# Delay Aversion*

Jean-Pierre Benoît
Department of Economics
New York University

Efe A. Ok
Department of Economics
New York University

November, 2004

**Abstract**

We address the following question: When can one person properly be said to be more averse to delay than another?

*JEL Classification*: D11, D90.

*Key words*: Delay aversion, impatience, consumption smoothing.

*We thank Juan Dubra, Ed Green, Faruk Gul, Jim Jordan, Boyan Jovanovic, Tom Sargent, Martin Schneider, and Ennio Stacchetti for various discussions about the content of this work. We are also grateful for the many comments of the participants of the seminars given at Columbia, NYU, Penn-State, Rutgers, UT-Austin, and University of Washington at St. Louis.

*The present always gets its rights.*
E. Böhm-Bawerk

# 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 time 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?

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 may not always be meaningful. Consider an environment

[1] See Frederick, Loewenstein and O'Donoghue (2002) for a recent survey on time preferences.

[2] Starting with Koopmans (1960) and Koopmans, Diamond and Williamson (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 Sections 2 and 5.

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 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 who evaluate an arbitrary consumption path $(x_0, x_1, ...)$ as $\sum_{t=0}^{\infty} \delta_1^t u_1(x_t)$ and $\sum_{t=0}^{\infty} \delta_2^t u_2(x_t)$. Here, unlike in the example considered above, the discount factor $\delta_i$ is uniquely determined by the $i$th person's intertemporal preferences, and the instantaneous utility function $u_i$ is uniquely determined up to a positive affine transformation. Would it now be reasonable to compare the relative delay aversion of these individuals by looking *only* at their discount factors? To see that a positive answer would be problematic, 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 will maximize by consuming the entire amount in the first period, while the second person will spread 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 induces – or 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 ex-

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

[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

ercises, 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 the different delay aversions of the citizens of the countries? Suppose that 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. Now, one might wish to argue that, provided that the utility functions do not differ too much, differences in discount factors will capture differences in delay aversions, so that it is an acceptable simplification to assume the utilities are in fact identical. 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 we depart 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 seek to risk theory suggests at least three constraints in this endeavor. First, the proposed methods should not, at least at the level of their primitive definitions, depend on the way in which intertemporal preferences are modeled. Second, these methods 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]

The present paper is organized in three major parts, each of which addresses one of the three concerns just listed. In Section 2 we introduce a very 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

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.

[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) which mesh extremely well with these definitions and their characterizations.

at an earlier date to receiving an increase at a later date, A does too, *ceteris paribus*. This approach to relative delay aversion 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 true 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 notion of delay aversion should be contrasted with a "pure" time bias, namely, a psychological preference for early gratification, which may or may not overwhelm other considerations. As we shall see, our approach easily modifies to provide an ordering which captures a pure time preference motive instead. We term this an *impatience* ordering.

In Section 3 we specialize to the case of separable intertemporal utilities, and provide various characterizations of our orderings. These characterizations yield further insight into the structure and appeal of the comparison methods introduced in Section 2. In particular, we find that in the exponential discounting model, if one agent is more delay averse (or impatient) than another, then her discount factor *must* be lower than the latter's, but not conversely. These characterizations are also used to show how to transform a given separable intertemporal utility function into a more delay averse (or more impatient) one.

In Section 4 we turn to some specific economic applications. First, we show that in the one-sector optimal growth problem the optimal capital stock of a country can never fall strictly below that of a more delay averse country. This generalizes a well-known result of growth theory (cf. Becker (1983)), which was obtained originally by varying the discount factors of agents with *identical* instantaneous utility functions. Second, we show that any stationary Nash equilibrium path of an infinitely repeated game with exponential utility maximizers remains an equilibrium path as players become more patient. Again, this type of result has previously been pursued in the literature by making alterations only in players' discount factors, thus considering a special manner in which players may change. We also derive a related result for infinitely repeated games in which the agents are not assumed to be exponential utility maximizers.

In our final application, we revisit a well-known result of Roth (1985) which shows that an agent's share in Rubinstein bargaining decreases as his utility function becomes more concave. Roth interprets this to mean that becoming more risk averse harms an agent, but since there is no risk in the model, this interpretation has been considered to be somewhat dubious. We clarify this result by showing that, in bargaining environments, a concave transformation of an agent's utility function renders him more delay averse. Consequently, Roth's result is properly understood as a result about delay aversion, not risk aversion.

The paper ends with a few concluding comments collected in Section 5. All proofs of the theoretical results are contained in the Appendix.

# 2 Comparative Delay Aversion: The General Case

In this section we lay out a general framework of analysis, and introduce three different methods of comparing the attitudes of decision makers towards time delay. Our first method allows one to compare the delay aversion of individuals, that is, loosely speaking, their tendency to prefer consumption streams that are skewed towards early periods. This ordering commingles the potential consumption smoothing motive of the individuals with their pure time preference. To concentrate on the latter effect alone, we then refine this ordering, thereby suggesting a method of comparing the relative impatience of individuals. Finally, in contexts where preferences are defined on the prize-time space (such as in bargaining games and/or preemptive investment models), we consider a further refinement of the resulting ordering.

## 2.1 Intertemporal Utility Functions

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 $(x_0, x_1, ...)$ with $x_t \geq 0$ for all $t$, and $\sup\{x_t : t = \mathbb{Z}_+\} < \infty$. 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 $\mathcal{X}$, and endow $\mathcal{X}$ with the product topology.

The generic members of $\mathcal{X}$ 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 $\mathbf{x} \equiv (x_0, x_1, ...)$. 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 $t_1, ..., t_k$, the expression $(a_{t_1}, ..., a_{t_k}, \mathbf{x}_{-(t_1,...,t_k)})$ stands for the sequence $\mathbf{y}$ where $y_{t_i} = a_{t_i}$, $i = 1, ..., k$, and $y_m = x_m$ for all $m \in \mathbb{Z}_+ \backslash \{t_1, ..., t_k\}$.

We work with strictly monotonic preferences over consumption streams, that is, preferences for which more consumption is preferred to less at any period. Moreover, we posit that any preference ordering can be represented by a utility function $U$ on $\mathcal{X}$ such that $U|_{[0,a]^\infty}$ is continuous for any $0 \leq a < \infty$. Such maps are referred to as *cube-continuous* in what follows.[6]

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

[6] Put differently, $U$ is cube-continuous if and only if the restriction of $U$ to any Hilbert cube in $\mathcal{X}$ is continuous. (Recall that a compact subset $[a, b]^\infty$ of $\mathcal{X}$ is called a *Hilbert cube* in $\mathcal{X}$ for any real numbers $a$ and $b$ with $0 \leq a < b$.) Hence, the term "cube-continuous."

Given that $\mathcal{X}$ is endowed with the product topology, the *continuity* of a real map on $\mathcal{X}$ is much more demanding than its cube-continuity. For example, the map $f : \mathcal{X} \to \mathbf{R}$ 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 $\delta$ and $\sigma$.

The requirement $U(0,0,...) = 0$ is merely a normalization that simplifies the exposition.

## 2.2 Comparative Delay Aversion

Consider an individual, $A$, facing a fixed endowment stream $\omega \in \mathcal{X}$. Suppose that he has won a prize which gives him the choice between an additional consumption of 100 in period $s$ or 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, in fact, he favors consuming an 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 will consider the larger amount of 120 to be insufficient compensation for waiting an extra period, and, instead, will prefer 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 Agent 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.

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

$$U(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\omega_t + b, \boldsymbol{\omega}_{-t}) \quad \textit{implies} \quad V(\omega_s + a, \boldsymbol{\omega}_{-s}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\omega_t + b, \boldsymbol{\omega}_{-t}) \tag{1}$$

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

$$U(\omega_t - b, \boldsymbol{\omega}_{-t}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\omega_s - a, \boldsymbol{\omega}_{-s}) \quad \textit{implies} \quad V(\omega_t - b, \boldsymbol{\omega}_{-t}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\omega_s - a, \boldsymbol{\omega}_{-s}) \tag{2}$$

*for all $\omega_s \geq a \geq 0$ and $\omega_t \geq b \geq 0$. We denote this situation by $U \precsim V$. We say that $V$ is* **strictly more delay averse than** *$U$, and write $U \prec V$, if $U \precsim V$ but not $V \precsim U$.*[7]

[7] Formally speaking, $\precsim$ should be viewed as the binary relation on $\mathfrak{U}$ defined by $(U,V) \in \precsim$ iff $V$ is at least as delay averse as $U$. In turn, $\prec$ is the asymmetric part of $\precsim$.

We should note that, from a formal point of view, Definition 2 contains a redundancy. Given that here we are working with *all* endowment streams $\omega$, the second part of this definition (i.e. the part that concerns (2)) is implied by its first part, and vice versa. (See Lemma 1 in the Appendix.) We

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.[8] The second condition says that the relatively delay averse $B$ prefers making a later payment whenever $A$ does.

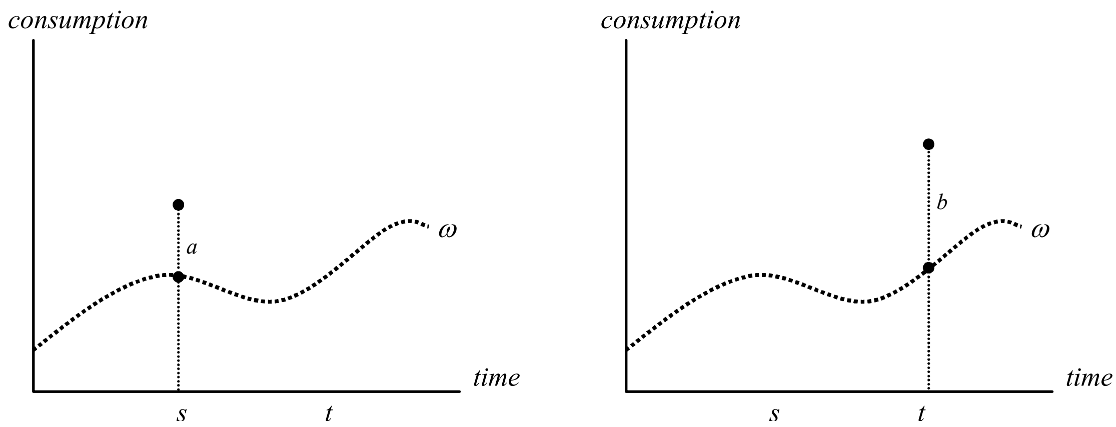


**Figure 1a:** If a person prefers the consumption path on the left, so would a more delay averse person.

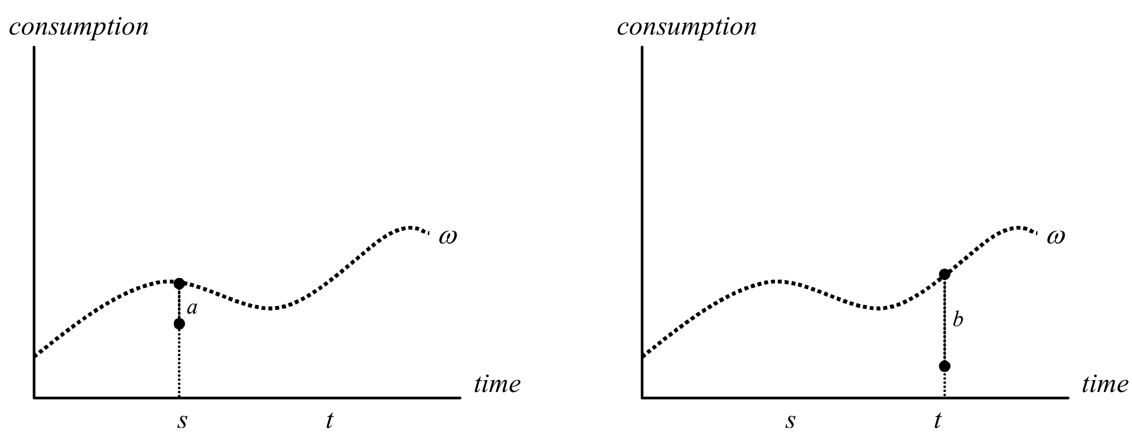


**Figure 1b:** If a person prefers the consumption path on the right, so would a more delay averse person.

maintain this redundancy here, because these two parts of the definition are conceptually distinct. Moreover, in the following subsections we will work with particular subsets of endowment streams, and this redundancy will disappear.

[8]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.

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 $U \precsim V$ iff $\lambda U + W \precsim \lambda V + W$ for all $\lambda > 0$ and $U, V, W \in \mathfrak{U}$.[9] 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.** *$\precsim$ 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 all $n$, 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 we will encounter a few interesting intertemporal utility functions that cannot be compared on the basis of this preorder.

## 2.3 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.[10] First we need to introduce the auxiliary concept of single-crossing streams.

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

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 $\mathbf{x}, \mathbf{y} \in \mathcal{X}$,*

$$U(\mathbf{y}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} U(\mathbf{x}) \quad \textit{implies} \quad V(\mathbf{y}) \begin{Bmatrix} \geq \\ > \end{Bmatrix} V(\mathbf{x})$$

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

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

[10] Horowitz' choice to model time continuously proves to be unfortunate in a number of respects. Most notably, within the standard exponential discounting model it results in a quite limited ordering, which applies *only* when the decision makers have the same instantaneous utility functions. As will become clear shortly, this contrasts markedly with the results we obtain. (More on this in Section 5.)

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 the ordering $\precsim$ is an obvious advantage – one which we exploit in deriving many of the results of the subsequent sections. Theorem 1 shows that this simplicity comes at no conceptual cost.[11]

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 $\boldsymbol{\omega}'$,

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

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 an $\boldsymbol{\omega} \in \mathcal{X}$ to an $\boldsymbol{\omega}'$ obtained from $\boldsymbol{\omega}$ through an investment, as above, then person B does as well. Then B must be more delay averse than A.[12]

[11] 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.

[12] 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 endowmendt stream $(\omega_s + a, \boldsymbol{\omega}_{-s})$.

## 2.4 Comparative Impatience

Consider an agent whose endowment stream is $(0, 10, 10, ...)$, 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 $(10, 0, 10, 10, ...)$ instead of $(0, 10, 10, ...)$.

The same ambiguity arises when making comparative statements. Consider agents A and B, both with the endowment stream $(10, 0, 10, ...)$. 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.

Our 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 may seem quite demanding, as we will see in Section 3, 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.[13] In line with this point of view, the notion of delay aversion commingles these two effects.

At the same time, many authors, including Friedman (1976), Olson and Bailey (1981), and Stigler (1987), wished to concentrate on agents' pure time preferences.[14]

[13] In the words of Böhm-Bawerk (1891) - also quoted by Olson and Bailey (1981) - "A *first* principal cause capable of producing a difference in value between present and future goods is ... if a person suffers in the present from appreciable lack of certain goods, or of goods in general, but has reason to hope to be more generously provided for at a future time, then that person will always place a higher value on a given quantity of immediately available goods than on the same quantity of future goods. ... a *second* phenomenon of human experience ... is the fact that we feel less concerned about future sensations of joy and sorrow simply because they do lie in the future, and the lessening of our concern is in proportion to the remoteness of that future."

[14] 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

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.

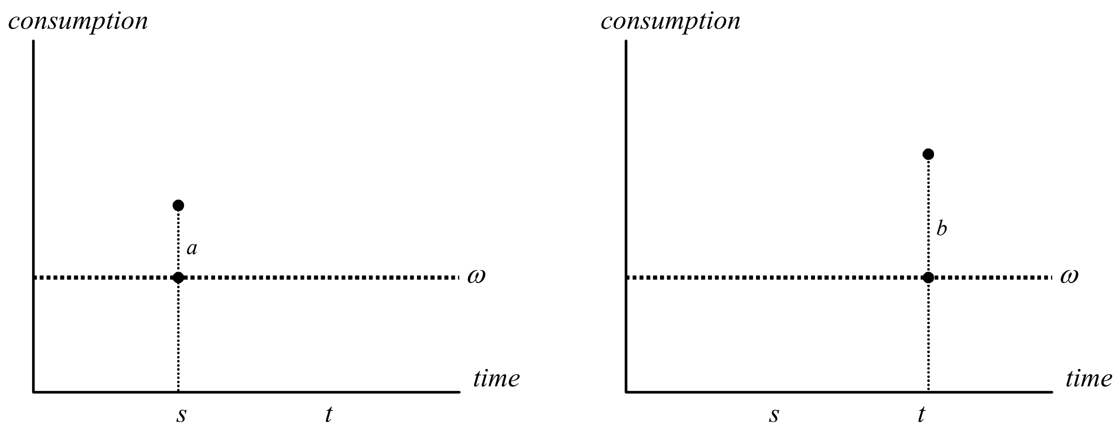


**Figure 2a:** If a person prefers the consumption path on the left, so would a more impatient person.

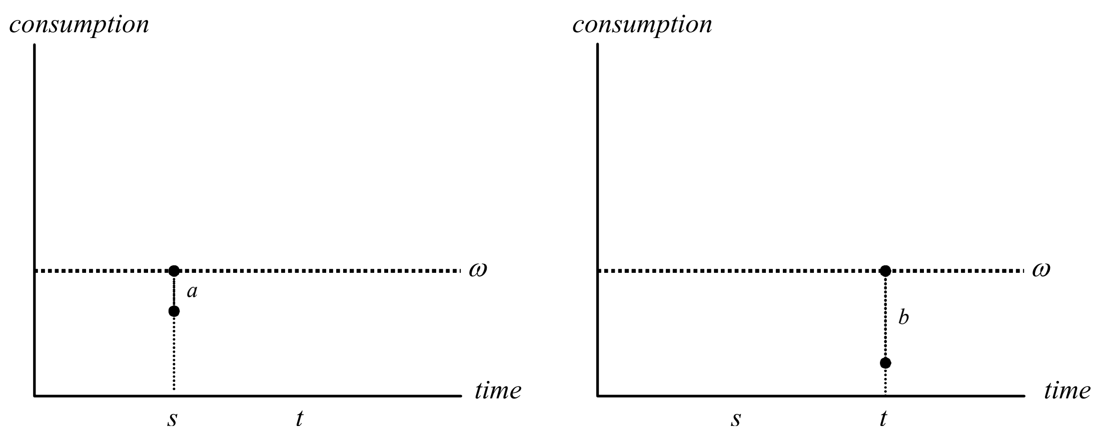


**Figure 2b:** If a person prefers the consumption path on the right, so would a more impatient person.

**Definition 5.** *Let $U$ and $V$ be two intertemporal utility functions. We say that $V$*

along constant endowment streams.

*is* **more impatient than** $U$*, if (1) and (2) hold for all endowment streams* $\boldsymbol{\omega} \in \mathcal{X}$ *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.[15] Like $\precsim$, this ordering is a vector preorder on $\mathcal{U}$ which is continuous (relative to the topology of uniform convergence).

## 2.5 Comparative Delay Aversion in Cryogenic Environments

In intertemporal decision theory, there are two basic 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.[16] 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. This setup 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 *cryogenic* 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 $\mathbf{0} := (0, 0, ...)$. 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 6.** *Let* $U$ *and* $V$ *be two intertemporal utility functions. We say that* $V$ *is* **cryogenically more delay averse than** $U$*, if for any given* $(s, t) \in \mathbb{Z}_+^2$ *with* $s < t$*,*

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

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

Note that this definition simply asks that condition (1) hold only when the initial endowment stream is identically zero.[17] Thus, while still partial, the "cryogenically more delay averse than" ordering $\precsim^{\mathbf{0}}$ is a significant refinement of the "more impatient

[15] We note that (1) may hold for all *constant* endowment streams, while (2) does not. Thus, in contrast to Definition 2, there is no redundancy in Definition 5.

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

[17] Condition (2) holds vacuously here, since no payments can be made from a starting point of 0.

than" ordering (which is itself a refinement of $\precsim$). We should note that, within the context of bargaining theory, a stationary version of this ordering is used by Osborne and Rubinstein (1994).

# 3 Comparative Delay Aversion: The Separable Case

In this section we specialize to the context of *separable* intertemporal utility functions, and obtain characterizations of the three comparison methods we introduced above. Most of the applications of Section 4 will derive from the results of this section.

## 3.1 Separable and Exponential Intertemporal Utility Functions

The most prevalent types of intertemporal utility functions in economic analysis posit the separability of the evaluation of time and outcomes, and presume an additively separable form. In the rest of 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 $\boldsymbol{\delta}(0)+\boldsymbol{\delta}(1)+\cdots < \infty$.[18] 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 $u(\infty) = \infty$.[19] The class of all discount functions is denoted by $\mathcal{D}$ and the class of all instantaneous utility functions by $\mathcal{U}$. At times we will work with differentiable members of $\mathcal{U}$. The following class will receive particular attention:

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

Note that any member $u$ of $\mathcal{V}$ has an inverse $u^{-1}$ which is continuously differentiable with a finite positive derivative on $(0, \infty)$.[20]

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

$$U(\mathbf{x}) = \sum_{t=0}^{\infty} \boldsymbol{\delta}(t)u(x_t) \quad \text{for all } \mathbf{x} \in \mathcal{X}. \tag{3}$$

[18] 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 $\mathcal{X}$ 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 would remain valid in the absence of this assumption as well.

[19] The assumption $u(\infty) = \infty$ considerably simplifies the subsequent analysis, but is not essential for it. In particular, the "if" part of every characterization theorem we report below remains valid without this assumption.

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

We denote the class of all separable intertemporal utility functions by $\mathfrak{U}_{\text{sep}}$. It is easy to verify that $\mathfrak{U}_{\text{sep}} \subseteq \mathfrak{U}$.[21] In what follows, by the tuple $(u, \boldsymbol{\delta})$ in $\mathcal{U} \times \mathcal{D}$, we mean the separable intertemporal utility function induced by $u$ and $\boldsymbol{\delta}$ by way of (3). We use the notation $(u, \boldsymbol{\delta})$ and $U$ interchangeably when (3) 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 $U : \mathcal{X} \to \mathbb{R}$ with

$$U(\mathbf{x}) = \sum_{t=0}^{\infty} \delta^t u(x_t) \quad \text{for all } \mathbf{x} \in \mathcal{X}, \tag{4}$$

where $u \in \mathcal{U}$ and $0 < \delta < 1$. In this case, $\delta$ 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 $(u, \delta)$ in $\mathcal{U} \times (0, 1)$, we mean the exponential intertemporal utility function induced by $u$ and $\delta$ through (4), and hence, with a slight abuse of terminology, we refer to any such pair $(u, \delta)$ as an "exponential intertemporal utility function."

In recent years, a large amount of experimental data has been gathered which 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.[22] 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 will thus be 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 will emphasize the exact nature of our subsequent results for this specific model.

## 3.2 Characterizations of the Delay Aversion Ordering

We begin by providing alternative characterizations of our "more delay averse than" ordering $\precsim$ in the case of separable intertemporal utility functions.

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

[21] The nontrivial part of the argument is to establish the cube-continuity of the map defined by (3) for some $(u, \boldsymbol{\delta}) \in \mathcal{U} \times \mathcal{D}$. 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 $\sum_{t=T+1}^{\infty} \boldsymbol{\delta}(t) < \varepsilon/2u(a)$. Moreover, since $u$ is continuous, there is a neighborhood $O$ of $\mathbf{x}$ such that $O$ is open in the product topology and $|u(x_t) - u(y_t)| < \varepsilon/2 \sum_{t=0}^{T} \boldsymbol{\delta}(t)$ for all $\mathbf{y} \in O$ and $t = 1, ..., T$. 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 $a > 0$ was arbitrary in this discussion, it follows that $U$ is cube-continuous.

[22] A very good survey about the recent developments in time preference theory is provided by Frederick, Loewenstein and O'Donoghue (2002).

*(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 $(s,t) \in \mathbb{Z}_+^2$ with $s < t$ and $x, y, z \geq 0$.*

*Moreover, if $u$ and $v$ belong to $\mathcal{V}$, 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*

$$\inf\{h'(x) : x > 0\} \geq \left(\frac{\boldsymbol{\beta}(t)/\boldsymbol{\beta}(s)}{\boldsymbol{\alpha}(t)/\boldsymbol{\alpha}(s)}\right) \sup\{h'(x) : x > 0\} \quad \textit{for all } (s,t) \in \mathbb{Z}_+^2 \textit{with } s < t. \tag{6}$$

*(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 $(s,t) \in \mathbb{Z}_+^2$ with $s < t$ and $x, y \geq 0$.*

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 $(u, \boldsymbol{\alpha})$ iff $v$ is a particular transformation of $u$. This 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.

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 $(v, \boldsymbol{\beta}) \precsim (u, \boldsymbol{\alpha})$ iff $(v, \boldsymbol{\beta})$ 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 $t$.[23]

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.

[23] 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.

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

$$(u, \boldsymbol{\alpha}) \precsim (v, \boldsymbol{\beta}) \quad \textit{only if} \quad \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} \textit{ for all } (s,t) \in \mathbb{Z}_+^2 \textit{ with } s < t.$$

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

As we argued 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 that of $B$.[24] 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 $\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}$ and $\frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}$ must also be considered.[25] The corollary also shows 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* $(v, \boldsymbol{\beta})$,

$$(u, \boldsymbol{\alpha}) \precsim (u, \boldsymbol{\beta}) \quad \textit{if and only if} \quad \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} \textit{ for all } (s,t) \in \mathbb{Z}_+^2 \textit{ with } s < t.$$

*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 trivial consequence –of Theorem 2 for the case of *exponential* intertemporal utility functions.[26]

[24] This assertion follows from the fact that $\frac{\alpha^t}{\alpha^s} \geq \frac{\beta^t}{\beta^s}$ for all $0 < s < t$ iff $\alpha \geq \beta$.

[25] In fact, it is sufficient that $\frac{\boldsymbol{\alpha}(t+1)}{\boldsymbol{\alpha}(t)} \geq \frac{\boldsymbol{\beta}(t+1)}{\boldsymbol{\beta}(t)}$ for all $t \geq 0$.

[26] The substance of the corollary was investigated by Horowitz (1992) in a continuous time framework. As noted earlier, he finds the rather limited result that two agents are comparable if and only if their instantaneous utilities are positive affine combinations of each other. We return to this issue in Section 5.

**Corollary 3.** *Let $(u,\alpha)$ and $(v,\beta)$ be exponential intertemporal utility functions with $u,v\in\mathcal{V}$. The following statements are equivalent:*[27]

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

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

$$\inf\{h'(x):x>0\}\geq\frac{\beta}{\alpha}\sup\{h'(x):x>0\} \tag{7}$$

*(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 $u,v\in\mathcal{V}$ be instantaneous utility functions such that $u'(0+)=\infty$ and $v'(0+)<\infty$. Regardless of the values of the discount factors $\alpha$ and $\beta$, $(u,\alpha)$ and $(v,\beta)$ *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 $\mathcal{V}$ 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 $(u,\alpha)$ and $(v,\beta)$ cannot be ranked by $\precsim$ . ‖

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

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

and let $\mathcal{U}_0:=\{u_\sigma:0\leq\sigma<1\}$. This class is widely used in intertemporal macroeconomic models. Let us take two exponential intertemporal utility functions $(u_{\sigma_1},\alpha)$ and $(u_{\sigma_2},\beta)$ in $\mathcal{U}_0\times(0,1)$. When can these intertemporal utility functions be ranked by $\precsim$? The answer is *only* when the agents have *identical* instantaneous utility functions.[28] More precisely, $(u_{\sigma_1},\alpha)\precsim(u_{\sigma_2},\beta)$ 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 $\sigma_1=\sigma_2$. (Recall the last observation made in Example 1.) ‖

[27] Corresponding to (b) of Theorem 2, we also have that $(v,\beta)$ is more delay averse than $(u,\alpha)$ iff there exists a map $h:\mathbb{R}_+\to\mathbb{R}_+$ such that $v=h\circ u$ and $h(x+\alpha y)\geq h(x)+\beta(h(y+z)-h(z))$ for all $(s,t)\in\mathcal{T}$ and $x,y,z\geq 0$.

[28] 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 $s$, $x_s\in[x_{\min},x_{\max}]$. Then it makes sense to restrict (1) of Definition 2, to $a,b$, and $\boldsymbol{\omega}$ such that $x_{\min}\leq\omega_s+c\leq x_{\max}$ for all $s$, where $c=a$, $b$, and similarly for (2). With this restricted definition, it will be possible to compare some members of this class with different instantaneous utilities.

The examples above 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 $(u, \alpha)$. We wish to find a $(v, \beta) \in \mathcal{U} \times (0,1)$ 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 $h'(\infty) \geq \frac{\beta}{\alpha} h'(0+)$.

## 3.3 Characterizations of the Impatience Ordering

We turn now to the impatience ordering of Section 2.4. 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*

$$h\left(\left(1 - \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}\right) x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} y\right) \geq \left(1 - \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}\right) h(x) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} h(y) \tag{8}$$

*and*

$$h\left(\left(1 - \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}\right) y + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} x\right) \leq \left(1 - \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}\right) h(y) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)} h(x) \tag{9}$$

*for all $(s,t) \in \mathbb{Z}_+^2$ with $s < t$ and $y \geq x \geq 0$.*

Technically speaking, functional inequalities (8) and (9) are much less binding than (5), because they depend on two variables ($x$ and $y$) while (5) depends on three variables ($x, y,$ and $z$).[29] Notice that since $\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \in (0,1)$, (8) 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, (9) 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 $\boldsymbol{\beta}$.

While the influence of the instantaneous utility functions is smaller for impatience comparisons than for delay aversion comparisons, this influence is nevertheless present. Since our impatience ranking is meant to reflect agents' *pure* time preferences, and these are habitually thought of as measured by the discount functions, this may at first seem surprising. Yet, there is no mystery here. In the separable

[29] Both (8) and (9) (when stated for all $y \geq x \geq 0$) are special cases of (5) (when stated for all $x, y, z \geq 0$), as can be shown by a suitable change of variables.

intertemporal utility model, the instantaneous utility functions inherently play a role that goes beyond modeling static preferences - they are not really *instantaneous* in character. It is thus hardly surprising that they take active roles in determining the *pure* time preferences of a decision maker.

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.

**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 $\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)} \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}$ for all $(s,t) \in \mathbb{Z}_+^2$ with $s < t$. Moreover, $(u, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$ iff $(u, \boldsymbol{\alpha}) \precsim (u, \boldsymbol{\beta})$, and $(v, \boldsymbol{\alpha})$ is more impatient than $(u, \boldsymbol{\alpha})$ iff $u = \theta v$ 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.[30]

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

$$h\left((1-\alpha)\, x + \alpha y\right) \geq (1-\beta)\, h(x) + \beta h(y) \tag{10}$$

*and*

$$h\left((1-\alpha)\, y + \alpha x\right) \leq (1-\beta)\, h(y) + \beta h(x) \tag{11}$$

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

$$\max\left\{\frac{\alpha}{\beta}, \frac{1-\beta}{1-\alpha}\right\} h'(y) \geq h'(x) \geq \min\left\{\frac{\beta}{\alpha}, \frac{1-\alpha}{1-\beta}\right\} h'(y), \tag{12}$$

*for all $y \geq x \geq 0$, then $(v, \beta)$ is at least as impatient as $(u, \alpha)$.*

[30] 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.

We now give two applications of this corollary. The first application supplies a transformation $h$ which yields more impatient utility functions. The second application provides examples of impatience ranked utility functions which 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.

**Example 4.** 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* $\theta \geq 1$, $h_\theta$ *satisfies* (10) *and* (11) *for all* $y \geq x \geq 0$ *if and only if* $\alpha^\theta \geq \beta$; and (2) *For any* $1 \geq \theta > 0$, $h_\theta$ *satisfies* (10) *and* (11) *for all* $y \geq x \geq 0$ *iff* $1 - \beta \geq (1-\alpha)^\theta$.

We prove only the first claim, the proof of the second one being analogous. Fix any $\theta \geq 1$. Since $h_\theta$ is a strictly increasing convex function, it readily satisfies (11) for all $y \geq x \geq 0$. Thus, all we need to show is that (10) holds for all $y \geq x \geq 0$ iff $\alpha^\theta \geq \beta$. 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 $y > x \geq 0$, 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 $\frac{d}{dy} f_0(y) < 0$ for small enough $y > 0$. Then, $f_0(y) < 0$ for small enough $y > 0$; that is, $h_\theta$ fails to satisfy (10) for $x = 0$ and small $y > 0$. $\|$

**Example 2.** [Continued] Consider the class $\mathcal{U}_0 \times (0,1)$ 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 $\theta := \frac{1-\sigma_2}{1-\sigma_1}$, and define $h : \mathbb{R}_+ \to \mathbb{R}_+$ by $h(x) := \frac{(1-\sigma_1)^\theta}{1-\sigma_2} x^\theta$. Clearly, $u_{\sigma_2} = h \circ u_{\sigma_1}$, so by Theorem 4, $(u_{\sigma_2}, \beta)$ is more impatient than $(u_{\sigma_1}, \alpha)$ iff $h$ satisfies (10) and (11) for all $y \geq x \geq 0$. It follows from the results of Example 4 that $h$ does indeed satisfy these two inequalities. $\|$

### 3.4 Characterizations of the Cryogenic Delay Aversion Ordering

We turn now to the "cryogenically more delay averse than" ordering $\precsim^{\mathbf{0}}$. It is evident from its definition that this preorder is substantially more complete than our delay aversion and impatience orderings. The following attests to this.

**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*

$$h\left(\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}x\right) \geq \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}h(x) \quad \text{for all} \quad (s,t) \in \mathbb{Z}_+^2 \text{ with } s < t \text{ and } x \geq 0. \qquad (13)$$

*For any exponential intertemporal utility functions $(u, \alpha)$ and $(v, \beta)$, 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*

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

The functional inequality (13) is a special case of the functional inequalities (5) and (8), 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 $(v, \beta)$ is cryogenically more delay averse than $(u, \alpha)$ *even* if $\beta > \alpha$.[31] 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 iff $f(0) \leq 0$ 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 $v \in \mathcal{U}$ is said to be **less convex than** $u \in \mathcal{U}$ 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 $u, v \in \mathcal{U}$, we have $(u, \boldsymbol{\delta}) \precsim^{\mathbf{0}} (v, \boldsymbol{\delta})$ for all $\boldsymbol{\delta} \in \mathcal{D}$ if and only if $v$ is less star-shaped than $u$. In particular, for any $u, v \in \mathcal{V}$*

$$(u, \boldsymbol{\delta}) \precsim^{\mathbf{0}} (v, \boldsymbol{\delta}) \text{ for all } \boldsymbol{\delta} \in \mathcal{D} \quad \text{if and only if} \quad \frac{u'(x)}{u(x)} \geq \frac{v'(x)}{v(x)} \text{ for all } x > 0. \qquad (14)$$

[31] *Example.* Let $1/4 < \alpha < 1/2$, 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 $(u, \alpha) \prec^{\mathbf{0}} (v, 1/2)$, even though $\alpha < 1/2$.

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

**Corollary 7.** *For any exponential intertemporal utility functions* $(u, \alpha)$ *and* $(v, \beta)$*, 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* $u$*.*

In the next section we will use this result in an application within the context of bargaining theory.[32]

# 4 Applications

This section is devoted to four applications of the delay aversion and impatience orderings introduced above, and the characterizations thereof. In our first application, we compare the optimal capital accumulation path of a given country with that of a more delay averse country in the one-sector optimal growth problem. In our second application, we compare the optimal consumption path of a given individual with that of a more delay averse individual, and a more impatient individual, in a standard investment problem. Finally, we turn to game theory, and explore the implications of increasing a player's delay aversion, and impatience, in the contexts of repeated games and bargaining environments.

## 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 $k_t = f(i_{t-1})$, $t = 1, 2, ....$ (There is no capital depreciation.) The preferences of Country 1 over consumption paths is represented by the exponential intertemporal utility function $(u, \alpha)$, and those of Country 2 by $(v, \beta)$. 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 $(c_0, c_1, ...)$ and $(i_0, i_1, ...)$ in order to

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

The problem of Country 2 is formulated analogously.

We denote the optimal consumption and investment paths of Country $\mathbf{j} = 1, 2$ by $(c_0^{\mathbf{j}}, c_1^{\mathbf{j}}, ...)$ and $(i_0^{\mathbf{j}}, i_1^{\mathbf{j}}, ...)$, respectively. The optimal capital accumulation path of

[32] The converse of Corollary 7 is false. For instance, define $u, v \in \mathcal{U}$ 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 $x \geq 0$ iff $\alpha \geq \sqrt{\beta}$. Hence, $(u, \alpha) \prec^{\mathbf{0}} (v, \beta)$ may hold even when $v$ is *more* convex than $v$.

Country $\mathbf{j}$ is denoted by $(k_0^{\mathbf{j}}, k_1^{\mathbf{j}}, ...)$, $\mathbf{j} = 1, 2$. The question we address here is this: *If Country 2 is more delay averse than Country 1, how do their capital accumulation paths compare?*

The optimal consumption and investment paths of Countries 1 and 2 are determined by the following Ramsey-Euler equations:

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

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

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

so that

$$\frac{u'(c_t^{\mathbf{1}}) f'(i_t^{\mathbf{2}})}{u'(c_t^{\mathbf{2}}) f'(i_t^{\mathbf{1}})} \geq \frac{u'(c_{t+1}^{\mathbf{1}})}{u'(c_{t+1}^{\mathbf{2}})}, \quad t = 0, 1, ... \tag{15}$$

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 $t = 0, ..., T-1$ 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 $\frac{u'(c_{T-1}^{\mathbf{1}}) f'(i_{T-1}^{\mathbf{2}})}{u'(c_{T-1}^{\mathbf{2}}) f'(i_{T-1}^{\mathbf{1}})} < 1$ and, by (15), also $\frac{u'(c_T^{\mathbf{1}})}{u'(c_T^{\mathbf{2}})} < 1$. 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 $(c_T^{\mathbf{1}}, c_{T+1}^{\mathbf{1}}, ...) > (c_T^{\mathbf{2}}, c_{T+1}^{\mathbf{2}}, ...)$. This contradicts the optimality of the consumption path $(c_0^{\mathbf{2}}, c_1^{\mathbf{2}}, ...)$ for Country 2, since, given that $k_T^{\mathbf{1}} < k_T^{\mathbf{2}}$, it is feasible for Country 2 to consume $(c_T^{\mathbf{1}}, c_{T+1}^{\mathbf{1}}, ...)$ instead of $(c_T^{\mathbf{2}}, c_{T+1}^{\mathbf{2}}, ...)$.

It follows that $k_T^{\mathbf{1}} < k_T^{\mathbf{2}}$ cannot hold for *any* $T \in \mathbb{N}$. Put differently: *In the standard optimal growth problem, the optimal capital stock of a country can never fall strictly below that of a more delay averse country.*

This generalizes a well-known result in the literature, in which the two countries are assumed to have the same instantaneous utility function, and to differ only in their discount factors (Becker (1983)). Moreover, as can be seen from the above argument, not only is the assumption that the two countries have the same utility function unnecessarily restrictive, it does not even simplify the analysis.

Note that the above result extends immediately 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

## 4.2 The Investment Problem

Consider two individuals, Persons 1 and 2, each with an initial wealth $w > 0$. The preferences of Person 1 (resp. Person 2) over consumption streams are represented by the exponential intertemporal utility function $(u, \alpha)$ (resp. $(v, \beta)$), which satisfies the assumptions made in the previous subsection. In period $t$, both individuals must decide how much of their wealth they should place in a savings account that yields an interest rate $r > 0$. The problem of Person 1, which we refer to as his standard investment problem, is to choose nonnegative sequences $(c_0, c_1, ...)$ and $(i_0, i_1, ...)$ in order to

$$\text{Maximize } \sum_{t=0}^{\infty} \alpha^t u(c_t) \quad \text{such that } c_0 + i_0 = w, \text{ and } c_t + i_t = (1 + r)i_{t-1}, \ t = 1, 2, ...$$

The standard investment problem of Person 2 is formulated analogously.

It follows immediately from the analysis of the previous subsection that, if Person 2 is more delay averse than Person 1, then in the optimum solution Person 1 will have at least as much wealth as Person 2 at every period. We can say more here: *In the standard investment problem, the optimal consumption path of an individual single crosses that of a less delay averse individual from above.* Again, this generalizes a result in which both people are assumed to have the same instantaneous utility (Horowitz (1992)).

To see this, let us again denote the optimal consumption paths of Person $\mathbf{j} = 1, 2$ by $(c_0^{\mathbf{j}}, c_1^{\mathbf{j}}, ...)$, and assume that $(u, \alpha) \precsim (v, \beta)$. In the present setting, (15) becomes

$$\frac{u'(c_t^{\mathbf{1}})}{u'(c_t^{\mathbf{2}})} \geq \frac{u'(c_{t+1}^{\mathbf{1}})}{u'(c_{t+1}^{\mathbf{2}})}, \quad t = 0, 1, ... \tag{16}$$

Now consider a period $T$ such that $c_T^{\mathbf{1}} \geq c_T^{\mathbf{2}}$. By concavity of $u$, we have $\frac{u'(c_T^{\mathbf{1}})}{u'(c_T^{\mathbf{2}})} \leq 1$, so it follows from (16) that $\frac{u'(c_{T+1}^{\mathbf{1}})}{u'(c_{T+1}^{\mathbf{2}})} \leq 1$, so that $c_{T+1}^{\mathbf{1}} \geq c_{T+1}^{\mathbf{2}}$. Proceeding inductively yields $c_t^{\mathbf{1}} \geq c_t^{\mathbf{2}}$ for all $t = T, T + 1, ....$ Put differently, once the optimal consumption of Person 1 exceeds that of Person 2, it must continue doing so in the subsequent periods as well, and hence the claim stated above.

A natural question is whether this claim remains true under the weaker assumption that Person 2 is more *impatient* than Person 1. Interestingly, the answer is no. In fact, the optimal consumption path of Person 2 may then even single cross that of Person 2 from below.

To illustrate, let us simplify the investment problem at hand by setting $w = 1$ and $r = 0$, and letting $u := u_{\sigma_1}$ and $v := u_{\sigma_2}$ for some $0 < \sigma_1, \sigma_2 < 1$, where $u_\sigma(x) := \frac{x^{1-\sigma}}{1-\sigma}$ as in Example 2 of Section 3.2. With this specification, the optimal consumption path of person $\mathbf{j} = 1, 2$ is found as $((1 - \gamma_{\mathbf{j}}), \gamma_{\mathbf{j}}(1 - \gamma_{\mathbf{j}}), \gamma_{\mathbf{j}}^2(1 - \gamma_{\mathbf{j}}), ...)$, where $\gamma_1 = \alpha^{1/\sigma_1}$ and $\gamma_2 = \beta^{1/\sigma_2}$. Therefore, the optimal consumption path of Person 2 single crosses that of Person 1 from below if $\alpha^{1/\sigma_1} < \beta^{1/\sigma_2}$. Notice that this implies $\sigma_2 > \sigma_1$. Recall from the continuation of Example 2 in Section 3.3 that when $\sigma_2 > \sigma_1$ Person 2 is more impatient than Person 1 if $(1 - \beta)^{\frac{1}{1-\sigma_2}} \geq (1 - \alpha)^{\frac{1}{1-\sigma_1}}$. Therefore, for any specification of $0<\beta<\alpha<1$ and $0<\sigma_1<\sigma_2<1$ such that

$$\alpha^{1/\sigma_1}<\beta^{1/\sigma_2} \qquad \text{and} \qquad (1-\beta)^{\frac{1}{1-\sigma_2}} \geq (1-\alpha)^{\frac{1}{1-\sigma_1}}, \tag{17}$$

the optimal consumption path of Person 2 single crosses that of Person 1 from below, even though Person 2 is the more impatient of the two.[33]

## 4.3 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 will be the equilibrium set. In this section we address this issue.

Let $n \in \{2,3,...\}$. Consider an arbitrary (single-shot) game $\mathcal{G} := (N, \{A_i, p_i\}_{i\in N})$, where $N := \{1,...,n\}$ 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 $\mathcal{G}$. For any $a \in A$ and $i \in N$, let $BR_i(a)$ denote a best response of player $i$ to $a \in A$. We assume that the set of minmax strategies and the sets of best responses are always nonempty.

Let $(\mathcal{G}, \{V_i\}_{i\in N})$ be the infinitely repeated game in which the stage game is $\mathcal{G}$ and $V_i \in \mathfrak{U}$ is the intertemporal utility function that player $i \in N$ uses to evaluate her monetary payoff streams. Note that although it is standard to model the players in repeated games as exponential utility maximizers, we do not impose that restriction here. On the other hand, we restrict ourselves to pure strategies as we have not considered individuals' attitudes towards risk in this paper. Let $S_i$ stand for the set of all pure strategies of player $i \in N$. Formally, for any $i \in N$, a strategy $s_i \in S_i$ is a sequence $(s_i^0, s_i^1, ...)$ 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 $(s_1^t, ..., s_n^t)$ which is a map from $A^t$ into $A$. Any strategy profile $s \in S$ inductively defines an outcome path $(a_0(s), a_1(s), ...)$ for the repeated game $(\mathcal{G}, \{V_i\}_{i\in N})$ as follows: $a_0(s) := s^0$ and $a_t(s) := s^t(a_0(s), ..., a_{t-1}(s))$, $t = 1, 2....$ A *Nash equilibrium* of this game is a strategy profile $s \in S$ such that

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

for all $s_i \in S_i$ and $i \in N$.[34] The corresponding outcome path is a called a *Nash equilibrium outcome path*.

[33] The inequalities in (17) are compatible. For instance, they are satisfied for $\sigma_1 = 1/4$, $\sigma_2 = 1/2$, $\alpha = 1/2$, and any $\beta \in (0.25, 0.35)$.

[34] 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.

The simple intuition that the equilibrium payoff set is larger with less delay averse players is not in general correct, even for exponential utility maximizers. 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.[35]

**Proposition 2.** *Suppose that* $(a_0, a_1, ...)$ *is a Nash equilibrium outcome path of the repeated game* $(\mathcal{G}, \{V_i\}_{i\in N})$ *such that*

$$p_i\left(a_t\right) \geq p_i\left(m_i\right) \qquad \textit{for all } i \in N, t = 0, 1, ...$$

*Then* $(a_0, a_1, ...)$ *is also a Nash equilibrium outcome path of the repeated game* $(\mathcal{G}, \{U_i\}_{i\in N})$*, where* $U_i \in \mathfrak{U}$ *is less delay averse than* $V_i$ *for each* $i \in N$.

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 $(a_0, a_1, ...)$ is an equilibrium of $(\mathcal{G}, \{V_i\}_{i\in N})$ if and only if for each player $i \in N$,

$$V_i\left(p_i\left(a_0\right), p_i\left(a_1\right), ...\right) \geq V_i\left(p_i\left(BR_i\left(a_0\right)\right), p_i\left(m_i\right), p_i\left(m_i\right), ...\right)$$

and

$$V_i\left(p_i\left(a_0\right), p_i\left(a_1\right), ...\right) \geq V_i\left(p_i\left(a_0\right), ..., p_i\left(a_{k-1}\right), p_i\left(BR_i\left(a_k\right)\right), p_i\left(m_i\right), p_i\left(m_i\right), ...\right)$$

for all $k \in \mathbb{N}$. Note that since $p_i\left(a_t\left(s\right)\right) \geq p_i\left(m_i\right)$ for all $t$, the path $(p_i\left(a_0\right), p_i\left(a_1\right), ...,$ $p_i\left(a_{k-1}\right), p_i\left(BR_i\left(a_k\right)\right), p_i\left(m_i\right), p_i\left(m_i\right), ...)$ single crosses $(p_i(a_0), p_i\left(a_1\right), ...)$ from above. It follows immediately from Theorem 1 that $U_i \precsim V_i$ implies

$$U_i\left(p_i\left(a_0\right), p_i\left(a_1\right), ...\right) \geq U_i\left(p_i\left(BR_i\left(a_0\right)\right), p_i\left(m_i\right), p_i\left(m_i\right), ...\right)$$

and

$$U_i\left(p_i\left(a_0\right), p_i\left(a_1\right), ...\right) \geq U_i\left(p_i\left(a_0\right), ..., p_i\left(a_{s-1}\right), p_i\left(BR_i\left(a_k\right)\right), p_i\left(m_i\right), p_i\left(m_i\right), ...\right)$$

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

[35] For exponential utility maximizers, it follows from Theorem 3 of Abreu, Pearce, and Stacchetti (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.

In the previous subsection, we saw that making changes in peoples' impatience, rather than their delay aversion, may result in radically different conclusions. In the present context, however, changes in impatience and delay aversion may yield similar results, at least within the exponential discounting model. 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.

**Proposition 3.** *If* $(a, a, ...)$ *is a Nash equilibrium outcome path of the repeated game* $(\mathcal{G}, \{(v_i, \beta_i)\}_{i\in N})$*, then* $(a, a, ...)$ *is also a Nash equilibrium outcome path of the repeated game* $(\mathcal{G}, \{(u_i, \alpha_i)\}_{i\in N})$*, 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 $(a, a, ...)$ is an equilibrium path for $(\mathcal{G}, \{(v_i, \beta_i)\}_{i\in N})$ if and only if, for each $i \in N$,

$$\frac{1}{1-\beta_i} v_i\left(p_i\left(a\right)\right) \geq v_i\left(p_i\left(BR_i\left(a\right)\right)\right) + \frac{\beta_i}{1-\beta_i} v_i\left(p_i\left(m_i\right)\right),$$

that is,

$$v_i\left(p_i\left(a\right)\right) \geq \left(1-\beta_i\right) v_i\left(p_i\left(BR_i\left(a\right)\right)\right) + \beta_i v_i\left(p_i\left(m_i\right)\right). \tag{18}$$

Now suppose that $(u_i, \alpha_i)$ is more patient than $(v_i, \beta_i)$, and let $h_i := v_i \circ u_i^{-1}$ for each $i \in N$. Then (18) and the functional inequality (11) of Corollary 5 yield

$$\begin{aligned} h_i(u_i\left(p_i\left(a\right)\right)) &\geq \left(1-\beta_i\right) h_i(u_i\left(p_i\left(BR_i\left(a\right)\right)\right)) + \beta_i h_i(u_i\left(p_i\left(m_i\right)\right)) \\ &\geq h_i(\left(1-\alpha_i\right) u_i(p_i(BR_i\left(a\right))) + \alpha_i u_i(p_i\left(m_i\right))) \end{aligned}$$

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

$$u_i\left(p_i\left(a\right)\right) \geq \left(1-\alpha_i\right) u_i(p_i(BR_i\left(a\right))) + \alpha_i u_i(p_i\left(m_i\right)) \quad \text{for all } i \in N.$$

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

## 4.4 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.[36] In this section, we argue that Roth's result is in fact properly understood as a result about delay aversion, not risk aversion.

[36] 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 left unexplored.

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 $u \in \mathcal{U}$, and his discount factor is $\alpha \in [0,1]$. The utility function of the second mover, player 2, is also a concave function $w \in \mathcal{U}$, and her discount factor is $\delta \in [0,1]$. Under this specification, there is a unique subgame perfect equilibrium of the game. 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 $v \in \mathcal{U}$ and discount factor is $\beta \in [0,1]$. Suppose that $(u,\alpha) \precsim^{\mathbf{0}} (v,\beta)$, that is, this new player is cryogenically more delay averse than the original player 1. 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 cryogenically more delay averse agent $(v,\alpha)$ should perform less successfully than the agent $(u,\alpha)$, 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.

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 $x' > x$. Then $y > y'$, so since the map $a \mapsto a - w^{-1}(\delta w(a))$ is increasing on $\mathbb{R}_+$,[37] we have

[37] The claim is not trivial only when $0 \leq \delta < 1$. In that case, define $\phi : \mathbb{R}_+ \to \mathbb{R}$ by $\phi(a) := a - w^{-1}(\delta w(a))$, and notice that $\phi(0) = 0$ and $\phi(a) > 0$ for all $a > 0$ (since $w(0) = 0$ and $w$ is strictly increasing). Take any $a > b > 0$. Since $w$ is strictly increasing, $w(a) > w(b)$, and therefore, convexity of $w^{-1}$ entails that

$$\frac{\phi(a)}{w(a) - \delta w(a)} = \frac{w^{-1}(w(a)) - w^{-1}(\delta w(a))}{w(a) - \delta w(a)} \geq \frac{w^{-1}(w(b)) - w^{-1}(\delta w(b))}{w(b) - \delta w(b)} = \frac{\phi(b)}{w(b) - \delta w(b)}.$$

But $(1-\delta)w(a) > (1-\delta)w(b)$ and $\phi(b) > 0$, so we have $\phi(b)/(w(b) - \delta w(b)) > \phi(b)/(w(a) - \delta w(a))$, and combining this with the previous inequality yields $\phi(a) > \phi(b)$.

$A \geq A'$. 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 $(v, \beta)$ being cryogenically more delay averse than $(u, \alpha)$, in view of Theorem 4. Conclusion: *In the Rubinstein bargaining model, a bargainer's share decreases as be becomes cryogenically 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 cryogenically 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.[38]

# 5 Concluding Comments

*The Decision Interval.* The analysis above makes it clear that the instantaneous utility functions may contribute substantially to an agent's aversion to delay. As one might suspect, the extent of this contribution, and hence the incompleteness of the preorder $\precsim$, depends upon the decision interval involved in the agent's choices.

To illustrate, take any exponential intertemporal utility function $(u, \alpha) \in \mathcal{V} \times (0,1)$ and assume that this function represents the preferences of a person who makes his consumption decisions once a year. Thus, $\alpha$ refers to the annual discount rate, while $u$ is best interpreted as the individual's annual flow payoff from consuming a constant amount throughout the year. Now suppose the individual makes his decisions every

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

$K$ years instead (with his consumption again being constant within a decision period). Then the appropriate discount factor to use is $\alpha^K$. (For instance, if she decides twice a year then $K = \frac{1}{2}$, and the accompanying discount factor is $\alpha^{\frac{1}{2}}$.) Let $u_K$ denote the corresponding instantaneous utility function, which must be modified to reflect the fact that the payoff flow is now over $K$ years.[39] Similarly, let $(v_K, \beta^K)$ represent an individual whose yearly intertemporal utility function is $(v, \beta) \in \mathcal{V} \times (0, 1)$, and who makes his decisions every $K$ years instead.

Using Corollary 3 one can show that $(u_K, \alpha^K) \precsim (v_K, \beta^K)$ can hold only if there exists a map $h$ such that $v = h \circ u$ and

$$\inf\{h'(x) : x > 0\} \geq \left(\frac{\beta}{\alpha}\right)^K \sup\{h'(x) : x > 0\}.$$

Clearly, when $K$ is small, this condition is difficult to satisfy. In particular, if $K$ is very small, then $\left(\frac{\beta}{\alpha}\right)^K \approx 1$, so that $h$ must be approximately linear. That is, for very small decision intervals, the two agents can be ranked only if their instantaneous preferences are essentially the same. In the limit, when $K = 0$, the ordering $\precsim$ is extremely incomplete; $(u_K, \alpha^K) \precsim (v_K, \beta^K)$ only if $u$ and $v$ are positive affine transformations of each other. Hence obtains the aforementioned result that in *continuous* time only instantaneously equivalent agents can be ordered (Horowitz (1992)).[40] At the other extreme, when $K$ is large (as when the decision intervals are generations), the above inequality becomes easy to satisfy. In fact, if the derivatives of $u$ and $v$ are bounded by strictly positive numbers, then, for large enough $K$ we have $(u_K, \alpha^K) \precsim (v_K, \beta^K)$ if and only if $\alpha \geq \beta$.

*Delay Aversion with Risk.* Our entire analysis has been confined to the context of intertemporal 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 cryogenic delay aversion are all applicable to intertemporal preferences over stochastic streams, so the present work provides a good starting point for such a study.[41]

*Delay Aversion with Time Consistency.* The general approach in this paper encompasses both preference structures that yield time consistent choices and those

[39] Formally, $u_K := \frac{(1-\alpha^K)}{(1-\alpha)} u$. Note that a constant consumption $x$ in every period then results in the total utility of $\frac{1}{1-\alpha^\alpha} \frac{(1-\alpha^\alpha)}{(1-\alpha)} u(x) = \frac{1}{1-\alpha} u(x)$, which is independent of $K$.

[40] At the other extreme, when $K$ becomes very large, the above inequality becomes easy to satisfy. In fact, if $u$ and $v$ are two instantaneous utility functions whose derivatives are uniformly bounded by strictly positive numbers, then, for large enough $K$: $(u_K, \alpha^K) \precsim (v_K, \beta^K)$ if and only if $\alpha \geq \beta$.

[41] Interesting issues will arise involving disentangling the effects of delay aversion and risk aversion. Similar issues arise in separating the consumption smoothing and risk aversion motives in intertemporal choice theory with risk (cf. Epstein and Zin (1989)).

that yield time inconsistent choices. As we saw, in the standard optimal growth problem (Section 4.1), the capital stock of a country will always be below the capital stock of a more delay averse country if (i) both countries are exponential utility maximizers, whether or not they can commit to their plans, or (ii) both countries have separable but not exponential preferences, and can commit to their plans. It would be interesting to see if the time consistent choices (i.e., the no-commitment choices) for countries that cannot commit, and have non-exponential separable preferences, have the same property.

Another worthwhile avenue of research is the extension of the theory we presented for separable time preferences to the context of recursive preferences which, per force, induce time consistent intertemporal plans.

# 6 Proofs

## 6.1 Proofs for Section 2

We begin with the following preliminary result which will facilitate some of the subsequent arguments.

**Lemma 1.** *For any $U, V \in U$, the following statements are equivalent:*

*(a) $U \precsim V$;*

*(b) For any given $\mathbf{x} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}_+^2$ 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 $\boldsymbol{\omega} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}_+^2$ 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 $\mathbf{x} \in \mathcal{X}$, $(s,t) \in \mathbb{Z}_+^2$ *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 $\boldsymbol{\omega} \in \mathcal{X}$, $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$, fix any $a, b \geq 0$, and assume first that $U(\omega_s + a, \boldsymbol{\omega}_{-s}) \left\{\substack{\geq \\ >}\right\} U(\omega_t + b, \boldsymbol{\omega}_{-t})$. 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 $V(\omega_s + a, \boldsymbol{\omega}_{-s}) \left\{\substack{\geq \\ >}\right\} V(\omega_t + b, \boldsymbol{\omega}_{-t})$. On the other hand, suppose that $\omega_s \geq a \geq 0$ and $\omega_t \geq b \geq 0$, and $U(\omega_t - b, \boldsymbol{\omega}_{-t}) \left\{\substack{\geq \\ >}\right\} U(\omega_s - a, \boldsymbol{\omega}_{-s})$. Defining $\mathbf{x} := (\omega_s - a, \boldsymbol{\omega}_{-s})$ and applying (b) we find $V(\omega_t - b, \boldsymbol{\omega}_{-t}) \left\{\substack{\geq \\ >}\right\} V(\omega_s - a, \boldsymbol{\omega}_{-s})$. Thus $U \precsim V$. ■

**Proof of Theorem 1.** We only need to prove 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 $T(t^*) := \{(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2 : U(\mathbf{y}) \geq U(\mathbf{x})$, $\mathbf{y}$ single crosses $\mathbf{x}$ from above and $|\{m \in \mathbb{Z}_+ : y_m > x_m\}| \leq t^*\}$. We wish to show that

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

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

Take any $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2$ and define $\sigma := \sup\{\max\{x_i, y_i\} : i = 1, 2, ...\}$. Since $U$ and $V$ are cube-continuous, $U|_{[0,\sigma]^\infty}$ and $V|_{[0,\sigma]^\infty}$ are continuous functions. We will 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 $s_m := \min S \backslash \{s_1, ..., s_{m-1}\}$, $m = 2, 3, ....$ 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 a $\xi_1 \in (0, y_0 - x_0)$ such that

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

By Lemma 1, then, we have

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

Similarly, there exists a $\xi_2 > 0$ 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 1 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 $(\xi_m)$ 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

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

Since $U(\mathbf{z}(m)) \leq U(\mathbf{y})$ for each $m$, the monotonicity of $U$ implies $z_0(m) \leq y_0$, $m = 1, 2, ...$ . 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 some $M_1 > 0$ 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 an $M_2 \in \mathbb{N}$ 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

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

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 a $k \in \mathbb{N}$ 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 an $M \in \mathbb{N}$ such that $y_m \geq x_m$ for all $m = 0, ..., M-1$, 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 a $\mathbf{w} \in \mathcal{X}$ such that $U(\mathbf{w}) = U(\mathbf{x})$, and

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

Notice that, by monotonicity of $U$, we have

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

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

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

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 $\{m \in \mathbb{Z}_+ : w_m > z_m\} = \{m \in \mathbb{Z}_+ : w_m > z_m\} \backslash \{0\}$, 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 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 $U \in \mathfrak{U}$, $\mathbf{x} \in \mathcal{X}$ and $(s, t) \in \mathbb{Z}_+^2$ *with* $s < t$, we define $\chi_{s,t,\mathbf{x}}^U : \mathbb{R}_+ \to \overline{\mathbb{R}}_+$ by

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

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

**Lemma 2.** *For any $U, V \in \mathfrak{U}$, we have $U \precsim V$ if and only if*

$$\chi^U_{s,t,\mathbf{x}} \geq \chi^V_{s,t,\mathbf{x}} \quad \text{for all } \mathbf{x} \in \mathcal{X} \text{ and } (s,t) \in \mathbb{Z}^2_+ \text{ with } s < t. \tag{21}$$

**Proof.** Let $U \precsim V$, fix any $\mathbf{x} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}^2_+$ *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 $\mathbf{x} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}^2_+$ *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 3.** *For any $U_n, U \in \mathfrak{U}$, $n = 1, 2, \ldots$, if $U_n \to U$ uniformly, then*

$$\chi^{U_n}_{s,t,\mathbf{x}} \to \chi^U_{s,t,\mathbf{x}} \quad \text{for all } \mathbf{x} \in \mathcal{X} \text{ and } (s,t) \in \mathbb{Z}^2_+ \text{ with } s < t.$$

**Proof.** Fix any $\mathbf{x} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}^2_+$ *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 an $N \in \mathbb{N}$ 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,

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

contradiction.

Without loss of generality, let $N = 1$, 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 an $N_1 \in \mathbb{N}$ 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 an $N_2 \in \mathbb{N}$ 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

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

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 2 and 3. ■

## 6.2 Proofs for Section 3

**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 $\mathbf{x} \in \mathcal{X}$ and $(s,t) \in \mathbb{Z}^2_+$ *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

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

Similarly,

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

By Lemma 2, (22) and (23), $(u, \boldsymbol{\alpha}) \precsim (v, \boldsymbol{\beta})$ if and only if

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

for all $\mathbf{x} \in \mathcal{X}$, $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$ and $b \geq 0$. Thus, letting $h := v \circ u^{-1}$, we find that $(u, \boldsymbol{\alpha}) \precsim (v, \boldsymbol{\beta})$ iff

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

for all $\mathbf{x} \in \mathcal{X}$, $(s,t) \in \mathbb{Z}_+^2$ *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 $(u, \boldsymbol{\alpha}) \precsim (v, \boldsymbol{\beta})$ iff

$$h\left(x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right) \geq h(x) + \frac{\boldsymbol{\beta}(t)}{\boldsymbol{\beta}(s)}\left(h(y+z) - h(z)\right) \quad \text{for all } (s,t) \in \mathbb{Z}_+^2 \text{ with } s < t \text{ and } x, y, z \geq 0,$$

as we sought.

Now suppose that $u$ and $v$ belong to $\mathcal{V}$, and take any $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$. If (b) holds, then we have

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

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 a $c \in \left[x, x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right]$ and a $d \in [y, y+z]$ such that

$$h\left(x + \frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y\right) - h(x) = h'(c)\frac{\boldsymbol{\alpha}(t)}{\boldsymbol{\alpha}(s)}y \quad \text{and} \quad h(y+z) - h(z) = h'(d)y. \tag{24}$$

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

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

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. ■

For any $(u, \boldsymbol{\delta}) \in \mathfrak{U}_{\text{sep}}$, $\omega \geq 0$, and $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$, define $\eta_{s,t,\omega}^{(u,\boldsymbol{\delta})} : \mathbb{R}_+ \to \mathbb{R}_+$ and $\varsigma_{s,t,\omega}^{(u,\boldsymbol{\delta})} : [0, \omega] \to [0, \omega]$ by

$$\boldsymbol{\delta}(s)u(\omega + \eta_{s,t,\omega}^{(u,\boldsymbol{\delta})}(b)) + \boldsymbol{\delta}(t)u(\omega) = \boldsymbol{\delta}(s)u(\omega) + \boldsymbol{\delta}(t)u(\omega + b)$$

and

$$\boldsymbol{\delta}(s)u(\omega - \varsigma_{s,t,\omega}^{(u,\boldsymbol{\delta})}(b)) + \boldsymbol{\delta}(t)u(\omega) = \boldsymbol{\delta}(s)u(\omega) + \boldsymbol{\delta}(t)u(\omega - b),$$

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

**Proof of Theorem 3.** It is an easy matter to verify that $(v, \boldsymbol{\beta})$ is more impatient than $(u, \boldsymbol{\alpha})$ iff

$$\eta_{s,t,\omega}^{(u,\boldsymbol{\alpha})} \geq \eta_{s,t,\omega}^{(v,\boldsymbol{\beta})} \quad \text{for all } \omega \geq 0 \text{ and } (s,t) \in \mathbb{Z}_+^2 \textit{ with } s < t \tag{25}$$

and

$$\varsigma_{s,t,\omega}^{(u,\boldsymbol{\alpha})} \geq \varsigma_{s,t,\omega}^{(v,\boldsymbol{\beta})} \quad \text{for all } \omega \geq 0 \text{ and } (s,t) \in \mathbb{Z}_+^2 \textit{ with } s < t. \tag{26}$$

Moreover, (25) holds iff

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

for all $\omega \geq 0$, $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$, and $b \geq 0$. Thus, letting $h := v \circ u^{-1}$, we find that (25) holds iff

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

for all $\omega \geq 0$, $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$, and $b \geq 0$. Making the change of variables $x := u(\omega)$ and $y := u(\omega + b)$, we have that (25) holds iff

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

for all $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$ and $y \geq x \geq 0$. One can similarly show that (26) holds iff (9) holds for all $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$and $y \geq x \geq 0$. ■

**Lemma 4.** *Let* $0 < \lambda < 1$. *If* $f : \mathbb{R}_+ \to \mathbb{R}_+$ *is continuous and*

$$f((1-\lambda)x + \lambda y) \geq (1-\lambda)f(x) + \lambda f(y) \quad \textit{for all } y \geq x \geq 0, \tag{27}$$

*then it is concave.*

**Proof.** We shall 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 $A_\infty := A_0 \cup A_1 \cup \cdots$ is dense in $[0, 1]$.

We only consider the case $1/2 \leq \lambda < 1$, the argument for the remaining case being analogous. Suppose that $cl(A_\infty) \neq [0, 1]$, that is, there exists a $\gamma \in (0, 1)\backslash cl(A_\infty)$. Since $(0, 1)\backslash cl(A_\infty)$ 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 $A_\infty$, 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 $1/2 \leq \lambda < 1$,

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

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

Lemma 4 can now be easily proved. Note first that one can easily show inductively that (27) holds iff

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

Since $f$ is continuous and $cl(A_\infty) = [0, 1]$, it follows that (27) holds iff

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

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 $(s,t) \in \mathbb{Z}_+^2$ *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

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

Combining this with (8) yields

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

Thus, by Lemma 4, $h$ is a concave function. But, for any fixed $y > x \geq 0$, (9) 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.

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 (8) and (9), and Lemma 4, 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 (10) holds for all $y \geq x \geq 0$, and suppose, as the induction hypothesis,

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

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

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

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 (8), holds for all $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$ and $y \geq x \geq 0$. Since one can similarly show (using this time (11)) that (9) also holds for all $(s,t) \in \mathbb{Z}_+^2$ *with* $s < t$ and $y \geq x \geq 0$, the claim follows from Theorem 3.

To prove the second assertion of Corollary 5, assume that (12) holds for all $y \geq x \geq 0$. Consider first the case $\frac{\beta}{\alpha} \leq \frac{1-\alpha}{1-\beta}$ so that

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

Observe that if $h'(x) < \frac{\beta}{\alpha}h'(y)$ for some $x, y \geq 0$, then we must have $x > y \geq 0$ and $\frac{\alpha}{\beta}h'(x) < h'(y)$ which contradicts (28). Thus $h'(x) \geq \frac{\beta}{\alpha}h'(y)$ for all $x, y \geq 0$, that is, (7) holds, so by Corollary 3, $(v, \beta) \precsim (u, \alpha)$. Hence, in particular, $(v, \beta)$ is more impatient than $(u, \alpha)$.

Finally, consider the case $\frac{\beta}{\alpha} > \frac{1-\alpha}{1-\beta}$ so that (12) becomes

$$\frac{1-\beta}{1-\alpha}h'(y) \geq h'(x) \geq \frac{1-\alpha}{1-\beta}h'(y) \quad \text{for all } y \geq x \geq 0. \tag{29}$$

Fix any $y \geq x \geq 0$ arbitrarily. Let $z := (1-\alpha)x + \alpha y$ and define $G : \mathbb{R}_+ \to \mathbb{R}_+$ by $G(\omega) := \frac{\omega}{1-\alpha} - \frac{\alpha}{1-\alpha}y$. Notice that $G(y) = y$, $G(z) = x$, $G(\omega) \leq \omega$ for all $\omega \in [0, y]$, and that $h'(G(\omega)) = (1-\alpha)\frac{d}{d\omega}h(G(\omega))$ for all $\omega$. By (29),

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

so by the Fundamental Theorem of Calculus, we have $h(y) - h(z) \leq (1-\beta)(h(y) - h(x))$, which is equivalent to (10).

Now let $w := (1-\alpha)y + \alpha x$ and define $H : \mathbb{R}_+ \to \mathbb{R}_+$ by $H(\omega) := \frac{\omega}{1-\alpha} - \frac{\alpha}{1-\alpha}x$. Notice that $H(x) = x$, $H(w) = y$, $H(\omega) \geq \omega$ for all $\omega \in [x, \infty)$, and that $h'(H(\omega)) = (1-\alpha)\frac{d}{d\omega}h(H(\omega))$ for all $\omega$. Then, by (29),

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

so by the Fundamental Theorem of Calculus, we have $h(w) - h(x) \leq (1-\beta)(h(y) - h(x))$, which is equivalent to (11).

We proved that (29) implies (10) and (11) for all $y \geq x \geq 0$. By the first part of Corollary 5, (29) implies that $(v, \beta)$ is more impatient than $(u, \alpha)$. ■

**Proof of Theorem 4.** The proof is analogous to those of Theorems 2 and 3; it 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 iff $\frac{d}{dt}\left(\frac{h(t)}{t}\right) \leq 0$ for all $t > 0$, or equivalently, $h'(t)t \leq h(t)$ for all $t > 0$. But, given any $u, v \in \mathcal{V}$, the first assertion of Corollary 6 says that $(u, \boldsymbol{\delta}) \precsim^{\mathbf{0}} (v, \boldsymbol{\delta})$ for all $\boldsymbol{\delta} \in \mathcal{D}$ iff $-(v \circ u^{-1})$ is star-shaped. Therefore, by the Inverse Function Theorem, $(u, \boldsymbol{\delta}) \precsim^{\mathbf{0}} (v, \boldsymbol{\delta})$ for all $\boldsymbol{\delta} \in \mathcal{D}$ iff

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

Since $u^{-1}(\mathbb{R}_{++}) = \mathbb{R}_{++}$, the latter statement is equivalent to (14). ■

# References


[1] Abreu, D., Pearce, D. and E. Stacchetti (1990), Toward a theory of discounted repeated games with imperfect monitoring, *Econometrica*, 58, 1041-1063.

[2] Arrow, K. (1964), The role of securities in the optimal allocation of risk-bearing, *Review of Economic Studies*, 31, 91-96.

[3] Becker, R. (1983), Comparative dynamics in the one sector optimal growth model, *Journal of Economic Dynamics and Control*, 6, 99 - 107.

[4] Böhm-Bawerk, E. (1891), *The Positive Theory of Capital*, [trans. by W. Smart] Freeport, Books for Libraries Press, 1971.

[5] Brock, W. and J. Scheinkman (1976), Global asymptotic stability of optimal control systems with applications to the theory of optimal growth, *Journal of Economic Theory*, 12, 164-190.

[6] Diamond, P. (1965), The evaluation of infinite utility streams, *Econometrica*, 33, 170-177.

[7] Epstein, L. (1987), Impatience, in *The New Palgrave: Dictionary of Economics*, ed. by J. Eatwell, M. Milgate and P. Newman, Macmillan Press, New York.

[8] Epstein, L. and S. Zin (1989), Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica*, 57 937-969.

[9] Fishburn, P. and A. Rubinstein (1982), Time preference, *International Economic Review*, 23, 677-694.

[10] Fisher, I. (1930), *The Theory of Interest*, New York, MacMillan.

[11] Frederick, S., G. Loewenstein and T. O'Donoghue (2002), Time discounting: A critical review," *Journal of Economic Literature*, 40, 351-401.

[12] Friedman, M. (1976), *Price Theory: A Provisional Text*, New York, Aldine de Gruyter.

[13] Horowitz, J. (1992), Comparative impatience, *Economics Letters*, 38, 25-29.

[14] Koopmans, T. C. (1960) Stationary ordinal utility and impatience, *Econometrica*, 28, 287-309.

[15] Koopmans, T. C. P. Diamond and R. Williamson (1964), Stationary utility and time perspective, *Econometrica*, 32, 82-100.

[16] Lawrence, E. (1991), Poverty and the rate of time preference: Evidence from panel data," *Journal of Political Economy*, 99, 54-77.

[17] Marinacci, M. (1998), An axiomatic approach to complete patience and time invariance, *Journal of Economic Theory*, 83, 105-144.

[18] Pratt, J. (1964), Risk aversion in the small and in the large, *Econometrica*, 32, 122-136.

[19] Osborne, M. and A. Rubinstein (1994), *A Course on Game Theory*, Cambridge, MIT Press.

[20] Olson, M. and M. Bailey (1981), Positive time preference, *Journal of Political Economy*, 89, 1-25.

[21] Roth, A. (1985), A note on risk aversion in a perfect equilibrium model of bargaining, *Econometrica*, 53, 207-211.

[22] Sorin, S. (1986), On repeated games with complete information, *Mathematics of Operations Research*, 11, 147-160.

[23] Stigler, G. (1987), *The Theory of Price*, New York, MacMillan.