z_0=w_0,\ z_m=x_m,\ m=1,\dots,M-1,\quad \text{and}\quad x_m\geq z_m\geq w_m,\ m=M,M+1,\dots.$$

Since $(\mathbf{x},\mathbf{z})\in T(1)$, we have $V(\mathbf{z})\geq V(\mathbf{x})$. Moreover, since $|\{m\in\mathbb{Z}_+ : w_m>x_m\}|\leq k+1$ and $\{m\in\mathbb{Z}_+ : w_m>z_m\}=\{m\in\mathbb{Z}_+ : w_m>z_m\}\setminus\{0\}$, we have $|\{m\in\mathbb{Z}_+ : w_m>z_m\}|\leq k$, that is, $(\mathbf{z},\mathbf{w})\in T(k)$. It follows that $V(\mathbf{w})\geq V(\mathbf{z})$ by the induction hypothesis. But, by the monotonicity of V, $V(\mathbf{y})\geq V(\mathbf{w})$, so we have $V(\mathbf{y})\geq V(\mathbf{x})$, as sought. □

For any intertemporal utility function $U\in\mathfrak{U}$, $\mathbf{x}\in\mathscr{X}$, and $s,\ t\in\mathbb{Z}_+$ with $s<t$, we define $\chi^U_{s,t,\mathbf{x}}:\mathbb{R}_+\to\overline{\mathbb{R}}_+$ by

$$\chi^U_{s,t,\mathbf{x}}(b):=\sup\{a\geq 0 : U(x_s+a,\mathbf{x}_{-s})\leq U(x_t+b,\mathbf{x}_{-t})\}.$$

Since U is cube-continuous, for any $b\geq 0$ we have $\chi^U_{s,t,\mathbf{x}}(b)<\infty$ if and only if $U(x_s+a,\mathbf{x}_{-s})=U(x_t+b,\mathbf{x}_{-t})$.

Lemma 4. *For any $U,V\in\mathfrak{U}$, we have $U\precsim V$ if and only if*

$$\chi^U_{s,t,\mathbf{x}}\geq\chi^V_{s,t,\mathbf{x}}\quad \textit{for all } \mathbf{x}\in\mathscr{X} \textit{ and } s,t\in\mathbb{Z}_+ \textit{ with } s<t. \tag{21}$$

Proof. Let $U\precsim V$, fix any $\mathbf{x}\in\mathscr{X}$ and $s,t\in\mathbb{Z}_+$ with $s<t$, and pick an arbitrary $b\geq 0$. Suppose first that $\chi^V_{s,t,\mathbf{x}}(b)=\infty$. This means that $V(x_s+a,\mathbf{x}_{-s})\leq V(x_t+b,\mathbf{x}_{-t})$ for all $a\geq 0$. Since $U\precsim V$, this is possible only if $U(x_s+a,\mathbf{x}_{-s})\leq U(x_t+b,\mathbf{x}_{-t})$ for all $a\geq 0$ as well, so it follows that $\chi^U_{s,t,\mathbf{x}}(b)=\infty$. Assume then that $\chi^V_{s,t,\mathbf{x}}(b)<\infty$. There is nothing to prove if $\chi^U_{s,t,\mathbf{x}}(b)=\infty$, so suppose $\chi^U_{s,t,\mathbf{x}}(b)$ is finite. Then $U\precsim V$ implies

$$V(x_s+\chi^U_{s,t,\mathbf{x}}(b),\mathbf{x}_{-s})\geq V(\mathbf{x})=V(x_s+\chi^V_{s,t,\mathbf{x}}(b),\mathbf{x}_{-s}).$$

Since V is increasing, we have $\chi^U_{s,t,\mathbf{x}}(b)\geq\chi^V_{s,t,\mathbf{x}}(b)$ as sought.

Conversely, assume that (21) holds, take any $\mathbf{x}\in\mathscr{X}$ and $s,t\in\mathbb{Z}_+$ with $s<t$, and pick any $a,b\geq 0$ such that $U(x_s+a,\mathbf{x}_{-s})\geq U(x_t+b,\mathbf{x}_{-t})$. In this case $\chi^U_{s,t,\mathbf{x}}(b)$ and $\chi^V_{s,t,\mathbf{x}}(b)$ are finite, and we have

$$U(x_s+a,\mathbf{x}_{-s})\geq U(x_t+b,\mathbf{x}_{-t})=U(x_s+\chi^U_{s,t,\mathbf{x}}(b),\mathbf{x}_{-s}).$$

It follows that $a\geq\chi^U_{s,t,\mathbf{x}}(b)\geq\chi^V_{s,t,\mathbf{x}}(b)$ by monotonicity of U and (21). So, $V(x_s+a,\mathbf{x}_{-s})\geq V(x_s+\chi^V_{s,t,\mathbf{x}}(b),\mathbf{x}_{-s})=V(\mathbf{x})$. (If $U(x_s+a,\mathbf{x}_{-s})>U(x_t+b,\mathbf{x}_{-t})$, then $a>\chi^U_{s,t,\mathbf{x}}(b)\geq\chi^V_{s,t,\mathbf{x}}(b)$, and hence $V(x_s+a,\mathbf{x}_{-s})>V(x_s+\chi^V_{s,t,\mathbf{x}}(b),\mathbf{x}_{-s})=V(\mathbf{x})$.) □

Lemma 5. *For any $U_n,U\in\mathfrak{U}$, $n=1,2,\dots$, if $U_n\to U$ uniformly, then*

$$\chi^{U_n}_{s,t,\mathbf{x}}\to\chi^U_{s,t,\mathbf{x}}\quad \textit{for all } \mathbf{x}\in\mathscr{X} \textit{ and } s,t\in\mathbb{Z}_+ \textit{ with } s<t.$$

PROOF. Fix any $\mathbf{x} \in \mathscr{X}$ and $s, t \in \mathbb{Z}_+$ with $s < t$. For each $n \in \mathbb{N}$ we define the real functions f_n and f on $\mathbb{R}_+$ by

$$f_n(a) := U_n(x_s + a, \mathbf{x}_{-s}) \quad \text{and} \quad f(a) := U(x_s + a, \mathbf{x}_{-s}).$$

Since each U_n and U are cube-continuous, each $f_n|_{[0,\sigma]}$ and $f|_{[0,\sigma]}$ are continuous for each $\sigma > 0$, which means that each f_n and f are continuous real maps on $\mathbb{R}_+$.

Finally, pick an arbitrary $b \geq 0$. In what follows, let $\mathbf{y} := (x_t + b, \mathbf{x}_{-t})$ and $c_n := \chi^{U_n}_{s,t,\mathbf{x}}(b)$ for each n. Assume first that $\chi^U_{s,t,\mathbf{x}}(b) < \infty$ (i.e. $U(\mathbf{y}) = f(\chi^U_{s,t,\mathbf{x}}(b))$). We claim that in this case there must exist $N \in \mathbb{N}$ such that $c_n < \infty$ for all $n \geq N$. If there does not exist such an N, then $c_n = \infty$ for infinitely many n. Without loss of generality, suppose this is the case for all n. Then, $\lim_{a\to\infty} f_n(a) = \sup f_n(\mathbb{R}_+) \leq U_n(\mathbf{y})$. Since $U_n(\mathbf{y}) \to U(\mathbf{y}) = f(\chi^U_{s,t,\mathbf{x}}(b))$, it follows that

$$\lim_{n\to\infty} \lim_{a\to\infty} f_n(a) \leq f(\chi^U_{s,t,\mathbf{x}}(b)).$$

Yet, since $f_n \to f$ uniformly and f is strictly increasing,

$$\lim_{n\to\infty} \lim_{a\to\infty} f_n(a) = \lim_{a\to\infty} \lim_{n\to\infty} f_n(a) = \lim_{a\to\infty} f(a) > f(\chi^U_{s,t,\mathbf{x}}(b)),$$

a contradiction.

Without loss of generality, we take $N = 1$, that is, $c_n < \infty$ for all n. Then, $f_n(c_n) = U_n(\mathbf{y}) \to U(\mathbf{y})$ so, for an arbitrarily fixed $\varepsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that $|f_n(c_n) - U(\mathbf{y})| < \varepsilon/2$ for all $n \geq N_1$. Moreover, since $U_n \to U$ uniformly, there exists $N_2 \in \mathbb{N}$ such that $|f(a) - f_n(a)| < \varepsilon/2$ for all $a \geq 0$ and $n \geq N_2$. Therefore,

$$|f(c_n) - U(\mathbf{y})| \leq |f(c_n) - f_n(c_n)| + |f_n(c_n) - U(\mathbf{y})| < \varepsilon$$

for all $n \geq \max\{N_1, N_2\}$. Since $\varepsilon > 0$ is arbitrary here, we conclude that $f(c_n) \to U(\mathbf{y}) = f(\chi^U_{s,t,\mathbf{x}}(b))$. Since f is continuous and strictly increasing, this is possible only if $c_n \to \chi^U_{s,t,\mathbf{x}}(b)$.

It remains to analyze the case $\chi^U_{s,t,\mathbf{x}}(b) = \infty$. If $c_n = \infty$ for all but finitely many n, there is nothing to prove here, so we assume, without loss of generality, that $c_n < \infty$ for each n. As shown in the previous paragraph, we have $f(c_n) \to U(\mathbf{y})$ in this case. But $\chi^U_{s,t,\mathbf{x}}(b) = \infty$ implies that $\sup f(\mathbb{R}_+) \leq U(\mathbf{y})$. So, since f is increasing, (c_n) must have an increasing subsequence, which we again denote by (c_n). Clearly, if $c_n \to c^*$ for some real number c^*, then

$$f(c^*) = f(\lim_{n\to\infty} c_n) = \lim_{n\to\infty} f(c_n) = \sup f(\mathbb{R}_+),$$

which is impossible since f is strictly increasing. It follows that $c_n \to \infty$, and the proof is complete. □

PROOF OF PROPOSITION 1. The first claim in Proposition 1 follows readily from the definitions. The second claim is an immediate consequence of Lemmas 4 and 5. □

PROOF OF THEOREM 2. Let U and V stand for the intertemporal utility functions that correspond to $(u,\boldsymbol{\alpha})$ and $(v,\boldsymbol{\beta})$, respectively. Given any $\mathbf{x}\in\mathscr{X}$ and $s,t\in\mathbb{Z}_+$ with $s<t$, $u(\infty)=\infty$ guarantees that

$$\boldsymbol{\alpha}(s)u(x_s+\chi^U_{s,t,\mathbf{x}}(b))+\boldsymbol{\alpha}(t)u(x_t)=\boldsymbol{\alpha}(s)u(x_s)+\boldsymbol{\alpha}(t)u(x_t+b),$$

whence

$$\chi^U_{s,t,\mathbf{x}}(b)=u^{-1}(u(x_s)+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(u(x_t+b)-u(x_t)))-x_s,\quad b\geq 0. \tag{22}$$

Similarly,

$$\chi^V_{s,t,\mathbf{x}}(b)=v^{-1}(v(x_s)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(v(x_t+b)-v(x_t)))-x_s,\quad b\geq 0. \tag{23}$$

By Lemma 4, (22), and (23), $(u,\boldsymbol{\alpha})\precsim(v,\boldsymbol{\beta})$ if and only if

$$u^{-1}(u(x_s)+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(u(x_t+b)-u(x_t)))\geq v^{-1}(v(x_s)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(v(x_t+b)-v(x_t)))$$

for all $\mathbf{x}\in\mathscr{X}$, $s,t\in\mathbb{Z}_+$ with $s<t$ and $b\geq 0$. Thus, letting $h:=v\circ u^{-1}$, we find that $(u,\boldsymbol{\alpha})\precsim(v,\boldsymbol{\beta})$ if and only if

$$h(u(x_s)+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(u(x_t+b)-u(x_t)))\geq v(x_s)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(v(x_t+b)-v(x_t))$$

for all $\mathbf{x}\in\mathscr{X}$, $s,t\in\mathbb{Z}_+$ with $s<t$ and $b\geq 0$. Making the change of variables $x:=u(x_s)$, $y:=u(x_t+b)-u(x_t)$, and $z:=u(x_t)$, we conclude that $(u,\boldsymbol{\alpha})\precsim(v,\boldsymbol{\beta})$ if and only if

$$h\left(x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right)\geq h(x)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(h(y+z)-h(z))$$

for all $s,t\in\mathbb{Z}_+$ with $s<t$ and $x,y,z\geq 0$, as we sought.

Now suppose that u and v belong to $\mathscr{V}$, and take any $s,t\in\mathbb{Z}_+$ with $s<t$. If (b) holds, then we have

$$\frac{h\big(x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\big)-h(x)}{\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y}\geq\left(\frac{\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)}{\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)}\right)\frac{h(y+z)-h(z)}{y}\quad\text{for all } x,y,z\geq 0.$$

Since u is differentiable, so is u^{-1}, and hence $h=v\circ u^{-1}$ is differentiable. Consequently, letting $y\to 0$ in the statement above, we find

$$h'(x)\geq\left(\frac{\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)}{\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)}\right)h'(z)\quad\text{for all } x,z\geq 0.$$

Conversely, assume that (c) holds, and fix any $x,y,z\geq 0$. Since h is differentiable, the Mean Value Theorem implies that there exist $c\in[x,x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y]$ and $d\in[y,y+z]$ such that

$$h\left(x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right)-h(x)=h'(c)\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\quad\text{and}\quad h(y+z)-h(z)=h'(d)y.$$

Moreover, by (c), we have $h'(c) \geq \frac{\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)}{\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)} h'(d)$. Combining these observations,

$$h\left(x+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right)-h(x)=h'(c)\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}h'(d)y=\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(h(y+z)-h(z))$$

as we sought.

Finally, the equivalence of (c) and (d) follows from the Inverse Function Theorem and the fact that $h = v \circ u^{-1}$. □

Proof of Corollary 1. This is a special case of Corollary 4. □

Proof of Corollary 2. The first claim follows immediately from Theorem 2 upon setting h in part (b) to be the identity function on $\mathbb{R}_+$. The second claim is a special case of the final assertion of Corollary 4. □

Proof of Corollary 3. The corollary follows immediately from Theorem 2. □

7.3 *Proofs for Section 4*

For any $(u, \boldsymbol{\delta}) \in \mathfrak{U}_{\text{sep}}$, $\omega \geq 0$, and $s, t \in \mathbb{Z}_+$ with $s < t$, define $\eta^{(u,\boldsymbol{\delta})}_{s,t,\omega} : \mathbb{R}_+ \to \mathbb{R}_+$ and $\varsigma^{(u,\boldsymbol{\delta})}_{s,t,\omega} : [0,\omega] \to [0,\omega]$ by

$$\boldsymbol{\delta}(s)u\left(\omega+\eta^{(u,\boldsymbol{\delta})}_{s,t,\omega}(b)\right)+\boldsymbol{\delta}(t)u(\omega)=\boldsymbol{\delta}(s)u(\omega)+\boldsymbol{\delta}(t)u(\omega+b)$$

and

$$\boldsymbol{\delta}(s)u\left(\omega-\varsigma^{(u,\boldsymbol{\delta})}_{s,t,\omega}(b)\right)+\boldsymbol{\delta}(t)u(\omega)=\boldsymbol{\delta}(s)u(\omega)+\boldsymbol{\delta}(t)u(\omega-b),$$

respectively. Since $0 < \boldsymbol{\delta} < 1$ and $u(\infty) = \infty$, both of these functions are well-defined.

Proof of Theorem 3. It is easy to verify that $(v, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$ if and only if

$$\eta^{(u,\boldsymbol{\alpha})}_{s,t,\omega} \geq \eta^{(v,\boldsymbol{\beta})}_{s,t,\omega} \quad \text{for all } \omega \geq 0 \text{ and } s,t \in \mathbb{Z}_+ \text{ with } s<t \tag{24}$$

and

$$\varsigma^{(u,\boldsymbol{\alpha})}_{s,t,\omega} \geq \varsigma^{(v,\boldsymbol{\beta})}_{s,t,\omega} \quad \text{for all } \omega \geq 0 \text{ and } s,t \in \mathbb{Z}_+ \text{ with } s<t. \tag{25}$$

Moreover, (24) holds if and only if

$$u^{-1}\left(u(\omega)+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(u(\omega+b)-u(\omega))\right) \geq v^{-1}\left(v(\omega)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(v(\omega+b)-v(\omega))\right)$$

for all $\omega \geq 0$, $s, t \in \mathbb{Z}_+$ with $s < t$, and $b \geq 0$. Thus, letting $h := v \circ u^{-1}$, we find that (24) holds if and only if

$$h\left(u(\omega)+\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(u(\omega+b)-u(\omega))\right) \geq v(\omega)+\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(v(\omega+b)-v(\omega))$$

for all $\omega \geq 0$, $s,t \in \mathbb{Z}_+$ with $s < t$, and $b \geq 0$. Making the change of variables $x := u(\omega)$ and $y := u(\omega + b)$, we have that (24) holds if and only if

$$h\left(x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}(y-x)\right) \geq h(x) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}(h(y) - h(x))$$

for all $s,t \in \mathbb{Z}_+$ with $s < t$ and $0 \leq x \leq y$. One can similarly show that (25) holds if and only if (11) holds for all $s,t \in \mathbb{Z}_+$ with $s < t$ and $0 \leq x \leq y$. □

LEMMA 6. *Let* $0 < \lambda < 1$. *If* $f : \mathbb{R}_+ \to \mathbb{R}_+$ *is continuous and*

$$(1-\lambda)f(x) + \lambda f(y) \leq f((1-\lambda)x + \lambda y) \quad \textit{for all } 0 \leq x \leq y, \tag{26}$$

then f *is concave.*

PROOF. We first prove an auxiliary fact. Let $A_0 := \{0,1\}$ and

$$A_m := \{(1-\lambda)a + \lambda b : a,b \in A_{m-1} \text{ and } a \leq b\},$$

$m = 1,2,\ldots$. We claim that $A_\infty := A_0 \cup A_1 \cup \ldots$ is dense in $[0,1]$.

We consider only the case $\frac{1}{2} \leq \lambda < 1$, the argument for the remaining case being analogous. Suppose that $cl(A_\infty) \neq [0,1]$, that is, there exists $\gamma \in (0,1) \setminus cl(A_\infty)$. Since $(0,1) \setminus cl(A_\infty)$ is an open set, we have

$$a := \sup([0,\gamma] \cap cl(A_\infty)) < \gamma \quad \text{and} \quad b := \inf([\gamma,1] \cap cl(A_\infty)) > \gamma.$$

(Obviously, $a,b \in cl(A_\infty)$ and $A_\infty \cap (a,b) = \varnothing$.) Define $\theta := (1-\lambda)(b-a) > 0$. Clearly, there exist $a', b' \in A_\infty$ such that

$$a - \theta < a' \leq a < b \leq b' < b + \theta.$$

By definition of A_∞, we have $(1-\lambda)a' + \lambda b' \in A_\infty$. However, since $(1-\lambda)a + \lambda b = b - \theta$, we have

$$(1-\lambda)a' + \lambda b' < (1-\lambda)a + \lambda(b+\theta) = b - \theta + \lambda\theta = b - (1-\lambda)\theta < b$$

and since $\frac{1}{2} \leq \lambda < 1$,

$$\begin{aligned}(1-\lambda)a' + \lambda b' &> (1-\lambda)(a-\theta) + \lambda b \\ &= a - \theta + \lambda(b - (a-\theta)) \\ &\geq a - \theta + (1-\lambda)(b - (a-\theta)) \\ &= a - \theta + \theta + (1-\theta)\theta \\ &> a.\end{aligned}$$

Thus, $(1-\lambda)a' + \lambda b' \in A_\infty \cap (a,b)$, a contradiction.

Lemma 6 is now easily proved. Note first that one can easily show inductively that (26) holds if and only if

$$(1-\mu)f(x)+\mu f(y)\leq f((1-\mu)x+\mu y) \quad \text{for all } 0\leq x\leq y \text{ and } \mu\in A_\infty.$$

Since f is continuous and $cl(A_\infty)=[0,1]$, it follows that (26) holds if and only if

$$(1-\mu)f(x)+\mu f(y)\leq f((1-\mu)x+\mu y) \quad \text{for all } 0\leq x\leq y \text{ and } 0\leq \mu\leq 1.$$

That is, f is concave. □

Proof of Corollary 4. Suppose that $(v,\boldsymbol{\beta})$ is more impatient than $(u,\boldsymbol{\alpha})$, but $\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)<\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)$ for some $s,t\in\mathbb{Z}_+$ with $s<t$. Let $h:=v\circ u^{-1}$, $\alpha:=\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)$ and $\beta:=\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)$. Since h is strictly increasing, $\alpha<\beta$ implies

$$h((1-\alpha)x+\alpha y)\leq h((1-\beta)x+\beta y) \quad \text{for all } 0\leq x\leq y.$$

Combining this with (10) yields

$$(1-\beta)h(x)+\beta h(y)\leq h((1-\beta)x+\beta y) \quad \text{for all } 0\leq x\leq y.$$

Thus, by Lemma 6, h is a concave function. But, for any fixed $0\leq x<y$, (11) and $\alpha<\beta$ entail that

$$h((1-\alpha)y+\alpha x)\leq (1-\beta)h(y)+\beta h(x)<(1-\alpha)h(y)+\alpha h(x),$$

which contradicts the concavity of h.

The second assertion of Corollary 4 follows from Theorem 3 (with h being the identity function) and Corollary 2. On the other hand, the "if" part of the final assertion of Corollary 4 is trivial. To prove its "only if" part, let $(v,\boldsymbol{\alpha})$ be more impatient than $(u,\boldsymbol{\alpha})$, $h:=v\circ u^{-1}$, and set $\alpha:=\boldsymbol{\alpha}(1)$. By (10) and (11), and Lemma 6, both h and $-h$ must be concave functions, so that h is affine. Since, $h(0)=0$, h is, in fact, a strictly increasing linear function. □

Proof of Corollary 5. The "only if" part of the first assertion here is immediate from Theorem 3. To prove its "if" part, assume that (12) holds for all $0\leq x\leq y$, and suppose, as the induction hypothesis,

$$(1-\beta^r)h(x)+\beta^r h(y)\leq h((1-\alpha^r)x+\alpha^r y) \quad \text{for all } 0\leq x\leq y,$$

where r is an arbitrary positive integer. Then, for any $0\leq x\leq y$, we have

$$\begin{aligned}(1-\beta^{r+1})h(x)+\beta^{r+1}h(y)&=(1-\beta^r)h(x)+\beta^r((1-\beta)h(x)+\beta h(y))\\ &\leq (1-\beta^r)h(x)+\beta^r h((1-\alpha)x+\alpha y)\\ &\leq h((1-\alpha^r)x+\alpha^r((1-\alpha)x+\alpha y))\\ &=h((1-\alpha^{r+1})x+\alpha^{r+1}y).\end{aligned}$$

It follows that $h((1-\alpha^{t-s})x+\alpha^{t-s}y)\geq(1-\beta^{t-s})h(x)+\beta^{t-s}h(y)$; that is, (10) holds for all $s,t\in\mathbb{Z}_+$ with $s<t$ and $y\geq x\geq 0$. Since one can similarly show (using this time (13)) that (11) also holds for all $s,t\in\mathbb{Z}_+$ with $s<t$ and $0\leq x\leq y$, the claim follows from Theorem 3.

To prove the second assertion of Corollary 5, assume that (14) holds for all $0<x<y$. Consider first the case $\beta/\alpha\leq(1-\alpha)/(1-\beta)$, so that

$$\frac{\beta}{\alpha}h'(y)\leq h'(x)\leq\frac{\alpha}{\beta}h'(y)\quad\text{for all } 0\leq x\leq y. \tag{27}$$

Observe that if $h'(x)<(\beta/\alpha)h'(y)$ for some $x,y\geq 0$, then we must have $x>y\geq 0$ and $(\alpha/\beta)h'(x)<h'(y)$, which contradicts (27). Thus $h'(x)\geq(\beta/\alpha)h'(y)$ for all $x,y\geq 0$, that is, (6) holds, so by Corollary 3, $(v,\beta)\precsim(u,\alpha)$. Hence, in particular, (v,β) is more impatient than (u,α).

Finally, consider the case $\beta/\alpha>(1-\alpha)/(1-\beta)$, so that (14) becomes

$$\frac{1-\alpha}{1-\beta}h'(y)\leq h'(x)\leq\frac{1-\beta}{1-\alpha}h'(y)\quad\text{for all } 0\leq x\leq y. \tag{28}$$

Fix any $y\geq x\geq 0$ arbitrarily. Let $z:=(1-\alpha)x+\alpha y$ and define $G:\mathbb{R}_+\to\mathbb{R}_+$ by $G(\omega):=\omega/(1-\alpha)-(\alpha/(1-\alpha))y$. Notice that $G(y)=y$, $G(z)=x$, $G(\omega)\leq\omega$ for all $\omega\in[0,y]$, and that $h'(G(\omega))=(1-\alpha)(d/d\omega)h(G(\omega))$ for all ω. By (28),

$$\int_z^y(1-\alpha)h'(\omega)\,d\omega\leq\int_z^y(1-\beta)h'(G(\omega))\,d\omega=(1-\beta)\int_z^y(1-\alpha)\frac{d}{d\omega}h(G(\omega))\,d\omega,$$

so by the Fundamental Theorem of Calculus, we have $h(y)-h(z)\leq(1-\beta)(h(y)-h(x))$, which is equivalent to (12).

Now let $w:=(1-\alpha)y+\alpha x$ and define $H:\mathbb{R}_+\to\mathbb{R}_+$ by $H(\omega):=\omega/(1-\alpha)-(\alpha/(1-\alpha))x$. Notice that $H(x)=x$, $H(w)=y$, $H(\omega)\geq\omega$ for all $\omega\in[x,\infty)$, and that $h'(H(\omega))=(1-\alpha)(d/d\omega)h(H(\omega))$ for all ω. Then, by (28),

$$\int_x^w(1-\alpha)h'(\omega)\,d\omega\leq\int_x^w(1-\beta)h'(H(\omega))\,d\omega=(1-\beta)\int_x^w(1-\alpha)\frac{d}{d\omega}h(H(\omega))\,d\omega,$$

so by the Fundamental Theorem of Calculus, we have $h(w)-h(x)\leq(1-\beta)(h(y)-h(x))$, which is equivalent to (13).

We have proved that (28) implies (12) and (13) for all $0\leq x\leq y$. By the first part of Corollary 5, (28) implies that (v,β) is more impatient than (u,α). □

The proof of Theorem 4 is analogous to those of Theorems 2 and 3 and is thus omitted.

Proof of Corollary 6. The first assertion is immediate from Theorem 4. To see the second, let $h:\mathbb{R}_+\to\mathbb{R}_+$ be any differentiable function with $h(0)=0$. Observe that $-h$ is star-shaped if and only if $(d/dt)(h(t)/t)\leq 0$ for all $t>0$, or equivalently, $h'(t)t\leq h(t)$ for all $t>0$. But, given any $u,v\in\mathscr{V}$, the first assertion of Corollary 6 says that $(u,\boldsymbol{\delta})\precsim^{\mathbf{0}}(v,\boldsymbol{\delta})$

for all $\boldsymbol{\delta} \in \mathscr{D}$ if and only if $-(v \circ u^{-1})$ is star-shaped. Therefore, by the Inverse Function Theorem, $(u, \boldsymbol{\delta}) \precsim^{\mathbf{0}} (v, \boldsymbol{\delta})$ for all $\boldsymbol{\delta} \in \mathscr{D}$ if and only if

$$\frac{v'(u^{-1}(a))a}{u'(u^{-1}(a))} \leq v(u^{-1}(a)) \text{ for all } a > 0.$$

Since $u^{-1}(\mathbb{R}_{++}) = \mathbb{R}_{++}$, the latter statement is equivalent to (18). □

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [86]

Arrow, Kenneth J. (1964), "The role of securities in the optimal allocation of risk-bearing." *Review of Economic Studies*, 31, 91–96. [72]

Asheim, Geir B. (1997), "Individual and collective time-consistency." *Review of Economic Studies*, 64, 427–443. [88]

Becker, Robert A. (1983), "Comparative dynamics in the one-sector optimal growth model." *Journal of Economic Dynamics and Control*, 6, 99–107. [74, 84]

Böhm-Bawerk, Eugen von (1891), *The Positive Theory of Capital*. Stechert, New York (translated by William Smart, 1923). [71, 91, 92]

Carillo, Juan D. and Thomas Mariotti (2000), "Strategic ignorance as a self-disciplining device." *Review of Economic Studies*, 67, 529–544. [88]

Epstein, Larry G. (1987), "Impatience." In *The New Palgrave* (John Eatwell, Murray Milgate, and Peter Newman, eds.), volume 2, 720–723, Macmillan, London. [72]

Epstein, Larry G. and Stanley E. Zin (1989), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica*, 57, 937–969. [101]

Fishburn, Peter C. and Ariel Rubinstein (1982), "Time preference." *International Economic Review*, 23, 677–694. [72, 97]

Fisher, Irving (1930), *The Theory of Interest as Determined by Impatience to Spend Income and Opportunity to Invest it*. Macmillan, New York. [82, 91]

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002), "Time discounting and time preference: A critical review." *Journal of Economic Literature*, 40, 351–401. [71, 77, 89]

Friedman, Milton (1972), *Price Theory: A Provisional Text (Revised Edition)*. Aldine, Chicago. [92]

Goldman, Steven M. (1980), “Consistent plans.” *Review of Economic Studies*, 47, 533–537. [88]

Horowitz, John K. (1992), “Comparative impatience.” *Economics Letters*, 38, 25–29. [72, 79, 83]

Kocherlakota, Narayana R. (1996), “Reconsideration-proofness: A refinement for infinite horizon time inconsistency.” *Games and Economic Behavior*, 15, 33–54. [88]

Koopmans, Tjalling C. (1960), “Stationary ordinal utility and impatience.” *Econometrica*, 28, 287–309. [72]

Koopmans, Tjalling C., Peter A. Diamond, and Richard E. Williamson (1964), “Stationary utility and time perspective.” *Econometrica*, 32, 82–100. [72]

Laibson, David (1997), “Golden eggs and hyperbolic discounting.” *Quarterly Journal of Economics*, 112, 443–477. [88]

Lancaster, Kelvin (1963), “An axiomatic theory of consumer time preference.” *International Economic Review*, 4, 221–231. [97]

Lawrance, Emily C. (1991), “Poverty and the rate of time preference: Evidence from panel data.” *Journal of Political Economy*, 99, 54–77. [73]

Marinacci, Massimo (1998), “An axiomatic approach to complete patience and time invariance.” *Journal of Economic Theory*, 83, 105–144. [72]

O’Donoghue, Ted and Matthew Rabin (1999), “Doing it now or later.” *American Economic Review*, 89, 103–124. [87, 88, 89]

Olson, Mancur and Martin J. Bailey (1981), “Positive time preference.” *Journal of Political Economy*, 89, 1–25. [92]

Osborne, Martin J. and Ariel Rubinstein (1994), *A Course on Game Theory*. MIT Press, Cambridge. [98, 99]

Peleg, Bezalel and Menahem Yaari (1973), “On the existence of a consistent course of action when tastes are changing.” *Review of Economic Studies*, 40, 391–401. [88]

Phelps, Edmund S. and Robert A. Pollak (1968), “On second-best national saving and game-equilibrium growth.” *Review of Economic Studies*, 35, 185–199. [88]

Pratt, John W. (1964), “Risk aversion in the small and in the large.” *Econometrica*, 32, 122–136. [72]

Prelec, Drazen (2004), “Decreasing impatience: A criterion for non-stationary time preference and ‘hyperbolic’ discounting.” *Scandinavian Journal of Economics*, 106, 511–532. [89, 97]

Rae, John (1834), *The Sociological Theory of Capital*. Macmillan, London, 1905. [71]

Roth, Alvin E. (1985), "A note on risk aversion in a perfect equilibrium model of bargaining." *Econometrica*, 53, 207–211. [75, 99]

Rubinstein, Ariel (2003), "'Economics and psychology'?: The case of hyperbolic discounting." *International Economic Review*, 44, 1207–1216. [97]

Sorin, Sylvain (1986), "On repeated games with complete information." *Mathematics of Operations Research*, 11, 147–160. [86]

Stigler, George J. (1966), *The Theory of Price*. Macmillan, New York, third edition. [92]

Strotz, Robert H. (1956), "Myopia and inconsistency in dynamic utility maximization." *Review of Economic Studies*, 23, 165–180. [88]

Submitted 2006-3-22. Final version accepted 2007-2-1. Available online 2007-2-5.