

al-Nowaihi, Ali; Dhami, Sanjit

**Working Paper**

## Foundations for Intertemporal Choice

CESifo Working Paper, No. 6913

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* al-Nowaihi, Ali; Dhami, Sanjit (2018) : Foundations for Intertemporal Choice, CESifo Working Paper, No. 6913, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/176932>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Foundations for Intertemporal Choice

*Ali al-Nowaihi, Sanjit Dhami*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl

[www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Foundations for Intertemporal Choice

## Abstract

We consider discounted-utility models with a reference stream of outcomes. We provide a common framework for the main empirically supported discount functions in terms of three underlying functions: The delay, speedup and generating functions. Each of the delay and speedup functions can be uniquely elicited from behavior and, hence, can be fitted to the data. These two functions determine whether the discount function is subadditive, additive or superadditive; and whether the discount function exhibits declining, constant or increasing impatience. The third function, the generating function, links the speedup function to the discount function. Our framework nests several important attribute-based models that are typically considered to be in a separate class. We also show that apparent intransitivities of time preferences can be accounted for by framing effects.

JEL-Codes: C600, D910.

Keywords: time discounting, framing effects, impatience, additivity, common difference effect, intransitive preferences.

*Ali al-Nowaihi*  
*Economics Division, School of Business*  
*University of Leicester*  
*University Road*  
*United Kingdom – Leicester, LE1 7RH*  
*aa10@le.ac.uk*

*Sanjit Dhami\**  
*Economics Division, School of Business*  
*University of Leicester*  
*University Road*  
*United Kingdom – Leicester, LE1 7RH*  
*Sanjit.Dhami@le.ac.uk*

\*corresponding author

12 February 2018

We are grateful to Clive Fraser, Itzhak Gilboa, Herbert Gintis, George Loewenstein, Larry Samuelson, Marc Scholten and Yingyu Guo for valuable comments. However, all errors are our own.

# 1 Introduction

Decisions that have a time dimension are central to economics. The *exponentially discounted utility* (EDU) model, due to Samuelson (1937), the workhorse model in economics, labours under the weight of assigning the entire psychology of intertemporal choice to a single parameter – the constant discount rate. The evidence now clearly shows that the EDU model is not able to account for the evidence from intertemporal choices. Several alternative models have been proposed (Loewenstein and Prelec, 1992; Frederick et al., 2002; Manzini and Mariotti, 2008); for a recent survey, see Dhami (2016). The refutations of the EDU model, commonly referred to as *anomalies*, are not simply mistakes (Frederick et al., 2002, section 4.3).

Several anomalies of the EDU model are explained by imposing suitable restrictions on the *instantaneous utility function* or *felicity*. In particular, these anomalies include the *magnitude effect* (larger magnitudes are discounted relatively less) and *gain-loss asymmetry* (losses are discounted relatively less). See, Loewenstein and Prelec (1992) and al-Nowaihi and Dhami (2006, 2008a, 2009).<sup>1</sup>

Another set of important anomalies of the EDU model have required suitable modifications of the discount function in EDU. Arguably, the most serious anomaly, and certainly the most discussed, is the rejection of the *stationarity* property of EDU, which is termed the *common difference effect*.

There are two competing explanations of the common difference effect. (1) The discount rate over a fixed time interval increases as the interval is gradually moved towards the present date. This is a violation of constant impatience, and it suggests *declining impatience*, which leads to the well-known phenomenon of *present-biased preferences*.<sup>2</sup> (2) EDU has the *additivity property*, i.e., for time periods  $0 \leq r < s < t$ , discounting a positive magnitude back from  $t$  to  $s$ , and then from  $s$  back to  $r$ , is equivalent to discounting from  $t$  to  $r$  in one step. Evidence indicates that discounting over the two sub-intervals, rather than over the entire interval, may result in relatively higher discounting (*subadditivity*) or relatively lower discounting (*superadditivity*). The Read-Scholten (RS) *interval discount function* explains the common difference effect either through declining impatience, subadditivity or both (Read, 2001; Scholten and Read, 2006a, 2016). Thus, it is the most general discount function, yet it is not widely used in economics.

Several concerns have been raised about the evidence on discount rates that arises from comparing outcome-time pairs. For instance, delayed payments may be considered

---

<sup>1</sup>See, however, al-Nowaihi and Dhami (2008b) who show that the gain-loss asymmetry can be explained by different discount rates in gains and losses. To the best of our knowledge, the two alternative explanations have not been disentangled by the empirical evidence.

<sup>2</sup>Experiments on animals and humans reveal a hyperbolically declining pattern of discount rates, giving rise to *hyperbolic discounting* (Ainslie, 1975, 1992; Thaler, 1981).

risky (Andreoni and Sprenger, 2012; Dhimi, 2016; Sections 2.12, 10.3.3); concave utility functions may impart caution that is mistakenly attributed to discount rates (Andersen et al., 2010; Andreoni and Sprenger, 2012); and, time preferences may not be stationary (Halevy, 2015). Dohmen et al. (2017) find that after controlling for all these potential confounds, and using large representative samples, temporal behavior is strongly consistent with subadditive discounting. This calls for greater theoretical explorations of the implications of subadditivity in economics.

A range of other anomalies of the EDU model have been identified. For a recent comprehensive survey and attempts to explain these anomalies, we refer the reader to Dhimi (2016; Part 3).

In this paper, we consider consumption streams under certainty.<sup>3</sup> We propose a *general discounted utility model* that differs from the standard EDU model in several respects that we outline more fully in Section 2. Here, we highlight some important differences.

**1.** The utility function takes the Köszegi-Rabin form, which nests the prospect theory gain-loss utility function, and also depends on a stream of date-contingent, but exogenous, reference outcomes.<sup>4</sup>

**2.** We allow the discount function to exhibit declining, constant, or increasing impatience; and subadditivity, additivity, or superadditivity.<sup>5</sup> We use an *interval discount function* that is more general than the Read-Scholten discount function discussed above. For dates  $0 \leq r \leq t$ , we allow outcomes at time  $t$  to be discounted back to time  $r$ , rather than just to time 0. Consider some simple examples of  $r > 0$ .

**Example 1 :** *A firm considers an investment in a production technology that becomes operational at some future time  $r > 0$ , but the benefit to the firm accrues at time  $t > r$ . Then, the firm might be interested in discounting the benefit to time  $r$  rather than to time 0. Or, consider a student deciding on University education with future returns, financed by loans that are repayable at time  $r > 0$  when the education is complete, as is the case in the UK. The student might wish to discount the benefits of the education to time  $r$ , rather than to time 0.*

**3.** We allow for *frame-dependent preferences*. A frame, in our model, consists very generally of the following elements: Streams of *actual outcome-time pairs*; *reference outcome-time pairs*; and a *sequence of increasing moments of time*.

---

<sup>3</sup>For this reason, our paper does not have a bearing on the exciting new literature that combines risk preferences with time preferences. For a recent survey of these issues, see Dhimi (2016, Sections 2.12 and 10.3.3).

<sup>4</sup>The issue of endogenous reference points arises in the presence of uncertainty, but here we only consider the case of certainty.

<sup>5</sup>Thus, we can account for the common difference effect either through declining impatience, subadditivity, or both.

We now describe the main contributions of our paper in the next four subsections, 1.1 to 1.4.

## 1.1 Contribution 1: Results for general discount functions

In Section 4 we extend the most commonly used time discount functions to allow for the case  $r \geq 0$ , and examine their *impatience* and *additivity* properties. In Section 6, we consider general discount functions and show that any additive discount function that exhibits constant impatience must be stationary. We also show that any discount function that is (i) additive but exhibits declining impatience, or (ii) is subadditive but exhibits constant impatience, or (iii) is subadditive and exhibits declining impatience, must exhibit the common difference effect.

Some of these results were already known for specific functional forms, such as for the hyperbolic discounting function (Loewenstein and Prelec, 1992), and for the Read-Scholten discount function (Scholten and Read, 2006a). Our contribution is to show that these results hold for quite general functional forms. Our second contribution is to show that these results hold not just for the case  $r = 0$ , but for any  $r \geq 0$  (recall that  $r$  is the time to which future outcomes are discounted back).

## 1.2 Contribution 2: Foundations for discount functions in terms of observable functions

Individual decision making over time depends jointly on the discount function and the utility function. In practice we do not directly observe either. The question then is: How can we empirically determine whether the behavior of a decision maker exhibits any of the important properties, such as stationarity, impatience and additivity, if the discount function is unobservable? To answer this question we define two functions in subsection 6.1: The *delay function*,  $D$ , and the *speedup function*,  $S$ . Each of these two functions can be uniquely elicited from behavior (Propositions 14 and 16, respectively). These two functions, together, determine whether the behavior of a decision maker exhibits any of the properties of interest. We define a third function, the *generating function*,  $G$ . The generating function,  $G$ , links the speedup function,  $S$ , to the discount function,  $\delta$ .

The functions,  $D$ ,  $S$ ,  $G$ , are defined more formally below but we give some intuition here. Suppose that a decision maker expresses indifference between a *smaller-sooner* (SS) and a *larger-later* (LL) outcome-time pair. The delay function,  $D$ , asks: If the SS reward is delayed, how much do we need to delay the LL reward to maintain indifference? The speedup function,  $S$ , asks: If the SS reward is brought forward to the present, then by how much should the LL reward be brought forward to maintain indifference? The generating

function,  $G$ , is identical to the discount function in the special case where we discount back to the current time (i.e.,  $r = 0$ ).

Among other results, we show the following (Sections 6.2 and 6.3): (1) The speedup and generating functions,  $S$  and  $G$ , jointly determine a unique discount function. Conversely, a continuous discount function determines unique speedup and generating functions,  $S$  and  $G$ . (2) The delay function  $D$  determines stationarity, the common difference effect and the converse of the common difference effect. (3) The speedup function  $S$  determines declining impatience, constant impatience and increasing impatience. (4) The delay function,  $D$ , and the speedup function,  $S$ , jointly determine subadditivity, additivity and superadditivity. In other words, our observable functions provide a solid foundation for understanding the microfoundations of time discount functions.

### 1.3 Contribution 3: Relation between attribute models and time-discounting models

There have been rapid developments in the class of *attribute-based models* that can also explain various anomalies of the EDU model.<sup>6</sup> The motivating principle in this literature is that evaluating streams of consumption profiles using the classical time-discounting models is cognitively too challenging. Hence, in evaluating smaller-sooner (SS) and larger-later (LL) outcome-time choices, decision makers employ a range of simple heuristics to directly compare the time dimension against the outcome dimension.

It is commonly believed that the attribute-based models is a different class of models to discounting-utility models of time-discounting (or delay-discounting models). For instance, Ericson et al. (2015) write: “Critically, however, our results show that, when confronted with intertemporal choices involving money, people use heuristics rather than systematic discounting to make their decisions.”

In Section 7 we show that the attribute-based models are in fact closely related to discounted-utility models. We show formally that there exists a discounted-utility model corresponding to most attribute-based models such that the predictions of the two models are observationally equivalent. If accepted, then this result not only combines disparate areas in the literature but also enriches the psychological foundations of existing time discounting models.

---

<sup>6</sup>For a partial list of papers, see Leland (2002), Rubinstein (2003), Manzini and Mariotti (2006), Read and Scholten (2010) and Ericson et al. (2015).



## 1.4 Contribution 4: Explaining apparent intransitivities through frame dependence

In Section 8 we show that apparent intransitivities in choices over outcome-time pairs can potentially be reconciled by changes in the ‘frame’ of a problem (recall that we allow for frame dependent preferences). As an example, we use a problem similar to that proposed by Ok and Masatlioglu (2017, p. 215) in their theory of intransitive choices. In contrast, we use an underlying model that respects transitivity within a specific frame. Apparent intransitivity may then arise if we do not recognize the change in frame. For an informal argument along these lines, see Gintis (2009). We do not carry out a stringent empirical test of our theoretical prediction. We leave this to future research.

More generally, we note that our framework can account for some of the well-known anomalies of the EDU model. In this paper we focus mainly on the common difference effect anomaly and intransitivity of time discounting. However, our utility function is general enough to also explain gain-loss asymmetry, the magnitude effect (Loewenstein and Prelec, 1992; al-Nowaihi and Dhami, 2009) and delay-speedup asymmetry (al-Nowaihi and Dhami, 2008b). Furthermore, our more general framework should increase the scope of applications by allowing for the case  $r \geq 0$  and by suggesting a useful definition of a *frame*.

All proofs are in the Appendix.

## 2 Formulation

In this section, we formulate a general delay-discount model. Our basic ingredients are a *frame*,  $\mathcal{F}$ , a *utility function*,  $u$ , and a *discount function*,  $\delta$ . We describe each of these below.

**Notation:** We shall use plain lowercase Latin letters, e.g.,  $r, s, t, x, y, z$ , for real numbers, bold lowercase Latin letters, e.g.,  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ , for real vectors, bold uppercase Latin letters, e.g.,  $\mathbf{Y}$  and  $\mathbf{Z}$ , for streams of dated real vectors and bold uppercase Greek letters, e.g.,  $\Sigma$ , for sets of streams of dated real vectors.

### 2.1 Frame, $\mathcal{F}$

A *frame* consists of the following.<sup>7</sup>

---

<sup>7</sup>In the literature, *frame* and *context* are used interchangeably, and neither is formally defined. In this paper, we use *frame* in the precise sense of this subsection. This is in agreement with common usage. For example, a change in the reference point for outcome is described as a change in frame. We shall continue to use *context* in its more general sense as used in the literature.

1. A sequence of increasing moments of time

$$0 \leq t_1 < t_2 < \dots < t_n, \quad (1)$$

where outcome  $\mathbf{y}_i \in \mathbb{R}_+^m$  is received at time  $t_i$ ,  $i = 1, 2, \dots, n$ .

2. A set,  $\Sigma$ , of streams of outcome-time pairs that the decision maker is interested in,

$$\Sigma = \{ \langle (\mathbf{y}_1, t_1), (\mathbf{y}_2, t_2), \dots, (\mathbf{y}_i, t_i), \dots, (\mathbf{y}_n, t_n) \rangle \}, \quad (2)$$

where  $\mathbf{y}_i \in \mathbb{R}_+^m$ ,  $i = 1, 2, \dots, n$  and  $t_i$ ,  $i = 1, 2, \dots, n$  are as in (1).

3. An exogenously given *reference stream* of outcome-time pairs

$$\mathbf{Z} = \langle (\mathbf{z}_1, t_1), (\mathbf{z}_2, t_2), \dots, (\mathbf{z}_i, t_i), \dots, (\mathbf{z}_n, t_n) \rangle, \quad (3)$$

where  $\mathbf{z}_i \in \mathbb{R}_+^m$ ,  $i = 1, 2, \dots, n$  and  $\mathbf{Z} \in \Sigma$ .

$\mathbf{Z}$  could be an individual's anticipated stream of future incomes based, perhaps, on one's own historical or the peer-observed stream of outcome-time pairs.

## 2.2 Utility function, $u$

**Definition 1** (*Material utility function,  $\mu$* ): An instantaneous, time-invariant, material utility function (or simply material utility) is a continuous function  $\mu : \mathbb{R}_+^m \rightarrow \mathbb{R}_+$  satisfying:

- (a)  $\mu(\mathbf{0}) = 0$ ,
- (b)  $\mu(x_1, x_2, \dots, x_m)$  is strictly increasing in each  $x_i$ ,  $i = 1, 2, \dots, m$ .

At any instant in time, we have the vectors of actual and reference consumption, respectively,  $\mathbf{y}$  and  $\mathbf{z}$  (see (2) and (3) above). As in prospect theory, particularly in its reformulation in Köszegi-Rabin, we now define a reference dependent utility function, defined over pairs of vectors  $(\mathbf{y}, \mathbf{z}) \in \mathbb{R}_+^{2m}$ , as follows.

**Definition 2** (*Utility function,  $u$* ): Let  $\gamma \in [0, 1]$ . We define the utility function,  $u : \mathbb{R}_+^{2m} \rightarrow \mathbb{R}$ , that is the object of a decision maker's optimization, as follows.

$$u(\mathbf{y}, \mathbf{z}) = (1 - \gamma) \mu(\mathbf{y}) + \gamma [\mu(\mathbf{y}) - \mu(\mathbf{z})], \text{ for } \mu(\mathbf{y}) \geq \mu(\mathbf{z}), \quad (4)$$

$$u(\mathbf{y}, \mathbf{z}) = (1 - \gamma) \mu(\mathbf{y}) + \lambda \gamma [\mu(\mathbf{y}) - \mu(\mathbf{z})], \text{ for } \mu(\mathbf{y}) < \mu(\mathbf{z}), \lambda \geq 1. \quad (5)$$

The term  $\mu(\mathbf{y}) - \mu(\mathbf{z})$  is called *gain-loss utility*. In case (4), we see that utility,  $u(\mathbf{y}, \mathbf{z})$ , is a convex combination of material utility,  $\mu(\mathbf{y})$ , and gain-loss utility,  $\mu(\mathbf{y}) - \mu(\mathbf{z})$ . In this case, the decision maker is said to be in the *domain of gains*. By contrast, in case

(5), the decision maker is said to be in the *domain of losses*.  $\lambda \geq 1$  is called the *coefficient of loss-aversion*. In Definition 1, material utility  $\mu$  is always non-negative. However, in Definition 2, utility can take positive or negative values on account of gain-loss utility.

Consider a frame  $\mathcal{F}$ , as defined above, and let  $\mathbf{Y} = \langle (\mathbf{y}_1, t_1), \dots, (\mathbf{y}_i, t_i), \dots, (\mathbf{y}_n, t_n) \rangle \in \Sigma$ . Suppose that at time  $t_i$ , the actual outcome is  $\mathbf{y}_i$  and the reference outcome is  $\mathbf{z}_i$ , which we represent by the pair of vectors  $(\mathbf{y}_i, \mathbf{z}_i)$ , then the utility function at time  $t_i$  is given by

$$u(\mathbf{y}_i, \mathbf{z}_i) = (1 - \gamma) \mu(\mathbf{y}_i) + \gamma [\mu(\mathbf{y}_i) - \mu(\mathbf{z}_i)], \text{ if } \mu(\mathbf{y}_i) \geq \mu(\mathbf{z}_i), \quad (6)$$

$$u(\mathbf{y}_i, \mathbf{z}_i) = (1 - \gamma) \mu(\mathbf{y}_i) + \lambda \gamma [\mu(\mathbf{y}_i) - \mu(\mathbf{z}_i)], \text{ if } \mu(\mathbf{y}_i) < \mu(\mathbf{z}_i). \quad (7)$$

**Remark 1 :** *The following special cases are of particular interest.*

- (a) For  $\gamma = 0$ , from (4) and (5), we get  $u(\mathbf{y}, \mathbf{z}) = \mu(\mathbf{y})$ , which is the classical case.
- (b) Suppose  $\mathbf{z} = \mathbf{0}$ . From Definition 1a,  $\mu(\mathbf{z}) = \mu(\mathbf{0}) = 0$ . From Definition 1b,  $\mu(\mathbf{y}) \geq \mu(\mathbf{0}) = 0$  for all  $\mathbf{y} \in \mathbb{R}_+^m$ . From (4), the decision maker is in the domain of gains. Hence, we get

$$\begin{aligned} u(\mathbf{y}, \mathbf{0}) &= (1 - \gamma) \mu(\mathbf{y}) + \gamma [\mu(\mathbf{y}) - \mu(\mathbf{0})], \\ &= (1 - \gamma) \mu(\mathbf{y}) + \gamma \mu(\mathbf{y}), \\ &= \mu(\mathbf{y}), \end{aligned} \quad (8)$$

which, again, is the classical case.

Thus, the classical case is a special case of the reference dependent case when  $\gamma = 0$  or when  $\mathbf{z} = \mathbf{0}$ .

- (c)  $\gamma = 1$  gives the case of pure gain-loss utility, which is the case considered in prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).

There is considerable evidence for loss-aversion,  $\lambda > 1$ , in humans and close primate relatives. A large number of observed phenomena that are problematic for standard decision theory ( $\gamma = 0$  or  $\mathbf{z} = \mathbf{0}$ ) can be explained by loss-aversion (Kahneman and Tversky, 2000; Dharni, 2016).  $\lambda$  can vary with context, mood, prior history of losses, and age (Dharni, 2016).

In their seminal work, Kahneman and Tversky (1979) and Tversky and Kahneman (1992) consider the one-dimensional case ( $m = 1$ ), and take the carrier of utility to be the difference between the actual outcome,  $y$ , and the reference outcome  $z$ . Thus, in our terminology,  $u(y, z) = \mu(y - z)$ . If  $y - z \geq 0$ , then the decision maker is in the *domain of gains*. If  $y - z < 0$ , then the decision maker is in the *domain of losses*. However, in more than one dimensions, the following problem occurs. Consider a two-dimensional case ( $m = 2$ ). Suppose  $y_1 - z_1 > 0$  and  $y_2 - z_2 < 0$ . Overall, is the decision maker in the domain

of gains or the domain of losses?<sup>8</sup> One solution to this issue, that we use in our paper, following Köszegi and Rabin (2006, 2007), is to take gain-loss utility to be the difference between material utility,  $\mu(\mathbf{y})$ , and a reference utility level,  $\mu(\mathbf{z})$ .

Our formulation in (4), (5) is inspired by Köszegi and Rabin (2006, 2007). They use the specification (in our terminology):  $u(\mathbf{y}, \mathbf{z}) = \mu(\mathbf{y}) + \gamma(\mu(\mathbf{y}) - \mu(\mathbf{z}))$ , where  $\gamma$  is a *function* (not a parameter) with appropriate properties to capture (among others) loss aversion. In this respect, their approach is more general than ours. However, they do not cover the special case of pure gain-loss utility (we do by setting our  $\gamma = 1$ , in (4), (5)). All our results extend to their framework.

### 2.3 Discount function, $\delta$

We begin by defining a general discount function. This is an *interval discount function* in the sense of Read (2001) and Scholten and Read (2006a).

**Definition 3** (*Discount functions*): Let

$$\Delta = \{(r, t) \in \mathbb{R} \times \mathbb{R} : 0 \leq r \leq t\}. \quad (9)$$

A discount function is a mapping,  $\delta : \Delta \rightarrow (0, 1]$ , satisfying:

- (a) For each  $r \in [0, \infty)$ ,  $\delta(r, t)$  is a strictly decreasing function of  $t \in [r, \infty)$  into  $(0, 1]$  with  $\delta(r, r) = 1$ .
- (b) For each  $t \in [r, \infty)$ ,  $\delta(r, t)$  is a strictly increasing function of  $r \in [0, t]$  into  $(0, 1]$ .
- (c) If  $\delta$  satisfies (a) with ‘into’ replaced with ‘onto’, then we call  $\delta$  a continuous discount function.

An example is the exponential discount function  $\delta(r, t) = e^{-\beta(t-r)}$ ,  $\beta > 0$ . Setting  $r = 0$ , we get the commonly used form  $\delta(r, t) = e^{-\beta t}$ , which is the special case when future outcomes are discounted back to the present, i.e., to time  $r = 0$ . However, our more general framework allows outcomes to be discounted back to a potentially future time  $r \geq 0$ , where  $r \leq t$ .

**Definition 4** (*Discounted utility*): Consider a frame  $\mathcal{F}$ , as defined above in subsection 2.1, and let  $\mathbf{Y} = \langle (\mathbf{y}_1, t_1), \dots, (\mathbf{y}_i, t_i), \dots, (\mathbf{y}_n, t_n) \rangle \in \Sigma$ . Let  $u$  be a utility function (subsection 2.2) and  $\delta$  a discount function (Definition 3). Then, the decision maker’s discounted utility, discounted back to time  $r$ , is

$$U_r(\mathbf{Y}, \mathbf{Z}) = \sum_{i=1}^n \delta(r, t_i) u(\mathbf{y}_i, \mathbf{z}_i). \quad (10)$$

---

<sup>8</sup>See Tversky and Kahneman (1991) for their suggested solution.

**Remark 2** (*Determination of the reference stream of outcome-time pairs*): None of the results in this paper depend on how the reference stream of outcome-time pairs,  $\mathbf{Z}$  in (3), is determined. However, a few comments may be in order. Usually a reasonable choice of the reference stream of outcome-time pairs suggests itself in applications. This could be, for example, the status quo, a legal or social entitlement, a fair outcome, or an expected outcome (Dhami, 2016; Section 2.4.4).

Kőszegi and Rabin (2006, 2007) consider the possibility of endogenous reference points. However, for this, they also require that subjects form rational expectations, which is a strong restriction. Experimentally, we do find that expectations influence the reference point, but we do not have evidence, yet, if these expectations are consistent with rational expectations (Dhami, 2016, Section 2.8). Furthermore, in this paper we are interested in riskless temporal choices, hence, the issue of an endogenous reference point in the sense of Kőszegi and Rabin (2006, 2007) does not arise. We leave an extension of our paper to risky choices for future work.

## 2.4 General properties of discount functions

We now derive general properties of discount functions (Definition 3). Our terminology suggests that a continuous discount function (Definition 3c) is continuous. That this is partly true, is established by the following Proposition.

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

**Proposition 2** (*Time sensitivity*): *Let  $\delta$  be a continuous discount function. Suppose  $r \geq 0$ . If  $0 < u(\mathbf{x}, \mathbf{z}) \leq u(\mathbf{y}, \mathbf{z})$ , or if  $u(\mathbf{y}, \mathbf{z}) \leq u(\mathbf{x}, \mathbf{z}) < 0$ , then  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$  for some  $t \in [r, \infty)$ .*

**Proposition 3** (*Existence of present values*): *Let  $\delta$  be a discount function. Let  $\mathbf{z} \in \mathbb{R}_+^m$  and  $0 \leq r \leq t$ . Then, for each  $\mathbf{y} \in \mathbb{R}_+^m$ , there is an  $\mathbf{x} \in \mathbb{R}_+^m$  such that  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ .*

## 3 Special properties of discount functions

In this section, we define nine properties that the behavior of a decision maker may exhibit: Stationarity, common difference effect, converse of common difference effect, declining impatience, constant impatience, increasing impatience, subadditivity, additivity and superadditivity.

To motivate these properties, consider the following example from Thaler (1981). Suppose that a decision maker is indifferent (today, at time 0) between one apple today and

two apples tomorrow. Assuming that the reference outcome is zero apples,  $z = 0$ , we may write this indifference in terms of our framework as  $u(1,0) = u(2,0)\delta(0,1)$ . Then, *stationarity* (Definition 5, below), implies that this decision maker must also be indifferent (today at time 0) between one apple in 50 days' time and two apples in 51 days' time, i.e.,  $u(1,0)\delta(0,0+50) = u(2,0)\delta(0,1+50)$ . Thus, the preferences are unaltered if, *ceteris-paribus*, we move both outcome-time pairs by the same fixed time distance; this feature plays a central role in the exponential discounting model.

**Definition 5** (*Stationarity*): *Stationarity holds if for all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^m$  and all  $r \geq 0, s \geq r, t \geq r, u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z})\delta(r, s)$  implies  $u(\mathbf{x}, \mathbf{z})\delta(r, t) = u(\mathbf{y}, \mathbf{z})\delta(r, s+t-r)$ ; where  $\mathbf{z} \in \mathbb{R}_+^m$  is a given reference point for outcomes.*

Noting that  $\delta(r, r) = 1$  (Definition 3), Definition 5 begins with the indifference  $u(\mathbf{x}, \mathbf{z})\delta(r, r) = u(\mathbf{y}, \mathbf{z})\delta(r, s)$ . It then requires that keeping the date  $r \geq 0$  to which the outcomes are discounted and the reference outcome  $\mathbf{z}$  fixed, but simultaneously moving the time forward by  $t - r$  for both outcomes  $\mathbf{x}$  and  $\mathbf{y}$ , leaves preferences between the two outcomes unaltered (as in the apples example, above). On the other hand, a decision maker violates stationarity if he is indifferent between one apple today and two apples tomorrow, but prefers two apples in 51 days' time to one apple in 50 days' time (i.e.,  $u(1,0)\delta(0,0) = u(2,0)\delta(0,1)$  but  $u(1,0)\delta(0,0+50) < u(2,0)\delta(0,1+50)$ ); this sometimes known as *Thaler's apples paradox*. In Definition 6 next, we term such a violation of stationarity as the *common difference effect*.

**Definition 6** (*Common difference effect*): *The common difference effect arises if, for all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^m$  and all  $r \geq 0, s > r, t > r, u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z})\delta(r, s)$  implies*

- (i)  $u(\mathbf{x}, \mathbf{z})\delta(r, t) < u(\mathbf{y}, \mathbf{z})\delta(r, s+t-r)$  for  $u(\mathbf{x}, \mathbf{z}) > 0$ , and
- (ii)  $u(\mathbf{x}, \mathbf{z})\delta(r, t) > u(\mathbf{y}, \mathbf{z})\delta(r, s+t-r)$  for  $u(\mathbf{x}, \mathbf{z}) < 0$ ,

*where  $\mathbf{z} \in \mathbb{R}_+^m$  is a given reference point for outcomes.*

As far as we know, the converse of the common difference effect has never been observed. However, for completeness, we give a definition.

**Definition 7** (*The converse of the common difference effect*): *The converse of the common difference effect holds if the first inequality in each of (i) and (ii) of Definition 6 is reversed.*

One interpretation of Definition 6 is that individuals who exhibit the common difference effect are impatient at short horizons, but relatively patient at long horizons. 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 8** (*Impatience*): A discount function,  $\delta(r, s)$ , exhibits<sup>9</sup>

$$\left\{ \begin{array}{ll} \text{Declining impatience if} & \delta(r, s) < \delta(r+t, s+t), \text{ for } t > 0 \text{ and } 0 \leq r < s, \\ \text{Constant impatience if} & \delta(r, s) = \delta(r+t, s+t), \text{ for } t \geq 0 \text{ and } 0 \leq r \leq s, \\ \text{Increasing impatience if} & \delta(r, s) > \delta(r+t, s+t), \text{ for } t > 0 \text{ and } 0 \leq r < s. \end{array} \right.$$

Does splitting an interval into subintervals influence the extent of discounting? The empirical evidence suggests that it does (Read, 2001; Dohmen et al., 2017). The next definition considers the various possibilities.

**Definition 9** (*Additivity*): A discount function,  $\delta(r, t)$ , is

$$\left\{ \begin{array}{ll} \text{Subadditive if} & \delta(r, s) \delta(s, t) < \delta(r, t), \text{ for } 0 \leq r < s < t, \\ \text{Additive if} & \delta(r, s) \delta(s, t) = \delta(r, t), \text{ for } 0 \leq r \leq s \leq t, \\ \text{Superadditive if} & \delta(r, s) \delta(s, t) > \delta(r, t), \text{ for } 0 \leq r < s < t. \end{array} \right.$$

In Definition 9, 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. This is the case for exponential discounting. However, in the other two cases, it does have an effect. Under subadditive discounting, there is more discounting over the subdivided intervals. The converse is true under superadditive discounting.

A possible psychological explanation of subadditivity is that splitting an interval into subintervals may make the passage of time more salient (counting the seconds makes a minute appear longer!).

## 4 Examples of discount functions

We now consider several important discount functions and examine some of their properties. We allow for the more general case  $r \geq 0$ . The standard forms of these functions, for  $r = 0$ , are easily recovered as special cases.

### 4.1 Exponential Discount Function

The *exponential discounting function* (EDF), introduced by Samuelson (1937), is the unique discount function that leads to time consistent choices (Strotz, 1956). It can be written as

$$\delta(r, t) = e^{-\beta(t-r)}, \beta > 0. \tag{11}$$

---

<sup>9</sup>Some authors use ‘present bias’ for what we call ‘declining impatience’. But other authors use ‘present bias’ to mean that the discount function,  $\delta(s, t)$  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’.

It follows that  $\delta(r+t, s+t) = e^{-\beta[(s+t)-(r+t)]} = e^{-\beta(s-r)} = \delta(r, s)$ . Thus, the EDF exhibits constant impatience. We can write  $\delta(r, t) = e^{-\beta(t-r)} = e^{-\beta[(t-s)-(r-s)]} = e^{-\beta(s-r)}e^{-\beta(t-s)} = \delta(r, s)\delta(s, t)$ . Thus, exponential discounting is additive. Finally, exponential discounting generates stationarity, as can be seen from the following argument. Let  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}_+^m$ ,  $r \geq 0, s \geq r, t \geq r$  and  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z})e^{-\beta(s-r)}$ . Then  $u(\mathbf{x}, \mathbf{z})\delta(r, t) = u(\mathbf{x}, \mathbf{z})e^{-\beta(t-r)} = u(\mathbf{y}, \mathbf{z})e^{-\beta(s-r)}e^{-\beta(t-r)} = u(\mathbf{y}, \mathbf{z})e^{-\beta s + \beta r - \beta t + \beta r} = u(\mathbf{y}, \mathbf{z})e^{-\beta((s+t-r)-r)} = u(\mathbf{y}, \mathbf{z})\delta(r, s+t-r)$ .

## 4.2 Quasi hyperbolic discounting (PPL discount function)

The *quasi hyperbolic discount function* is popular in applied work (Phelps and Pollack, 1968; Laibson, 1997); we use the acronym PPL discount function after the names of its authors. While the PPL is normally presented in discrete time, it is convenient to present the continuous time analogue, which is given by

$$\delta(r, t) = \begin{cases} 1 & \text{if } r = t = 0 \\ e^{-(\alpha+\beta t)}, \beta > 0, \alpha > 0, & \text{if } 0 = r < t \\ e^{-\beta(t-r)}, \beta > 0 & \text{if } 0 < r \leq t \end{cases} \quad (12)$$

The EDF and PPL embody an identical degree of impatience for future choices, but different degrees of impatience when one makes a choice between an immediate reward and a future reward.

I. The case  $0 < r \leq t$ . From (11) and (12) note that, for  $0 < r \leq t$ , PPL is identical to EDF. Hence, when discounting back to a future date,  $r > 0$ , PPL is additive and exhibits constant impatience.

II. The case  $r = 0$ . Let  $0 < s < t$ .

Then  $\delta(0, s)\delta(s, t) = e^{-(\alpha+\beta s)}e^{-\beta(t-s)} = e^{-(\alpha+\beta t)} = \delta(0, t)$ . Thus, PPL is also additive when discounting back to the present. Since additivity holds for the cases  $r > 0$  and,  $r = 0$ , hence, PPL is additive.

From (12), we get  $\delta(0, s) = e^{-(\alpha+\beta s)} < e^{-\beta s} = \delta(t, s+t)$ , since  $\alpha > 0$ . Hence, PPL exhibits declining impatience when discounting back to the present ( $r = 0$ ), but exhibits constant impatience when discounting back to a future time ( $0 < r \leq t$ ).

## 4.3 The generalized hyperbolic discount function (LP discount function)

We use the acronym LP for the *generalized hyperbolic discount function* proposed by Loewenstein and Prelec (1992). It is axiomatically founded (Loewenstein and Prelec, 1992, al-Nowaihi and Dhami, 2006, 2008a), and is given by

$$\delta(r, t) = \left( \frac{1 + \alpha t}{1 + \alpha r} \right)^{-\frac{\beta}{\alpha}}, \quad t \geq 0, r \geq 0, \alpha > 0, \beta > 0. \quad (13)$$



Unlike the quasilinear form of hyperbolic discounting that exhibits declining impatience for  $r = 0$  and constant impatience thereafter, the LP discount function exhibits *declining impatience* throughout. To see this, note that for  $0 \leq r < s \leq t$  we have  $\frac{1+\alpha(s+t)}{1+\alpha(r+t)} < \frac{1+\alpha s}{1+\alpha r}$ . Since  $-\frac{\beta}{\alpha} < 0$ , it follows that  $\delta(r, s) < \delta(r+t, s+t)$ . Thus, the LP function exhibits declining impatience throughout.

By a simple calculation, for  $0 \leq r \leq s \leq t$ , we have  $\left(\frac{1+\alpha s}{1+\alpha r}\right)^{-\frac{\beta}{\alpha}} \left(\frac{1+\alpha t}{1+\alpha s}\right)^{-\frac{\beta}{\alpha}} = \left(\frac{1+\alpha t}{1+\alpha r}\right)^{-\frac{\beta}{\alpha}}$ , or  $\delta(r, s)\delta(s, t) = \delta(r, t)$ . Thus, the LP function, like the PPL function, is *additive* throughout.

#### 4.4 The interval discount function (RS discount function)

The three discount functions that we have considered so far are *additive* (Definition 9). We use the acronym RS for the particular interval discount function introduced by Read (2001) and Scholten and Read (2006a). It is given by

$$\delta(r, t) = [1 + \alpha(t^\tau - r^\tau)^\rho]^{-\frac{\beta}{\alpha}}, \quad 0 \leq r \leq t, \alpha > 0, \beta > 0, \rho > 0, \tau > 0. \quad (14)$$

The RS discount function is particularly flexible. Depending on the parameter values, it allows for constant, increasing or decreasing impatience, as well as subadditivity; the next proposition makes this precise.

**Proposition 4 :** *Let  $\delta(r, t)$  be the RS-discount function (14), then:*

- (a) *If  $0 < \rho \leq 1$ , then  $\delta$  is subadditive.*
- (b) *If  $\rho > 1$ , then  $\delta$  is neither subadditive, additive nor superadditive.*
- (ci) *If  $0 < \tau < 1$ , then  $\delta$  exhibits declining impatience.*
- (cii) *If  $\tau = 1$ , then  $\delta$  exhibits constant impatience.*
- (ciii) *If  $\tau > 1$ , then  $\delta$  exhibits increasing impatience.*

In the light of Proposition 4, we can now see the interpretation of the parameters  $\tau$  and  $\rho$  in the RS discount function.<sup>10</sup>  $\tau$  controls impatience, independently of the values of the other parameters  $\alpha, \beta$  and  $\rho$ .  $0 < \tau < 1$ , 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  $\alpha, \beta$  and  $\tau$ . However, if  $\rho > 1$ , then the RS discount function can be neither subadditive, additive nor superadditive.<sup>11</sup>

<sup>10</sup>Scholten and Read (2006a), bottom of p1425, state:  $\alpha > 0$  implies subadditivity (incorrect),  $\rho > 1$  implies superadditivity (incorrect) and  $0 < \tau < 1$  implies declining impatience (correct but incomplete).

<sup>11</sup>In this case, depending on the particular values of  $r, s$  and  $t$ , we may have  $\delta(r, s)\delta(s, t) < \delta(r, t)$ ,  $\delta(r, s)\delta(s, t) = \delta(r, t)$  or  $\delta(r, s)\delta(s, t) > \delta(r, t)$ .

## 4.5 The generalized Read-Scholten discount function (GRS)

Scholten and Read (2006b, 2010, 2016) present a critique of the psychological basis for discounting models. They develop an *attribute model* that is based on firmer psychological foundations.<sup>12</sup> al-Nowaihi and Dhami (2008b, subsection 5.1, pp. 38-40) proposed the GRS discount function (GRS is an acronym for *generalized* Read-Scholten) and argued that the Scholten and Read (2006b, 2010) trade-off model is equivalent to a discounted utility model with the GRS discount function (see Section 7 for a generalization of this result). Hence, the arguments in Scholten and Read (2006b, 2010) in support of their trade-off model also lend further support to their own discount function, the RS discount function (14), and its generalization, GRS.

The GRS discount function is given by

$$\delta(r, t) = e^{-Q(\phi(t) - \phi(r))}, 0 \leq r \leq t, \text{ where} \quad (15)$$

$$Q : [0, \infty) \xrightarrow{\text{onto}} [0, \infty) \text{ is strictly increasing and } \phi : [0, \infty) \xrightarrow{\text{onto}} [0, \infty) \text{ is strictly increasing.} \quad (16)$$

We only use the GRS discount function in Section 7.

## 4.6 Further remarks on the discount functions

Note that LP approaches the exponential as  $\alpha \rightarrow 0$ . In general, neither LP nor RS is a special case of the other. However, for  $r = 0$  (and only for  $r = 0$ ), RS reduces to LP when  $\rho = \tau = 1$ .

It is straightforward to check that each of the discount functions, the exponential, LP, RS, and GRS, is a continuous discount function in the sense of Definition 3. It is also straightforward to check that PPL is a discount function, but it is not a *continuous* discount function because  $\lim_{t \rightarrow 0+} \delta(0, t) = e^{-\alpha} < 1 = \delta(0, 0)$ . This arises on account of the sudden and discontinuous increase in impatience in this discount function, as noted above.

Note that the restrictions  $r \geq 0$  and  $t \geq 0$  are needed in LP and RS. From LP we see that the further restriction  $r \leq t$  is needed.<sup>13</sup> From PPL we see that the ‘into’ in Definition 3(b) cannot be strengthened to ‘onto’.

Exponential discounting exhibits constant impatience while LP exhibits declining impatience, but both are additive. The RS discount function allows for additivity, subadditivity and all the three cases in Definition 8, hence, it is of great practical importance.

<sup>12</sup>In outcome-time pairs of the form  $(y_i, z_i)$ , there are three attributes. The actual outcome and the reference outcomes,  $y_i, z_i$ , respectively, and the time  $t_i$  at which the outcome is received. Attribute based models describe choices between outcome-time pairs in terms of comparing the attributes of time or outcome in some order, using various threshold rules of comparison.

<sup>13</sup>One alternative is to *define*  $\delta(t, s)$  to be  $1/\delta(s, t)$ . 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.

## 5 Explaining the common difference effect

In Section 4, above, we saw that the exponential discount function (EDF) is additive, exhibits constant impatience and generates stationarity. We also saw that the generalized hyperbolic discount function (LP), which is additive, explains the common difference effect through declining impatience. Finally, we saw that the interval discount function (RS) can explain the common difference effect through either declining impatience, subadditivity or a combination of both. We now extend these properties of these specific discount functions to general discount functions that allow for discounting back to a future date,  $r > 0$ .

**Proposition 5** : *An additive discount function that exhibits constant impatience is stationary.*

Of the discount functions that we have considered in Section 4, only the exponential function exhibits additivity, constant impatience, and stationarity. Our contribution here is to show that ‘any’ additive discount function that exhibits constant impatience is stationary.

**Proposition 6** : *An additive discount function that exhibits declining impatience exhibits the common difference effect.*

The LP discount function (Section 4.3) exhibits additivity, declining impatience and the common difference effect, hence, it can explain Thaler’s apples paradox that we outlined above. Proposition 6 proves that ‘any’ discount function that exhibits additivity and declining impatience also exhibits the common difference effect. The reader can check that the PPL discount function can also explain Thaler’s apples paradox.

**Proposition 7** : *A subadditive discount function that exhibits constant impatience also exhibits the common difference effect.*

**Proposition 8** : *A subadditive discount function that exhibits declining impatience also exhibits the common difference effect.*

It is well known that for some parameter values, the RS discount function can exhibit the common difference effect either through subadditivity, declining impatience, or a combination of both (Read, 2001; Scholten and Read, 2006a). Propositions 7, 8 extend this result to ‘any’ discount function: Any subadditive discount function that exhibits constant or declining impatience must also exhibit the common difference effect. In particular, Proposition 4 showed that the RS discount function is subadditive if  $0 < \rho \leq 1$  and exhibits constant or declining impatience if  $0 < \tau \leq 1$ . Hence, by Propositions 7, 8, the RS discount function exhibits the common difference effect if  $0 < \rho \leq 1$  and  $0 < \tau \leq 1$ .

For completeness, we also state the following three propositions that give the conditions for the converse of the common difference effect (Definition 7); we are not aware of any empirical evidence for this effect.

**Proposition 9** : *An additive discount function that exhibits increasing impatience exhibits the converse of the common difference effect.*

**Proposition 10** : *A superadditive discount function that exhibits constant impatience exhibits the converse of the common difference effect.*

**Proposition 11** : *A superadditive discount function that exhibits increasing impatience exhibits the converse of the common difference effect.*

## 6 Foundations for discount functions in terms of observable functions

In Section 3, the nine properties, stationarity, common difference effect, converse of common difference effect, declining impatience, constant impatience, increasing impatience, subadditivity, additivity and superadditivity were all defined in terms of the discount function,  $\delta$ . However, individual decision making over time depends jointly on the discount function and the utility function. In practice, we do not directly observe either. The question then is: How can we empirically determine whether the behavior of a decision maker exhibits any of these nine properties, if the discount function is unobservable? To answer this question, we define two functions in subsection 6.1: The *delay function*,  $D$ , and the *speedup function*,  $S$ , which can both be uniquely elicited from behavior. These two functions, together determine whether the behavior of a decision maker exhibits any of the nine properties of interest. We also define a third function, the generating function,  $G$ . The generating function,  $G$ , links the speedup function,  $S$ , to the discount function,  $\delta$ .

### 6.1 Delay, speedup and generating functions

#### 6.1.1 Delay Function, $D$

Consider a fixed reference point,  $\mathbf{z}$ , for outcomes. Suppose that a decision maker reveals the following indifference:  $\mathbf{x}$  received at time  $r \geq 0$  is equivalent to  $\mathbf{y}$  received at time  $t \geq r$ . This gives

$$u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t). \quad (17)$$

Now suppose that, keeping  $\mathbf{z}$  fixed, the receipt of  $\mathbf{x}$  is *delayed* to time  $s$ ; see Figure 1. We ask, at what time,  $T$ , will  $\mathbf{y}$  received at time  $T$  be equivalent to  $\mathbf{x}$  received at time  $s$ , i.e.,

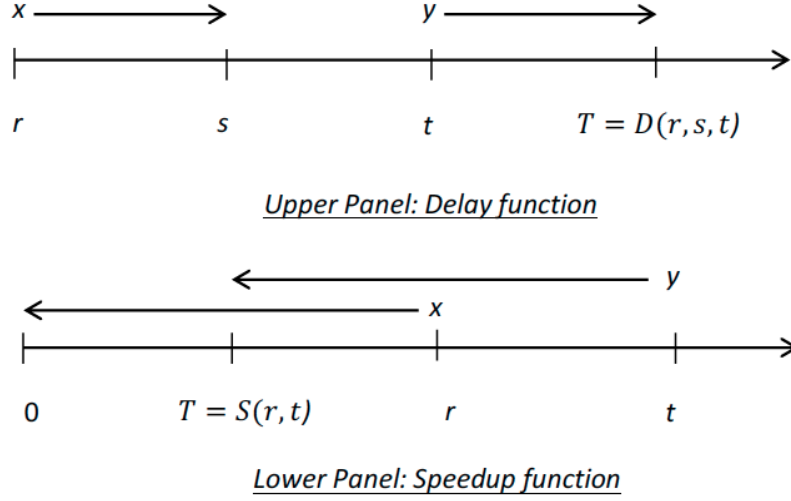


Figure 1: The delay function and the speedup functions.

for what  $T$  does the following hold?

$$u(\mathbf{x}, \mathbf{z}) \delta(r, s) = u(\mathbf{y}, \mathbf{z}) \delta(r, T). \quad (18)$$

Let us conjecture that  $T$  depends on  $r, s, t$  through a functional relation, say,  $T = D(r, s, t)$ . By considering an  $\mathbf{x} \in \mathbb{R}_+^m$  for which  $u(\mathbf{x}, \mathbf{z}) \neq 0$  (and, hence,  $u(\mathbf{y}, \mathbf{z}) \neq 0$ ) we get, from (17) and (18), that  $D(r, s, t)$  must satisfy

$$\delta(r, s) \delta(r, t) = \delta(r, D(r, s, t)). \quad (19)$$

We shall call the function  $D(r, s, t)$ , if it exists, a delay function (see Definition 10, below).

**Example 2 :** For the exponential discount function in (11), we have, from (19),  $e^{-\beta(s-r)}e^{-\beta(t-r)} = e^{-\beta(T-r)}$ , which gives the solution  $T = D(r, s, t) = s + t - r$ . We show below in Proposition 18 that this is a necessary and sufficient condition for stationarity.

More generally, we shall show that such a delay function exists, is unique, and depends on  $r, s, t$ . We shall also examine its properties (see Propositions 12, 13). The delay function is shown in the upper panel of Figure 1.

**Definition 10 (Delay functions):** Let  $\delta$  be a discount function. Suppose that the function,  $D$ , has the property  $\delta(r, s) \delta(r, t) = \delta(r, D(r, s, t))$ ,  $s \geq r, t \geq r$ . Then we call  $D$  a delay function corresponding to the discount function,  $\delta$ .

The next proposition establishes properties of a delay function, if it exists.

**Proposition 12** (*Properties of a delay function*): Let  $\delta$  be a discount function and  $D$  a corresponding delay function. Then  $D$  has the following properties:

- (a)  $D$  is unique,
- (b)  $D(r, s, t)$  is strictly increasing in each of  $s$  and  $t$ ,
- (c)  $D(r, s, t) = D(r, t, s)$ ,
- (d)  $D(r, r, t) = D(r, t, r) = t$ ,
- (e)  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$  if, and only if,  $u(\mathbf{x}, \mathbf{z}) \delta(r, s) = u(\mathbf{y}, \mathbf{z}) \delta(r, D(r, s, t))$ .

The next proposition establishes that a continuous discount function has a delay function.

**Proposition 13** (*Existence of a delay function*): A continuous discount function has a delay function,  $D(r, s, t)$ . Furthermore,  $D(r, s, t)$  is unique and continuous in  $s, t$ .

From (17)–(19), the delay function is easily elicited, say, by using experiments. One possible method is as follows. First, fix  $\mathbf{x}, \mathbf{z}, r$  and  $t$ , where  $0 \leq r \leq t$ . Next, elicit a  $\mathbf{y}$  for which the decision maker expresses indifference between  $\mathbf{x}$  received at time  $r$  and  $\mathbf{y}$  received at time  $t$ . Finally, fix  $s$  and elicit the value  $T$  for which the decision maker expresses indifference between  $\mathbf{x}$  received at time  $s$  and  $\mathbf{y}$  received at time  $T$ . Then  $D(r, s, t) = T$ . This argument, and in the light of Proposition 12a, has established the following proposition.

**Proposition 14** : The delay function,  $D$ , can be uniquely elicited from behavior.

### 6.1.2 Speedup function, $S$

Consider a fixed reference point,  $\mathbf{z}$ , for outcomes. Suppose that  $\mathbf{x}$  received at time  $r$  is equivalent to  $\mathbf{y}$  received at time  $t$ ,  $0 \leq r \leq t$ . This gives

$$u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t). \quad (20)$$

Suppose that, keeping  $\mathbf{z}$  fixed, the receipt of  $\mathbf{x}$  is brought forward from time  $r$  to time 0; see Figure 1. We ask, at what time,  $T$ , will  $\mathbf{y}$  received at time  $T$  be equivalent to  $\mathbf{x}$  received at time 0? Or, for what time,  $T$ , will the following hold?

$$u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(0, T). \quad (21)$$

For the exponential discount function, stationarity implies a clear answer:  $T = t - r$ . More generally, let us conjecture that  $T$  depends on  $r, t$ , so that we can write  $T = S(r, t)$ , where  $S : \Delta \rightarrow [0, \infty)$  will be called the *speedup function* ( $\Delta$  is as in Definition 3); see the lower

panel of Figure 1. By considering an  $\mathbf{x} \in \mathbb{R}_+^m$  for which  $u(\mathbf{x}, \mathbf{z}) \neq 0$  (hence,  $u(\mathbf{y}, \mathbf{z}) \neq 0$ ) we get, from (20) and (21), that the speedup function  $S(r, t)$  must satisfy

$$\delta(r, t) = \delta(0, S(r, t)). \quad (22)$$

Definition 11 formally defines a speedup function.

**Definition 11** *Let  $\delta$  be a discount function. Let  $\Delta$  be as in Definition 3. Suppose that the function  $S : \Delta \rightarrow [0, \infty)$  has the property  $\delta(r, t) = \delta(0, S(r, t))$ ,  $0 \leq r \leq t$ . Then we call  $S$  a speedup function corresponding to the discount function,  $\delta$ .*

We now give some properties of the speedup function in the next proposition. These are analogous to Propositions 12 and 13.

**Proposition 15** : *Let  $\delta$  be a discount function and  $S$  a corresponding speedup function. Then  $S$  has the following properties:*

- (a)  *$S$  is unique.*
- (b) *For each  $r \in [0, \infty)$ ,  $S(r, t)$  is a strictly increasing function of  $t \in [r, \infty)$  into  $[0, \infty)$ , with  $S(r, r) = 0$ . If, ‘into’ is replaced with ‘onto’, then we call  $S$  a continuous speedup function.*
- (c) *For each  $t \in [0, \infty)$ ,  $S(r, t)$  is a strictly decreasing function of  $r \in [0, t]$  into  $[0, t]$ , with  $S(0, t) = t$ .*
- (d)  *$u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$  if, and only if,  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(0, S(r, t))$ .*
- (e) *A continuous discount function has a unique speedup function,  $S(r, t)$  that is continuous in  $t$ .*

From (20), (21), the speedup function may be easily elicited in experiments. For instance, first fix  $\mathbf{x}, \mathbf{z}, r$  and  $t$ , where  $0 \leq r \leq t$ . Next, elicit  $\mathbf{y}$  such that the decision maker is indifferent between receiving  $\mathbf{x}$  at time  $r$  and  $\mathbf{y}$  at time  $t$ . Finally, use these values of  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ , and elicit a value of  $T$  such that the decision maker is indifferent between receiving  $\mathbf{x}$  at time 0 and  $\mathbf{y}$  at time  $T$ . Then  $S(r, t) = T$ . This argument, in the light of Proposition 15a, establishes the following proposition.

**Proposition 16** : *The speedup function,  $S$ , can be uniquely elicited from behavior.*

### 6.1.3 The generating function, $G$

We now define a third useful function, the *generating function*,  $G$ . We will show, below, that the generating function coincides with the discount function in the special case when discounting back to time  $r = 0$  (Proposition 17).

**Definition 12** (*The generating function*): Let  $G : [0, \infty) \rightarrow (0, 1]$  be a strictly decreasing function with  $G(0) = 1$ . Then we call  $G$  a generating function. If, in addition,  $G$  is onto, we call  $G$  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.

## 6.2 The relation between discount, speedup and generating functions

Having completed the description of the three underlying functions (delay,  $D$ , speedup,  $S$ , and generating,  $G$ ) we will show how a discount function can be elicited. We prove that  $G$  and  $S$  jointly determine a unique discount function  $\delta(r, t) = G(S(r, t))$ . Thus, once we know the generating and the speedup functions, we can immediately recover the corresponding discount function. Conversely, a given continuous discount function determines unique generating and speedup functions  $G$  and  $S$  (which are then continuous). We also prove that the generating function is nothing but  $\delta(0, t)$ .

**Proposition 17** (a) Suppose  $G$  is a generating function (Definition 12) and  $S$  is a speedup function (Definition 11). Then,  $\delta = G \circ S$  is a discount function (Definition 3). In particular,  $\delta(0, t) = G(t)$ . If  $G$  and  $S$  are continuous, then so is  $\delta$ .

(b) If  $\delta$  is a continuous discount function, then there are unique generating and speedup functions,  $G$  and  $S$ , such that  $\delta = G \circ S$ . Moreover,  $G$  and  $S$  are also continuous.

(c)  $\delta$  is additive if, and only if,  $\delta(r, t) = \frac{G(t)}{G(r)}$  for all  $0 \leq r \leq t$ .

In Table 1 below, we shall use Proposition 17 to derive the generating, delay and speedup functions for each of the discount functions in Section 4. In general, suppose that we are given a discount function  $\delta(r, t)$ , then we can find the unique generating function by using the result  $\delta(0, t) = G(t)$  (Proposition 17(a)). The unique speedup function,  $S$ , can then be found by inverting the relation  $\delta = G \circ S$  (Proposition 17(b)). From (19), we have  $\delta(r, s)\delta(r, t) = \delta(r, D(r, s, t))$ , which can be solved out for the unique delay function,  $D$ .

## 6.3 Properties of discount functions in terms of the delay and speedup functions

In this section, we show that the properties of the delay and speedup functions, such as impatience (Definition 8) and additivity (Definition 9), determine stationarity (Definition 5) and the common difference effect (Definition 6). In what follows,  $D$  and  $S$  are the delay and speedup functions corresponding to the discount function,  $\delta$ .



In our first result, we show that the properties of stationarity and its violation, the common difference effect, for any continuous discount function, can be expressed in a simpler manner in terms of the associated delay function. Hence, the delay function accounts for these properties of the discount function.

**Proposition 18 :** *A continuous discount function with the delay function,  $D$ , exhibits:*

- (a) *stationarity if, and only if,  $D(r, s, t) = s + t - r$ , for all  $r \geq 0, s \geq r, t \geq r$ ,*
- (b) *the common difference effect if, and only if,  $D(r, s, t) > s + t - r$ , for all  $r \geq 0, s > r, t > r$ .*
- (c) *the converse of the common difference effect if, and only if,  $D(r, s, t) < s + t - r$ , for all  $r \geq 0, s > r, t > r$ .*

Recall that the exponential discounting function exhibits constant impatience. In actual practice, we may observe declining or increasing impatience (Definition 8). The next proposition shows that different degrees of impatience for a continuous discount function are associated with simple conditions on the speedup function. Hence, impatience is a property of the speedup function.

**Proposition 19 :** *A continuous discount function with the speedup function,  $S$ , exhibits:*

- (a) *declining impatience if, and only if,  $S(r, s) > S(r + t, s + t)$ , for all  $t > 0, s > r$ ,*
- (b) *constant impatience if, and only if,  $S(r, s) = S(r + t, s + t)$ , for all  $t \geq 0, s \geq r$ ,*
- (c) *increasing impatience if, and only if,  $S(r, s) < S(r + t, s + t)$ , for all  $t > 0, s > r$ .*

The exponential discounting function is additive, yet human temporal behavior may exhibit subadditivity or superadditivity (Definition 9). We now show how these concepts relate to the underlying delay and speedup functions.

**Proposition 20 :** *A continuous discount function with delay function,  $D$ , and speedup function,  $S$ , exhibits:*

- (a) *subadditivity if, and only if,  $S(r, t) < D(0, S(r, s), S(s, t))$ , for all  $0 \leq r < s < t$ ,*
- (b) *additivity if, and only if,  $S(r, t) = D(0, S(r, s), S(s, t))$ , for all  $0 \leq r \leq s \leq t$ ,*
- (c) *superadditivity, if and only if,  $S(r, t) > D(0, S(r, s), S(s, t))$ , for all  $0 \leq r < s < t$ .*

**Remark 3 :** *Propositions 19 and 20 unpack the discount function to identify which components are responsible for determining impatience and additivity. While this result is a novel theoretical result, it also has implications for future empirical work. Since the discount function is not observable, one may not be able to directly check the conditions in Definitions 8 and 9. Yet, the delay and speedup functions can be easily fitted from empirical data, and individual points on these functions are easily elicited (Propositions 14, 16). It*

is then straightforward to check empirically the conditions in Propositions 19 and 20 that determine impatience and additivity.<sup>14</sup>

In Proposition 21 below, we give our three fundamental functions, the delay, speedup and the generating functions corresponding to each of the discount functions  $\delta(r, t)$ , given in Section 4. Let us first consider the familiar case of the exponential discount function as an illustrative example; the calculations for the remaining discount functions are analogous.

**Example 3** (*Exponential function*): Consider the exponential function,  $\delta(r, t) = e^{-\beta(t-r)}$ ,  $\beta > 0$ , in (11). From Proposition 17,  $\delta(0, t) = G(t)$ . Thus,  $G(t) = e^{-\beta t}$ . From (19), we have  $\delta(r, s)\delta(r, t) = \delta(r, D(r, s, t))$ , hence,  $e^{-\beta(s-r)}e^{-\beta(t-r)} = e^{-\beta(D(r, s, t)-r)}$ . Simplifying, we get  $e^{-\beta((s+t-r)-r)} = e^{-\beta(D(r, s, t)-r)}$ . Comparing both sides, we get  $D(r, s, t) = s + t - r$ ; in particular,  $D(0, s, t) = s + t$ . From Proposition 17, we have  $\delta = G \circ S$ . Substituting  $\delta, G$  we have  $e^{-\beta(t-r)} = e^{-\beta S}$ . Solving out for the speedup function, we get  $S(r, t) = t - r$ .  $D(r, s, t) = s + t - r$ , hence, the exponential discount function is stationary (Proposition 18(a)).  $S(r, s) = S(r + t, s + t) = s - r$ , hence, exponential discounting exhibits constant impatience (Proposition 19(b)). For the exponential discount function, we have  $D(0, S(r, s), S(s, t)) = S(r, s) + S(s, t) = (s - r) + (t - s) = t - r = S(s, t)$ , hence, exponential discounting is additive (Proposition 20(b)).

**Proposition 21** : Table 1, below, gives the generating function,  $G$ , the speedup function,  $S$ , and the delay function,  $D$ , for each of the discount functions  $\delta(r, t)$ , given in Section 4.

Discount function	$G(t)$	$S(r, t)$	$D(r, s, t)$
Exponential	$e^{-\beta t}$	$t - r$	$s + t - r$
PPL ( $r = t = 0$ )	1	0	$s$
PPL ( $0 = r < t$ )	$e^{-(\alpha+\beta t)}$	$t$	$\frac{\alpha}{\beta} + s + t$
PPL ( $0 < r \leq t$ )	$e^{-\beta t}$	$t - r$	$s + t - r$
LP	$(1 + \alpha t)^{-\frac{\beta}{\alpha}}$	$\frac{t-r}{1+\alpha r}$	$s + t + \alpha st$
RS	$[1 + \alpha t^{\tau\rho}]^{-\frac{\beta}{\alpha}}$	$(t^{\tau} - r^{\tau})^{\frac{1}{\tau}}$	$\left[ \left[ \frac{[1+\alpha(s^{\tau}-r^{\tau})^{\rho}][1+\alpha(t^{\tau}-r^{\tau})^{\rho}]-1}{\alpha} \right]^{\frac{1}{\rho}} + r^{\tau} \right]^{\frac{1}{\tau}}$
GRS	$e^{-Q[\phi(t)]}$	$\phi^{-1}[\phi(t) - \phi(r)]$	$\phi^{-1}[Q^{-1}(A+B) + \phi(r)];$ $A = Q[\phi(s) - \phi(r)],$ $B = Q[\phi(t) - \phi(r)]$

Table 1: Generating, speedup and delay functions for various discount functions.

<sup>14</sup>This forms a part of the authors' ongoing research agenda in this area.

If we know the discount function<sup>15</sup>, then, by using the delay and discounting functions give in Table 1, it is straightforward to use Propositions 18, 19 and 20 to verify the claimed impatience and additivity properties of various discount functions in Section 4. Example 3 already provides an illustrative template for the exponential function. However, if we do not know the discount function, then Remark 3 applies, and we can use fitted delay and speedup functions to infer the stationarity, impatience and additivity properties of the underlying discount function.

## 7 Reinterpreting attribute-models as time-discounting models

The anomalies of the EDU model have led to a search for solutions in two quite distinct directions, which are commonly believed to be unrelated. In the first direction are attempts to consider alternative discount functions such as the hyperbolic and the subadditive discount functions (see Section 4). For instance, the common difference effect may be explained by hyperbolic discounting, subadditive discounting, or both (Propositions 6, 7, and 8).

In the second direction is the class of *attribute-based models* that tries to achieve cognitive simplification in temporal choices through the use of simple heuristics.<sup>16</sup> The typical choice in these models is between outcome-time pairs  $(x, s)$  and  $(y, t)$  that are to be discounted back to time zero, where  $0 \leq s < t$  and  $x < y$ . Thus, this is a choice between a smaller-sooner reward  $(x, s)$  and a larger-later reward  $(y, t)$ . This choice has two outcome attributes  $\{x, y\}$  and two time attributes  $\{s, t\}$ . Suppose that there is some vector of parameters,  $\theta$ , which could include variables such as reference outcome levels, interest rates, and relative weights placed on the different elements in the sets  $\{s, t\}$  and  $\{x, y\}$ . Thus,  $\theta$  captures aspects of the frame (recall subsection 2.1) or, more generally, the context. The attribute models then establish preference relationships,  $\preceq$ , between pairs of choices such that

$$(x, s) \preceq (y, t) \Leftrightarrow f(s, t, \theta) \leq g(x, y, \theta), \quad (23)$$

where  $f$  is the advantage of the smaller-sooner reward (it is available earlier); this depends on the time attributes and  $\theta$  only.  $g(x, y, \theta)$  is the advantage of the larger-later reward (it contains a higher outcome), and this depends on the outcome attributes and  $\theta$  only. We consider two different classes of attribute models in this section.

---

<sup>15</sup>The reason is that the condition  $\delta(0, t) = G(t)$  requires knowledge of the true discount function to infer the corresponding generating function, which in turn helps us to recover the corresponding delay and speedup functions, as shown in Example 3.

<sup>16</sup>See, for instance Leland (2002), Rubinstein (2003), Manzini and Mariotti (2006), Read and Scholten (2010), Ericson et al. (2015).

## 7.1 The trade-off model of intertemporal choice

An influential model in the class of attribute-based models is the *trade-off model* of intertemporal choice (Scholten and Read, 2006b, 2010, 2016).<sup>17</sup> We argue below that the trade-off model of Scholten and Read (2006b) is equivalent to a discounting model with the GRS discount function (see (15), (16)). If this is accepted, then the arguments in Scholten and Read (2006b, 2010) in support of their trade-off model also lend further support to their own discount function, the RS discount function (see (14)), and its generalization, GRS (see (15)).

We proceed by first recasting the model of Scholten and Read (2006b, 2010) in a more general form and indicate how their model is to be obtained as a special case. However, there should be no presumption that they would agree with our reformulation. We focus on the fourth and final version of the model in Scholten and Read (2006b).<sup>18</sup>

Let us discount back to time  $r \geq 0$  (in Read and Scholten, 2006b,  $r = 0$ ). This enables us to use the interval discount function, GRS (see (15), which is a generalization of the interval discount function, RS (see (14)).

The trade-off model establishes preference relationships,  $\prec_r$  and  $\sim_r$ , between outcome-time pairs  $(x, s)$  and  $(y, t)$  when both outcomes are discounted back to time,  $r$ . Thus,  $(x, s) \prec_r (y, t)$  if, and only if,  $y$  received at time  $t$  is strictly preferred to  $x$  received at time  $s$ . Similarly,  $(x, s) \sim_r (y, t)$  if, and only if, the decision maker is indifferent to  $y$  received at time  $t$  and  $x$  received at time  $s$ . These relationships are established using three functions, a *utility function*,  $u$  (recall subsection 2.2, but here  $m = 1$ ), a *trade-off function*  $Q$ , and a *delay-perception function*,  $\phi$  (see (15), (16)).

We assume that  $Q : [0, \infty) \xrightarrow{\text{onto}} [0, \infty)$  is strictly increasing and  $\phi : [0, \infty) \xrightarrow{\text{onto}} [0, \infty)$  is also strictly increasing.<sup>19</sup>

In all that follows in this section, let  $s \geq r \geq 0$  and  $t \geq r \geq 0$ .

From Definitions 1 and 2, it follows that, given  $z$ , there exists a unique  $x_0$  such that

---

<sup>17</sup>To quote from Scholten and Read (2006b): “Research on intertemporal judgement and choices between a smaller-sooner and a larger-later outcome has revealed many anomalies to the discounted-utility model. Attempts to account for these anomalies within the discounting paradigm have resulted in convoluted and psychologically opaque models. We therefore develop a new model of intertemporal choice, the trade-off model, in which choice results from a trade-off between the perceived time difference (interval) and the perceived outcome difference (compensation). This model is both more parsimonious and more intuitive than any rival discounting model of comparable scope. Moreover, it accurately describes archival data as well as data from new experiments.”

<sup>18</sup>In terms of Scholten and Read (2010), we focus on their equation (5) on page 932, which they write as  $Q_{X|T}(z) = z$ .

<sup>19</sup>Scholten and Read (2006b) explicitly state two assumptions:  $Q' > 0$ ,  $Q'' < 0$ . However, in the next paragraph, they say that  $Q'' > 0$  for sufficiently small intervals. So, we make no assumptions on  $Q''$ . They explicitly state no further assumptions on  $Q$  and  $w$ . However, we believe our other assumptions on  $Q$  and  $\phi$  are in line with what they intend (see their equations (2) and (5) for the earlier, and simpler, versions of their model).

$u(x_0, z) = 0$ . Thus,  $x_0$  is the unique outcome at which utility is zero. We may term  $x_0$  as the reference outcome. Note that from Definitions 1 and 2, it also follows that  $u(x, z) < 0$  for  $x < x_0$  (losses) and  $u(x, z) > 0$  for  $x > x_0$  (gains).

First, consider the case of gains:  $x_0 < x < y$ . Then, given the reference outcome  $z$ , the decision criteria in this model are given analogously to the decision criterion in (23):

$$(x, s) \sim_r (y, t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] = u(y, z) - u(x, z), \quad (24)$$

$$(x, s) \succ_r (y, t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] > u(y, z) - u(x, z). \quad (25)$$

To understand (24) and (25), suppose that  $s < t$ , so there is a choice between a *smaller-sooner* (SS) reward,  $(x, s)$ , and a *larger later* (LL) reward,  $(y, t)$ . The two *attributes*, outcome and time, are used to consider the relative advantages and disadvantages of the SS and LL rewards, as follows:

1. The *advantage of the LL reward* (alternatively, the *disadvantage of the SS reward*) is along the outcome attribute because it offers a higher outcome; this is indicated by the RHS of the equality in (24) and RHS of the inequality (25),  $u(y, z) - u(x, z) > 0$ .
2. The *advantage of the SS reward* is along the time dimension because it is available at an earlier date. This is indicated by the LHS of the equality in (24) and the LHS of the inequality in (25). For instance, when  $r = 0$  (as in Read and Scholten, 2006a), then the LHS is simply  $Q(\phi(t)) - Q(\phi(s)) > 0$ . Alternatively this may be termed as the *disadvantage of the LL reward*.

Thus, according to (25), the SS reward is strictly preferred to the LL reward if the advantage of the SS reward is greater than the advantage of the LL reward. This can always be achieved if the LL reward is sufficiently delayed.

Recall that the outcome  $x_0$  is defined as the unique outcome at which  $u(x_0, z) = 0$ , so that for any  $x \leq x_0$  we have  $u(x, z) < 0$ . We now state the analogue of the decision criteria in (24), (25) when the outcomes are losses relative to the reference point  $x_0$ :  $y < x < x_0$ . In this case, the LL reward has a relative disadvantage along the outcome dimension, and a relative advantage along the time dimension because future losses are discounted. The decision criteria are:

$$(x, s) \sim_r (y, t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] = u(x, z) - u(y, z), \quad (26)$$

$$(x, s) \prec_r (y, t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] > u(x, z) - u(y, z). \quad (27)$$

Thus, the RHS of the equality in (26) and the RHS in the inequality in (27) is the relative advantage of the SS reward. The relative advantage of the LL reward is given on the LHS

of the two expressions.<sup>20</sup>

To complete the set of decision criteria, we now specify the decision criteria for the cases  $x \leq x_0 < y$ ,  $x < x_0 \leq y$  and  $x = y = x_0$ .

$$x \leq x_0 < y \Rightarrow (x, s) \prec_r (y, t) \text{ for all } s \geq r \text{ and all } t \geq r, \quad (28)$$

$$x < x_0 \leq y \Rightarrow (x, s) \prec_r (y, t) \text{ for all } s \geq r \text{ and all } t \geq r, \quad (29)$$

$$(x_0, s) \sim_r (x_0, t) \text{ for all } s \geq r \text{ and all } t \geq r. \quad (30)$$

Thus, irrespective of the dates when the outcomes are realized, the decision maker always strictly prefers an outcome that generates positive utility to an outcome that generates zero or negative utility (see (28)); the decision maker always strictly prefers an outcome that generates non-negative utility to an outcome that generates negative utility (see (29)); and the decision maker is always indifferent between the times when the zero-utility outcome is received (see (30)).

To get the trade-off model of Read and Scholten, set  $r = s$  in the above equations. They explicitly state only (24) and (26) (but restricted to  $r = s$ ). We believe that our other equations are consistent with the spirit of their framework. However, we generalize their theory by allowing for  $r \leq s$ .

To show that these preferences can be expressed in a discounted-utility model, we first define a function,  $v(x)$ , which is a strict monotonic transformation of the utility function  $u(x)$ . Hence,  $v(x)$  also a utility function.<sup>21</sup>

$$v(x) = \begin{cases} e^{u(x)} & \text{for } x > x_0 \\ 0 & \text{for } x = x_0 \\ -e^{-u(x)} & \text{for } x < x_0 \end{cases}. \quad (31)$$

Next, we will show that we can take our discount function to be the GRS function.<sup>22</sup> Noting that  $v$  is defined in (31), it is obvious that the above relations, (24) to (30), are observationally equivalent to the following decision criteria:

$$(x, s) \sim_r (y, t) \Leftrightarrow v(x) e^{-Q(\phi(s)-\phi(r))} = v(y) e^{-Q(\phi(t)-\phi(r))}, \quad (32)$$

$$(x, s) \prec_r (y, t) \Leftrightarrow v(x) e^{-Q(\phi(s)-\phi(r))} < v(y) e^{-Q(\phi(t)-\phi(r))}. \quad (33)$$

---

<sup>20</sup>Here is another example. Suppose that initial income is  $x_0$ . The decision maker has a choice between paying a small fine,  $x_0 - x$  at time  $s$ , or a larger fine  $x_0 - y > x_0 - x$  at a later date,  $t > s$ . The inequality (27) then says a decision maker will always prefer to pay the larger fine if it can be sufficiently delayed.

<sup>21</sup>Note that this transformation is carried out *before* representing the model as a discounted-utility model. After that, only positive affine transformations are allowed. To preserve the normalization  $v(0) = 0$ , only positive ratio transformations are allowed.

<sup>22</sup>Note that  $v$  as defined by (31) is continuous for  $x > x_0$  and for  $x < x_0$  but is discontinuous at  $x_0$ . In particular,  $v(x) \downarrow e^{u(x_0)} > 0$  as  $x \downarrow x_0$  and  $v(x) \uparrow -e^{-u(x_0)} < 0$  as  $x \uparrow x_0$ . However, it can be verified that this discontinuity at  $x_0$  causes no problems.

Recall from (15), (16) that  $e^{-Q[\phi(t)-\phi(r)]}$  is the GRS discount function, which is a generalization of the interval discount function RS of Scholten and Read (2006a) (with the appropriate restrictions, as described in (16)). From the decision criteria in (32) and (33), it follows that the trade-off model is a discounting model with the GRS discount function.

## 7.2 The DRIFT and the ITCH models

A newer set of attribute-based models, for instance, the DRIFT model of Read et al. (2012) and the ITCH model of Ericson et al. (2015), formulate (23) by assuming that there is linear combination of different factors that define  $f(s, t, \boldsymbol{\theta}), g(x, y, \boldsymbol{\theta})$ .<sup>23</sup> The empirical evidence in both papers leads to the claim that the attribute-based models outperform the commonly known time-discounting models.

The two models (DRIFT and ITCH) are closely related, but here it is pedagogically more convenient to choose the formulation of Read et al. (2012). Consider two outcome-time pairs  $(x, s), (y, t)$  (perhaps it might help the reader to consider these, respectively, as a SS and LL rewards). Then, (23) is formalized by defining

$$f(s, t, \boldsymbol{\theta}) = \omega_T (t - s); \quad g(x, y, \boldsymbol{\theta}) = \omega_D (y - x) + \omega_R \left( \frac{y - x}{x} \right) + \omega_I I + \omega_F F,$$

where  $\omega_T, \omega_D, \omega_R, \omega_I, \omega_F$  are non-negative weights;  $I$  is a dummy variable that captures an interest rate context (i.e., whether an experimental interest rate is highlighted in the experiment);  $F$  is a dummy variable that captures framing in the experiment (e.g., whether the outcomes are presented as an opportunity to invest or to consume). Then, the relevant decision criterion in (23) can be written as

$$(x, s) \preceq (y, t) \Leftrightarrow \omega_T (t - s) \leq \omega_D (y - x) + \omega_R \left( \frac{y - x}{x} \right) + \omega_I I + \omega_F F.$$

Simplifying, we get

$$(x, s) \preceq (y, t) \Leftrightarrow \omega_T (t - s) \leq \omega_D y - \omega_D x + \varphi; \quad \varphi = \omega_R \left( \frac{y}{x} - 1 \right) + \omega_I I + \omega_F F > 0. \quad (34)$$

Let  $u(x) = \omega_D x$  for all  $x$ . Then, using (23), (31), (34), we get

$$(x, s) \preceq (y, t) \Leftrightarrow v(x) e^{-\omega_T s} \leq v(y) (e^{-\omega_T t + \varphi}). \quad (35)$$

The decision rule in (35) clearly suggests that choices are consistent with a discounted-utility model, although it is not of the additively separable form that is given in (10) with

---

<sup>23</sup>DRIFT is short for the *Difference, Ratio, Interest, and Finance model*, while ITCH is short for *intertemporal choice heuristic model*.

$r = 0$ . The reason is that the outcomes influence the extent of discounting through the term  $\varphi$ .<sup>24</sup> This interaction between discounting and the magnitudes of the outcomes is supported by the emerging empirical evidence; see section 2.9.3 in Dhimi (2016), but it is not formally considered in any of the discount functions in Section 4.

The decision rule (35) bears the following interpretation. When given smaller-sooner choices  $(x, s)$  versus larger-later choices  $(y, t)$ , the decision maker follows the following rule. The smaller-sooner choice is discounted as under exponential discounting, using the discount factor  $e^{-\omega_T s}$  ( $\omega_T$  plays the role of discount rate). The larger-later choice is discounted by the discount factor  $e^{-\omega_T t + \varphi}$ , where  $\varphi$  is defined in (34). The definition of  $\varphi$  suggests that larger amounts are discounted less, hence, this can explain the magnitude effect.

### 7.3 Summary

The formulation in (32), (33), and (35) shows that the criticism of attribute-based models of the time-discounting models is too harsh. Indeed, most attribute-based models can be shown to be equivalent to time-discounting models. However, in criticizing time-discounting models, the attribute-based models focus on too narrow a class of time-discounting models. In each of the cases in (32), (33), and (35), we identify time-discounting models that can explain the evidence as well as the attribute-based models.

## 8 Intransitivity as a potential framing effect: An illustrative example

We now investigate if some observed intransitive choices over time can potentially be explained by framing effects. We shall show that this is indeed the case, provided that the discount function is non-additive (recall that exponential discounting is additive). It is a plausible idea that changes in reference points for outcomes may induce apparent intransitivities in choices. Here, we wish to propose another possible source of intransitivity, namely, the change in the time back to which payoffs are discounted,  $r \geq 0$ . Our proposal is in the nature of a hypothesis that is purely based on an illustrative example from Ok and Masatlioglu (2007); we do not make a stronger claim than this.

**Example 4 :** *Consider a UK household that was planning on investing in a buy-to-let property in Spain in a year's time, so that, in making its decision, it was discounting all future cash flows, a year from now ( $r = 1$ ). However, it is now worried with the outcome*

---

<sup>24</sup>This would also be the case if we had used the ITCH model of Ericson et al. (2015) for exposition purposes.



of Brexit negotiations, and wishes to postpone its decision to buy by another year. Hence, it now discounts future cash flows, two years from now ( $r = 2$ ).

We shall assume, in this Section, that all reference outcomes are zero, hence, under this assumption, there are no changes in reference outcomes. Further, under this condition, utility,  $u$ , and material utility,  $\mu$ , are identical (recall Remark 1b).

Consider the following hypothetical situation (Ok and Masatlioglu, 2007, p.215). A decision maker prefers a payoff of 1 now (period 0) to a payoff of 2 in the next period (period 1). The decision maker also prefers a payoff of 2 next period to a payoff of 3, two periods from now (period 2). Finally, the same decision maker prefers a payoff of 3, two periods from now, to a payoff of 1, now. Schematically:

$$(1, 0) \succ (2, 1) \succ (3, 2) \succ (1, 0). \quad (36)$$

We propose that one may view (36) as only ‘apparent intransitivity in preferences’ due to framing effects, but the underlying preferences are transitive (unlike the intransitive preferences approach taken in Ok and Masatlioglu, 2007).

Towards this end, we define three different frames:  $\mathcal{F}_1$ ,  $\mathcal{F}_2$ ,  $\mathcal{F}_3$  (recall subsection 2.1). Frame  $\mathcal{F}_1$  enables us to compare  $(1, 0)$  with  $(2, 1)$ ; frame  $\mathcal{F}_2$  enables us to compare  $(2, 1)$  with  $(3, 2)$ ; and frame  $\mathcal{F}_3$  enables us to compare  $(1, 0)$  with  $(3, 2)$ . We shall use a utility function,  $u$ , and a non-additive discount function,  $\delta$ , such that  $(1, 0) \succ (2, 1)$  in  $\mathcal{F}_1$ ,  $(2, 1) \succ (3, 2)$  in  $\mathcal{F}_2$ , and  $(3, 2) \succ (1, 0)$  in  $\mathcal{F}_3$ . However, since  $\mathcal{F}_1$ ,  $\mathcal{F}_2$ ,  $\mathcal{F}_3$  are different frames, we will be unable to derive the contradiction  $(1, 0) \succ (1, 0)$ .

## 8.1 The frames

We describe Frame  $\mathcal{F}_1$  that allows us to compare the outcome-time pairs  $(1, 0)$  and  $(2, 1)$ . The set,  $\Sigma_1$  (recall (2)), of streams of outcome-time pairs of interest to the decision maker in this frame contains three such streams,  $\Sigma_1 = \{\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Z}_1\}$ , where  $\mathbf{Y}_{11}$  is “receive 1 in period 0 and receive 0 in period 1”, i.e.,  $\mathbf{Y}_{11} = \langle (1, 0), (0, 1) \rangle$ .  $\mathbf{Y}_{12}$  is “receive 0 in period 0 and receive 2 in period 1”, i.e.,  $\mathbf{Y}_{12} = \langle (0, 0), (2, 1) \rangle$ .  $\mathbf{Z}_1$  (recall (3)) is the reference stream “receive 0 in period 0 and receive 0 in period 1”, i.e.,  $\mathbf{Z}_1 = \langle (0, 0), (0, 1) \rangle$ .

The construction of the other two frames is analogous. We summarize the three frames in Table 2; each column corresponds to a distinct frame.

## 8.2 Utility function, $u$

As noted earlier, for simplicity, we have taken all reference outcomes to be zero. Hence, in this case, utility,  $u$ , is identical to material utility,  $\mu$  (recall Definitions 1 and 2 and Remark

Frame $\mathcal{F}_1$	Frame $\mathcal{F}_2$	Frame $\mathcal{F}_3$
$\Sigma_1 = \{\mathbf{Y}_{11}, \mathbf{Y}_{12}, \mathbf{Z}_1\}$	$\Sigma_2 = \{\mathbf{Y}_{21}, \mathbf{Y}_{22}, \mathbf{Z}_2\}$	$\Sigma_3 = \{\mathbf{Y}_{31}, \mathbf{Y}_{32}, \mathbf{Z}_3\}$
$\mathbf{Y}_{11} = \langle (1, 0), (0, 1) \rangle$	$\mathbf{Y}_{21} = \langle (2, 1), (0, 2) \rangle$	$\mathbf{Y}_{31} = \langle (1, 0), (0, 2) \rangle$
$\mathbf{Y}_{12} = \langle (0, 0), (2, 1) \rangle$	$\mathbf{Y}_{22} = \langle (0, 1), (3, 2) \rangle$	$\mathbf{Y}_{32} = \langle (0, 0), (3, 2) \rangle$
$\mathbf{Z}_1 = \langle (0, 0), (0, 1) \rangle$	$\mathbf{Z}_2 = \langle (0, 1), (0, 2) \rangle$	$\mathbf{Z}_3 = \langle (0, 0), (0, 2) \rangle$

Table 2: The description of three different frames

1b). While our results are not dependent on a particular utility function, it is convenient to take the utility function to be a *simple increasing elasticity* (SIE) utility function<sup>25</sup>

$$u(x) = \mu(x) = x^{\frac{1}{2}}(1+x)^{\frac{1}{2}}, \quad x \geq 0. \quad (37)$$

Thus, (working to five significant figures),

$$u(0) = 0, u(1) = 1.4142, u(2) = 2.4495 \text{ and } u(3) = 3.4641. \quad (38)$$

### 8.3 Discount function, $\delta$

As our discount function, we take the Read-Scholten discount function, RS (see (14)), the most flexible discount function, among the ones that we have considered. Recall that this discount function is non-additive (see Section 4.4). In (14), we use the illustrative parameter values  $\alpha = \beta = 1$  and  $\rho = \tau = \frac{1}{2}$ , so the RS discount function can be written as

$$\delta(r, t) = \left(1 + \left(t^{\frac{1}{2}} - r^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)^{-1}. \quad (39)$$

Thus,

$$\delta(0, 1) = \frac{1}{2}, \delta(1, 2) = 0.60842 \text{ and } \delta(0, 2) = 0.45679. \quad (40)$$

### 8.4 Discounted utility, $U$

From (38) and (40), Definition 4 and Table 2, we can now make the comparisons between each pair of outcome-time pairs, in each frame. We distinguish frames by the time to which future outcomes are discounted. In particular, we assume that, fixing the frame, in a pair of outcomes, the outcomes are discounted back to the earlier of the two times in the pair. For instance, in Frame  $\mathcal{F}_1$ , the earlier of the two times is 0, so both outcomes are discounted back to time zero. Analogously, this assumption implies that in frames  $\mathcal{F}_2$  and  $\mathcal{F}_3$ , the outcomes are discounted back to time 1 and 0, respectively. The calculations for discounted utility in each frame are as follows.

---

<sup>25</sup>The SIE class of utility functions were introduced by al-Nowaihi and Dhami (2009) to provide a model for the Loewenstein and Prelec (1992) theory. In (37), we have used the parameter values  $\mu = 1$ ,  $\theta_+ = 0.5$ ,  $\lambda = 2$ ,  $0 < \sigma = \gamma = 0.5$  to parametrize the SIE function; these values are purely for illustrative purposes.

$$\text{Frame } \mathcal{F}_1: \begin{cases} (1, 0) : U_0(\mathbf{Y}_{11}, \mathbf{Z}_1) = u(1) \delta(0, 0) = 1.4142 \\ (2, 1) : U_0(\mathbf{Y}_{12}, \mathbf{Z}_1) = u(2) \delta(0, 1) = 1.2248 \\ (1, 0) \text{ vs } (2, 1) : (1, 0) \succ (2, 1) \end{cases} \quad (41)$$

$$\text{Frame } \mathcal{F}_2: \begin{cases} (2, 1) : U_1(\mathbf{Y}_{21}, \mathbf{Z}_2) = u(2) \delta(1, 1) = 2.4495 \\ (3, 2) : U_1(\mathbf{Y}_{22}, \mathbf{Z}_2) = u(3) \delta(1, 2) = 2.1076 \\ (2, 1) \text{ vs } (3, 2) : (2, 1) \succ (3, 2) \end{cases} \quad (42)$$

$$\text{Frame } \mathcal{F}_3: \begin{cases} (3, 2) : U_0(\mathbf{Y}_{32}, \mathbf{Z}_3) = u(3) \delta(0, 2) = 1.5824 \\ (1, 0) : U_0(\mathbf{Y}_{31}, \mathbf{Z}_3) = u(1) \delta(0, 0) = 1.4142 \\ (3, 2) \text{ vs } (1, 0) : (3, 2) \succ (1, 0) \end{cases} \quad (43)$$

If we ignore the fact that the comparisons (41), (42) and (43) are carried out in different frames, then we would conclude that (36) holds. Putting together the last rows in each of (41), (42) and (43), here the intransitivity is only ‘apparent’ and, taking account of frames, there is no inconsistency with maximizing a discounted utility model, which is exactly the exercise we carry out above (see Definition 4). It remains to be seen if our hypothesis can be empirically supported with future experiments.

## 9 Conclusions

Discounting over time is central to the subject matter of economics. The exponentially discounted utility (EDU) model has been the workhorse model in economics, yet it is not supported by the evidence. There are now well documented and robust anomalies of the EDU model that have led to many diverse explanations.

We employ a general delay-discounting framework that encompasses reference-dependent utility, frame dependence, and allows for outcomes to be discounted back to any future time  $r \geq 0$ . We believe that all these features are better descriptors of reality.

The aim of our paper is to provide a unified approach to time discounting that nests the main time-discounting models. Our unified approach identifies three basic building blocks of temporal choice: the delay, speedup, and the generating functions. The delay and speedup functions can be uniquely elicited from behavior. These two functions determine whether the behavior of a decision maker exhibits any of the following nine important properties: Stationarity, common difference effect, converse of common difference effect, declining impatience, constant impatience, increasing impatience, subadditivity, additivity and superadditivity.

We show that there is a close relation between time-discounting and attribute-based models. Indeed, attribute-based models can be expressed either in terms of time-discounting models that use known discount functions, or discount functions that require interaction between time and outcome terms. Empirical evidence is supportive of such interactions and

exploring these newer discount functions is one of the important research agendas thrown up by our paper. We also suggest by an illustrative example that apparent intransitivities in time preferences can be explained by shifts in the frame of the problem.

Finally, the arguments for an endogenous reference point, in the tradition of Köszegi and Rabin (2006, 2007) do not apply to our paper because we do not deal with risky temporal choices. However, the interaction between risk and time preferences is a growing and interesting area of new research. We anticipate that future research that extends our framework to allow for risky temporal choices may benefit from an endogenous reference point formulation.

## 10 Appendix: Proofs

**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  $\{\delta(r, t_n)\}_{n=1}^\infty$  converges to  $\delta(r, t)$ . Since  $\delta(r, t)$  is a strictly decreasing function of  $t \in [r, \infty)$  (Definition 3), it is sufficient to show that any monotone subsequence of  $\{\delta(r, t_n)\}_{n=1}^\infty$  converges to  $\delta(r, t)$ . In particular, let  $\{\delta(r, t_{n_i})\}_{i=1}^\infty$  be a decreasing subsequence of  $\{\delta(r, t_n)\}_{n=1}^\infty$ . Since  $\{\delta(r, t_{n_i})\}_{i=1}^\infty$  is bounded below by  $\delta(r, t)$ , it must converge to, say,  $q$ , where  $\delta(r, t) \leq q \leq \delta(r, t_{n_i})$ , for all  $i$ . Since  $\delta$  is onto, there is a  $p \in [r, \infty)$  such that  $\delta(r, p) = q$ . Moreover,  $t_{n_i} \leq p \leq t$ , for each  $i$ . Suppose  $\delta(r, t) < q$ . 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,  $\delta(r, t) = q$  and  $\{\delta(r, t_{n_i})\}_{i=1}^\infty$  converges to  $\delta(r, t)$ . Similarly, we can show that any increasing subsequence of  $\{\delta(r, t_n)\}_{n=1}^\infty$  converges to  $\delta(r, t)$ . Hence,  $\{\delta(r, t_n)\}_{n=1}^\infty$  converges to  $\delta(r, t)$ . Thus,  $\delta(r, t)$  is continuous in  $t$ . ■

**Proposition 2:** Let  $\delta(r, t)$  be a continuous discount function and  $r \geq 0$ . Suppose  $0 < u(\mathbf{x}, \mathbf{z}) \leq u(\mathbf{y}, \mathbf{z})$ . Then,  $0 < \frac{u(\mathbf{x}, \mathbf{z})}{u(\mathbf{y}, \mathbf{z})} \leq 1$ . Since, by Definition 3(c),  $\delta(r, t) : [r, \infty) \xrightarrow{\text{onto}} (0, 1]$ , it follows that  $\frac{u(\mathbf{x}, \mathbf{z})}{u(\mathbf{y}, \mathbf{z})} = \delta(r, t)$  for some  $t \in [r, \infty)$ . A similar argument applies if  $u(\mathbf{y}, \mathbf{z}) \leq u(\mathbf{x}, \mathbf{z}) < 0$ . ■

**Proposition 3:** We consider several cases.

(a) Suppose  $\mu(\mathbf{y}) \geq \mu(\mathbf{z})$ .

From (4) we get  $u(\mathbf{y}, \mathbf{z}) = (1 - \gamma)\mu(\mathbf{y}) + \gamma[\mu(\mathbf{y}) - \mu(\mathbf{z})] \geq (1 - \gamma)\mu(\mathbf{y}) + \gamma[\mu(\mathbf{z}) - \mu(\mathbf{z})] = (1 - \gamma)\mu(\mathbf{y}) \geq 0$ . Since  $\delta(r, t) \in (0, 1]$  and since  $u(\mathbf{y}, \mathbf{z}) \geq 0$  we get

$$0 \leq u(\mathbf{y}, \mathbf{z})\delta(r, t) \leq u(\mathbf{y}, \mathbf{z}). \quad (44)$$

(ai) If  $\mu(\mathbf{z}) = 0$ , then  $u(\mathbf{0}, \mathbf{z}) = (1 - \gamma)\mu(\mathbf{0}) + \gamma[\mu(\mathbf{0}) - \mu(\mathbf{0})] = 0$  because  $\mu(\mathbf{0}) = 0$  from Definition 1.

(aii) If  $\mu(\mathbf{z}) > 0$  then, from (5),  $u(\mathbf{0}, \mathbf{z}) = (1 - \gamma)\mu(\mathbf{0}) + \lambda\gamma[\mu(\mathbf{0}) - \mu(\mathbf{z})] = -\lambda\gamma\mu(\mathbf{z}) < 0$ .

From (44), (ai) and (aia), we get

$$u(\mathbf{0}, \mathbf{z}) \leq u(\mathbf{y}, \mathbf{z}) \delta(r, t) \leq u(\mathbf{y}, \mathbf{z}). \quad (45)$$

We now apply the intermediate value theorem.

Since  $\mathbb{R}_+^m$  is connected there exists a continuous curve,  $\mathbf{x} : [0, 1] \rightarrow \mathbb{R}_+^m$ , such that  $\mathbf{x}(0) = \mathbf{0}$  and  $\mathbf{x}(1) = \mathbf{y}$ . Let  $f : [0, 1] \rightarrow \mathbb{R}_+$  be the function defined by  $f(s) = u(\mathbf{x}(s), \mathbf{z})$ . Then  $f$  is continuous with  $f(0) = u(\mathbf{x}(0), \mathbf{z}) = u(\mathbf{0}, \mathbf{z})$  and  $f(1) = u(\mathbf{x}(1), \mathbf{z}) = u(\mathbf{y}, \mathbf{z})$ . From (45) and the intermediate value theorem it then follows that there exists an  $s \in [0, 1]$  such that  $f(s) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ . From the definition of  $f$  we get  $u(\mathbf{x}(s), \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ . Thus, we have shown that there exists an  $\mathbf{x} \in \mathbb{R}_+^m$  such that  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ .

(b) Suppose  $\mu(\mathbf{y}) < \mu(\mathbf{z})$ .

From (5) we get  $u(\mathbf{y}, \mathbf{z}) = (1 - \gamma) \mu(\mathbf{y}) + \lambda \gamma [\mu(\mathbf{y}) - \mu(\mathbf{z})]$ .

(bi) Suppose  $u(\mathbf{y}, \mathbf{z}) \leq 0$ .

Since  $\delta(r, t) \in (0, 1]$  we get that  $u(\mathbf{y}, \mathbf{z}) \leq u(\mathbf{y}, \mathbf{z}) \delta(r, t) \leq 0 \leq (1 - \gamma) \mu(\mathbf{y}) = u(\mathbf{z}, \mathbf{z})$ .

As before, the intermediate value theorem gives that there exists an  $\mathbf{x} \in \mathbb{R}_+^m$  such that  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ .

(bii) Suppose  $u(\mathbf{y}, \mathbf{z}) > 0$ .

Hence,  $0 \leq u(\mathbf{y}, \mathbf{z}) \delta(r, t) \leq u(\mathbf{y}, \mathbf{z})$ . We consider two cases.

If  $\mu(\mathbf{z}) = \mathbf{0}$ , then  $u(\mathbf{0}, \mathbf{z}) = (1 - \gamma) \mu(\mathbf{0}) + \gamma [\mu(\mathbf{0}) - \mu(\mathbf{z})] = 0$ .

If  $\mu(\mathbf{z}) > \mathbf{0}$ , then  $u(\mathbf{0}, \mathbf{z}) = (1 - \gamma) \mu(\mathbf{0}) + \lambda \gamma [\mu(\mathbf{0}) - \mu(\mathbf{z})] = -\lambda \gamma \mu(\mathbf{z}) < 0$ .

In either case,  $u(\mathbf{0}, \mathbf{z}) \leq 0$ .

Thus, we have  $u(\mathbf{0}, \mathbf{z}) \leq u(\mathbf{y}, \mathbf{z}) \delta(r, t) \leq u(\mathbf{y}, \mathbf{z})$ . As before, by the intermediate value theorem there exists an  $\mathbf{x} \in \mathbb{R}_+^m$  such that  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ . ■

**Proposition 4:** We shall use the following two simple mathematical results (for proofs, see al-Nowaihi and Dhami, 2008b, p49):

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

**Result 2 :** Let  $\tau > 0$ ,  $0 \leq s < t$  and  $r > 0$ . Let  $F(r) = (t + r)^\tau - (s + r)^\tau - (t^\tau - s^\tau)$ . Then:

(i)  $0 < \tau < 1 \Rightarrow F(r) < 0$ ,

(ii)  $\tau = 1 \Rightarrow F(r) = 0$ ,

(iii)  $\tau > 1 \Rightarrow F(r) > 0$ .

(a) Suppose  $0 < \rho \leq 1$ . Let  $0 \leq r < s < t$ . For RS discount function, we get:

$$\begin{aligned} \delta(r, s) \delta(s, t) &= [1 + \alpha (s^\tau - r^\tau)^\rho]^{-\frac{\beta}{\alpha}} [1 + \alpha (t^\tau - s^\tau)^\rho]^{-\frac{\beta}{\alpha}} \\ &= [1 + \alpha (t^\tau - s^\tau)^\rho + \alpha (s^\tau - r^\tau)^\rho + \alpha^2 (s^\tau - r^\tau)^\rho (t^\tau - s^\tau)^\rho]^{-\frac{\beta}{\alpha}}. \end{aligned}$$

However,  $\alpha^2 (s^\tau - r^\tau)^\rho (t^\tau - s^\tau)^\rho > 0$  and  $-\frac{\beta}{\alpha} < 0$ . Hence, we have:

$\delta(r, s) \delta(s, t) < [1 + \alpha [(t^\tau - s^\tau)^\rho + (s^\tau - r^\tau)^\rho]]^{-\frac{\beta}{\alpha}}$ . From Result 1 we know that  $(t^\tau - s^\tau)^\rho + (s^\tau - r^\tau)^\rho \geq (t - s + s - r)^\rho = (t - r)^\rho$ . Using this fact and  $-\frac{\beta}{\alpha} < 0$ , we get that:

$\delta(r, s) \delta(s, t) < [1 + \alpha (t^\tau - r^\tau)^\rho]^{-\frac{\beta}{\alpha}} = \delta(r, t)$ . Hence, the RS discount function is subadditive for  $0 < \rho \leq 1$ .

(b) It is sufficient to give an example. Let  $\alpha = \tau = 1$  and  $\rho = 2$ . Hence,  $\delta(0, 1) \delta(1, 2) = 4^{-\beta} > 5^{-\beta} = \delta(0, 2)$ . Hence, for  $\alpha = \tau = 1$  and  $\rho = 2$ ,  $\delta$  is superadditive, but it cannot be additive or subadditive. However, for the same parameter values, we have  $\delta(0, 10) \delta(10, 20) = 10201^{-\beta} < 401^{-\beta} = \delta(0, 20)$ . Hence,  $\delta$  cannot be superadditive either.<sup>26</sup>

(c) Similar to part a, except that we use Result 2. ■

**Proposition 5:** Assume that  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$  (or  $u(\mathbf{x}, \mathbf{z}) \delta(r, r) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$ ) because  $\delta(r, r) = 1$ ) and that  $\delta$  exhibits constant impatience and additivity. From this, we get

$$\begin{aligned} & u(\mathbf{x}, \mathbf{z}) \delta(r, t) \\ &= u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) \\ &= u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r + (s - r), t + (s - r)), \text{ by constant impatience,} \\ &= u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, t + s - r) \\ &= u(\mathbf{y}, \mathbf{z}) \delta(r, t + s - r), \text{ by additivity,} \\ &= u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r). \end{aligned}$$

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Thus, stationarity (Definition 5) holds. ■

**Proposition 6:** Assume  $r \geq 0$ ,  $s > r$ ,  $t > r$  and  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$  (or  $u(\mathbf{x}, \mathbf{z}) \delta(r, r) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$ ). Let  $\delta$  be additive and exhibit declining impatience. Note that  $\delta(r + (s - r), t + (s - r)) = \delta(s, s + t - r)$ .

(i) Assume  $u(\mathbf{x}, \mathbf{z}) > 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) > 0$  (because from Definition 3,  $\delta \in (0, 1]$ ).

Applying declining impatience, then additivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) < u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) = u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) < u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

(ii) Assume  $u(\mathbf{x}, \mathbf{z}) < 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) < 0$ .

Applying declining impatience, then additivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) > u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) = u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) > u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence, the common difference effect (Definition 6) holds. ■

**Proposition 7:** Assume  $r \geq 0$ ,  $s > r$ ,  $t > r$  and  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$ . Note that  $\delta(r + (s - r), t + (s - r)) = \delta(s, s + t - r)$ .

---

<sup>26</sup>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  $\rho > 1$ .

(i) Assume  $u(\mathbf{x}, \mathbf{z}) > 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) > 0$ .

Applying constant impatience, then subadditivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) < u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) < u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

(ii) Assume  $u(\mathbf{x}, \mathbf{z}) < 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) < 0$ .

Applying constant impatience, then subadditivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) > u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) > u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence, the common difference effect holds. ■

**Proposition 8:** Assume  $r \geq 0$ ,  $s > r$ ,  $t > r$  and  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, s)$ . Note that  $\delta(r + (s - r), t + (s - r)) = \delta(s, s + t - r)$ .

(i) Assume  $u(\mathbf{x}, \mathbf{z}) > 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) > 0$ .

Applying declining impatience, then subadditivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) < u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) < u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) < u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

(ii) Assume  $u(\mathbf{x}, \mathbf{z}) < 0$ . Hence,  $u(\mathbf{y}, \mathbf{z}) < 0$ .

Applying declining impatience, then subadditivity, we get  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) > u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(s, s + t - r) > u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence,  $u(\mathbf{x}, \mathbf{z}) \delta(r, t) > u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ .

Hence, the common difference effect holds. ■

**Propositions 9, 10, 11:** The proofs of Propositions 9, 10, 11 are similar to those of 6, 7, 8, respectively. ■

**Proposition 12:** Let  $\delta$  be a discount function and  $D$  and  $\Phi$  two corresponding delay functions. Let  $s, t \in [r, \infty)$ . Then, using Definition 10,  $\delta(r, \Phi(r, s, t)) = \delta(r, s) \delta(r, t) = \delta(r, D(r, s, t))$ . Since  $\delta(r, \tau)$  is strictly decreasing in  $\tau$  (Definition 3), we must have  $\Phi(r, s, t) = D(r, s, t)$ . This establishes (a). Using Definition 3, it is straightforward to check that properties (b) to (d) follow from Definition 10. Now, suppose  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ . Multiply both sides by  $\delta(r, s)$  to get  $u(\mathbf{x}, \mathbf{z}) \delta(r, s) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, D(r, s, t))$ . Conversely, suppose  $u(\mathbf{x}, \mathbf{z}) \delta(r, s) = u(\mathbf{y}, \mathbf{z}) \delta(r, D(r, s, t))$ . Then  $u(\mathbf{x}, \mathbf{z}) \delta(r, s) = u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t)$ . Since  $\delta(r, s) > 0$ , we can cancel it to get  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z}) \delta(r, t)$ . This establishes (e). ■

**Proposition 13:** Let  $\delta$  be a continuous discount function. Let  $s, t \in [r, \infty)$ . Then  $\delta(r, s), \delta(r, t) \in (0, 1]$ . Since  $\tau \mapsto \delta(r, \tau)$  is onto  $(0, 1]$ , there is some  $T \in [0, \infty)$  such that  $\delta(r, s) \delta(r, t) = \delta(r, T)$ . Since  $\delta(r, \tau)$  is strictly decreasing in  $\tau$ , this  $T$  is unique. Set  $T = D(r, s, t)$ . The function,  $D(r, s, t)$ , thus defined, is a delay function corresponding to  $\delta$ . Let  $\psi(t) = \delta(r, t)$ . Then  $\psi : [r, \infty) \xrightarrow{\text{onto}} (0, 1]$  is strictly decreasing and continuous. Hence it has an inverse,  $\psi^{-1}$ , which is also strictly decreasing and continuous (proof as in

Proposition 1). From Definition 10, we get  $\psi(s)\psi(t) = \psi(D(r, s, t))$ . Hence,  $D(r, s, t) = \psi^{-1}(\psi(s)\psi(t))$  is continuous. ■

**Proposition 14:** The proof follows from the argument immediately above the statement of the proposition. ■

**Proposition 15:** The proofs are similar to those of Propositions 12 and 13. ■

**Proposition 16:** The proof follows from the argument immediately above the statement of the proposition. ■

**Proposition 17:** (a) Since  $S$  maps  $\Delta$  into  $[0, \infty)$  and  $G$  maps  $[0, \infty)$  into  $(0, 1]$ , it follows that  $\delta = G \circ S$  maps  $\Delta$  into  $(0, 1]$ . Since  $S(r, r) = 0$  and  $G(0) = 1$ , it follows that  $\delta(r, r) = G(S(r, r)) = 1$ . Since  $S(r, t)$  is strictly decreasing in  $r$  and strictly increasing in  $t$ , and since  $G$  is strictly decreasing, it follows that  $\delta(r, t) = G(S(r, t))$  is strictly increasing in  $r$  and strictly decreasing in  $t$ . Hence,  $\delta = G \circ S$  is a discount function. In particular,  $\delta(0, t) = G(S(0, t)) = G(t)$ , by Definition 11. Let  $r \in [0, \infty)$ . Suppose  $S(r, \cdot)$  maps  $[r, \infty)$  onto  $[0, \infty)$  and  $G$  maps  $[0, \infty)$  onto  $(0, 1]$ , it follows that  $\delta = G \circ S$  maps  $\Delta$  onto  $(0, 1]$ . Hence, if  $G$  and  $S$  are continuous, then so is  $\delta$ .

(b) Let  $\delta$  be a continuous discount function. Define  $G(t) = \delta(0, t)$ . Then  $G$  maps  $[0, \infty)$  onto  $(0, 1]$ . Since  $\delta(0, t)$  is strictly decreasing in  $t$ , so is  $G$ . Finally,  $G(0) = \delta(0, 0) = 1$ . Hence,  $G$  is a continuous generating function. Since  $G : [0, \infty) \xrightarrow{\text{onto}} (0, 1]$  is strictly decreasing, it follows that it has a unique inverse,  $G^{-1} : (0, 1] \xrightarrow{\text{onto}} [0, \infty)$ , which is also strictly decreasing. Define  $S = G^{-1} \circ \delta$ . Since  $\delta$  is strictly increasing in  $r$  and strictly decreasing in  $t$ , and since  $G^{-1}$  is strictly decreasing, it follows that  $S$  is strictly decreasing in  $r$  and strictly increasing in  $t$ . Furthermore,  $S(r, r) = G^{-1}(\delta(r, r)) = G^{-1}(1) = 0$  and  $S(0, t) = G^{-1}(\delta(0, t)) = G^{-1}(G(t)) = t$ . Since  $\delta : \Delta \xrightarrow{\text{onto}} (0, 1]$  and since  $G^{-1} : (0, 1] \xrightarrow{\text{onto}} [0, \infty)$ , it follows that  $S = G^{-1} \circ \delta$  maps  $\Delta$  onto  $[0, \infty)$ . Hence,  $S$  is a continuous generating function. Suppose,  $\delta = G' \circ S'$  for generating and speedup functions  $G'$  and  $S'$ . Then  $G(t) = \delta(0, t) = G'(S'(0, t)) = G'(t)$ . Hence,  $G = G'$ , which implies that  $\delta = G \circ S'$ , so  $S' = G^{-1} \circ \delta = S$ . Hence,  $G$  and  $S$  are unique.

(c) If  $\delta(r, t) = \frac{G(t)}{G(r)}$  for all  $0 \leq r \leq t$ , then  $\delta(r, s)\delta(s, t) = \frac{G(s)}{G(r)} \frac{G(t)}{G(s)} = \frac{G(t)}{G(r)} = \delta(r, t)$  for all  $0 \leq r \leq s \leq t$ . Conversely, suppose  $\delta(r, s)\delta(s, t) = \delta(r, t)$  for all  $0 \leq r \leq s \leq t$ , then, in particular,  $\delta(0, r)\delta(r, t) = \delta(0, t)$  for all  $0 \leq r \leq t$ , i.e.,  $G(r)\delta(r, t) = G(t)$  for all  $0 \leq r \leq t$ . Hence,  $\delta(r, t) = \frac{G(t)}{G(r)}$  for all  $0 \leq r \leq t$ . ■

**Proposition 18:** Consider a continuous discount function,  $\delta$ . We first prove part (a).

(i) Assume that  $D(r, s, t) = s + t - r$ , for all  $r \geq 0, s \geq r, t \geq r$ . Let  $u(\mathbf{x}, \mathbf{z}) = u(\mathbf{y}, \mathbf{z})\delta(r, s)$ . Then  $u(\mathbf{x}, \mathbf{z})\delta(r, t) = u(\mathbf{y}, \mathbf{z})\delta(r, s)\delta(r, t) = u(\mathbf{y}, \mathbf{z})\delta(r, D(r, s, t)) = u(\mathbf{y}, \mathbf{z})\delta(r, s + t - r)$ . Hence, stationarity holds.

(ii) Assume  $\delta$  is stationary. Let  $r \geq 0, s \geq r, t \geq r$ . Choose any outcome  $y$  such that  $u(y) > 0$ . From Proposition 3 it follows that  $u(\mathbf{x}, \mathbf{z})\delta(r, r) = u(\mathbf{y}, \mathbf{z})\delta(r, s)$  for some outcome  $x$  (where  $\delta(r, r) = 1$ ). Hence, by stationarity (Definition 5),  $u(\mathbf{x}, \mathbf{z})\delta(r, t) =$



$u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ . It follows that,  $u(\mathbf{y}, \mathbf{z}) \delta(r, s) \delta(r, t) = u(\mathbf{y}, \mathbf{z}) \delta(r, s + t - r)$ . Hence,  $\delta(r, s) \delta(r, t) = \delta(r, s + t - r)$ . By Definition 10 and Proposition 13, the continuous discount function,  $\delta$ , has a unique delay function,  $D$ , and  $\delta(r, s) \delta(r, t) = \delta(r, D(r, s, t))$ . Hence,  $\delta(r, D(r, s, t)) = \delta(r, s + t - r)$ . Since  $\delta(r, \tau)$  is strictly decreasing in  $\tau$  (Definition 3), we get  $D(r, s, t) = s + t - r$ .

The proof of part (b) is similar to that of part (a) except, where appropriate, equality is replaced with strict inequality. The proof of part (c) is similar to that of part (b). ■

**Proposition 19:** Follows from Definition 8 (impatience) and the facts that  $\delta(r, t) = G(S(r, t))$  (Proposition 17) and  $G$  is strictly decreasing (Definition 12). ■

**Proposition 20:** By Definition 10,  $\delta(0, s) \delta(0, t) = \delta(0, D(0, s, t))$ ,  $s \geq 0, t \geq 0$ . Under the given conditions, we have from Proposition 17b that  $\delta = G \circ S$  (where  $G, S$  are unique), and from Proposition 15(c) we have  $S(0, t) = t$ , hence,

$$\delta(0, s) \delta(0, t) = \delta(0, D(0, s, t)) \Rightarrow G(s) G(t) = G(D(0, s, t)), \quad s \geq 0, t \geq 0.$$

Since this holds for all  $s \geq 0, t \geq 0$ , it also holds for the two dates  $S(r, s) \geq 0$  and  $S(s, t) \geq 0$ . Hence, we get  $G(S(r, s)) G(S(s, t)) = G(D(0, S(r, s), S(s, t)))$ ,  $0 \leq r < s < t$ .

By Definition 9, a discount function,  $\delta(r, t)$ , is subadditive if, and only if,  $\delta(r, s) \delta(s, t) < \delta(r, t)$  for all  $0 \leq r < s < t$ , i.e., if, and only if,  $G(S(r, s)) G(S(s, t)) < G(S(r, t))$ , for all  $0 \leq r < s < t$  (using Proposition 17b), i.e., if, and only if,  $G(D(0, S(r, s), S(s, t))) < G(S(r, t))$ , for all  $0 \leq r < s < t$  (using what has been established above), i.e., if, and only if,  $D(0, S(r, s), S(s, t)) > S(r, t)$ , for all  $0 \leq r < s < t$  (since  $G$  is strictly decreasing, Definition 12). This establishes part (a). Parts (b) and (c) are similar. ■

**Proposition 21:** All the claims can be verified by straightforward calculations. ■

## References

- 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. 08/22.
- al-Nowaihi, A., Dhami, S. 2009. A value function that explains the magnitude and sign effects. *Economics Letters* 105, 224-229.

- Andersen, S., Harrison, G., Morten, I. L. and Rutstroem, E. (2010) Discounting behavior: A reconsideration. *European Economic Review*. 71, 15-33.
- Andreoni, J., and Sprenger, C. (2012) Estimating Time Preferences from Convex Budgets. *American Economic Review*. 102(7): 3333-3356.
- Dhami, S. (2016) *Foundations of behavioral economic analysis*. Oxford University Press: Oxford.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2005) The Robustness and Pervasiveness of Sub-Additivity in Intertemporal Choice. Mimeo University of Bonn.
- Ericson, K.M.M., White J.M, Laibson D., Cohen J.D. (2015). Money Earlier or Later? Simple Heuristics Explain Intertemporal Choices Better Than Delay Discounting Does. *Psychological Science*. 26(6):826-833.
- Fishburn, P.C., and Rubinstein, A. (1982). Time preference. *International Economic Review*. 23(3): 677–694.
- Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preferences: a critical review. *Journal of Economic Literature*. 40(2): 351-401.
- Gintis, H. (2009) *The bounds of reason: Game theory and the unification of the behavioral sciences*. Princeton University Press: Princeton.
- Gigliotti, G., & Sopher, B. (1997). Violations of present-value maximization in income choice. *Theory and Decision*, 43, 45–69.
- Halevy, Y. (2015) Time consistency: Stationarity and time invariance. *Econometrica*. 83(1): 335-352.
- 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.
- Kőszegi, Botond and Matthew Rabin. 2006. A model of reference-dependent preferences. *The Quarterly Journal of Economics* 121 (4): 1133-1165.
- Kőszegi, Botond and Matthew Rabin. 2007. Reference-dependent risk attitudes. *The American Economic Review* 97 (4): 1047-1073.

- Laibson, D. 1997. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112, 443-477.
- Leland, J. W. (2002) Similarity judgements and anomalies in intertemporal choice. *Economic Inquiry*, 40, 574-81
- 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.
- Manzini, P. and M. Mariotti, 2006. A vague theory of choice over time. *Advances in Theoretical Economics*. Berkeley Electronic Press Vol. 6(1), Article 6.
- Manzini, P., Mariotti, M., & Mittone, L. (2010). Choosing monetary sequences: Theory and experimental evidence. *Theory and Decision*, 69, 327–354. <http://dx.doi.org/10.1007/s11238-010-9214-7>
- Ok, E. A., and Y. Masatlioglu, 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.
- Read, D., Frederick, S., & Scholten, M. (2013). DRIFT: An Analysis of Outcome Framing in Intertemporal Choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(2): 573-588.
- Rubinstein, A. 2003. Is it economics and psychology? The case of hyperbolic discounting, *International Economic Review*, 44 (4), 1207-1216.
- 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., 2006b. Beyond Discounting: The Trade-off Model of Intertemporal Choice. Working Paper LSE-OR 06.88. London: London School of Economics and Political Science.

Scholten, M., & Read, D. (2010). The Psychology of Intertemporal Trade-offs. *Psychological Review*, 117 (3): 925-944.

Scholten, M., Read, D., and Sanborn, A. (2016) Cumulative weighting of time in intertemporal payoffs. *Journal of Experimental Psychology-General*. 145(9): 1177-1205.

Strotz, R.H. (1955-1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*. 23(3): 165–180.

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

Tversky, A., and Kahneman, D. (1991). Loss aversion in riskless choice: A reference dependent model. *Quarterly Journal of Economics*, 106, 1039-1061.

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

Wakker, P. P. 2010. *Prospect Theory for Risk and Ambiguity*. Cambridge University Press.

Wu, C. Y., & He, G.-B. (2012). The effects of time perspective and salience of possible monetary losses on intertemporal choice. *Social Behavior and Personality*, 40, 1645–1653. <http://dx.doi.org/10.2224/sbp.2012.40.10.1645>