

Non-hyperbolic time inconsistency[☆]

Han Bleichrodt, Kirsten I.M. Rohde, Peter P. Wakker^{*}

Department of Economics, Erasmus University, Rotterdam, The Netherlands

Received 21 June 2007

Available online 11 July 2008

Abstract

The commonly used hyperbolic and quasi-hyperbolic discount functions have been developed to accommodate decreasing impatience, which is the prevailing empirical finding in intertemporal choice, in particular for aggregate behavior. However, these discount functions do not have the flexibility to accommodate increasing impatience or strongly decreasing impatience. This lack of flexibility is particularly disconcerting for fitting data at the individual level, where various patterns of increasing impatience and strongly decreasing impatience will occur for a significant fraction of subjects. This paper presents discount functions with constant absolute (CADI) or constant relative (CRDI) decreasing impatience that can accommodate any degree of decreasing or increasing impatience. In particular, they are sufficiently flexible for analyses at the individual level. The CADI and CRDI discount functions are the analogs of the well-known CARA and CRRA utility functions for decision under risk.

© 2008 Elsevier Inc. All rights reserved.

JEL classification: D90; C60

Keywords: Hyperbolic discounting; CADI discounting; CRDI discounting; Decreasing impatience; Time preference

1. Introduction

Under stationarity, indifference between a small outcome received soon and a large outcome received later is preserved if both outcomes are equally delayed. Stationarity reflects constant impatience, and is equivalent to time consistency under common assumptions (“stopwatch time,” resetting the zero of time to the moment of decision). Empirical studies have found that stationarity is usually violated (Frederick et al., 2002), with impatience mostly decreasing and not constant. That is, delaying the aforementioned outcomes makes the decision maker less impatient and more willing to wait for the (large and) late outcome. Thus, the indifference turns into a preference for the late outcome, and stationarity is violated.

The most popular discount functions today, the generalized hyperbolic (Loewenstein and Prelec, 1992) and quasi-hyperbolic (Phelps and Pollak, 1968; Laibson, 1997) discount functions, were introduced so as to accommodate

[☆] Han Bleichrodt’s research was made possible through a VICI research grant, and Kirsten Rohde’s research through a VENI research grant, both from the Netherlands Organization for Scientific Research (NWO).

^{*} Corresponding author at: Econometric Institute, Erasmus University, PO Box 1738, 3000 DR, Rotterdam, The Netherlands. Fax: +31 (0)10 408 91 62.

E-mail address: Wakker@few.eur.nl (P.P. Wakker).

URL: <http://people.few.eur.nl/wakker/index.html> (P.P. Wakker).

decreasing impatience. A drawback is that they do not have enough flexibility to accommodate increasing or strongly decreasing impatience, as we will demonstrate. These restrictions make it impossible to fit data at the individual level because there will always be significant fractions of subjects with increasing or strongly decreasing impatience (Abdellaoui et al., 2007; Harrison et al., 2002; Barsky et al., 1997). The impossibility to fit data at the individual level is particularly disconcerting in view of recent advances in neuroeconomics, where typically only a few individuals can be analyzed.

To further clarify the above points, we compare intertemporal choice to choice under risk. Even while risk aversion is the prevailing empirical phenomenon, commonly found for aggregate behavior, at the individual level some risk seekers are always found. Hence, to fit data at the individual level we need utility functions that are flexible enough to accommodate risk seeking. The CARA (constant absolute risk averse) and CRRA (constant relative risk averse) functions can do so, and these functions have, accordingly, been widely used. The logarithmic function¹ $U(x) = b(\ln(a + x))$ cannot accommodate risk seeking and is, therefore, considerably less popular. Without utilities such as CARA and CRRA available, no quantitative analysis of risk attitude at the individual level would be possible. Such a problematic situation exists at present for intertemporal choice. Indeed, some studies that attempted to fit intertemporal preferences at the individual level using (quasi-)hyperbolic discounting typically had to discard some 15 percent of their subjects (Kirby, 1997; Green et al., 1999).

The problem just mentioned is aggravated by some recent empirical studies that even found increasing impatience at the aggregate level (Attema et al., 2008; Gigliotti and Sopher, 2004; Read et al., 2005; Sayman and Öncüler, 2006). Onay and Öncüler (2007) and Chesson and Viscusi (2003) even found concave discounting at the aggregate level, implying strongly increasing impatience. Hence, also at the aggregate level there is an interest in new discount functions that can flexibly accommodate increasing impatience.

The aforementioned findings show that there are problems with hyperbolic and quasi-hyperbolic discounting. This paper introduces two classes of discount functions that avoid the problems mentioned. These classes can accommodate any degree of increasing impatience, and also any degree of decreasing impatience. Hence, they cover all degrees covered by hyperbolic discounting and allow additional degrees on top of those, giving increased flexibility at no cost. Our classes are the intertemporal counterparts of the CARA and CRRA utility functions from decision under risk (or, more precisely, of state-dependent generalizations thereof). They generalize classes of discount functions introduced by Prelec (1989) and Ebert and Prelec (2007). The latter were of a similar mathematical form as our families, but were not rich enough to cover all degrees of increasing or decreasing impatience.

The paper is organized as follows. Basic definitions are in Section 2. Section 3 analyzes the properties of generalized hyperbolic and quasi-hyperbolic discounting. Section 4 presents our classes of discount functions, while their main empirical implications are used to provide preference foundations in Sections 5 and 6. Section 7 derives further properties of our discount functions, demonstrating in particular that they can accommodate any degree of increasing or decreasing impatience. We further generalize the model by allowing for zero and negative utility. Section 8 concludes, and proofs are in Appendix A.

2. Discounting

Throughout this paper we study preferences \succsim over *timed outcomes* $(t : x) \in T \times X$, where T is a nondegenerate subinterval of $[0, \infty]$ and X is the outcome set. A timed outcome $(t : x)$ is interpreted as the receipt of *outcome* x at *time point* t . The preference notation $\succ, \sim, \preccurlyeq$, and \prec is as usual.

We assume that preferences \succsim can be represented by *discounted utility* (DU):

$$DU(t : x) = \varphi(t)U(x).$$

That is, $(t : x) \succ (s : y)$ if and only if $DU(t : x) \geq DU(s : y)$. DU is the *discounted utility*, U is the *utility*, and φ is the *discount function*. Preference conditions for DU were given by Roskies (1965), Krantz et al. (1971, Chapter 7), and Fishburn and Rubinstein (1982) (either only for gains or with φ for losses different than for gains). For brevity, we do not state these conditions but assume DU throughout. Time points are often denoted by s and t , where we usually

¹ We will often use the fact that minus the logarithm of a discount function plays a role in intertemporal choice that in a mathematical sense is similar to utility in decision under risk. Decreasing impatience is the analog of risk aversion and increasing impatience is the analog of risk seeking. Logarithmic utility is the analog of hyperbolic discounting.

have $s < t$ with s abbreviating “soon.” Throughout this paper our main interest will concern the discount function. We assume $\varphi > 0$, with φ continuous and strictly decreasing. The discount function is often normalized through $\varphi(0) = 1$ in the literature, but in this paper it is more convenient not to commit to such a normalization, the more so as we allow T not to contain time point 0. The reason for allowing T not to contain 0 is, besides the general increased flexibility in domains, that it gives rise to new discount functions with more flexibility regarding increasing and decreasing impatience. We will elaborate on this in Section 4.

We assume that the utility image $U(X)$ contains a nondegenerate interval $(0, \epsilon)$ for an $\epsilon > 0$, implying that X must be infinite. Other than that, X and $U(X)$ can be general. For example, with U continuous, outcomes can be monetary with X containing a subinterval of the reals, or outcomes can be commodity bundles with X a convex subset of \mathbb{R}^m . With a slight abuse of notation, for outcomes x, y we write $x \succcurlyeq y$ if there exists a time point t such that $(t : x) \succcurlyeq (t : y)$. Because of positivity of φ , U represents \succcurlyeq over the outcomes, and the aforementioned preference at one time point t automatically implies it at all time points t . *Gains* are outcomes with positive utility and *losses* are outcomes with negative utility.

Because φ is strictly decreasing, *impatience* follows: for all $s < t$ and gains x , $(s : x) \succ (t : x)$. The condition implies a corresponding condition, with the implied preference reversed, for losses: for all $s < t$ and losses x , $(s : x) \prec (t : x)$. Impatience is assumed throughout this paper. In what follows, many preference conditions will be imposed only on gains. The corresponding preference conditions for losses, with preferences reversed, then automatically follow from DU.

Stationarity, also called *constant impatience*, holds if

$$(s : x) \sim (t : y) \iff (s + \tau : x) \sim (t + \tau : y) \text{ for every } \tau > 0.$$

Thus, the indifference depends only on t and s through their difference $t - s$. The condition is defined only if τ and the other variables are such that all time points considered are contained in T . As a convention, for all preference conditions throughout this paper it will always be assumed implicitly that the time points considered are contained in T .

Decreasing impatience holds if for all $s < t$, $\sigma > 0$, and gains $x < y$,

$$(s : x) \sim (t : y) \implies (s + \sigma : x) \preccurlyeq (t + \sigma : y).$$

Increasing impatience holds if the implied preference always is the reverse. Under decreasing impatience, if we consider two equivalent timed outcomes, then delaying both outcomes equally long results in less distinction between the times of receipt so that the best outcome then decides. In this sense, decreasing impatience reflects diminishing sensitivity: a time difference becomes less important as it lies farther ahead in the future. This interpretation applies to both gains and losses. Prelec and Loewenstein (1991) discussed this interpretation, and Ebert and Prelec (2007) provided empirical evidence supporting it.

We summarize the assumptions made.

Assumption 2.1 (*Structural Assumption*). $DU(t : x) = \varphi(t)U(x)$ represents \succcurlyeq on $T \times X$ with T a subinterval of $[0, \infty)$, $\varphi : T \rightarrow (0, \infty)$, $U : X \rightarrow \mathbb{R}$, $U(X) \supset (0, \epsilon)$ for an $\epsilon > 0$, and φ continuous and strictly decreasing.

3. Hyperbolic and quasi-hyperbolic discounting

Loewenstein and Prelec (1992) proposed generalized hyperbolic discounting to accommodate the various findings of decreasing impatience in the literature. There has been some confusion in the literature regarding the term hyperbolic discounting. To avoid confusion, Prelec (2004) used the term actually-hyperbolic instead of Loewenstein and Prelec’s (1992) generalized hyperbolic. We will use the term hyperbolic discounting to designate generalized hyperbolic discounting. The *hyperbolic discount function* is given by

$$\varphi(t) = (1 + \alpha t)^{-\beta/\alpha} \text{ for some } \alpha \geq 0 \text{ and } \beta > 0. \tag{1}$$

The function with $\alpha = 0$ is to be taken as exponential discounting (e^{-rt}). If α tends to 0 then the function indeed tends to this limit (Loewenstein and Prelec, 1992). Herrnstein (1961, p. 270) used the discount function $\varphi(t) = 1/t$.

Harvey (1986) studied hyperbolic discounting with $\alpha = 1$, and Mazur (1987) and Harvey (1995) discussed hyperbolic discounting with $\alpha = \beta$. Harvey called the latter *proportional discounting*.

Phelps and Pollak (1968) introduced *quasi-hyperbolic discounting*, where the discount function is

$$\varphi(t) = \beta\delta^t \quad \text{for some } \beta \leq 1 \text{ and } \delta > 0 \text{ for all } t > 0, \text{ and where } \varphi(0) = 1. \quad (2)$$

Laibson (1997) demonstrated the usefulness of quasi-hyperbolic discounting for economic applications. Quasi-hyperbolic discounters have decreasing impatience only at present, and constant impatience throughout the future.

We will now demonstrate that hyperbolic and quasi-hyperbolic discounting impose serious restrictions on the degrees of increasing and decreasing impatience. It is well known that both models imply decreasing impatience, as can be verified from substitution. There is, unfortunately, no easy way to generalize these models to accommodate increasing impatience. Because of its importance, we display this point as an observation.

Observation 3.1. *If hyperbolic discounting (Eq. (1)) holds without the restrictions on α and β , then either impatience is decreasing or constant, or impatience is increasing with $\alpha < 0$ and $T \subset [0, -1/\alpha)$. If quasi-hyperbolic discounting (Eq. (2)) holds without the restrictions on β and δ , then impatience still is decreasing or constant.*

Thus, increasing impatience can occur only for a particular subcase of the extension of the hyperbolic family. This subcase allows only for a bounded time interval and imposes a restriction on the parameter $\alpha < 0$ of increasing impatience, depending on the time domain considered ($-1/\alpha > t$ for all t). Another drawback of hyperbolic and quasi-hyperbolic discounting is that these models can accommodate decreasing impatience only to a limited degree. We analyze this claim in some detail, using indifferences as in Eq. (3) in a way similar to Prelec (1989, 2004). Consider gains $x < y$, time points $0 < s < t$, and a “delay” $\sigma > 0$, with

$$(s : x) \sim (t : y) \ \& \ (s + \sigma : x) \sim (t + \sigma + \tau : y). \quad (3)$$

In the second indifference, the receipt of both outcomes has first been delayed by σ . Then τ is the extra time the decision maker is willing to wait for the better y . The larger τ , the more impatience has decreased by the σ increase in time. The extra delay τ for the last timed outcome can be taken as a measure of decreasing impatience, and τ can be interpreted as a time premium similar to risk premiums for decision making under risk. Lower bounds for τ naturally result from impatience.

Lemma 3.2. *Assume Structural Assumption 2.1 and Eq. (3).*

- (i) *Impatience implies $\tau > -(t - s)$ and $\tau > -\sigma$.*
- (ii) *Decreasing impatience implies $\tau \geq 0$.*
- (iii) *Constant impatience (stationarity) implies $\tau = 0$.*
- (iv) *Increasing impatience implies $\tau \leq 0$.*

We next demonstrate that hyperbolic and quasi-hyperbolic discounting can only accommodate moderate degrees of decreasing impatience.

Observation 3.3. *Assume Structural Assumption 2.1 and Eq. (3). Then*

- (i) *Hyperbolic discounting implies $0 < \tau < \frac{\sigma(t-s)}{s}$.*
- (ii) *Quasi-hyperbolic discounting implies $\tau = 0$.*

The general model of Section 2 does not impose an upper bound on τ in Eq. (3), as we will see in Section 7, but hyperbolic discounting and quasi-hyperbolic discounting do. The following example illustrates some limitations resulting from these upper bounds.

Example 3.4. Assume (next week : \$100) \sim (in 2 weeks : \$105). Under hyperbolic discounting, (in 4 weeks : \$100) \sim (in 9 weeks : \$105) cannot hold, because it would imply, with $\sigma = 3$ weeks, $t - s = 1$ week, $\tau = 4 >$

$3 = \frac{\sigma(t-s)}{s}$ in violation of Observation 3.3.² Although the above indifferences reflect a strong degree of decreasing impatience, they can be exhibited by some subjects, and hyperbolic discounting cannot describe such subjects.

There is another way to see that hyperbolic discounting imposes strong limitations on the deviations from stationarity, to which we now turn. Prelec (2004) defined comparative decreasing impatience by comparing the values of τ in Eq. (3) for different persons. He showed that these comparisons reflect deviations from stationarity and vulnerability to arbitrage. Mathematically, they concern convexity of the logarithmic transform of the discount function. That is, more decreasing impatience corresponds to dominance at each time point t in terms of the following index of convexity (Prelec, 2004, Proposition 2b)³:

$$\gamma(t) = -\frac{[\ln \varphi(t)]''}{[\ln \varphi(t)]'}. \tag{4}$$

By straightforward algebraic manipulations, not given here, we obtain the following result.

Observation 3.5. *For quasi-hyperbolic discounting, $\gamma(t) = 0$ for all $t > 0$, and $\gamma(0)$ is undefined. For hyperbolic discounting (with $\alpha > 0$):*

$$\gamma(t) = \frac{\alpha}{1 + \alpha t} \in \left[0, \frac{1}{t}\right].$$

The result shows once again that there can be no increasing impatience (because $\gamma(t) \geq 0$), and that there is an upper bound to the degree of decreasing impatience with decreasing impatience vanishing if t tends to infinity.

Note that all limitations on discounting established in this section were irrespective of utility, i.e. they hold whatever the utility function is. They were, in fact, also independent of the actual discount rate, i.e. the power of the discount function. Prelec (2004, p. 515) argued for the plausibility of such independencies when studying deviations from stationarity.

4. Constant absolute and constant relative decreasing impatience for discount functions

This section defines the two new discount functions of this paper that will overcome the limitations of hyperbolic and quasi-hyperbolic discounting demonstrated in the preceding section. Their logarithms will be the well-known CARA and CRRA utility families from expected utility. For the first family, the index of Eq. (4) is constant, as we will see, and it will be the parameter c defined next.

Definition 4.1. The discount function φ is a *CADI function* if there exist constants $r > 0$ and c , and a normalization constant $k > 0$, such that:

$$\begin{aligned} \text{for } c > 0, \quad \varphi(t) &= ke^{re^{-ct}}; \\ \text{for } c = 0, \quad \varphi(t) &= ke^{-rt}; \\ \text{for } c < 0, \quad \varphi(t) &= ke^{-re^{-ct}}. \end{aligned}$$

Lemma 4.2. *For the CADI discount functions, Prelec’s (2004) measure is constant and is given by $\gamma(t) = c$.*

The proof follows from substitution and is omitted. The CADI family could also be called the double exponential family. The parameter c is the constant that indicates the convexity of $\ln(\varphi)$. Our family allows for any degree of decreasing impatience, simply because c can be any real number.

The discount function for $c = 0$ is the constant (exponential) discounting introduced by Samuelson (1937). The parameter k is a scaling factor without empirical meaning in the sense that it does not affect preferences. The parameter

² By Observation 3.3, hyperbolic discounting always implies that (in 4 weeks : \$100) > (in 9 weeks : \$105). Quasi-hyperbolic discounting implies the same, and even (in 4 weeks : \$100) ~ (in 5 weeks : \$105).

³ The index measures convexity of $\ln(\varphi)$, a decreasing function. For increasing functions, the index would measure concavity.

r , which is the power of the discount function (given that we can set $k = 1$), determines the degree of discounting and is empirically relevant when streams of two or more outcomes are received. In this paper we restrict attention to receipts of single outcomes. Then the joint power of discounting and utility can be chosen freely without affecting preference. That is, if preferences over timed outcomes can be represented by $\varphi(t)U(x)$, then they can also be represented by $\varphi(t)^\lambda U(x)^\lambda$ whenever $\lambda > 0$. Thus, r has no empirical meaning as long as utility has not been specified.

Prelec (1989) introduced the subpart of the CADI family for $c > 0$, i.e. the part with decreasing impatience. For $c > 0$, the function is bounded, being ke^r at $t = 0$ and converging to k for t tending to infinity (Prelec, 1989). No matter how far remote the future is, it always receives some minimal weight. To what extent this property is desirable will depend on the context where it is applied. For within-human decision making with bounded lifetime, the limiting behavior at infinity is not important. For societal decisions there are pros and cons. If the property is deemed undesirable, then the CRDI family defined later may be more suited.

The most convenient normalization of the parameters k and r , yielding $\varphi(0) = 1$, is as follows:

$$\begin{aligned} \text{for } c > 0, & \quad \varphi(t) = e^{(e^{-ct}-1)}; \\ \text{for } c = 0, & \quad \varphi(t) = e^{-t}; \\ \text{for } c < 0, & \quad \varphi(t) = e^{1-e^{-ct}}. \end{aligned}$$

We next turn to our second family. Outside of $t = 0$, dominance at each time point t in terms of the index in Eq. (4) corresponds, obviously, with dominance at each time point t in terms of the following index:

$$\delta(t) = -t \frac{[\ln \varphi(t)]''}{[\ln \varphi(t)]'}. \quad (5)$$

Hence, this index can also serve as a measure of decreasing impatience. It can be seen that it corresponds to relative, instead of absolute, concavity (“risk aversion”) of the logarithmic transform of the discount function. This index will be constant for our second family, and it will be the parameter d .

The first two cases below are defined only for $t > 0$ because φ tends to infinity for t tending to 0. A similar complication arises for CRRA functions, which also are not defined at 0 for negative powers and for power zero (which designates the logarithmic function). To incorporate the present time point into such models, we should not set $t = 0$ for the present but $t = \epsilon$ for some positive ϵ . This term ϵ can be compared to the initial wealth w that is often incorporated in CRRA utility functions $x \rightarrow \frac{(x+w)^r}{r}$, so that the functions are defined at $x = 0$ also for negative and zero power.

Definition 4.3. The discount function φ is a *CRDI function* if there exist constants $r > 0$ and d , and a normalization constant $k > 0$, such that:

$$\begin{aligned} \text{for } d > 1, & \quad \varphi(t) = ke^{rt^{1-d}} \text{ (only if } 0 \notin T); \\ \text{for } d = 1, & \quad \varphi(t) = kt^{-r} \text{ (only if } 0 \notin T); \\ \text{for } d < 1, & \quad \varphi(t) = ke^{-rt^{1-d}}. \end{aligned}$$

Lemma 4.4. For the CRDI discount functions, Prelec’s (2004) measure $\gamma(t)$ is d/t . The measure $\delta(t)$ is constant and is d .

Derivation is, again, left to the reader. Samuelson’s (1937) constant impatience results for $d = 0$. The functions exhibit decreasing impatience for $d > 0$ and increasing impatience for $d < 0$.

The CADI family results from the CRDI family when applied to e^t instead of to t (and with c for $d - 1$). Herrnstein (1961) considered the CRDI discount function for $d = 1$ with $r = 1$. Hyperbolic discounting and proportional discounting evolved as generalizations of Herrnstein’s discounting. The CRDI discount functions are also natural generalizations of Herrnstein’s function. Ebert and Prelec (2007) introduced the subpart of the CRDI family for $d < 1$, i.e. the part with increasing impatience or moderately decreasing impatience, for the domain $[0, \infty)$. If $t = 0$ is contained in the domain, then only $d < 1$ is possible, so that our family only generalizes Ebert and Prelec’s (2007) family in the absence of $t = 0$. Only then can the CRDI family accommodate strong degrees of decreasing impatience. This

can again be compared to constant relative risk aversion, where logarithmic utility and negative powers (risk aversion index 1 or more) are only possible if the 0 outcome is not contained in the domain.

A convenient normalization of the parameters k and r , yielding $\varphi(0) = 1$ when $t = 0$ is contained in the domain (in the case of $d < 1$), is as follows:

$$\begin{aligned} \text{for } d > 1, \quad \varphi(t) &= e^{t^{1-d}} \text{ (only if } t = 0 \notin T); \\ \text{for } d = 1, \quad \varphi(t) &= t^{-1} \text{ (only if } t = 0 \notin T); \\ \text{for } d < 1, \quad \varphi(t) &= e^{-t^{1-d}}. \end{aligned}$$

5. Constant absolute decreasing impatience: A preference condition

This section considers constant absolute decreasing impatience in terms of preferences. The resulting preference condition was introduced by Prelec (1989). He provided a preference foundation for CADI discount functions with a positive c under the assumption of linear utility. We generalize his foundation to general CADI functions and general utility.

Definition 5.1. *Constant absolute decreasing impatience (CADI) holds for preferences if for all $s < t < l$, all τ , and all outcomes⁴ x, y, z*

$$\begin{aligned} (s : x) \sim (t : y) \text{ and } (s + \tau : y) \sim (t + \tau : z) \text{ and} \\ (t : x) \sim (l : y) \\ \text{imply } (t + \tau : y) \sim (l + \tau : z). \end{aligned}$$

Constant absolute decreasing impatience has the following interpretation, assuming gains. The first indifference suggests that a delay from s to t is compensated by the improvement from x to y . The third indifference, the one below the first, suggests that the delay from t to l (l refers to “late”) requires exactly the same compensation. Under CADI, this equality of delays in terms of required compensation should not be affected if there is a common extra delay of τ for everything involved.

The next lemma translates the definition of CADI into a condition for discount function ratios: equality of such ratios is not affected if a common delay is added to all time points. The proof of the lemma clarifies the CADI condition and, hence, is given in the main text.

Lemma 5.2. *The following two statements are equivalent.*

- (i) *CADI holds for preferences.*
- (ii) *For all $s \leq t \leq l$ and τ*

$$\frac{\varphi(s)}{\varphi(t)} = \frac{\varphi(t)}{\varphi(l)} \iff \frac{\varphi(s + \tau)}{\varphi(t + \tau)} = \frac{\varphi(t + \tau)}{\varphi(l + \tau)}. \tag{6}$$

Proof. We will consider logical relations between the following equalities:

$$\frac{U(y)}{U(x)} \stackrel{1}{=} \frac{\varphi(s)}{\varphi(t)} \stackrel{2}{=} \frac{\varphi(t)}{\varphi(l)} \stackrel{3}{=} \frac{U(y)}{U(x)}; \quad \frac{U(z)}{U(y)} \stackrel{4}{=} \frac{\varphi(s + \tau)}{\varphi(t + \tau)} \stackrel{5}{=} \frac{\varphi(t + \tau)}{\varphi(l + \tau)} \stackrel{6}{=} \frac{U(z)}{U(y)}.$$

Equalities $\stackrel{1}{=}$, $\stackrel{3}{=}$, $\stackrel{4}{=}$, and $\stackrel{6}{=}$ are equivalent to the upper left, lower left, upper right, and lower right equivalences in the definition of CADI, respectively.

First assume that CADI holds. We assume the left equality in Eq. (6) and derive the right one (the reversed implication is symmetric). Because the image of U contains an interval $(0, \epsilon)$ for $\epsilon > 0$, we can find outcomes $x < y < z$ such that equalities $\stackrel{1}{=}$ and $\stackrel{4}{=}$ hold. Because $\stackrel{2}{=}$ is assumed, $\stackrel{3}{=}$ follows. We have $\stackrel{1}{=}$, $\stackrel{3}{=}$, and $\stackrel{4}{=}$, i.e. the three antecedent

⁴ Impatience implies $x < y < z$ for gains and $x > y > z$ for losses. It is sufficient to impose the CADI condition only for gains.

indifferences in the definition of CADI. The fourth indifference there follows from CADI, implying $\stackrel{6}{=}$ and, hence, $\stackrel{5}{=}$. This establishes Eq. (6).

Next assume Eq. (6). To derive CADI, assume the three antecedent indifferences there, implying $\stackrel{1}{=}$, $\stackrel{3}{=}$, and $\stackrel{4}{=}$. Then $\stackrel{2}{=}$ follows and then, by Eq. (6), $\stackrel{5}{=}$. This and $\stackrel{4}{=}$ imply $\stackrel{6}{=}$ and, thus, the fourth indifference in CADI. CADI has been established. \square

We now obtain the main result of this section.

Theorem 5.3. *Under Structural Assumption 2.1, the discount function is a CADI function if and only if CADI holds for preferences.*

6. Constant relative decreasing impatience: A preference condition

This section presents a preference condition for constant relative decreasing impatience. It is somewhat weaker than the preference condition used by Ebert and Prelec (2007), and it generalizes their preference foundation to the full CRDI family.

Definition 6.1. *Constant relative decreasing impatience (CRDI) holds for preferences if for all $s < t < l$, all $\lambda > 0$, and all outcomes⁵ x, y, z*

$$\begin{aligned} (s : x) \sim (t : y) \text{ and } (\lambda s : y) \sim (\lambda t : z) \text{ and} \\ (t : x) \sim (l : y) \\ \text{imply } (\lambda t : y) \sim (\lambda l : z). \end{aligned}$$

The interpretation is similar to that of CADI, but with scalar multiplication of the time points by a factor $\lambda > 0$ rather than the addition of a constant τ . In other words, if the unit of time is changed from a day into a week without changing the numbers, then equality of time delays in terms of required outcome-compensation should be maintained.

Lemma 6.2. *The following two statements are equivalent.*

- (i) *CRDI holds for preferences.*
- (ii) *For all $0 \leq s \leq t \leq l$ and $\lambda > 0$*

$$\frac{\varphi(s)}{\varphi(t)} = \frac{\varphi(t)}{\varphi(l)} \iff \frac{\varphi(\lambda s)}{\varphi(\lambda t)} = \frac{\varphi(\lambda t)}{\varphi(\lambda l)}. \quad (7)$$

The proof of the lemma is very similar to the one of Lemma 5.2 and will not be given. It can be derived from the aforementioned proof through the transformation $t \rightarrow \ln(t)$.⁶

Theorem 6.3. *Under Structural Assumption 2.1, the discount function is a CRDI function if and only if CRDI holds for preferences.*

7. Properties of CADI and CRDI discounting

We saw before that Prelec's (2004) measure $\gamma(t)$ is constant and can take any value for CADI discount functions. For CRDI, Prelec's measure can also be arbitrarily large or small, but always tends to plus or minus infinity for t tending to 0, while tending to 0 for t tending to infinity. The index $\delta(t) = t\gamma(t)$ is constant for CRDI functions and can take any value.

⁵ Impatience implies $x < y < z$ for gains and $x > y > z$ for losses. It is sufficient to impose the CRDI condition only for gains.

⁶ Now the values $\ln(t)$ play the role of t for CADI. The former can be negative (for $t < 1$), unlike the latter. This possible negativity does not require any modification of the analysis.

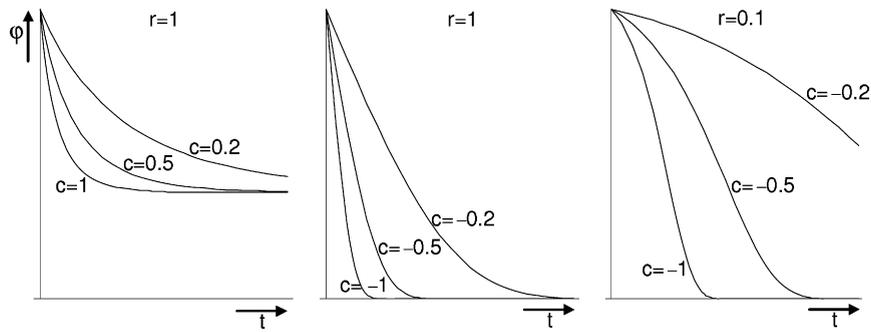


Fig. 1. CADI discount functions.

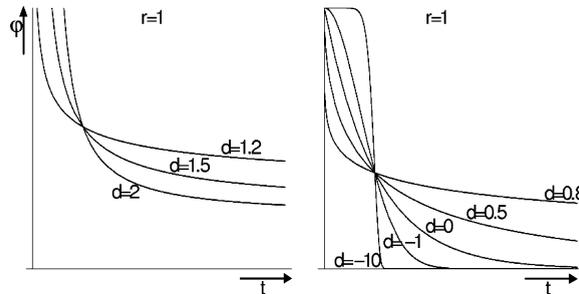


Fig. 2. CRDI discount functions.

Figs. 1 and 2 display some shapes of ϕ . Most shapes in the figures are convex, but some have concave parts initially, implying strongly increasing impatience there. The CRDI function in Fig. 2 has almost become a step function for $d = -10$, with no discounting prior to a horizon and complete discounting thereafter.

The following observation shows that the CADI and CRDI discount functions can account for all values of τ in Eq. (3). Thus, contrary to the conventional discounting functions, they can handle any degree of increasing or decreasing impatience. The proof will also illustrate that larger parameters c for CADI discounting and larger parameters d for the CRRI model lead to more strongly decreasing impatience.

Observation 7.1. *Under Structural Assumption 2.1, for all $s < t, \sigma, \tau$ and $(s : x) \sim (t : y)$, there exists a CADI discount function, and also a CRDI discount function, such that Eq. (3) is satisfied.*

A CADI discount function with $c < 0$ and appropriate r , and a CRDI discount function with $d < 1$ and appropriate r , can explain the results of Onay and Öncüler (2007) with a concave discount function on the relevant domain. When assuming discounted expected utility their results can only be explained by a concave discount function, which implies not only increasing impatience but even strongly so. Chesson and Viscusi (2003) similarly found that many of their subjects violate convex discounting. No traditional discount model can account for concave discounting.

In the same way as our CADI and CRDI functions are the analogs of CARA and CRRI utility for decision making under risk, the hyperbolic discount function is the analog of logarithmic utility $U(x) = \beta \ln(x + \alpha)$ for decision under risk. Quasi-hyperbolic discounting is the analog of linear utility $U(x) = x \ln(\beta)$ with a degeneracy at $x = 0$. The latter two families for decision under risk do not have the flexibility to accommodate risk seeking or extreme risk aversion, which explains why they are less popular than the CARA and CRRI utility functions. We have presented a similar case for discount functions.

8. Conclusion

This paper has introduced two families of discount functions (CADI and CRDI) that naturally extend Samuelson’s (1937) constant discounting and Herrnstein’s (1961) hyperbolic discounting. They serve to flexibly fit various patterns of intertemporal choice better than hyperbolic and quasi-hyperbolic discounting can do, by allowing any degree of

increasing or decreasing impatience. Thus, the CADI and CRDI discount families are the first that can be used to fit data at the individual level. They can also accommodate findings at the aggregate level that could not be explained up to now, such as the increasing impatience that has sometimes been found there. Hyperbolic discounting encounters the same restrictions for intertemporal choice as logarithmic utility does for decision making under risk. Similarly, the CADI and CRDI discount functions provide the same flexibility for intertemporal choice as the CARA and CRRA utility functions do for decision under risk.

Acknowledgments

Arthur E. Attema and an anonymous referee made helpful comments.

Appendix A. Proofs

Proof of Observation 3.1. First we consider Eq. (1) and derive decreasing or constant impatience. If $\alpha < 0$ then φ would be undefined or nonpositive for $t \geq -1/\alpha$ (unless $-\beta/\alpha$ is even which, if positive, violates impatience and, if negative, implies decreasing impatience), so that the domain restriction in the observation follows. Increasing impatience follows from substitution. $\alpha = 0$ implies constant discounting. So, assume $\alpha > 0$. Then $\beta > 0$ implies decreasing impatience. $\beta \leq 0$ cannot be because it would violate impatience.

Next we consider Eq. (2) and derive decreasing or constant impatience. For $\delta < 0$ the power δ^t would not be defined (imaginary number) for many t 's and φ would switch sign between odd and even integers t , so that it would not always be positive. $\delta = 0$ is also excluded by positivity of φ . Hence, $\delta > 0$. $\beta > 0$ is required because φ is positive. $\beta \leq 1$ implies constant or decreasing impatience. Finally, $\beta > 1$ would imply increasing and nonconstant impatience, but is excluded because it violates our assumption of impatience. \square

Proof of Lemma 3.2. The implication of impatience follows because we must have $t + \sigma + \tau > s + \sigma$ and $t + \sigma + \tau > t$. The other results in the observation are straightforward. \square

Proof of Observation 3.3. For hyperbolic discounting it follows from algebraic manipulations that $\tau = \frac{\alpha\sigma(t-s)}{1+\alpha s}$. Because $\alpha > 0$, the bounds in the observation follow. The bound for quasi-hyperbolic discounting follows because all time points considered are positive, and there quasi-hyperbolic discounting implies constant impatience. \square

Proof of Theorem 5.3. Eq. (6) entails that midpoints of $-\ln(\varphi(t))$ are invariant with respect to the addition of a common term to all arguments. According to Miyamoto (1983, Lemma 1), this is equivalent to $-\ln(\varphi(t))$ being from the CARA family. A useful feature of Miyamoto's lemma is that it allows for general intervals as domain. \square

Proof of Theorem 6.3. Eq. (7) implies that $-\ln(\varphi(e^t))$ -midpoints are invariant with respect to the addition of a common term to all arguments. According to Miyamoto (1983, Lemma 1), this is equivalent to $-\ln(\varphi(e^t))$ being from the CARA family. It is equivalent to $-\ln(\varphi)$ being from the CRRA family. \square

Proof of Observation 7.1. Assume Eq. (3). It implies that $\frac{U(y)}{U(x)}$ equals

$$\frac{\varphi(s)}{\varphi(t)} = \frac{\varphi(s + \sigma)}{\varphi(t + \sigma + \tau)}. \quad (\text{A.1})$$

As an aside, for all time points and discount functions we can take outcomes/utility such that the first indifference in Eq. (3) holds and, hence, the equality concerning the utility ratio above can be verified from preferences as in Eq. (3). To simplify upcoming formulas, we rewrite the equality as

$$\ln(\varphi(s)) - \ln(\varphi(t)) = \ln(\varphi(s + \sigma)) - \ln(\varphi(t + \sigma + \tau)). \quad (\text{A.2})$$

We assume all variables in Eq. (3) fixed except τ . We show that for each τ proper discount functions can be given that generate that τ . Unfortunately, there exist no analytical solutions to the proper choices of parameters c or d . We, therefore, give a proof based on continuity. We may assume that $t \leq s + \sigma$, because otherwise we can interchange these two symbols in Eq. (A.2)

We first consider CRDI discounting. The parameters k, r do not affect preference and we assume that they are 1. For any smooth function f , a difference $f(y) - f(x)$ can be written as $y - x$ times the average derivative of f over the interval $[x, y]$. Thus, because of Eq. (A.2), the ratio $\frac{(t+\sigma+\tau)-(s+\sigma)}{t-s}$ equals the [average derivative of $\ln(\varphi)$ over the interval $[s, t]$] divided by [the average derivative of $\ln(\varphi)$ over the interval $[(s + \sigma), (t + \sigma + \tau)]$].

We first show that, for arbitrarily large d , τ becomes arbitrarily large. The above ratio of average derivatives, $\frac{(s^{1-d}-t^{1-d})(s-t-\tau)}{(s-t)((s+\sigma)^{1-d}-(t+\sigma+\tau)^{1-d})}$, tends to ∞ for large d by dividing numerator and denominator by s^{1-d} , making the denominator tend to 0. For $\frac{(t+\sigma+\tau)-(s+\sigma)}{t-s}$ to tend to infinity, τ must tend to infinity.

We next show that for very small (very negative) d , τ becomes arbitrarily close to $-(t - s)$, which, given $t < s + \sigma$ and Lemma 3.2, is the lowest value τ can take. The ratio of average derivatives, $\frac{(s^{1-d}-t^{1-d})(s-t-\tau)}{(s-t)((s+\sigma)^{1-d}-(t+\sigma+\tau)^{1-d})}$, tends to 0 for very negative d by dividing numerator and denominator by $(t + \sigma + \tau)^{1-d}$, making the numerator tend to 0. For $\frac{(t+\sigma+\tau)-(s+\sigma)}{t-s}$ to tend to 0, τ must tend to $-(t - s)$.

We have seen that large d generate arbitrarily large τ and very negative d imply τ to get arbitrarily close to its lower bound $-(t - s)$. The ratio of average derivatives over the two intervals mentioned depends on d in a continuous manner (as does any normalized φ difference), also at $d = 1$. Hence, for each τ between its supremum and its infimum there exists a d that generates that τ .⁷

The proof for CADI discount functions is completely analogous, the only change being that all time points are replaced by their exponents. The exponent of $t + \sigma + \tau$ tending to infinity or to its infimum corresponds with τ doing so. \square

References

- Abdellaoui, Mohammed, Attema, Arthur E., Bleichrodt, Han, 2007. Intertemporal tradeoffs for gains and losses: An experimental measurement of discounted utility. iMTA/iBMG, Erasmus University, Rotterdam, The Netherlands.
- Attema, Arthur E., Bleichrodt, Han, Rohde, Kirsten I.M., Wakker, Peter P., 2008. Time-tradeoff sequences for analyzing discounting and time inconsistency. Economic Department, Erasmus University, Rotterdam, The Netherlands.
- Barsky, Robert B., Juster, F. Thomas, Kimball, Miles S., Shapiro, Matthew D., 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Quart. J. Econ.* 112, 537–579.
- Chesson, Harrell W., Viscusi, W. Kip, 2003. Commonalities in time and ambiguity aversion for long-term risks. *Theory Dec.* 54, 57–71.
- Ebert, Jane E.J., Prelec, Drazen, 2007. The fragility of time: Time-insensitivity and valuation of the near and far future. *Manage. Sci.* 53, 1423–1438.
- Fishburn, Peter C., Rubinstein, Ariel, 1982. Time preference. *Int. Econ. Rev.* 23, 677–694.
- Frederick, Shane, Loewenstein, George F., O'Donoghue, Ted, 2002. Time discounting and time preference: A critical review. *J. Econ. Lit.* 40, 351–401.
- Gigliotti, Gary, Sopher, Barry, 2004. Analysis of intertemporal choice: A new framework and experimental results. *Theory Dec.* 55, 209–233.
- Green, Leonard, Myerson, Joel, Ostaszewski, Pawel, 1999. Discounting of delayed rewards across the life span: Age differences in individual discounting functions. *Behavioural Processes* 46, 89–96.
- Harrison, Glenn W., Lau, Morten I., Williams, Melonie B., 2002. Estimating individual discount rates in Denmark: A field experiment. *Amer. Econ. Rev.* 92, 1606–1617.
- Harvey, Charles M., 1986. Value functions for infinite-period planning. *Manage. Sci.* 32, 1123–1139.
- Harvey, Charles M., 1995. Proportional discounting of future costs and benefits. *Math. Operations Res.* 20, 381–399.
- Herrnstein, R.J., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exper. Anal. Behav.* 4, 267–272.
- Kirby, Kris N., 1997. Bidding on the future: Evidence against normative discounting of delayed rewards. *J. Exper. Psych.: General* 126, 54–70.
- Krantz, David H., Luce, R. Duncan, Suppes, Patrick, Tversky, Amos, 1971. *Foundations of Measurement*, vol. I, Additive and Polynomial Representations. Academic Press, New York.
- Laibson, David I., 1997. Golden eggs and hyperbolic discounting. *Quart. J. Econ.* 112, 443–477.
- Loewenstein, George F., Prelec, Drazen, 1992. Anomalies in intertemporal choice: Evidence and an interpretation. *Quart. J. Econ.* 107, 573–597.
- Mazur, James E., 1987. An adjusting procedure for studying delayed reinforcement. In: Mazur, James E., Commons, Michael L., Nevin, John A., Rachlin, Howard (Eds.), *Quantitative Analyses of Behavior*, vol. 5: The Effect of Delay and of Intervening Events on Reinforcement Value. Erlbaum, Hillsdale, NJ, pp. 55–73.
- Miyamoto, John M., 1983. An axiomatization of the ratio/difference representation. *J. Math. Psych.* 27, 439–455.
- Onay, Selçuk, Öncüler, Ayse, 2007. Intertemporal choice under timing risk: An experimental approach. *J. Risk Uncertainty* 34, 99–121.
- Phelps, Edmund S., Pollak, Robert A., 1968. On second-best national saving and game-equilibrium growth. *Rev. Econ. Stud.* 35, 185–199.

⁷ It can be seen that the existing d is unique, because the ratio of average derivatives over the two intervals mentioned depends monotonically on d . This can be seen by dividing all derivatives by the derivative at t . In the interval $[s, t]$ the derivatives divided by that at t are increasing in d , and in the interval $[(s + \sigma), (t + \sigma + \tau)]$ they are decreasing. This monotonicity also shows that greater values of d lead to stronger decreasing impatience.

Prelec, Drazen, 1989. Decreasing impatience. Mimeo.

Prelec, Drazen, 2004. Decreasing impatience: A criterion for non-stationary time preference and ‘hyperbolic’ discounting. *Scand. J. Econ.* 106, 511–532.

Prelec, Drazen, Loewenstein, George F., 1991. Decision making over time and under uncertainty: A common approach. *Manage. Sci.* 37, 770–786.

Read, Daniel, Airoldi, Mara, Loewenstein, George, 2005. Intertemporal tradeoff priced in interest rates and amounts: A study of method variance. Working Paper.

Roskies, Ralph, 1965. A measurement axiomatization for an essentially multiplicative representation of two factors. *J. Math. Psych.* 2, 266–276.

Samuelson, Paul A., 1937. A note on measurement of utility. *Rev. Econ. Stud.* 4 (2), 155–161.

Sayman, Serdar, Öncüler, Ayşe, 2006. Reverse time-inconsistency. Working Paper.