University of Leicester

Department of Economics

Foundations and Properties of Time Discount Functions

Ali al-Nowaihi, University of Leicester

Sanjit Dhami, University of Leicester

Working Paper No. 13/27

November 2013

Foundations and Properties of Time Discount Functions

Ali al-Nowaihi* **Sanjit Dhami**†

9 November 2013

Abstract

A critical element in all discounted utility models is the specification of a discount function. We extend the standard model to allow for reference points for both outcomes and time. We consider the axiomatic foundations and properties of two main classes of discount functions. The first, the Loewenstein-Prelec discount function, accounts for declining impatience but cannot account for the evidence on subadditivity. A second class of discount functions, the Read-Scholten discount function accounts for declining impatience and subadditivity. We derive restrictions on an individual's preferences to expedite or to delay an outcome that give rise to the discount functions under consideration. As an application of our framework we consider the explanation of the common difference effect.

Keywords: Hyperbolic discounting, Subadditive and superadditive discounting; Prospect theory, γ-delay, α-*subadditivity*.

JEL Classification Codes: C60(General: Mathematical methods and programming); D91(Intertemporal consumer choice).

*Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522898. Fax: +44-116-2522908. E-mail: aa10@le.ac.uk.

†Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: Sanjit.Dhami@le.ac.uk.

1 Introduction

Consider a decision maker who, at time $t = 0$, takes an action that results in a dated outcome profile $w_1, w_2, ..., w_n$ at times $t_1, t_2, ..., t_n$, respectively. Thus, we have the sequence of outcome-time pairs: $(w_1, t_1), (w_2, t_2), ..., (w_n, t_n)$. Under *discounted-utility models* (DU) in which preferences are additively separable over time, the utility from the sequence of outcome-time pairs is given by

$$U = \Sigma_{i=1}^{n} u(w_i) D(0, t_i), \tag{1}$$

where u is the utility in any period (also known as felicity) and $D(0,t)$ is a discount function that discounts the utility at time $t \geq 0$ back to the present, time $t = 0$. A special case of the DU model is the *exponential discounted utility model* (EDU) in which $D(0,t) = e^{-\rho t}$, where $\rho > 0$ is the constant discount rate.

Extensive evidence indicates that all plant and animal life respond to changes in stimuli from a status-quo, or *reference level*. The *reference dependence* of human preferences is widely documented.[1] In a seminal paper, Loewenstein and Prelec (1992), henceforth LP, introduced reference dependence for outcomes in the DU model. They replaced the utility function, u, in (1) by the prospect theory utility function, v, due to Kahneman and Tversky (1979). v is defined over outcomes relative to a reference point, $x_i = w_i - w_0$, where w_0 is the reference outcome level.

Two properties of the discount function have received particular attention in the literature.

(a) *Degree of impatience*: Under EDU, for $0 \leq s < t$ and $\tau > 0$, $D(s,t) = D(s + \tau, t + \tau)$. This is also known as *stationarity of the discount factor* and arises from the discount factor being the same for intervals of identical lengths. By contrast, one often observes that as one shifts a future reward closer to the present, the *degree of impatience* increases in the following sense. Suppose that a decision maker reveals at time 0, indifference between the two outcome time pairs (w_1, s) and (w_2, t) so that $u(w_1) D(0,s) = u(w_2) D(0,t)$. But when both the rewards are shifted towards the current date by s periods, he prefers the earlier reward, i.e., $u(w_1) D(0,0) > u(w_2) D(0, t-s)$, where $D(0,0) = 1$. LP referred to this finding as the *common difference effect*. They proposed a general form of *hyperbolic discounting* through a discount function that we call the "*LP discount function*", which accounts for the common difference effect.[2] Under hyperbolic discounting

[1] Since the evidence is well known we refer the reader to Kahneman and Tversky (2000) and Loewenstein (1988), Loewenstein and Prelec (1992).

[2] The common difference effect can also be explained by the $\beta - \delta$ form of hyperbolic discounting (quasi-hyperbolic discounting) due to Phelps and Pollak (1968) and Laibson (1997).

one, therefore, allows for *declining impatience* as one moves into the future, unlike EDU.

(b) *Subadditivity of the discount function*: A very different explanation of the common difference effect arises from the work of Read (2001) and Scholten and Read (2006). Following the work of Read (2001) they argued that, for $r < s < t$, discounting a positive magnitude back from t to s then from s to r results in a lower present value than discounting back from t to r in one step, i.e., $D(r,s)D(s,t) < D(r,t)$.[3] These authors propose their own discount function, the "RS discount function" that can account for the common difference effect either through declining impatience, subadditivity or both. Thus, the RS discount function is the most general discount function available.

This sets the stage for an explanation of the threefold contribution of our paper.

Our first contribution is the introduction of a *reference time* in addition to the reference outcome level (Section 2). In LP, there is a reference outcome level but, by default, the time period 0 is the reference time to which all outcomes are discounted back. In many cases it may be desirable to have a reference time, $r > 0$. Consider the following two examples.

Example 1 *: At time* 0 *a business may be contemplating an investment in a new plant that will become operational at time* $t_1 > 0$*. In calculating the viability of the investment, the business may wish to compare the future profits and costs from the investment discounted back to time* t_1*, which then becomes the reference time.*

Example 2 *: At time* 0*, a student starting a three year degree may wish to use a government loan to finance the degree. The government may require loan repayments to be made only upon completion of the degree (as in the UK). The student may wish to calculate the future benefits and costs of this action, discounted back to time* t_1 *at which the degree is completed, which then becomes the reference time.*

Our second contribution is to provide axiomatic foundations of the LP and RS discount functions: LP gave axiomatic foundations for their LP discount function.[4] No axiomatic foundations have been given for the RS discount function. We first provide generalizations of the LP and RS discount functions taking account of reference time, $r \geq 0$.[5]

[3] The converse of this phenomena is superadditivity. Scholten and Read (2011a,b) report evidence for both subaddivity and superadditivity.

[4] See al-Nowaihi and Dhami (2006, 2008a).

[5] For $r = 0$ our generalization reduces to the standard forms of these functions.

We specify our axioms in terms of four important functions, the *generating function*, the *delay function*, the *speedup function* and the *seed function* (Section 4). The explanation of the common difference effect, additivity and declining/constant/increasing impatience can be stated in terms of the properties of these four functions.

The *delay function* is motivated by the following question. Suppose that the decision maker is indifferent between the outcome-time pairs $(x, 0)$ and (y, t), $t \geq 0$. If the receipt of x is delayed by s time periods, for what time period T will the decision maker be indifferent between (x, s) and (y, T)? By contrast, the *speedup function* is motivated by the following question. Suppose that the decision maker is indifferent between the outcome-time pairs (x, s) and (y, t), $0 < s \leq t$. If the receipt of x is brought forward by s periods, for what time period T will the decision maker be indifferent between $(x, 0)$ and (y, T)?

In their axiomatization, LP added the extra assumption of *linear delay* to that of the common difference effect. While there is considerable empirical evidence for the common difference effect, the assumption of linear delay is added purely for convenience. We extend the LP derivation as follows. At the most general level, which requires neither linear delay nor the common difference effect, we have our Representation Theorem 2 (Proposition 13, below) that gives a more general form of the LP discount function.

Our more general approach also allows us to derive the RS-discount function (Proposition 21, below). In particular, as in the RS discount function, we can explain the common difference effect as due to either declining impatience, subadditivity or a combination of both. We leave it to empirical evidence to select the correct explanation.

Our third contribution is to apply our framework to the understanding of the common difference effect in Section 8. We introduce a weaker notion of subadditivity, which we call α-subadditivity (Definition 11, below).[6] According to our Characterization Theorem 4 (Proposition 22, below), preferences exhibit the common difference effect if, and only if, α-subadditivity holds. We also introduce a generalization of the concept of linear delay of LP. We call this γ-delay. Our Proposition 19, below, shows that γ-delay implies the common difference effect. Imposing additivity, as well as γ-delay, gives our Proposition 20, below. The special case of the latter with $\gamma = 1$ gives the LP-discount function.

All proofs are contained in the appendix.

2 A reference-time theory of intertemporal choice (RT)

In this section we outline the *reference time theory* (RT) of al-Nowaihi and Dhami (2008b). Consider a decision maker who, at time t_0, takes an action that results in the outcome w_i

[6] This is not a property of discount functions and neither is it related to the property of subadditivity of the discount function.

at time t_i, $i = 1, 2, ..., n$, where

$$t_0 \leq r \leq t_1 < ... < t_n. \tag{2}$$

Time r is the *reference time*. It is the time back to which all values are to be discounted, using a discount function, $D(r, t)$; r need not be the same as t_0 (recall Examples 1, 2). Without loss of generality, we normalize the time at which the decision is made to be $t_0 = 0$.

We assume that the decision maker has a *reference outcome level*, w_0, relative to which all outcomes are to be evaluated using the prospect theory utility function, $v(x_i)$, introduced by Kahneman and Tversky (1979), where $x_i = w_i - w_0$.

The utility function, v, has four main properties: *reference dependence*, *monotonicity*, *declining sensitivity*, and *loss aversion*. There is good empirical support for these features; see, for instance, Kahneman and Tversky (2000). In particular, v satisfies:

$$v : (-\infty, \infty) \to (-\infty, \infty) \text{ is strictly increasing (monotonicity)}, \tag{3}$$

$$v(0) = 0 \text{ (reference dependence)}, \tag{4}$$

$$\text{For } x > 0\text{:} \quad -v(-x) > v(x) \text{ (loss aversion)}, \tag{5}$$

$$v \text{ is concave for gains, } x \geq 0 \text{, but convex for losses, } x \leq 0 \text{ (declining sensitivity)}. \tag{6}$$

Furthermore, it is assumed that

$$v \text{ is continuous, and twice differentiable except at 0.} \tag{7}$$

For each *reference outcome* and *reference time* pair $(w_0, r) \in (-\infty, \infty) \times [0, \infty)$, the decision maker has a complete and transitive preference relation, $\preceq_{w_0,r}$ on $(-\infty, \infty) \times [r, \infty)$ given by

$$(w_1, t_1) \preceq_{w_0,r} (w_2, t_2) \Leftrightarrow v(w_1 - w_0) D(r, t_1) \leq v(w_2 - w_0) D(r, t_2). \tag{8}$$

Let $\mathbf{S}$ be a non-empty set of outcome-time sequences from $(-\infty, \infty) \times [0, \infty)$ of the form $(x_1, t_1), (x_2, t_2), ..., (x_i, t_i), ...$ Using (8), we extend $\preceq_{w_0,r}$ to a complete transitive preference relation on sequences in $\mathbf{S}$, as follows[7]:

$$\begin{aligned} &((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \preceq_{w_0,r} ((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)) \\ &\Leftrightarrow \Sigma_{i=1}^{m} v(x_i) D(r, t_i) \leq \Sigma_{i=1}^{n} v(y_i) D(r, t_i). \end{aligned} \tag{9}$$

Thus, the decision maker's intertemporal utility function is given by:

$$V_r((w_1, t_1), (w_2, t_2), ..., (w_n, t_n), w_0) = \Sigma_{i=1}^{n} v(x_i) D(r, t_i), \tag{10}$$

[7]The following also holds for infinite sequences, provided the sums in (9) converge.

Comparing (10) to the utility function proposed by LP, and successfully employed to explain several anomalies in intertemporal choice, we see that LP implicitly set $r = 0$. Also, LP implicitly assume that the discount function is additive (Definition 2, below), in which case there is no loss in assuming $r = 0$. To accommodate the empirical evidence, however, we allow the discount function to be non-additive, in which case the choice of reference time does matter.

2.1 Determination of the reference time

We now consider the determination of the reference point for time. Let

$$T = \{t \in [0, \infty) : t = t_i \text{ for some sequence } \{(x_1, t_1), (x_2, t_2), ..., (x_i, t_i), ...\} \text{ in } \mathbf{S}\}. \quad (11)$$

Since T is bounded below by 0 and non-empty, it follows that T has a greatest lower bound, r. We make the following assumption for reference time.

A0: Given S, T, r, as described just above, the decision maker takes r (i.e., the greatest lowest bound of T) as the reference point for time.

Example 3 *: Suppose that a decision maker wants to compare (x, s) with (y, t), $s \leq t$, then $\mathbf{S} = \{(x, s), (y, t)\}$ and $T = \{s, t\}$. Thus, A0 implies that $r = s$. If $v(x) < v(y) D(s, t)$ then the decision maker chooses (y, t) over (x, s).*

A0 should be regarded as a tentative assumption, whose implications are to be explored. Its motivation comes from the status-quo justification of reference points; see, for instance, Kahneman and Tversky (2000).

3 Discount functions and their properties

We now give a formal definition of a *discount function*.

Definition 1 *(Discount functions): Let*

$$\Delta = \{(r, t) \in \mathbb{R} \times \mathbb{R} : 0 \leq r \leq t\}. \quad (12)$$

A discount function is a mapping, $D : \Delta \to (0, 1]$, satisfying:
(a) For each $r \in [0, \infty)$, $D(r, t)$ is a strictly decreasing function of $t \in [r, \infty)$ into $(0, 1]$ with $D(r, r) = 1$.
(b) For each $t \in [0, \infty)$, $D(r, t)$ is a strictly increasing function of $r \in [0, t]$ into $(0, 1]$.
(c) Furthermore, if D satisfies (a) with 'into' replaced with 'onto', then we call D a continuous discount function.

Outcomes that are further out into the future are less salient, hence, they are discounted more, thus, $D(r,t)$ is strictly decreasing in t. For a fixed t, if the reference point, r, becomes closer to t then an outcome at time t, when discounted back to r, is discounted less. Hence, $D(r,t)$ is strictly increasing in r.

In theories of time discounting that do not have a reference time, the discount functions are stated under the implicit assumption that $r = 0$. Hence, we first need to restate the main discount functions for the case $r > 0$. We extend four common discount functions to the case $r > 0$. The standard versions of these functions can simply be obtained by setting $r = 0$; the reason for the choice of the acronymns corresponding to these functions will become clear below.

$$\textbf{Exponential}\text{: } D(r,t) = e^{-\beta(t-r)},\ \beta > 0. \tag{13}$$

$$\textbf{PPL}\text{: } D(r,t) = \left\{ \begin{array}{ccc} 1 & \text{when} & r = t = 0 \\ e^{-(\delta+\beta t)} & \text{when} & r = 0, t > 0 \\ e^{-\beta(t-r)} & \text{when} & 0 < r \le t \end{array} \right\},\ \beta > 0,\ \delta > 0. \tag{14}$$

$$\textbf{LP}\text{: } D(r,t) = \left(\frac{1+\alpha t}{1+\alpha r} \right)^{-\frac{\beta}{\alpha}},\ t \ge 0,\ r \ge 0,\ \alpha > 0,\ \beta > 0. \tag{15}$$

$$\textbf{RS}\text{: } D(r,t) = \left[1 + \alpha \left(t^{\tau} - r^{\tau} \right)^{\rho} \right]^{-\frac{\beta}{\alpha}},\ 0 \le r \le t,\ \alpha > 0,\ \beta > 0,\ \rho > 0,\ \tau > 0. \tag{16}$$

In addition, we propose the following generalization of (16),

$$\begin{aligned} \textbf{Generalized RS}\text{:} \qquad & D(r,t) = e^{-Q(\phi(t)-\phi(r))},\ 0 \le r \le t, \\ Q \ : \ & [0,\infty) \overset{\text{onto}}{\rightarrow} [0,\infty) \text{ is strictly increasing}, \\ \phi \ : \ & [0,\infty) \overset{\text{onto}}{\rightarrow} [0,\infty) \text{ is strictly increasing}. \end{aligned} \tag{17}$$

The *exponential discount function* (13) was introduced by Samuelson (1937). The main attraction of EDU is that it is the unique discount function that leads to time-consistent choices. The $\beta - \delta$ or *quasi-hyperbolic discount function* (14) was proposed by Phelps and Pollak (1968) and Laibson (1997) and is popular in applied work (we use the acronym PPL for it).[8] The *generalized hyperbolic discount function* (15) was proposed by Loewenstein and Prelec (1992) (we use the acronymn LP for it). These three discount functions are *additive* (Definition 2, below). They can account for the common difference effect through declining impatience (Definition 3, below) but they cannot account for either non-additivity or intransitivity. The interval discount function (16) was introduced by Read (2001) and Scholten and Read (2006) (we use the acronym RS for it). It can account

[8] It can be given the following psychological foundation. The decision maker essentially uses exponential discounting. But in the short run is overcome by *visceral influences* such as temptation or procrastination; see for instance Loewenstein et al. (2001).

for both non-additivity and intransitivity. It can account for the common difference effect though declining impatience, subadditivity or a combination of both.

Note that (15) approaches (13) as $\alpha \to 0$. In general, neither of (15) or (16) is a special case of the other. However, for $r = 0$ (and only for $r = 0$), (16) reduces to (15) when $\rho = \tau = 1$. While ρ, τ are parameters, r is a variable. Hence, neither discount function is a special case of the other.[9]

Our terminology suggests that a continuous discount function (Definition 1(c)) is continuous. That this is partly true, is established by the following Proposition.

Proposition 1 : *A continuous discount function, $D(r,t)$, is continuous in t.*

It is straightforward to check that each of (13), (15), (16) and (17) is a continuous discount function in the sense of Definition 1. It is also straightforward to check that (14) is a discount function. The reason that the latter is not a *continuous* discount function is that $\lim_{t\to 0+} D(0,t) = e^{-\delta} < 1 = D(0,0)$.

From (15) and (16) we see that the restrictions $r \geq 0$ and $t \geq 0$ are needed. From (16) we see that the further restriction $r \leq t$ is needed.[10] From (13) we see that the 'into' in Definition 1(b) cannot be strengthened to 'onto'.

Proposition 2 *(Time sensitivity): Let D be a continuous discount function. Suppose $r \geq 0$. If $0 < x \leq y$, or if $y \leq x < 0$, then $v(x) = v(y) D(r,t)$ for some $t \in [r, \infty)$.*[11]

Proposition 3 *(Existence of present values): Let D be a discount function. Let $r \leq t$ and $y \geq 0$ $(y \leq 0)$. Then, for some x, $0 \leq x \leq y$ $(y \leq x \leq 0)$, $v(x) = v(y) D(r,t)$.*

Definition 2 *(Additivity): A discount function, $D(r,t)$, is*

$$\begin{cases} \textit{Additive if} & D(r,s)\,D(s,t) = D(r,t), \quad \textit{for } r \leq s \leq t, \\ \textit{Subadditive if} & D(r,s)\,D(s,t) < D(r,t), \quad \textit{for } r < s < t, \\ \textit{Superadditive if} & D(r,s)\,D(s,t) > D(r,t), \quad \textit{for } r < s < t. \end{cases}$$

In Definition 2, additivity implies that discounting a quantity from time t back to time s and then further back to time r is the same as discounting that quantity from time t back to time r in one step. In other words, breaking an interval into subintervals has no effect on discounting. However, in the other two cases, it does have an effect. Under subadditive

[9] Scholten and Read (2006a) report incorrectly that the LP-discount function is a special case of the RS-discount function.

[10] One alternative is to *define* $D(t,s)$ to be $1/D(s,t)$. But we do not know if people, when compounding forward, use the inverse of discount function (as they should, from a normative point of view). Fortunately, we have no need to resolve these issues in this paper.

[11] We have chosen the phrase 'time sensitivity' to conform with the terminology of Ok and Masatlioglu (2007), Axiom A1, p219, and Claim 3, p235.

discounting, there is more discounting over the subdivided intervals (future utilities are shrunk more), while the converse is true under superadditive discounting.[12]

Example 4 *: Consider the following example that illustrates the importance of subadditive discounting. Suppose that time is measured in years. Consider three dates,* $0 \leq r < s < t$ *and the following sequential financial investment opportunity. The investor is given a choice to invest* $\$a$*, at date* r *for the next* $s-r$ *years in return for a promised receipt of* $\$b$ *at date* s*. At date* s *the amount* $\$b$ *is automatically invested for the next* $t-s$ *years, so at date* t *the investor receives an amount* $\$c$*, which we assume equals* $\frac{\$a}{D(r,t)}$*. Suppose also that at date* r *the investor has no other alternative use of his funds for the next* $t-r$ *years. Should the investor take up this opportunity?*
There are the following two kinds of investors whose behavior under additive discounting is identical. Investor 1 *discounts over the whole interval* $t-r$ *so he invests if* $a \geq cD(r,t)$*. Since* $\$c = \frac{a}{D(r,t)}$*, investor 1 invests. By contrast, investor 2 first discounts the final amount from time* t *to* s *and then back from time* s *to* r *and. In order to illustrate the contrast with additive discounting, suppose that investor 2 uses subadditive discounting. Thus, for investor 2,* $D(r,s)\,D(s,t) < D(r,t) = \frac{a}{c}$ *so*

$$a > cD(r,s)\,D(s,t)\,,$$

thus, he does not accept the financial investment opportunity.

Example 5 *Consider the RS discount function in (16),* $D(r,t) = [1+\alpha(t^{\tau}-r^{\tau})^{\rho}]^{-\frac{\beta}{\alpha}}$*, with* $\tau = 1$*,* $\rho = 0.6$*,* $\alpha = \beta = 0.5$*. Let the utility function be* $v(x) = x$ *for* $x \geq 0$*. We can then compute*

$$D(0,4) = \left(1+0.5(4)^{0.6}\right)^{-1},\ D(4,10) = \left(1+0.5(6)^{0.6}\right)^{-1},\ D(0,10) = \left(1+0.5(10)^{0.6}\right)^{-1},$$

which, respectively, equal 0.46540, 0.40567, 0.33439*. As expected,* $D(0,t)$ *is declining in* t *and* $D(r,10)$ *is increasing in* r *(see Definition 1). Let the unit of time be one month. Then a decision maker who discounts over the entire interval* $[0,10]$ *is indifferent between receiving \$15 in 10 months time and \$5.0159 at date* $r=0$ *because*

$$5.0159 = 15D(0,10)\,.$$

However, unlike the additivity assumption made in the exponential, PPL and LP discount functions, we have

$$D(0,4)\,D(4,10) = 0.1888 < D(0,10) = 0.33439.$$

[12]For empirical evidence on subadditive discounting, see Read (2001).

It follows that \$15 discounted back to time $r = 0$ using the two subintervals $(0, 4)$ and $(4, 10)$ lead to a lower present value $15D(0,4)D(4,10) = 2.832 < 5.0159$. Thus, under subadditive discounting, two different decision makers, one who discounts over the full interval and another who discounts over two subintervals, may make very different decisions.

We now formalize the sense in which an individual may exhibit various degrees of impatience. The basic idea is to shift a time interval of a given size into the future and observe if this leads to a smaller, unchanged or larger discounting of the future.

Definition 3 *(Impatience): A discount function, $D(r,s)$, exhibits*[13]

$$\begin{cases} \textit{declining impatience if} & D(r,s) < D(r+t, s+t), \quad \textit{for } t > 0 \textit{ and } r < s, \\ \textit{constant impatience if} & D(r,s) = D(r+t, s+t), \quad \textit{for } t \geq 0 \textit{ and } r \leq s, \\ \textit{increasing impatience if} & D(r,s) > D(r+t, s+t), \quad \textit{for } t > 0 \textit{ and } r < s. \end{cases}$$

Exponential discounting exhibits constant impatience while hyperbolic discounting exhibits declining impatience. The RS-discount function (16) allows for additivity, subadditivity and all the three cases in Definition 3, hence, it is of great practical importance. The next definition describes the properties of this function.

Proposition 4 *: Let $D(r,t)$ be the RS-discount function (16), then:*
(a) If $0 < \rho \leq 1$, then D is subadditive.
(b) If $\rho > 1$, then D is neither subadditive, additive nor superadditive.
(ci) If $0 < \tau < 1$, then D exhibits declining impatience.
(cii) If $\tau = 1$, then D exhibits constant impatience.
(ciii) If $\tau > 1$, then D exhibits increasing impatience.

In the light of Proposition 4, we can now see the interpretation of the parameters τ and ρ in the RS-discount function (16).[14] τ controls impatience, independently of the values of the other parameters α, β and ρ. $0 < \tau < 1$, gives declining impatience, $\tau = 1$ gives constant impatience and $\tau > 1$ gives increasing impatience. If $0 < \rho \leq 1$, then we get subadditivity, irrespective of the values of the other parameters α, β and τ. However, if $\rho > 1$, then (16) can be neither subadditive, additive nor superadditive[15].

[13] Some authors use 'present bias' for what we call 'declining impatience'. But other authors use 'present bias' to mean that the discount function, $D(s,t)$ is declining in t. So we prefer 'declining impatience' to avoid confusion. It is common to use 'stationarity' for what we call 'constant impatience'. We prefer the latter, to be in conformity with 'declining impatience' and 'increasing impatience'.

[14] Scholten and Read (2006), bottom of p1425, state: $\alpha > 0$ implies subadditivity (incorrect), $\rho > 1$ implies superadditivity (incorrect) and $0 < \tau < 1$ implies declining impatience (correct but incomplete).

[15] In this case, depending on the particular values of r, s and t, we may have $D(r,s) < D(r+t, s+t)$, $D(r,s) = D(r+t, s+t)$ or $D(r,s) > D(r+t, s+t)$).

4 Some special functions and their properties

In order to explore the foundations of some of the popular discount functions we now introduce a set of functions. This section explains the motivation for the functions and outlines some of their basic properties and their relation with the interval discount function. These are intermediate results in our paper but are also of independent interest.

4.1 Generating Function, φ

From Definition 1, we know that for a fixed r, $D(r,t)$ is a strictly decreasing function of t and for a fixed t, $D(r,t)$ is a strictly increasing function of r (see also Example 5). So we may wonder if there exists a decreasing function, φ, which under 'certain restrictions' (see Definition 4) allows us to write $D(r,t) = \varphi(t)/\varphi(r)$? If yes, then what are the properties of such a discount function? For instance, is it additive or subadditive (see Proposition 5)?

Definition 4 *(The generating function): Let* $\varphi : [0,\infty) \to (0,1]$ *be a strictly decreasing function with* $\varphi(0) = 1$*. Then we call* φ *a generating function. If, in addition,* φ *is onto, we call* φ *a continuous generating function.*

A 'continuous generating function' is continuous. The proof is the same as that of Proposition 1 and, therefore, will be omitted.

Proposition 5 *: (a) Let* $D(r,t) = \frac{\varphi(t)}{\varphi(r)}$ *for some strictly decreasing real valued function,* $\varphi : [0,\infty) \to (0,1]$*. Then the following hold:*
(i) D *is an additive discount function (see Definition 2).*
(ii) If $\varphi(0) = 1$ *(so that* φ *is a generating function), then* $D(0,t) = \varphi(t)$*.*
(iii) If φ *is onto (so that* φ *is a continuous generating function), then* D *is an additive continuous discount function and* $D(0,t) = \varphi(t)$*.*
(b) Let D *be an additive discount function. Then the following hold:*
(i) $D(r,t) = \frac{\varphi(t)}{\varphi(r)}$ *for some strictly decreasing real valued function,* $\varphi : [0,\infty) \to (0,1]$ *with* $\varphi(0) = 1$ *(hence,* φ *is a generating function).*
(ii) If D *is a continuous discount function, then* φ *is onto (hence,* φ *is a continuous generating function).*
(iii) $D(0,t) = \varphi(t)$*.*

Proposition 5 justifies the following definition.

Definition 5 *(Additive extensions): Let* φ *be a generating function. Let* $D(r,t) = \frac{\varphi(t)}{\varphi(r)}$*. Then:*
(i) We call φ *the generating function of the additive discount function,* D*.*
(ii) We call D *the additive extension of* φ*.*
(iii) We also refer to $D(r,t)$ *as the additive extension of* $D(0,t)$*.*

Proposition 6 *: The discount functions (13), (14) and (15) are additive. In each case,* $D(r,t)$ *is the additive extension of* $D(0,t)$ *and* $\varphi(t) = D(0,t)$ *is the generating function for* $D(r,t)$*. However, (16) is not additive. (17) is additive if, and only if,* $Q[w(t) - w(r)] = Q[w(t)] - Q[w(r)]$*, in which case* $D(r,t) = e^{Q[w(r)]-Q[w(t)]}$ *is the additive extension of* $D(0,t) = e^{-Q[w(t)]}$ *and* $e^{-Q[w(t)]}$ *is the generating function.*

4.2 Delay Function, Ψ

Let the reference time be $r = 0$. Suppose that a decision maker reveals the following indifference: x received at time 0 is equivalent to y received at time t, thus,

$$v(x) = v(y) D(0,t). \tag{18}$$

Now suppose that the receipt of x is *delayed* to time s. We ask, at what time, T, will y received at time T be equivalent to x received at time s, i.e., for what T does the following hold?

$$v(x) D(0,s) = v(y) D(0,T). \tag{19}$$

Let us conjecture that T depends on s,t through a functional relation, say, $T = \Psi(s,t)$ where $\Psi(s,t)$ is a function that maps from $\mathbb{R}^2$ to $\mathbb{R}$. From (18), (19) we get that $\Psi(s,t)$ must satisfy

$$D(0,s) D(0,t) = D(0, \Psi(s,t)). \tag{20}$$

We shall call the function $\Psi(s,t)$, if it exists, a delay function (see Definition 6). This situation is shown in the upper panel of Figure 1. For the exponential discount function (13), the answer is clear: $\Psi(s,t) = s + t$. More generally, we show that such a delay function exists, is unique and depends on s,t. We shall also examine its properties (see Propositions 7, 8).

Definition 6 *(Delay functions): Let* D *be a discount function. Suppose that the function,* Ψ*, has the property* $D(0,s) D(0,t) = D(0, \Psi(s,t))$*,* $s \geq 0$*,* $t \geq 0$*. Then we call* Ψ *a delay function corresponding to the discount function,* D*. We also say that the discount function,* D*, exhibits* Ψ*-delay.*

Proposition 7 *(Properties of a delay function): Let* D *be a discount function and* Ψ *a corresponding delay function. Then* Ψ *has the following properties:*
(a) Ψ *is unique,*
(b) $\Psi(s,t)$ *is strictly increasing in each of* s *and* t*,*
(c) $\Psi(s,t) = \Psi(t,s)$*,*
(d) $\Psi(0,t) = \Psi(t,0) = t$*,*
(e) $v(x) = v(y) D(0,t)$ *if, and only if,* $v(x) D(0,s) = v(y) D(0, \Psi(s,t))$*.*

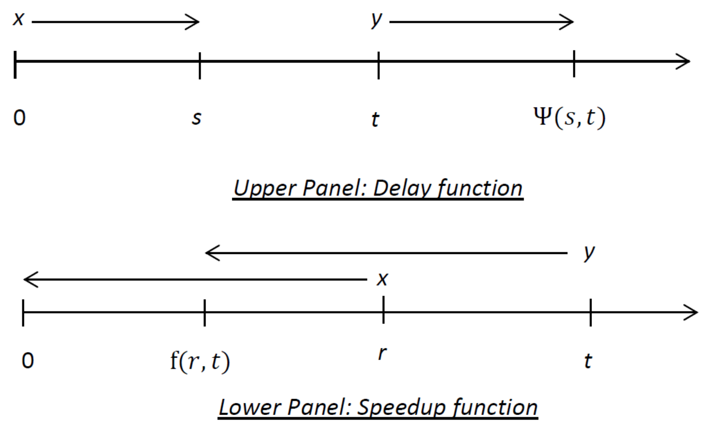

Figure 1: A diagrammatic illustration of the delay and speedup functions

Proposition 8 *(Existence of a delay function): A continuous discount function has a unique delay function.*

4.3 Speedup function, f

Suppose that x received at time r is equivalent to y received at time t, $0 \leq r \leq t$, time r being the reference time; so that (recall that $D(r,r) = 1$)

$$v(x) = v(y) D(r,t). \tag{21}$$

Suppose that the receipt of x is brought forward from time r to time 0, where time 0 is the new reference time. We ask, at what time, T, will y received at time T be equivalent to x received at time 0? Or, for what time, T, will the following hold?

$$v(x) = v(y) D(0,T). \tag{22}$$

For the exponential discount function (13) the answer is clear: $T = t - r$. More generally, let us conjecture that T depends on r, t, so that we can write $T = f(r,t)$ where $f : \Delta \to [0,\infty)$ will be called the *speedup function*. This situation is shown in the lower panel of Figure 1. Definition 7 formally defines a speedup function.

Definition 7 *Let $f : \Delta \to [0,\infty)$ satisfy:*
(a) For each $r \in [0,\infty)$, $f(r,t)$ is a strictly increasing function of $t \in [r,\infty)$ into $[0,\infty)$, with $f(r,r) = 0$.
(b) For each $t \in [0,\infty)$, $f(r,t)$ is a strictly decreasing function of $r \in [0,t]$ into $[0,t]$, with $f(0,t) = t$.
Then we call f a speedup function. If, in (a), 'into' is replaced with 'onto', then we call f a continuous speedup function.

A 'continuous speedup function', $f(r,t)$, is continuous in t. The proof is the same as that of Proposition 1 and, therefore, will be omitted. From (21), (22) we get that

$$D(0, f(r,t)) = D(r,t). \tag{23}$$

Since $f(r,r) = 0$ we get from (23) that $D(0,0) = D(r,r) = 1$. These observations motivate the next definition.

Definition 8 *Let D be a discount function. Let $f : \Delta \to [0,\infty)$ satisfy $D(0, f(r,t)) = D(r,t)$, then*
(a) we call f a speedup function corresponding to D,
(b) we refer to $D(r,t)$ as an f-extension of $D(0,t)$, or just an f-extension.

Definition 7 defines speedup functions independently of any discount function. By contrast, Definition 8 defines a speedup function corresponding to a give discount function. Our terminology suggests that 'a speedup function corresponding to a given discount function' in the sense of Definition 8 is, in fact, 'a speedup function' in the sense of Definition 7. That this is indeed the case, is established in the following proposition.

Proposition 9 *: Let D be a discount function. Let f be a corresponding speedup function in the sense of Definition 8(a). Then:*
(a) f is unique.
(b) f is a speedup function in the sense of Definition 7.
(c) $v(x) = v(y) D(r,t)$ if, and only if, $v(x) = v(y) D(0, f(r,t))$.

Proposition 10 *Let the discount function, D, be continuous. Then there exists a speedup function, f, corresponding to D in the sense of Definition 8. Moreover, f is unique and is a continuous speedup function in the sense of Definition 7.*

Proposition 11 *: Let φ be a generating function and f a speedup function. Then:*
(a) $D(r,t) = \varphi(f(r,t))$ is a discount function.
(b) f is the speedup function corresponding to D and $D(r,t)$ is the f-extension of $D(0,t)$.
(c) If φ is a continuous generating function and f a continuous extension function, then D is a continuous discount function.

To summarize, given a generating function, φ, and a speedup function, f, by Proposition 11, we can construct a discount function D so that $D(r,t)$ is the f-extension of $\varphi(t) = D(0,t)$. Proposition 10 tells us that *all* continuous discount functions are obtainable in this way from continuous generating functions and continuous speedup functions.

4.4 Seed Function, σ

Our final special function is a *seed function*, which we define next. The rationale for the seed function is examined in subsequent sections. For instance, it will turn out that the delay function can be expressed uniquely in terms of the seed function.

Definition 9 *(The seed function): Let $\sigma : [0,\infty) \to [0,\infty)$ be strictly increasing, with $\sigma(0) = 0$. We call σ a seed function. If, in addition, σ is onto, we call σ a continuous seed function.*

A 'continuous seed function' is continuous. The proof is similar to that of Proposition 1 and, therefore, will be omitted.

5 Representation theorems

The following definition gives a useful representation for discount functions.

Definition 10 *(Representations): Let $\alpha > 0$ and $\beta > 0$. We call*

$$D(0,t) = [1 + \alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$$

an (α,β)-representation of the discount function $D(r,t)$.

Proposition 12, below, establishes the existence of (α,β)-representations for continuous discount functions and shows their connection to delay and seed functions.

Proposition 12 *(Representation Theorem 1): Let D be a continuous discount function. Then, for each $\alpha > 0$ and each $\beta > 0$, D has the unique (α,β)-representation $D(0,t) = [1 + \alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$. Moreover, the function σ has the properties:*
(a) $\sigma : [0,\infty) \overset{onto}{\to} [0,\infty)$ is strictly increasing. Hence $\sigma(0) = 0$ and σ is a continuous seed function.
(b) σ^{-1} exists and $\sigma^{-1} : [0,\infty) \overset{onto}{\to} [0,\infty)$ is strictly increasing with $\sigma^{-1}(0) = 0$.
(c) The delay function can be expressed in terms of the seed function

$$\Psi(s,t) = \sigma^{-1}[\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)]. \tag{24}$$

From Proposition 12, we see that if Ψ is to be the delay function of some continuous discount function, then it must satisfy (24). In the light of this, when considering possible delay functions, we can restrict ourselves, without loss of generality, to the class of functions of the form (24), where σ is a continuous seed function.

The next proposition is a generalization of LP's derivation of their generalized hyperbolic discount function. Its motivation can be explained as follows. According to Propositions 7(a) and 8, a continuous discount function determines a unique delay function. Hence, we can partition the set of all continuous discount functions into equivalence classes. Two continuous discount functions being in the same equivalence class if, and only if, they have the same delay function. However, many different discount functions could have the same delay function. In fact, according to Representation Theorem 1 (Proposition 12), all (the different) continuous discount functions, D, for which $D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$ (fixed α and σ, different β's) have the same delay function and, hence, lie in the same equivalence class. But does an equivalence class contain other continuous discount functions? Representation Theorem 2 (Proposition 13) that we derive next, gives the answer in the negative. Consider an arbitrary equivalence class. Choose some member of that class. Let it have the (α,β)-representation $D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$. Then *all* members of its class can be obtained by varying β, keeping α and σ fixed.

Proposition 13 *(Representation Theorem 2): Let $\alpha > 0$ and let σ be a continuous seed function. Let D be a continuous discount function with delay function, Ψ, given by (24). Then, for some $\beta > 0$, $D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$.*

In the standard sense, a function, σ, is *subadditive* if, for all s and t for which σ is defined: $\sigma(s+t) \leq \sigma(s) + \sigma(t)$. We now introduce the concept of an α-subadditive function.

Definition 11 *(α-subadditivity): Let $\alpha > 0$. A function, σ, is α-subadditive if, for all s and t for which σ is defined and non-zero: $\sigma(s+t) < \sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)$.*

A function that is α-subadditive, for some $\alpha > 0$, need not be subadditive. However, a function is subadditive if, and only if, it is α-subadditive for all $\alpha > 0$.

Remark 1 *: Consider the exponential discount function, $D(r,t) = e^{-\beta(t-r)}$, $\beta > 0$. Then $\ln D(0,t)$ is additive in this sense. And, of course, $D(r,t)$ is additive in the sense of Definition 2. Also note that α-subadditivity, as in Definition 11, neither implies nor is implied by subadditivity of the discount function, as in Definition 2.*

Proposition 14 *(Representation Theorem 3): Let D be a continuous discount function. Let $\alpha > 0$ and $\beta > 0$. Let $D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$ be an (α,β)-representation of D. Then the following are equivalent:*
(a) σ is α-subadditive continuous seed function.
(b) Let Ψ be the corresponding delay function. If $s > 0$ and $t > 0$, then $\Psi(s,t) > s+t$.
(c) If $s > 0$ and $t > 0$, then $D(0,s)D(0,t) < D(0,s+t)$.

From (20), we have $D(0,s)\,D(0,t) = D(0,\Psi)$ where $D(0,t)$ is decreasing in t. Hence, if $\Psi(s,t) > s+t$ then Proposition 14(c) follows immediately. The relationship of Proposition 14(a) with the other parts of the proposition follows from the (α,β)-representation $D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$.

6 Characterization theorems

We now combine the representation theorems (Propositions 12, 13 and 14) with the results on our special functions in subsection 4 to produce characterization theorems. The next proposition shows that continuous discount functions can be uniquely expressed in terms of their seed and speedup functions.

Proposition 15 *(Characterization Theorem 1): D is a continuous discount function if, and only if, $D(r,t) = [1+\alpha\sigma(f(r,t))]^{-\frac{\beta}{\alpha}}$, $\alpha > 0$, $\beta > 0$, where σ is a continuous seed function and f is a continuous speedup function. Furthermore, f is uniquely determined by D.*

Definition 3 above, expressed the important concepts of constant, declining and increasing impatience in terms of the discount function. The next proposition shows that these concepts can also be expressed in terms of the speedup function.

Proposition 16 *(Characterization Theorem 2): A continuous discount function with the speedup function, f, exhibits:*
(a) declining impatience if, and only if, $f(r,s) > f(r+t,s+t)$, for $t > 0$ and $r < s$,
(b) constant impatience if, and only if, $f(r,s) = f(r+t,s+t)$, for $t \geq 0$ and $r \leq s$,
(c) increasing impatience if, and only if, $f(r,s) < f(r+t,s+t)$, for $t > 0$ and $r < s$.

Proposition 17 *(Characterization Theorem 3): A continuous discount function, D, with speedup function f and seed function σ, is additive if, and only if, $f(r,t) = \sigma^{-1}\left(\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right)$, in which case $D(r,t) = \left[\frac{1+\alpha\sigma(t)}{1+\alpha\sigma(r)}\right]^{-\frac{\beta}{\alpha}}$.*

Proposition 18 *: The following two Tables give a seed function, σ, the generating function, φ, the speedup function, f, and the delay function, Ψ, of each of the discount functions $D(r,t)$, given in (13) to (16).*

The discount functions for the various parameters used in Table 1 are shown in Table 2.

	$\sigma(t)$	$\varphi(t)$	$f(r,t)$	$\Psi(s,t)$
Exponential	$\frac{e^{\alpha t}-1}{\alpha}$	$e^{-\beta t}$	$t-r$	$s+t$
LP	t	$(1+\alpha t)^{-\frac{\beta}{\alpha}}$	$\frac{t-r}{1+\alpha r}$	$s+t+\alpha st$
RS	$t^{\tau\rho}$	$[1+\alpha t^{\tau\rho}]^{-\frac{\beta}{\alpha}}$	$(t^{\tau}-r^{\tau})^{\frac{1}{\tau}}$	$[s^{\tau\rho}+t^{\tau\rho}+\alpha(st)^{\tau\rho}]^{\frac{1}{\tau\rho}}$
Generalized RS	$\frac{e^{\frac{\alpha}{\beta}Q[w(t)]}-1}{\alpha}$	$e^{-Q[w(t)]}$	$w^{-1}[w(t)-w(r)]$	$w^{-1}Q^{-1}[Q(w(s))+Q(w(t))]$
PPL $(r=t=0)$	0	1	0	s
PPL $(0=r<t)$	$\frac{e^{\frac{\alpha\delta}{\beta}+\alpha t}-1}{\alpha}$	$e^{-(\delta+\beta t)}$	t	$\frac{\delta}{\beta}+s+t$
PPL $(0<r\leq t)$	$\frac{e^{\alpha t}-1}{\alpha}$	$e^{-\beta t}$	$t-r$	N.A.

Table 1: Special functions corresponding to various discount functions

	$D(r,t)$
Exponential	$e^{-\beta(t-r)}$
LP	$\left(\frac{1+\alpha t}{1+\alpha r}\right)^{-\frac{\beta}{\alpha}}$
RS	$[1+\alpha(t^{\tau}-r^{\tau})^{\rho}]^{-\frac{\beta}{\alpha}}$
Generalized RS	$e^{-Q[w(t)-w(r)]}$
PPL $(r=t=0)$	1
PPL $(0=r<t)$	$e^{-(\delta+\beta t)}$
PPL $(0<r\leq t)$	$e^{-\beta(t-r)}$

Table 2: Various discount functions for alternative parameter values

Starting with a continuous seed function, σ, an $\alpha>0$ and a $\beta>0$, we can 'grow' from them a unique generating function, $\varphi(t)=[1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$ (which turns out to be continuous). Given this generating function and a continuous speedup function, $f(r,t)$, we obtain a unique discount function $D(r,t)=\varphi(f(r,t))=[1+\alpha\sigma(f(r,t))]^{-\frac{\beta}{\alpha}}$ (which also turns out to be continuous). This discount function determines a unique delay function, $\Psi(s,t)=\sigma^{-1}[\sigma(s)+\sigma(t)+\alpha\sigma(s)\sigma(t)]$.

Conversely, a continuous discount function, D, determines a unique (continuous) generating function, $\varphi(t)=D(0,t)$ and a unique (continuous) speedup function, f, so that D is the f-extension of φ: $D(r,t)=\varphi(f(r,t))$.

Although a continuous discount function, D, determines unique generating, speedup and delay functions, φ, f and Ψ, it does *not* determine unique α, β or σ in the representation $D(0,t)=[1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$. For example, the LP-discount function $D(r,t)=(1+\alpha r)^{\frac{\beta}{\alpha}}(1+\alpha t)^{-\frac{\beta}{\alpha}}$, $\alpha>0$, $\beta>0$, has, obviously, the representation $D(0,t)=(1+\alpha t)^{-\frac{\beta}{\alpha}}$ (with $\sigma(t)=t$) and, hence, the delay function $\Psi(s,t)=s+t+\alpha st$. But it also has many other representations: $D(0,t)=[1+a\sigma(t)]^{-\frac{b}{a}}$, $\sigma(t)=\frac{(1+\alpha t)^{\frac{\beta a}{\alpha b}}-1}{a}$, for all $a>0$ and all $b>0$. However, it can easily be check that $[1+a\sigma(t)]^{-\frac{b}{a}}=(1+\alpha t)^{-\frac{\beta}{\alpha}}$ and $\sigma^{-1}[\sigma(s)+\sigma(t)+a\sigma(s)\sigma(t)]=s+t+\alpha st$. Since the delay function, Ψ, but not the seed function, σ, is uniquely determined by D, it is better to say that D exhibits Ψ delay rather than σ delay.

7 Foundations for of some discount functions

In this section we ask what properties of the special functions introduced in section 4 give us the LP and the RS discount functions. A particularly tractable form of the delay function that aids in the understanding of several discount functions is given next.

Definition 12 *(γ-delay): Preferences exhibit γ-delay, if the delay function is*

$$\Psi(s,t) = \left(s^\gamma + t^\gamma + \alpha s^\gamma t^\gamma\right)^{\frac{1}{\gamma}}, \ \alpha > 0, \ 0 < \gamma \leq 1. \tag{25}$$

In particular, if $\gamma = 1$, we say that preferences exhibit linear delay.

A delay function, if it exists, is unique (Proposition 7) and it always exists for a continuous discount function (Proposition 8). Hence, Definition 12 is a sound definition. However, it should be remembered that γ-delay is a property of the delay function, Ψ, not of the seed function, σ.

In the next proposition we generalize the hyperbolic discount function in Loewenstein and Prelec (1992) when the reference time is $r = 0$.

Proposition 19 *: If preferences, with a continuous discount function, D, exhibit γ-delay, then, necessarily, $D(0,t) = (1 + \alpha t^\gamma)^{-\frac{\beta}{\alpha}}$, where α are β are positive.*

We now give a generalization of the general hyperbolic discounting function in LP for an arbitrary reference time, $r \geq 0$.

Proposition 20 *(Generalization of LP): If preferences, with an additive continuous discount function, D, exhibits γ-delay, then, necessarily, $D(r,t) = \left(\frac{1+\alpha t^\gamma}{1+\alpha r^\gamma}\right)^{-\frac{\beta}{\alpha}}$, where α and β are positive.*
In particular, $\gamma = 1$ gives the LP-discount function (15).

Unlike the LP discount function, γ-delay is not sufficient to derive the RS discount function. We need, in addition, a particular kind of speedup function; this is the content of the next proposition.

Proposition 21 *(RS-discount functions): Let $\tau > 0$, $\rho = \frac{\gamma}{\tau}$ and $f(r,t) = (t^\tau - r^\tau)^{\frac{1}{\tau}}$. Let preferences with continuous discount function, D, exhibit γ-delay. Let the speedup function be f. Then D is the RS-discount functions, $D(r,t) = \left[1 + \alpha (t^\tau - r^\tau)^\rho\right]^{-\frac{\beta}{\alpha}}$,*
where α and β are positive.

8 An Application to the common difference effect

In this section we apply our framework to the common difference effect. Let us first reconsider the common difference effect, using Thaler's apples example (Thaler, 1981). A decision maker prefers one apple today to two apples tomorrow, so that

$$v(1) > v(2)\, D(0,1). \tag{26}$$

However, the decision maker, today, prefers to receive two apples in 51 days' time to receiving one apple in 50 days' time, so that

$$v(1)\, D(0,50) < v(2)\, D(0,51). \tag{27}$$

From (26) and (27) we get

$$D(0,50)D(0,1) < D(0,51). \tag{28}$$

Under exponential discounting we would get $D(0,50)D(0,1) = D(0,51)$ which violates (28) and, hence, the assumption of constant discounting (or the stationarity axiom) under exponential discounting. Following Loewenstein and Prelec (1992) we refer to a violation of the stationarity axiom as the common difference effect.

Definition 13 *(Common difference effect): The common difference effect is said to arise if for some outcomes $0 < z < z'$, and some time periods $t, t' > 0$ we observe the following pattern of preference. $u(z) = D(t)\, u(z')$ and $D(t')\, u(z) < D(t'+t)\, u(z')$.*

If we assume additivity of the discount function (Definition 2), as did LP, so that $D(0,51) = D(0,50)D(50,51)$, we get, from (28), $D(0,1) < D(50,51)$. So the decision maker exhibits declining impatience (Definition 3). However, and as Read (2001) pointed out, subadditivity could be an alternative explanation. To see this, assume constant impatience, so that $D(0,1) = D(50,51)$. In conjunction with (28) we then get $D(0,50)D(50,51) < D(0,51)$, which is the definition of subadditivity (Definition 2). Thus, the common difference effect is consistent with constant impatience if the discount function is sufficiently subadditive.

Example 6 *: (Thaler's apples example) Take the reference point to be the status quo, i.e., no apple. Use the prospect theory utility function*

$$v(y) = \begin{cases} y^{\gamma} & \text{if } \; y \geq 0 \\ -\lambda(-y)^{\gamma} & \text{if } \; y < 0 \end{cases} \tag{29}$$

where γ, λ are constants such that $0 < \gamma < 1$, and $\lambda > 1$ is known as the coefficient of loss aversion.[16] *Tversky and Kahneman (1992) estimated that $\gamma = 0.88$ and $\lambda = 2.25$.*[17] *So let us assume in this Example that $\lambda = 2.25$, $\gamma = 0.88$. Thus (working to four decimal figures),*

$$v(1) = 1; v(2) = 1.8404. \tag{30}$$

We compare the resolution of the 'common difference effect' anomaly under the LP-discount function (15) and the RS-discount function (16). To simplify as much as possible, choose the parameters: $\alpha = \beta = \tau = \rho = 1$. We tabulate the relevant magnitudes below:

	LP: $D(s,t) = (1+s)(1+t)^{-1}$	RS: $D(s,t) = (1+t-s)^{-1}$
$D(0,1)$	1/2	1/2
$D(0,50)$	1/51	1/51
$D(0,51)$	1/52	1/52
$D(50,51)$	51/52	1/2
$D(0,50)D(50,51)$	1/52	1/102

Table 3: Tabulation of selected values for the LP and RS discount functions

Substituting from Table 3 and (30) into (26) and (27) gives, respectively,

$$v(1) = 1 > 0.9202 = v(2)D(0,1), \tag{31}$$

$$v(1)D(0,50) = 0.0196 < 0.0354 = v(2)D(0,51), \tag{32}$$

for both the LP-discount function and the RS-discount function. This illustrates that both approaches can explain the common difference effect. However, they explain it in very different ways. Using Table 3, and the second column we see that the LP-discount function exhibits declining impatience because $D(0,1) = \frac{1}{2} < \frac{51}{52} = D(50,51)$. By contrast, from row 3 we see that the RS-discount function exhibits constant impatience because $D(0,1) = \frac{1}{2} = D(50,51)$.

On the other hand, from the second column of Table 3, we see that the LP-discount function is additive because $D(0,51) = \frac{1}{52} = D(0,50)D(50,51)$. By contrast, the RS-discount function is subadditive because $D(0,51) = \frac{1}{52} > \frac{1}{102} = D(0,50)D(50,51)$.

Thus, the LP-discount function explains the common difference effect as exclusively due to declining impatience, while, in this example, the RS-discount function explains this effect as due exclusively to subadditivity.[18]

[16] For the axiomatic foundations of this utility function, see al-Nowaihi et al. (2008).

[17] The high value of γ, approximately close to 1, is confirmed in subsequent work; see, for instance, Bruhin et al. (2010).

[18] More generally, and provided $0 < \rho \leq 1$, the RS-discount function can combine subadditivity with declining impatience ($0 < \tau < 1$), constant impatience ($\tau = 1$) or increasing impatience ($\tau > 1$).

Of course, and as Read (2001) pointed out, the common difference effect could be due to both declining impatience and subadditivity. Read (2001) found support for the common difference effect and for subadditivity but rejected declining impatience in favour of constant impatience. This result needs to be replicated and studied further.

We now state two propositions that characterize the common difference effect in terms of the underlying preferences.

From Proposition 14 we know (part (a) that α-subadditivity of the seed function, σ, is equivalent to $D(0,s)D(0,t) < D(0,s+t)$, $s>0$, $t>0$. From (28) we know that this is precisely the condition that is required to explain Thaler's apples example. The next proposition states this result formally.

Proposition 22 *(Characterization Theorem 4): Let D be a continuous discount function. Then preferences exhibit the common difference effect if, and only if, the seed function, σ, is α-subadditive.*[19]

From Propositions 19, 20 we know that for a continuous discount function D, γ-delay gives rise to the LP discount function. Example 6 shows that the common difference effect can be explained by the LP function. In conjunction these two observations motivate the next proposition.

Proposition 23 *: If preferences exhibit γ-delay, then they also exhibit the common difference effect.*

9 Summary and conclusions

9.1 Technical summary

A discount function, $D(r,t)$, that is continuous in t, can always be expressed in the form

$$D(r,t) = \varphi(f(r,t)) = [1 + \alpha\sigma(f(r,t))]^{-\frac{\beta}{\alpha}},$$

where φ is a (continuous) generating function, f is a speedup function (continuous in t), σ is a (continuous) seed function and $\alpha > 0$, $\beta > 0$. The corresponding delay function is given by $\Psi(s,t) = \sigma^{-1}[\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)]$. Ψ, φ and f (but not α, β or σ) are uniquely determined by D. The main properties of preferences over time are determined by φ, Ψ, f and σ as follows. Preferences exhibits the common difference effect if, and only if, σ is α-subadditive; which will be the case if, and only if, $\Psi(s,t) > s+t$ for $s>0$

[19] We noted earlier that the common difference effect can be explained by a combination of declining impatience and subadditivity *of the discount function*. However, α-subadditivity is not to be confused with subadditivity of the discount function.

and $t > 0$. Whether preferences exhibit declining, constant or increasing impatience is determined by f. D is additive if, and only if, $f(r,t) = \sigma^{-1}\left(\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right)$; in which case $D(r,t) = \frac{\varphi(t)}{\varphi(r)}$.

9.2 Conclusions

We explore the foundations of some commonly used discount functions in discounted utility models. We generalize the framework of Loewenstein and Prelec (1992) and introduce a *reference time* in addition to a *reference outcome* level. Our two main tools are a *delay function* and a *speedup function*. The delay function captures temporal preferences of individuals when two outcomes are delayed by the same time interval and the speedup function captures preferences when both outcomes are speeded up by the same interval.

We focus on three main classes of discount functions and explore their axiomatic foundations. The first category of discount functions rely on declining impatience, commonly associated with hyperbolic discounting. The second category relies on subadditive discounting that is sensitive to how an interval of time is partitioned. The third relies on both channels and provides the most general explanation of anomalies of the exponential discounted utility model. We not only provide the generalizations of these discount functions when there is a reference time but also provide a more general form of the general hyperbolic function.

As an application of our framework we discussed the common difference effect and gave necessary and sufficient conditions for its existence.

10 Appendix: Proofs

Proof of Proposition 1: Let $r \in [0,\infty)$ and $t \in [r,\infty)$. Let $\{t_n\}_{n=1}^{\infty}$ be a sequence in $[r,\infty)$ converging to t. We want to show that $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. It is sufficient to show that any monotone subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. In particular, let $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ be a decreasing subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$. Since $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ is bounded below by $D(r,t)$, it must converge to, say, q, where $D(r,t) \leq q \leq D(r,t_{n_i})$, for all i. Since D is onto, there is a $p \in [r,\infty)$ such that $D(r,p) = q$. Moreover, $t_{n_i} \leq p \leq t$, for each i. Suppose $D(r,t) < q$. Then $t_{n_i} < p$, for each i. Hence also $t_{n_i} < t$, for each i. But this cannot be, since $\{t_{n_i}\}_{i=1}^{\infty}$, being a subsequence of the convergent sequence $\{t_n\}_{n=1}^{\infty}$, must also converge to the same limit, t. Hence, $D(r,t) = q$. Hence, $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ converges to $D(r,t)$. Similarly, we can show that any increasing subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. Hence, $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. Hence, $D(r,t)$ is continuous in t. ■.

Proof of Proposition 2 (Time sensitivity): Let $D(r,t)$ be a continuous discount function and $r \geq 0$. Suppose $0 < x \leq y$. From (3) and (4), it follows that $0 < v(x) \leq v(y)$ and, hence, $0 < \frac{v(x)}{v(y)} \leq 1$. Since, by Definition 1(c), $D(r,t) : [r,\infty) \overset{\text{onto}}{\rightarrow} (0,1]$, it follows that $\frac{v(x)}{v(y)} = D(r,t)$ for some $t \in [r,\infty)$. A similar argument applies if $y < x < 0$. $\blacksquare$.

Proof of Proposition 3 (Existence of present values): Let $r \leq t$ and $y \geq 0$. Then, $0 < D(r,t) \leq 1$. Hence, $0 = v(0) \leq v(y) D(r,t) \leq v(y)$. Since v is continuous, (7), and strictly increasing, (3), it follows that $v(y) D(r,t) = v(x)$ for some $x \in [0,y]$. Similarly, if $y \leq 0$, then $v(y) D(r,t) = v(x)$ for some $x \in [y,0]$. $\blacksquare$.

To facilitate the proof of Propositions 4, below, and 23, later, we first establish Lemmas 24 and 25.

Lemma 24 *: Let $x \geq 0$ and $y \geq 0$. Then:*
(a) $\rho \geq 1 \Rightarrow x^\rho + y^\rho \leq (x+y)^\rho$.
(b) $0 < \rho \leq 1 \Rightarrow x^\rho + y^\rho \geq (x+y)^\rho$.

Proof of Lemma 24: Clearly, the results hold for $x = 0$. Suppose $x > 0$. Let $z = \frac{y}{x}$ and $f(z) = (1+z)^\rho - 1 - z^\rho$. Then $f(z) = 0$ and $f'(z) = \rho\left[(1+z)^{\rho-1} - z^{\rho-1}\right]$, for $z > 0$. Suppose $\rho \geq 1$. Then $f'(z) \geq 0$. Since f is continuous, it follows that $f(z) \geq 0$ for $z \geq 0$. Part (a) follows from this. Now suppose $0 < \rho \leq 1$. Then $f'(z) \leq 0$. Since f is continuous, it follows that $f(z) \leq 0$ for $z \geq 0$. Part (b) follows from this. $\blacksquare$.

Lemma 25 *: Let $\tau > 0$, $0 \leq s < t$ and $r > 0$. Let $f(r) = (t+r)^\tau - (s+r)^\tau - (t^\tau - s^\tau)$. Then:*
(a) $\tau > 1 \Rightarrow f(r) > 0$.
(b) $0 < \tau < 1 \Rightarrow f(r) < 0$.

Proof of Lemma 25: Clearly, $f(0) = 0$. Also, $f'(r) = \tau\left[(t+r)^{\tau-1} - (s+r)^{\tau-1}\right]$. If $\tau > 1$, then $f'(r) > 0$ for $r > 0$. Since f is continuous, it follows that $f(r) > 0$ for $r > 0$. This establishes part (a). If $0 < \tau < 1$, then $f'(r) < 0$ for $r > 0$. Since f is continuous, it follows that $f(r) < 0$ for $r > 0$. This establishes part (b). $\blacksquare$.

Proof of Proposition 4: (a) Suppose $0 < \rho \leq 1$. Let $0 \leq r < s < t$. From (16), we get:

$$D(r,s) = \left[1 + \alpha\left(s^\tau - r^\tau\right)^\rho\right]^{-\frac{\beta}{\alpha}},$$

$$D(s,t) = \left[1 + \alpha\left(t^\tau - s^\tau\right)^\rho\right]^{-\frac{\beta}{\alpha}},$$

$$D(r,t) = \left[1 + \alpha\left(t^\tau - r^\tau\right)^\rho\right]^{-\frac{\beta}{\alpha}},$$

and, hence,

$$\begin{aligned} D(r,s)\,D(s,t) &= \left\{1+\alpha\left[(s^\tau-r^\tau)^\rho+(t^\tau-s^\tau)^\rho\right]+\alpha^2(s^\tau-r^\tau)^\rho(t^\tau-s^\tau)^\rho\right\}^{-\frac{\beta}{\alpha}}, \\ &< \left\{1+\alpha\left[(s^\tau-r^\tau)^\rho+(t^\tau-s^\tau)^\rho\right]\right\}^{-\frac{\beta}{\alpha}}, \text{ since } \alpha^2(s^\tau-r^\tau)^\rho(t^\tau-s^\tau)^\rho>0 \text{ and } -\frac{\beta}{\alpha}<0, \\ &\le \left[1+\alpha(s^\tau-r^\tau+t^\tau-s^\tau)^\rho\right]^{-\frac{\beta}{\alpha}}, \text{ by Lemma 24b and since } -\frac{\beta}{\alpha}<0, \\ &= \left[1+\alpha(t^\tau-r^\tau)^\rho\right]^{-\frac{\beta}{\alpha}}, \\ &= D(r,t). \end{aligned}$$

(b) It is sufficient to give an example. Let $\alpha=\tau=1$ and $\rho=2$. Hence, $D(0,1)\,D(1,2)=4^{-\beta}>5^{-\beta}=D(0,2)$. Hence, for $\alpha=\tau=1$ and $\rho=2$, D cannot be additive or subadditive. However, for the same parameter values, we have $D(0,10)\,D(10,20)=10201^{-\beta}<401^{-\beta}=D(0,20)$. Hence, D cannot be superadditive either.[20]

(c) Let $0\le s<t$. (ii) is obvious from inspecting (16). Let $r>0$. (iii) For $\tau>1$, Lemma 25(a) gives $D(s+r,t+r)=\left\{1+\alpha\left[(t+r)^\tau-(s+r)^\tau\right]^\rho\right\}^{-\frac{\beta}{\alpha}}<\left\{1+\alpha\left[t^\tau-s^\tau\right]^\rho\right\}^{-\frac{\beta}{\alpha}}=D(s,t)$. (i) For $0<\tau<1$, Lemma 25(b) gives $D(s+r,t+r)=\left\{1+\alpha\left[(t+r)^\tau-(s+r)^\tau\right]^\rho\right\}^{-\frac{\beta}{\alpha}}>\left\{1+\alpha\left[t^\tau-s^\tau\right]^\rho\right\}^{-\frac{\beta}{\alpha}}=D(s,t)$. ■.

Proof of Proposition 5: (a) Suppose that $D(r,t)=[\varphi(r)]^{-1}\varphi(t)$ for some strictly decreasing real valued function, $\varphi:[0,\infty)\to(0,1]$. We first check that D satisfies (a) and (b) of Definition 1. Clearly, $D(r,t)$ is strictly decreasing in t and strictly increasing in r. Let $r\in[0,\infty)$. $D(r,r)=[\varphi(r)]^{-1}\varphi(r)=1$. Hence, for fixed $r\in[0,\infty)$, $t\mapsto D(r,t)$ maps $[r,\infty)$ into $(0,1]$ and Definition 1(a) holds. Next, let $t\in[0,\infty)$. $D(t,t)=[\varphi(t)]^{-1}\varphi(t)=1$. Hence, for fixed $t\in[0,\infty)$, $r\longmapsto D(r,t)$ maps $[0,t]$ into $(0,1]$. Thus Definition 1(b) also holds. Hence, D is a discount function. If, in addition, $\varphi(0)=1$, then $D(0,t)=[\varphi(0)]^{-1}\varphi(t)=\varphi(t)$.

Now suppose that φ is onto. Let $p\in(0,1]$. Hence, also, $\varphi(r)\,p\in(0,1]$. Since φ is onto $(0,1]$, we get $\varphi(t)=\varphi(r)\,p$ for some $t\in[0,\infty)$. But $\varphi(t)=\varphi(r)\,p\le\varphi(r)$. Hence, $t\ge r$. We also have $D(r,t)=[\varphi(r)]^{-1}\varphi(t)=[\varphi(r)]^{-1}\varphi(r)\,p=p$. Hence, for each $r\in[0,\infty)$, $t\longmapsto D(r,t)$ maps $[r,\infty)$ onto $(0,1]$. Thus Definition 1(a) holds with 'into' replaced by 'onto'. Hence, D is a continuous discount function. Since $\varphi:[0,\infty)\overset{\text{onto}}{\to}(0,1]$ is strictly decreasing, we must have $\varphi(0)=1$. Hence, $D(0,t)=[\varphi(0)]^{-1}\varphi(t)=\varphi(t)$.

For all r, s and t, $D(r,s)\,D(s,t)=[\varphi(r)]^{-1}\varphi(s)\,[\varphi(s)]^{-1}\varphi(t)=[\varphi(r)]^{-1}\varphi(t)=D(r,t)$. Hence, D is additive.

(b) Suppose that D is an additive discount function. Then, for all r, s, t, where $0\le r\le s\le t$, $D(r,s)\,D(s,t)=D(r,t)$. From this, it follows that, for any s and any t $(0\le s\le t)$: $\frac{D(r,t)}{D(r,s)}=D(s,t)$, which is independent of r, for all $r\in[0,s]$. Similarly, for any

[20] Other examples can be given to show that there is nothing special about $r=0$, $\alpha=1$, $\tau=1$, or $\rho=2$, as long as $\rho>1$.

r and any s $(0 \leq r \leq s)$: $\frac{D(r,t)}{D(s,t)} = D(r,s)$, which is independent of t, for all $t \in [s,\infty)$. This can only hold if $D(r,t) = F(r)\Phi(t)$, for all r and t $(0 \leq r \leq t)$. In particular, $F(r)\Phi(r) = D(r,r) = 1$. Hence, $F(r) = [\Phi(r)]^{-1}$. Hence, $D(r,t) = [\Phi(r)]^{-1}\Phi(t)$. Set $\varphi(t) = \Phi(t)/\Phi(0)$. Then $D(r,t) = [\varphi(r)]^{-1}\varphi(t)$. In particular, $D(0,t) = [\varphi(0)]^{-1}\varphi(t) = \varphi(t)$. Hence, φ is a strictly decreasing function from $[0,\infty)$ into $(0,1]$. If D is continuous, so that $D(0,t)$ is onto, then φ is also onto. ■.

Proof of Proposition 6: Exponential: $\varphi(t) = e^{-\beta t}$. PPL: $\varphi(0) = 1$ and $\varphi(t) = e^{-\delta-\beta t}$ for $t > 0$. LP: $\varphi(t) = (1+\alpha t)^{-\frac{\beta}{\alpha}}$. Generalized RS: $\varphi(t) = e^{-Q[w(t)]}$. ■.

Proof of Proposition 7 (Properties of a delay function): Let D be a discount function and Ψ and Φ two corresponding delay functions. Let $s,t \in [0,\infty)$. Then $D(0,\Phi(s,t)) = D(0,s)D(0,t) = D(0,\Psi(s,t))$. Since $D(0,r)$ is strictly decreasing in r, we must have $\Phi(s,t) = \Psi(s,t)$. This establishes (a). Using Definition 1, it is straightforward to check that properties (b) to (d) follow from Definition 6. Now, suppose $v(x) = v(y)D(0,t)$. Multiply both sides by $D(0,s)$ to get $v(x)D(0,s) = v(y)D(0,s)D(0,t) = v(y)D(0,\Psi(s,t))$. Conversely, suppose $v(x)D(0,s) = v(y)D(0,\Psi(s,t))$. Then $v(x)D(0,s) = v(y)D(0,s)D(0,t)$. Since $D(0,s) > 0$, we can cancel it to get $v(x) = v(y)D(0,t)$. This establishes (e) and completes the proof. ■.

Proof of Proposition 8 (Existence of a delay function): Let D be a continuous discount function. Let $s,t \in [0,\infty)$. Then $D(0,s), D(0,t) \in (0,1]$. Since $r \mapsto D(0,r)$ is onto $(0,1]$, there is some $T \in [0,\infty)$ such that $D(0,s)D(0,t) = D(0,T)$. Since $D(0,r)$ is strictly decreasing in r, this T is unique. Set $T = \Psi(s,t)$. The function, $\Psi(s,t)$, thus defined, is a delay function corresponding to D. ■.

Proof of Proposition 9: (a) Let f and g be speedup functions corresponding to the discount function D. Let $r \in [0,\infty)$ and $t \in [r,\infty)$. Then, by Definition 8, $D(0,f(r,t)) = D(r,t) = D(0,g(r,t))$. Since $D(r,s)$ is strictly decreasing in s, it follows that $f(r,t) = g(r,t)$. Hence, $f = g$. (b) For $r = 0$ we have $D(0,t) = D(0,f(0,t))$. Since $D(0,s)$ is strictly decreasing in s, it follows that $f(0,t) = t$. We also have $D(0,f(r,r)) = D(r,r) = 1 = D(0,0)$, i.e., $D(0,f(r,r)) = D(0,0)$. Again, since $D(0,t)$ is strictly decreasing in t, it follows that $f(r,r) = 0$. Since $D(0,f(r,t)) = D(r,t)$ and $D(r,t)$ is strictly decreasing in t, it follows that $f(r,t)$ must be strictly increasing in t. Since $D(r,t)$ is strictly increasing in r, it follows that $f(r,t)$ must be strictly decreasing in r. This completes the proof that f is a speedup function. (c) Since $D(r,t) = D(0,f(r,t))$, it follows that $v(x) = v(y)D(r,t)$ if, and only if, $v(x) = v(y)D(0,f(r,t))$. ■.

Proof of Proposition 10: Let $r \in [0,\infty)$ and $t \in [r,\infty)$. Then $D(r,t) \in (0,1]$. Since D is a continuous discount function, it follows that $s \longmapsto D(0,s)$ is onto $(0,1]$. Hence, for some $T \in (0,1]$, $D(r,t) = D(0,T)$. Since $D(r,s)$ is strictly decreasing in s, it follows that this T is unique. Set $f(r,t) = T$. The function, f, thus defined has the property $D(r,t) = D(0,f(r,t))$. Hence, f is a speedup function corresponding to D. Let

$r \in [0,\infty)$. Let $t \in [0,\infty)$. Then $D(0,t) \in (0,1]$. Since, D is a continuous discount function, it maps $[r,\infty)$ onto $(0,1]$. Hence $D(r,s) = D(0,t)$, for some $s \in [r,\infty)$. But since f is a speedup function corresponding to D, we have $D(r,s) = D(0, f(r,s))$. Hence, $D(0, f(r,s)) = D(0,t)$. Since $D(0,q)$ is strictly decreasing in q, it follows that $f(r,s) = t$. Hence, $t \longmapsto f(r,t)$ maps $[r,\infty)$ onto $[0,\infty)$. ■.

Proof of Proposition 11: (a) Since $f : \Delta \to [0,\infty)$ and $\varphi : [0,\infty) \to (0,1]$, it follows that $\varphi \circ f : \Delta \to (0,1]$. Let $r \in [0,\infty)$. Then $t \mapsto \varphi(f(r,t))$ maps $[r,\infty)$ into $(0,1]$ and is strictly decreasing with $D(r,r) = \varphi(f(r,r)) = \varphi(0) = 1$. Let $t \in [0,\infty)$. Then $r \mapsto \varphi(f(r,t))$ maps $[0,t]$ into $(0,1]$ and is strictly decreasing. Thus, $D(r,t) = \varphi(f(r,t))$ satisfies parts (a), (b) of Definition 1 and, hence, is a discount function. (b) $D(0, f(r,t)) = \varphi(f(0, f(r,t))) = \varphi(f(r,t)) = D(r,t)$. Hence, f is the extension function corresponding to D and $D(r,t)$ is the f-extension of $D(0,t)$. (c) Let $r \in [0,\infty)$. Let $p \in (0,1]$. Since φ is onto, there exists $t \in [0,\infty)$ such that $p = \varphi(t)$. Since $s \mapsto f(r,s)$ maps $[r,\infty)$ onto $[0,\infty)$, there exists $s \in [r,\infty)$ such that $f(r,s) = t$. Hence, $D(r,s) = \varphi(f(r,s)) = \varphi(t) = p$. Hence, $s \mapsto D(r,s)$ maps $[r,\infty)$ onto $[0,\infty)$. Hence, D is a continuous discount function (Definition 1 (c). ■.

Proof of Proposition 12: (Representation Theorem 1) Let $\alpha > 0$ and $\beta > 0$. It is easy to verify that $D(0,t) = [1 + \alpha\sigma(t)]^{-\frac{\beta}{\alpha}}$ is an (α,β)-representation of the discount function $D(r,t)$ if, and only if, $\sigma(t) = \frac{1}{\alpha}[D(0,t)]^{-\frac{\alpha}{\beta}} - 1$. From this, using Definition 1, it is straightforward to verify that part (a) holds. It follows that the inverse of σ exists and $\sigma^{-1} : [0,\infty) \overset{\text{onto}}{\to} [0,\infty)$ is strictly increasing with $\sigma(0) = 0$, this establishes part (b). We now turn to part (c). We have $[1 + \alpha\sigma(\Psi(s,t))]^{-\frac{\beta}{\alpha}} = D(0, \Psi(s,t)) = D(0,s)\, D(0,t) = [1 + \alpha\sigma(s)]^{-\frac{\beta}{\alpha}} [1 + \alpha\sigma(t)]^{-\frac{\beta}{\alpha}} = [1 + \alpha\sigma(s) + \alpha\sigma(t) + \alpha^2\sigma(s)\sigma(t)]^{-\frac{\beta}{\alpha}}$. Hence, $[1 + \alpha\sigma(\Psi(s,t))]^{-\frac{\beta}{\alpha}} = [1 + \alpha\sigma(s) + \alpha\sigma(t) + \alpha^2\sigma(s)\sigma(t)]^{-\frac{\beta}{\alpha}}$. From the latter it follows that $\sigma(\Psi(s,t)) = \sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)$. Hence, $\Psi(s,t) = \sigma^{-1}[\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)]$. ■.

Proof of Proposition 13 (Representation Theorem 2): Since σ is a continuous seed function,

$$\sigma : [0,\infty) \overset{\text{onto}}{\to} [0,\infty) \text{ be strictly increasing (hence, } \sigma(0) = 0\text{)}, \tag{33}$$

and let

$$\Psi(s,t) = \sigma^{-1}(\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)),\ \alpha > 0. \tag{34}$$

Let D be a continuous discount function with delay function, Ψ. Then

$$D(0,s)\, D(0,t) = D(0, \Psi(s,t)),\ s \geq 0,\ t \geq 0. \tag{35}$$

From (34) and (35), we get

$$D(0,s)\, D(0,t) = D\left(0, \sigma^{-1}(\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t))\right),\ s \geq 0,\ t \geq 0. \tag{36}$$

From (33), we get that σ^{-1} exists and $\sigma^{-1} : [0, \infty) \overset{\text{onto}}{\rightarrow} [0, \infty)$ is strictly increasing with $\sigma^{-1}(0) = 0$. Bear these facts in mind for when the functions G and h are defined, below ((39) and (42)).

Let

$$X = 1 + \alpha\sigma(s), \; Y = 1 + \alpha\sigma(t). \tag{37}$$

Hence

$$\sigma(s) = \frac{X-1}{\alpha}, \; \sigma(t) = \frac{Y-1}{\alpha}, \; s = \sigma^{-1}\left(\frac{X-1}{\alpha}\right), \; t = \sigma^{-1}\left(\frac{Y-1}{\alpha}\right),$$

$$\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t) = \frac{XY-1}{\alpha},$$

$$\sigma^{-1}(\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)) = \sigma^{-1}\left(\frac{XY-1}{\alpha}\right). \tag{38}$$

Define the function $G : [1, \infty) \rightarrow (0, \infty)$ by

$$G(X) = D\left(0, \sigma^{-1}\left(\frac{X-1}{\alpha}\right)\right). \tag{39}$$

Hence,

$$G(Y) = D\left(0, \sigma^{-1}\left(\frac{Y-1}{\alpha}\right)\right), \; G(XY) = D\left(0, \sigma^{-1}\left(\frac{XY-1}{\alpha}\right)\right). \tag{40}$$

From (36), (38), (39), (40)

$$\begin{aligned} G(XY) &= D\left(0, \sigma^{-1}\left(\frac{XY-1}{\alpha}\right)\right) \\ &= D\left(0, \sigma^{-1}(\sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t))\right) \\ &= D(0, s)\, D(0, t) \\ &= D\left(0, \sigma^{-1}\left(\frac{X-1}{\alpha}\right)\right) D\left(0, \sigma^{-1}\left(\frac{Y-1}{\alpha}\right)\right) \\ &= G(X)\, G(Y). \end{aligned} \tag{41}$$

Define[21] the function $h : [0, \infty) \rightarrow (0, \infty)$ by

$$h(y) = G(e^y), y \geq 0 \tag{42}$$

Hence, and in the light of Definition 1(i), h satisfies[22]:

$$h : [0, \infty) \rightarrow (0, \infty) \text{ is strictly decreasing and } h(x+y) = h(x)\, h(y). \tag{43}$$

[21] It is tempting, at this stage, to take a shortcut and conclude, from the fact that $G(XY) = G(X)\,G(Y)$, that, necessarily, $G(X) = X^c$. However, the relevant theorem (Theorem 1.9.13 in Eichhorn (1978) or Theorem 3, page 41, in Aczel (1966)) requires that $G(X)$ be defined for *all* $X > 0$. However, $\psi^{-1}(t)$ is not defined for $t < 0$ and, hence, $G(X)$ is not defined for $X < 1$.

[22] It is sufficient that h be strictly decreasing in some interval: $(a, a+\delta)$, $a \geq 0, \delta > 0$.

As is well known, see for example Corollary 1.4.11 in Eichhorn (1978) or Theorem 1, page 38, of Aczel (1966), the unique solution to (43) is the exponential function

$$h(y) = e^{cy},\ y \geq 0,\ c < 0, \tag{44}$$

(37), (38), (40), (42), (44) give, in succession,

$$\begin{aligned} h(y) &= (e^{y})^{c} \\ G(e^{y}) &= (e^{y})^{c} \\ G(Y) &= Y^{c} \\ D\left(0, \sigma^{-1}\left(\frac{Y-1}{\alpha}\right)\right) &= Y^{c} \\ D\left(0, \sigma^{-1}\left(\sigma(t)\right)\right) &= [1+\alpha\sigma(t)]^{c} \\ D(0,t) &= [1+\alpha\sigma(t)]^{c}. \end{aligned} \tag{45}$$

Let

$$\beta = -\alpha c, \tag{46}$$

(45), (46) give

$$D(0,t) = [1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}},\ \alpha > 0,\ \beta > 0,\ t \geq 0, \tag{47}$$

where $\beta > 0$ because $\alpha > 0$ and $c < 0$. ■.

Proof of Proposition 14 (Representation Theorem 3): From Definition 10 and Proposition 12a, it follows that σ is a continuous seed function. Let $s > 0$ and $t > 0$, then $D(0,s)\,D(0,t) < D(0,s+t) \Leftrightarrow [1+\alpha\sigma(s)]^{-\frac{\beta}{\alpha}}[1+\alpha\sigma(t)]^{-\frac{\beta}{\alpha}} < [1+\alpha\sigma(s+t)]^{-\frac{\beta}{\alpha}}$ $\Leftrightarrow [1+\alpha\sigma(s)][1+\alpha\sigma(t)] > 1+\alpha\sigma(s+t) \Leftrightarrow \sigma(s+t) < \sigma(s)+\sigma(t)+\alpha\sigma(s)\sigma(t)$ $\Leftrightarrow s+t < \sigma^{-1}[\sigma(s)+\sigma(t)+\alpha\sigma(s)\sigma(t)] \Leftrightarrow s+t < \Psi(s,t)$. From this chain it follows that (c) $\Leftrightarrow$ (a) $\Leftrightarrow$ (b). ■.

Proof of Proposition 15 (Characterization Theorem 1): Let D be a continuous discount function. Then, by Representation Theorem 1 (Proposition 12) and Proposition 10, $D(r,t) = D(0, f(r,t)) = [1+\alpha\sigma(f(r,t))]^{-\frac{\beta}{\alpha}}$, where $\alpha > 0$, $\beta > 0$, σ is a continuous seed function and f is a continuous extension function. Conversely, if $D(r,t) = [1+\alpha\sigma(f(r,t))]^{-\frac{\beta}{\alpha}}$, where $\alpha > 0$, $\beta > 0$, σ is a continuous seed function and f is a continuous extension function, then D satisfies all the conditions of Definition 1, with 'onto' replacing 'into' in part (a). Hence, D is a continuous discount function. Uniqueness of f follows from Extension Theorem 1 (Proposition 9). ■.

Proof of Proposition 16 (Characterization Theorem 2): Follows from Definition 3 and Characterization Theorem 1 (Proposition 15). ■.

Proof of Proposition 17 (Characterization Theorem 3): Let $f(r,t)=\sigma^{-1}\left(\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right)$. Then $D(r,t)=\left[1+\alpha\sigma(f(r,t))\right]^{-\frac{\beta}{\alpha}}=\left[1+\alpha\sigma\left(\sigma^{-1}\left(\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right)\right)\right]^{-\frac{\beta}{\alpha}}=\left[1+\alpha\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right]^{-\frac{\beta}{\alpha}}$ $=\left[\frac{1+\alpha\sigma(t)}{1+\alpha\sigma(r)}\right]^{-\frac{\beta}{\alpha}}$, which is additive, by Proposition 5(aiii). Conversely, suppose D is additive. Then, by Proposition 5(b), $D(r,t)=[\varphi(r)]^{-1}\varphi(t)$ for some strictly decreasing real valued function, $\varphi:[0,\infty)\overset{\text{onto}}{\to}(0,1]$, $\varphi(0)=1$. Let $D(r,t)=\left[1+\alpha\sigma(f(r,t))\right]^{-\frac{\beta}{\alpha}}$. Then $\varphi(t)=D(0,t)=\left[1+\alpha\sigma(t)\right]^{-\frac{\beta}{\alpha}}$. Hence, $\left[1+\alpha\sigma(f(r,t))\right]^{-\frac{\beta}{\alpha}}=\left[\frac{1+\alpha\sigma(t)}{1+\alpha\sigma(r)}\right]^{-\frac{\beta}{\alpha}}$ and, hence, $1+\alpha\sigma(f(r,t))=\frac{1+\alpha\sigma(t)}{1+\alpha\sigma(r)}$. From which it follows that $f(r,t)=\sigma^{-1}\left(\frac{\sigma(t)-\sigma(r)}{1+\alpha\sigma(r)}\right)$. ■.

Proof of Proposition 18: All the claims can be verified by straightforward calculations. However, when dealing with PPL, do not use Proposition 12, as PPL is not continuous. So, for example, instead of using part b of Proposition 12, check directly that $D(0,s)D(0,t)=D(0,\Psi(s,t))$. ■.

Proof of Proposition 19: Let preferences with the continuous discount function, D, exhibit γ-delay. Then, by Definition 12, its delay function is $\Psi(s,t)=\left(s^{\gamma}+t^{\gamma}+\alpha s^{\gamma}t^{\gamma}\right)^{\frac{1}{\gamma}}$, $\alpha>0$, $0<\gamma\leq 1$. Hence, by Proposition 13, we must have $D(0,t)=\left(1+\alpha t^{\gamma}\right)^{-\frac{\beta}{\alpha}}$, for some $\beta>0$. ■.

Proof of Proposition 20: By Proposition 19 we have, necessarily, $D(0,t)=\left(1+\alpha t^{\gamma}\right)^{-\frac{\beta}{\alpha}}$, for some $\beta>0$. Hence, by Characterization Theorem 3 (Proposition 17) we must have $D(r,t)=\left(\frac{1+\alpha t^{\gamma}}{1+\alpha r^{\gamma}}\right)^{-\frac{\beta}{\alpha}}$. ■.

Proof of Proposition 21: By Proposition 19 we have, $D(0,t)=\left(1+\alpha t^{\gamma}\right)^{-\frac{\beta}{\alpha}}$, for some $\beta>0$. Hence, $D(r,t)=D(0,f(r,t))=\left[1+\alpha\left[f(r,t)\right]^{\gamma}\right]^{-\frac{\beta}{\alpha}}=\left[1+\alpha\left[\left(t^{\tau}-r^{\tau}\right)^{\frac{1}{\tau}}\right]^{\gamma}\right]^{-\frac{\beta}{\alpha}}$ $=\left[1+\alpha\left(t^{\tau}-r^{\tau}\right)^{\frac{\gamma}{\tau}}\right]^{-\frac{\beta}{\alpha}}=\left[1+\alpha\left(t^{\tau}-r^{\tau}\right)^{\rho}\right]^{-\frac{\beta}{\alpha}}$. ■.

Proof of Proposition 22 (Characterization Theorem 4): Assume σ is α-subadditive. Let $0<x<y$, $v(x)=v(y)D(0,t)$ and $s>0$. Then $0<v(x)<v(y)$. It follows that $D(0,t)<1$ and, hence, $t>0$. From Representation Theorem 3(c) (Proposition 14), it follows that $D(0,s)D(0,t)<D(0,s+t)$, thus, $v(y)D(0,t)D(0,s)<v(y)D(0,s+t)$. Hence, $v(x)D(0,s)<v(y)D(0,s+t)$, so the common difference effect holds.

Conversely, assume that the common difference effect holds. Let $s>0$ and $t>0$, thus, $0<s<s+t$. Hence, by Proposition 3, for some x, $0\leq x\leq 1$, $v(x)=v(1)D(0,t)$, so, $x>0$. Therefore, $v(x)D(0,s)<v(1)D(0,s+t)$, so $v(1)D(0,s)D(0,t)<v(1)D(0,s+t)$. Hence, $D(0,s)D(0,t)<D(0,s+t)$. From Representation Theorem 3 (Proposition 14), it follows that σ is α-subadditive.

It follows that preferences exhibit the common difference effect if, and only if, the seed function for gains, σ, is α-subadditive. ■.

Proof of Proposition 23: Let $\sigma(t)=t^{\gamma}$, where $0<\gamma\leq 1$. Then, for $s>0$ and

$t > 0$, $\sigma(s+t) = (s+t)^{\gamma} \leq t^{\gamma} + s^{\gamma} < s^{\gamma} + t^{\gamma} + \alpha s^{\gamma} t^{\gamma} = \sigma(s) + \sigma(t) + \alpha\sigma(s)\sigma(t)$. Hence, by Characterization Theorem 4 (Proposition 22), preferences that exhibit γ-delay (Definition 12), also exhibit the common difference effect. $\blacksquare$.

References

Aczel, J. 1966. *Lectures on Functional Equations and their Applications*. Academic Press, New York and London.

al-Nowaihi, A., Dhami, S. 2006. A Note On The Loewenstein-Prelec Theory Of Intertemporal Choice. *Mathematical Social Sciences*, Volume 52, Issue 1, Pages 99-108.

al-Nowaihi, A., Dhami, S. 2008a. A note on the Loewenstein-Prelec theory of intertemporal choice: Corrigendum. *Mathematical Social Sciences*. 52, 99-108.

al-Nowaihi, A., Dhami, S. 2008b. A general theory of time discounting: The reference-time theory of intertemporal choice. *University of Leicester Working Paper* No. tba.

al-Nowaihi, A., Bradley, I., Dhami, S. 2008. The Utility Function Under Prospect Theory. *Economics Letters* 99, p.337–339.

Bruhin, A., Fehr-Duda, H., Epper, T. 2010. Risk and Rationality: Uncovering Heterogeneity in Probability Distortion, *Econometrica*, 78: 1375-1412.

Eichhorn, W. 1978. *Functional Equations in Economics*, Addison-Wesley.

Kahneman, D., Tversky, A. 1979. Prospect theory : an analysis of decision under risk. *Econometrica* 47, 263-291.

Kahneman D., Tversky A. 2000. *Choices, Values and Frames, Cambridge University Press*, Cambridge.

Laibson, D. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112, 443-477.

Loewenstein G. 1988. Frames of mind in intertemporal choice. *Management Science* 34, 200-214.

Loewenstein G., Prelec D. 1992. Anomalies in intertemporal choice : evidence and an interpretation. *The Quarterly Journal of Economics* 107, 573-597.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, E. S. (2001) Risk as feelings, *Psychological Bulletin*, 127: 267–286.

Ok, E. A., and Masatlioglu, Y. 2007. A theory of (relative) discounting. *Journal of Economic Theory* 137, 214-245.

Phelps, E. S., Pollak, R. 1968. On second-best national saving and game-equilibrium growth. *Review of Economic Studies* 35, 185-199.

Read, D. 2001. Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, 23, 5-32.

Samuelson, P. 1937. A note on measurement of utility. *Review of Economic Studies*, 4, 155-161.

Scholten, M., and Read, D. 2006a. Discounting by intervals: A generalized model of intertemporal choice, *Management Science*, 52, 1426-1438.

Scholten, M., and Read, D. 2011a. Descriptive Models of Intertemporal Choice Part 1: Anomalies in Choices Between Less Sooner and More Later. mimeo.

Scholten, M., and Read, D. 2011b. Descriptive Models of Intertemporal Choice Part 2: The Delay-Speedup Asymmetry and Other Anomalies. mimeo.

Thaler, R.H. 1981. Some empirical evidence on dynamic consistency. *Economic Letters* 8, 201-207.

Tversky, A., and Kahneman, D. 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.