University *of* Leicester

Department of Economics

A Theory of Reference Time

Ali al-Nowaihi, University of Leicester

Sanjit Dhami, University of Leicester

Working Paper No. 13/26

November 2013

A Theory of Reference Time

Ali al-Nowaihi* **Sanjit Dhami†**

8 November 2013

Abstract

We consider a discounted utility model that has two components. (1) The instantaneous utility is of the prospect theory form, thus, allowing for reference dependent outcomes. (2) The discount function embodies a 'reference time' to which all future outcomes are discounted back to, hence, the name, reference time theory. We allow the discount function to exhibit declining impatience, as in hyperbolic discounting models, subadditivity or both. We show that if the discount function is non-additive, then the presence of a reference time has important effects on intertemporal choices. For instance, this helps to explain apparently intransitive choices over time. We also show how several recent approaches to time discounting can be incorporated within our proposed framework; these include attribute models and models of uncertainty.

Keywords: Discounted utility models; Reference time theory; Prospect theory; Hyperbolic discounting, Subadditive discounting.

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

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

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

1 Introduction

In a seminal paper, Loewenstein and Prelec (1992), henceforth LP, introduced reference dependence for outcomes in the discounted utility model (DU). Suppose that a decision maker, at time $t = 0$, takes an action that gives rise to the sequence of outcome-time pairs: $(w_1, t_1), (w_2, t_2), ..., (w_n, t_n)$, where w_i is the outcome at time t_i, $i = 1, 2, 3, ...$. LP proposed the following intertemporal utility function for the decision maker:

$$U = \Sigma_{i=1}^{n} v (w_i - w_0) D(0, t_i). \tag{1}$$

In (1), v is the utility function under prospect theory, due to Kahneman and Tversky (1979). v is defined over outcomes relative to a reference point, $x_i = w_i - w_0$, where w_0 is the reference outcome level. When $x_i \geq 0$ the individual is said to be in the *domain of gains*, otherwise, if $x_i < 0$, the individual is in the *domain of losses*. Extensive evidence indicates that all plant and animal life responds to changes in stimuli from a status-quo, or *reference level*. The *reference dependence* of human preferences is widely documented.[1] $D(0, t)$ is a discount function that discounts the utility at time $t \geq 0$ *back to the present, time* $t = 0$. LP proposed the general hyperbolic discounting function for $D(0, t)$.[2]

Our main contribution is the introduction of a *reference time* in addition to the reference outcome level (Section 3). In LP, one has a reference outcome level but, by default, the time period 0 is the reference time to which all outcomes are discounted back. In many cases it may be desirable to have a reference time, $r \geq 0$. Consider the following two examples.

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

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

The exponentially discounted utility model (EDU) was suggested by Samuelson (1937) as a purely technical device because it turned out to be the unique functional form that

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

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

gave rise to time consistent choices.[3] Samuelson was under no illusion that such a model captures human behavior over time.[4] Indeed, the subsequent evidence has rejected the EDU model on several counts. This has led to several well known anomalies of the EDU model.[5] The anomalies of the EDU model are briefly outlined in section 2. These include the magnitude effect, gain-loss asymmetry, common difference effect, delay-speedup asymmetry and intransitivity of preferences.

LP gave conditions on the utility function in (1), v, that could explain the magnitude effect and the gain-loss asymmetry. al-Nowaihi and Dhami (2009) proposed the *simple increasing elasticity* (SIE) class of utility functions (SIE) that satisfy the LP conditions (see section 2.1 below). The delay-speedup asymmetry was explained using (1) by Loewenstein (1988). The common difference effect was explained by LP using *declining impatience* (Definition 4 below) through their general hyperbolic discounting function. Scholten and Read (2006) showed that it could also be explained by *subadditivity* of the discount function (Definition 3 below). Intransitivity of time preferences can be incorporated through a variety of attribute models (Rubinstein, 2003; Manzini and Mariotti, 2006). Yet another explanation relies on models that do not assume transitivity of preferences, such as Ok and Masatlioglu (2007). However, it cannot account for either gain-loss asymmetry or delay-speedup asymmetry.

We alter the LP framework in two ways in Section 3. First, we augment it to allow for a reference time as well as a reference outcome level. Second, we consider a general class of discount functions, D, that allow for declining impatience, subaddivity or a combination of both. Our main results are as follows. In Section 5 we show that if the discount function is *additive* (Definition 3 below) then preferences do not depend on the reference time. Then we show that some (not all) of the intransitivities may be due to reference time shifts: One's reference time may depend on the different binary choices under consideration. We do not claim that this would account for all intransitivities that may be observed in time preferences.

In Section 6, we compare our reference time theory with other recent approaches in the literature. We show that the tradeoff model of Scholten and Read (2010) and Read and Scholten (2006) can be accommodated within the reference time theory. Using examples we also show that the treatment of uncertainty in Halevy (2007) and the vague time preference model of Manzini and Mariotti (2006) can also be incorporated within the reference time time. Finally we compare our model with the model of intransitive preferences proposed in Ok and Masatlioglu (2007).

[3] In (1) under EDU, $D(0,t) = e^{-\beta t}$, $\beta > 0$.

[4] Samuelson (1937) wrote: "It is completely arbitrary to assume that the individual behaves so as to maximize an integral of the form envisaged in [the EDU model]".

[5] We refer the reader to Frederick et al. (2002) and Loewenstein and Prelec (1992).

Proofs are in the Appendix.

2 Anomalies of the exponentially discounted utility model (EDU)

The EDU model has been heavily refuted by the evidence.[6] We provide a brief outline of some of the anomalies of the EDU model in this section. Unlike the assumption of constant discounting in EDU, empirically one finds that larger magnitudes are discounted less (*magnitude effect*) and losses are discounted less relative to gains (*gain-loss asymmetry*). One also finds that unlike the assumption of constant discount rate in EDU, as one considers shorter time horizons, the discount rate increases, i.e., individuals appear to be less patient for immediate rewards (*common difference effect*). One also finds that people require a larger premium to delay a reward as compared to the payment that they are willing to make to speedup a reward by the same time interval (*delay-speedup asymmetry*). These anomalies are stated below as A1-A4 and, we believe, they must form an essential core of any plausible theory of time discounting.

A1 *Magnitude effect.* If $0 < x < y$, $v(x) = v(y) D(0,t)$ and $a > 1$, then $v(ax) < v(ay) D(0,t)$. If $y < x < 0$, $v(x) = v(y) D(0,t)$ and $a > 1$, then $v(ax) > v(ay) D(0,t)$.

A2 *Gain-loss asymmetry.* If $0 < x < y$ and $v(x) = v(y) D(0,t)$, then $v(-x) > v(-y) D(0,t)$.

A3 *Common difference effect.* If $0 < x < y$, $v(x) = v(y) D(0,t)$ and $s > 0$, then $v(x) D(0,s) < v(y) D(0,s+t)$. If $y < x < 0$, $v(x) = v(y) D(0,t)$ and $s > 0$, then $v(x) D(0,s) > v(y) D(0,s+t)$.[7]

A4 *Delay-speedup asymmetry.* For $c > 0$, $s > 0$ and $t > 0$, $v(c) D(0,s) + v(-c) D(0,s+t) < -[v(-c) D(0,s) + v(c) D(0,s+t)]$, i.e., the loss, $-[v(-c) D(0,s) + v(c) D(0,s+t)]$, from a delay, t, in receiving a real reward, c, is greater than the gain from bringing the reward forward. (Neo)classically, these quantities should be the same.[8]

[6] See, for instance, Frederick et al. (2002) and Loewenstein and Prelec (1992).

[7] The following is a well known example of the common difference effect due to Richard Thaler. Most people would prefer one apple today to two apples tomorrow. However, people would also typically prefer two apples in 52 days to one apple in 51 days. Since the time interval between the two choices is one day, exponential discounted utility that makes the assumption of constant discount rates, cannot account for this example. Models of hyperbolic discounting (see below) can easily account for these sorts of results on account of declining impatience (see Definition 4).

[8] Delay-speedup asymmetry is of great importance in the online purchase of items. For instance, in most online orders, one can pay extra to speedup delivery of an item or get a price reduction to downgrade one's order from an express delivery to a standard delivery.

For future reference, we now state some results from Loewentein and Prelec (1992).

Proposition 1 *(LP, p583): For a continuous discount function, A1 gives:*
(a) $0 < x < y \Rightarrow \frac{v(x)}{v(y)} > \frac{v(-x)}{v(-y)}$,
(b) $x > 0 \Rightarrow \epsilon_v(-x) > \epsilon_v(x)$, *where* $\epsilon_v(x) = \frac{x}{v}\frac{dv}{dx}$, $x \neq 0$.

Proposition 2 *(LP, p584): For a continuous discount function, A2 gives*
(a) subproportionality: $(0 < x < y$ *or* $y < x < 0) \Rightarrow \frac{v(x)}{v(y)} > \frac{v(ax)}{v(ay)}$, *for* $a > 1$,
(b) $(0 < x < y$ *or* $y < x < 0) \Rightarrow \epsilon_v(x) < \epsilon_v(y)$.

2.1 The Simple increasing elasticity utility function

al-Nowaihi and Dhami (2009) showed that several popular classes of utility functions violate the conclusions of Propositions 1, 2 and, hence, they violate the assumptions, A1 and A2, from which these Propositions are derived. These include CRRA (constant relative risk aversion), HARA (hyperbolic absolute risk aversion), CARA (constant absolute risk aversion), logarithmic and quadratic[9] Therefore, they proposed the class of *simple increasing elasticity (SIE) utility functions* that satisfy the relevant conditions required to explain A1, A2.

Definition 1 *(al-Nowaihi and Dhami, 2009): The simple increasing elasticity utility function is given by:*

$$v(x) = \begin{cases} \frac{\gamma x^{\sigma}}{1-\gamma}\left(\mu + \frac{\theta_+}{\gamma}x\right)^{1-\gamma}, & x \geq 0, \\ -\lambda\frac{\gamma(-x)^{\sigma}}{1-\gamma}\left(\mu - \frac{\theta_-}{\gamma}x\right)^{1-\gamma}, & x < 0, \end{cases} \tag{2}$$

where $\mu > 0$, $\theta_- > \theta_+ > 0$, $\lambda \geq 1$, $0 < \sigma \leq \gamma < 1$.

It may be interesting to note that (2) is a product of a CRRA function, x^{σ}, and a HARA function, $\frac{\gamma}{1-\gamma}\left(\mu + \frac{\theta_\pm}{\gamma}x\right)^{1-\gamma}$.

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

In this section we outline the *reference time theory* (RT) of al-Nowaihi and Dhami (2008b). Consider a decision maker who, at time t_0, takes an action that results in the outcome w_i at time t_i, $i = 1, 2, ..., n$, where

$$t_0 \leq r \leq t_1 < ... < t_n. \tag{3}$$

[9] The latter three classes of functions are also regarded as members of the HARA family.

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

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

The utility function, v, has four main properties: *reference dependence*, *monotonicity*, *declining sensitivity*, and *loss aversion*.[10] There is good empirical support for these features; see, for instance, Kahneman and Tversky (2000). In particular, v satisfies:

$$v : (-\infty,\infty) \to (-\infty,\infty) \text{ is increasing (monotonicity),} \tag{4}$$

$$v(0) = 0 \text{ (reference dependence),} \tag{5}$$

$$\text{For } x > 0\text{: } -v(-x) > v(x) \text{ (loss aversion),} \tag{6}$$

$$v \text{ is concave for gains, } x \geq 0\text{, but convex for losses, } x \leq 0 \text{ (declining sensitivity).} \tag{7}$$

Furthermore, it is assumed that

$$v \text{ is continuous, and twice differentiable except at } 0. \tag{8}$$

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

$$(w_1,t_1) \preceq_{w_0,r} (w_2,t_2) \Leftrightarrow v(w_1 - w_0) D(r,t_1) \leq v(w_2 - w_0) D(r,t_2). \tag{9}$$

Let $\mathbf{S}$ be a non-empty set of outcome-time sequences from $(-\infty,\infty) \times [0,\infty)$ of the form $(x_1,t_1),(x_2,t_2),...,(x_i,t_i),....$ Using (9), we extend $\preceq_{w_0,r}$ to a complete transitive preference relation on sequences in $\mathbf{S}$, as follows[11]:

$$\begin{aligned}((x_1,s_1),(x_2,s_2),...,(x_m,s_m)) &\preceq_{w_0,r} ((y_1,t_1),(y_2,t_2),...,(y_n,t_n)) \\ &\Leftrightarrow \Sigma_{i=1}^{m} v(x_i) D(r,t_i) \leq \Sigma_{i=1}^{n} v(y_i) D(r,t_i).\end{aligned} \tag{10}$$

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

$$V_r((w_1,t_1),(w_2,t_2),...,(w_n,t_n),w_0) = \Sigma_{i=1}^{n} v(x_i) D(r,t_i), \tag{11}$$

Comparing (11) to the utility function (1), proposed by LP, and successfully employed to explain several anomalies in intertemporal choice, we see that LP implicitly set $r = 0$.

[10]Furthermore, in prospect theory, an important role is played by *non-linear transformation of probabilities*. We consider this in subsection 6.2, below.

[11]The following also holds for infinite sequences, provided the sums in (10) converge.

Also, LP implicitly assume that the discount function is additive (Definition 3, below), in which case there is no loss in assuming $r = 0$. To accommodate the empirical evidence, however, we allow the discount function to be non-additive, in which case the choice of reference time does matter.

3.1 Determination of the reference time

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

$$T = \{t \in [0, \infty) : t = t_i \text{ for some sequence } \{(x_1, t_1), (x_2, t_2), ..., (x_i, t_i), ...\} \text{ in } \mathbf{S}\}. \quad (12)$$

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

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

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

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

4 Discount functions and their properties

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

Definition 2 *(Discount functions): Let*

$$\Delta = \{(r, t) \in \mathbb{R} \times \mathbb{R} : 0 \leq r \leq t\}. \quad (13)$$

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

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

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

$$\textbf{Exponential}\text{: } D(r,t) = e^{-\beta(t-r)},\ \beta > 0. \tag{14}$$

$$\textbf{PPL}\text{: } D(r,t) = \left\{ \begin{array}{ccc} 1 & \text{when} & r = t = 0 \\ e^{-(\delta+\beta t)} & \text{when} & r = 0, t > 0 \\ e^{-\beta(t-r)} & \text{when} & 0 < r \leq t \end{array} \right\},\ \beta > 0,\ \delta > 0. \tag{15}$$

$$\textbf{LP}\text{: } D(r,t) = \left(\frac{1+\alpha t}{1+\alpha r}\right)^{-\frac{\beta}{\alpha}},\ t \geq 0,\ r \geq 0,\ \alpha > 0,\ \beta > 0. \tag{16}$$

$$\textbf{RS}\text{: } D(r,t) = \left[1 + \alpha\left(t^{\tau} - r^{\tau}\right)^{\rho}\right]^{-\frac{\beta}{\alpha}},\ 0 \leq r \leq t,\ \alpha > 0,\ \beta > 0,\ \rho > 0,\ \tau > 0. \tag{17}$$

In addition, we propose the following generalization of (17),

$$\begin{array}{rcl} \textbf{Generalized RS}\text{:} & & D(r,t) = e^{-Q(\phi(t)-\phi(r))},\ 0 \leq r \leq t, \\ Q & : & [0,\infty) \overset{\text{onto}}{\rightarrow} [0,\infty) \text{ is strictly increasing,} \\ \phi & : & [0,\infty) \overset{\text{onto}}{\rightarrow} [0,\infty) \text{ is strictly increasing.} \end{array} \tag{18}$$

The *exponential discount function* (14) was introduced by Samuelson (1937). The main attraction of EDU is that it is the unique discount function that leads to time-consistent choices. The $\beta - \delta$ or *quasi-hyperbolic discount function* (15) was proposed by Phelps and Pollak (1968) and Laibson (1997) and is popular in applied work (we use the acronym PPL for it).[12] The *generalized hyperbolic discount function* (16) was proposed by Loewenstein and Prelec (1992) (we use the acronymn LP for it). These three discount functions are *additive* (Definition 3, below). They can account for the common difference effect through declining impatience (Definition 4, below) but they cannot account for either non-additivity or intransitivity. The interval discount function (17) was introduced by Read (2001) and Scholten and Read (2006) (we use the acronym RS for it). It can account

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

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

Note that (16) approaches (14) as $\alpha \to 0$. In general, neither of (16) or (17) is a special case of the other. However, for $r = 0$ (and only for $r = 0$), (17) reduces to (16) when $\rho = \tau = 1$. While ρ, τ are parameters, r is a variable. Hence, neither discount function is a special case of the other.[13]

Our terminology suggests that a continuous discount function (see the last sentence in Definition 2) is continuous. That this is partly true, is established by the following Proposition.

Proposition 3 : *A continuous discount function, $D(r,t)$, is continuous in t.*

It is straightforward to check that each of (14), (16), (17) and (18) is a continuous discount function in the sense of Definition 2. It is also straightforward to check that (15) is a discount function. The reason the latter is not a *continuous* discount function is that $\lim_{t \to 0+} D(0,t) = e^{-\delta} < 1 = D(0,0)$.

From (16) and (17) we see that the restrictions $r \geq 0$ and $t \geq 0$ are needed. From (17) we see that the further restriction $r \leq t$ is needed.[14] From (14) we see that the 'into' in Definition 2(b) cannot be strengthened to 'onto'.

Proposition 4 *(Time sensitivity): Let D be a continuous discount function. Suppose $r \geq 0$. If $0 < x \leq y$, or if $y \leq x < 0$, then $v(x) = v(y) D(r,t)$ for some $t \in [r, \infty)$.*[15]

Proposition 5 *(Existence of present values): Let D be a discount function. Let $r \leq t$ and $y \geq 0$ ($y \leq 0$). Then, for some x, $0 \leq x \leq y$ ($y \leq x \leq 0$), $v(x) = v(y) D(r,t)$.*

Definition 3 *(Additivity): A discount function, $D(r,t)$, is*

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

In Definition 3, 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

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

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

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

on discounting. However, in the other two cases, it does have an effect. Under subadditive discounting, there is more discounting over the subdivided intervals (future utilities are shrunk more), while the converse is true under superadditive discounting.[16]

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

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

thus, he does not accept the financial investment opportunity.

Example 5 *Consider the RS discount function in (17),* $D(r, t) = [1 + \alpha (t^{\tau} - r^{\tau})^{\rho}]^{-\frac{\beta}{\alpha}}$*, with* $\tau = 1$*,* $\rho = 0.6$*,* $\alpha = \beta = 0.5$*. Let the utility function be* $v(x) = x$ *for* $x \geq 0$*. We can then compute*

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

which, respectively, equal 0.46540, 0.40567, 0.33439*. As expected,* $D(0, t)$ *is declining in* t *and* $D(r, 10)$ *is increasing in* r *(see Definition 2). Let the unit of time be one month. Then a decision maker who discounts over the entire interval* $[0, 10]$ *is indifferent between receiving \$15 in 10 months time and \$5.0159 at date* $r = 0$ *because*

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

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

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

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

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

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

Definition 4 *(Impatience): A discount function,* $D(r, s)$*, exhibits*[17]

$$\begin{cases} \text{Declining impatience if} & D(r,s) < D(r+t, s+t), \quad \text{for } t > 0 \text{ and } r < s, \\ \text{Constant impatience if} & D(r,s) = D(r+t, s+t), \quad \text{for } t \geq 0 \text{ and } r \leq s, \\ \text{Increasing impatience if} & D(r,s) > D(r+t, s+t), \quad \text{for } t > 0 \text{ and } r < s. \end{cases}$$

Exponential discounting exhibits constant impatience while hyperbolic discounting exhibits declining impatience.[18] The RS-discount function (17) allows for additivity, subadditivity and all the three cases in Definition 4, hence, it is of great practical importance. The next definition describes the properties of this function.

Proposition 6 *: Let* $D(r, t)$ *be the RS-discount function (17), then:*
(a) If $0 < \rho \leq 1$*, then* D *is subadditive.*
(b) If $\rho > 1$*, then* D *is neither subadditive, additive nor superadditive.*
(ci) If $0 < \tau < 1$*, then* D *exhibits declining impatience.*
(cii) If $\tau = 1$*, then* D *exhibits constant impatience.*
(ciii) If $\tau > 1$*, then* D *exhibits increasing impatience.*

In the light of Proposition 6, we can now see the interpretation of the parameters τ and ρ in the RS-discount function (17).[19] τ controls impatience, independently of the values of the other parameters α, β and ρ. $0 < \tau < 1$, gives declining impatience, $\tau = 1$ gives constant impatience and $\tau > 1$ gives increasing impatience. If $0 < \rho \leq 1$, then we get subadditivity, irrespective of the values of the other parameters α, β and τ. However, if $\rho > 1$, then (17) can be neither subadditive, additive nor superadditive.[20]

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

[18] Hyperbolic discounting, through declining impatience, explains many important phenomena that are difficult to explain under exponential discounting. A small sample includes the following. Under-savings, sharp drop in consumption at retirement, procrastination, choice between annual gym memberships and paying on a pay-per-visit basis, addictions, obesity, failure of New Year resolutions etc.

[19] Scholten and Read (2006a), bottom of p. 1425, state: $\alpha > 0$ implies subadditivity (incorrect), $\rho > 1$ implies superadditivity (incorrect) and $0 < \tau < 1$ implies declining impatience (correct but incomplete).

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

4.1 The common difference effect: Declining impatience or subadditivity?

Read (2001), conducted a series of experiments that tested for the common difference effect and could also discriminate between subadditivity and declining impatience. He found support for the common difference effect and for subadditivity but rejected declining impatience in favour of constant impatience. Read (2001) also discusses the psychological foundation for subadditivity.

5 Results on reference time theory

5.1 Additivity of discount function and reference time theory

Proposition 7, below, establishes that, for an additive discount function, discounting back to time r is equivalent to discounting back to time 0. Thus the choice of reference time can only make a difference if the discount function is non-additive.

Proposition 7 *(Invariance to the choice of reference time) If $D(r,t)$ is additive, then*

$$\begin{aligned} &((x_1,s_1),(x_2,s_2),...,(x_m,s_m)) \preceq_{w0,r} ((y_1,t_1),(y_2,t_2),...,(y_n,t_n)) \\ \Leftrightarrow\ &((x_1,s_1),(x_2,s_2),...,(x_m,s_m)) \preceq_{w0,0} ((y_1,t_1),(y_2,t_2),...,(y_n,t_n)). \end{aligned}$$

5.2 Are intransitivities due to shifts in reference time?

We now investigate if some observed intransitive choices can be explained by shifts in the reference time. Consider the following hypothetical situation. A decision maker prefers a payoff of 1 now to a payoff of 2 next period. The decision maker also prefers a payoff of 2 next period to a payoff of 3 two periods from now. Finally, the same decision maker prefers a payoff of 3 two periods from now to a payoff of 1 now. Schematically:

$$(1, \text{now}) \succ (2, \text{next period}) \succ (3, 2 \text{ two periods from now}) \succ (1, \text{now}). \tag{19}$$

Ok and Masatlioglu (2007, p215) use a similar example to motivate their intransitive theory of relative discounting.

Alternatively, we may view (19) as due to a framing effect resulting in a shift in the reference point for time. Assume that the choice of reference time in each pairwise comparison is the sooner of the two dates, in conformity with Assumption A0. Then (19) can be formalized as follows:

$$V_0(1,0) > V_0(2,1); \ V_1(2,1) > V_1(3,2); \ V_0(3,2) > V_0(1,0). \tag{20}$$

If this view is accepted, then the apparent intransitivity in (19) arises from conflating $V_0(3,2)$ with $V_1(3,2)$ and $V_1(2,1)$ with $V_0(2,1)$. The following example shows that (20) is consistent with a reference-time theory of intertemporal choice.

Example 6 : *Take the reference point for wealth to be the current level of wealth, so each payoff is regarded as a gain to current wealth. Take the utility function to be the SIE utility function, (2), with* $\mu = 1$, $\theta_+ = 0.5$, $\lambda = 2$, $0 < \sigma = \gamma = 0.5$, *so that*

$$v(x) = x^{\frac{1}{2}} (1+x)^{\frac{1}{2}}, \; x \geq 0. \tag{21}$$

Thus (working to five significant figures),

$$v(1) = 1.4142, \; v(2) = 2.4495 \text{ and } v(3) = 3.4641. \tag{22}$$

As our discount function we take the Read-Scholten discount function (17) with $\alpha = \beta = 1$ *and* $\rho = \tau = \frac{1}{2}$*:*

$$D(r,t) = \left(1 + \left(t^{\frac{1}{2}} - r^{\frac{1}{2}}\right)^{\frac{1}{2}}\right)^{-1}. \tag{23}$$

Thus,

$$D(0,1) = \frac{1}{2}, \; D(1,2) = 0.60842 \text{ and } D(0,2) = 0.45679. \tag{24}$$

From (22) and (24) we get

$$V_0(1,0) = v(1) D(0,0) = 1.4142, \tag{25}$$

$$V_0(3,2) = v(3) D(0,2) = 1.5824, \tag{26}$$

$$V_1(3,2) = v(3) D(1,2) = 2.1076, \tag{27}$$

$$V_1(2,1) = v(2) D(1,1) = 2.4495, \tag{28}$$

$$V_0(2,1) = v(2) D(0,1) = 1.2248. \tag{29}$$

From (25) to (29), we get

$$V_0(1,0) > V_0(2,1), \; V_1(2,1) > V_1(3,2), \; V_0(3,2) > V_0(1,0), \tag{30}$$

confirming (20).

A consequence of Proposition 7 is that no additive discount function (e.g., exponential (14), PPL (15) or LP (16)) can explain (apparently) intransitive choices as exhibited in (19). The reason is that, under the conditions of that proposition, all utilities can be discounted back to time zero and, hence, can be compared and ordered.

5.3 Explanation of the anomalies of EDU by RT theory

RT can explain the magnitude effect (A1) and gain-loss asymmetry (A2) using the SIE class of utility functions in Section 2.1; indeed this has been shown formally in al-Nowaihi and Dhami (2009). The common difference effect can be explained under RT by using either the PPL or LP discount function, given in (15), (16) respectively, hence, relying on declining impatience. However, under RT one may also use the RS discount function in (17) to explain the common difference effect; this can be on account of subadditivity, declining impatience or a combination of both. The delay-speedup asymmetry was explained using the insights of Loewenstein (1988) by Dhami and al-Nowaihi (2008b) using RT theory. Finally, the explanation of some types of intransitive preferences is explained in the current paper (see Section 5.2). Clearly an even richer model may be needed to explain other sorts of observed intransitivies.

6 The relation of RT with other theories

In this section, we compare the reference-time theory (RT) of section 3 with four recent developments.

First, we consider the tradeoff model of intertemporal choice of Scholten and Read (2010) and Read and Scholten (2006), SR for short. We will argue that SR's tradeoff criterion can be represented by a discount function. Hence, it can be incorporated within RT. The gain is that their psychological arguments for their tradeoff model gives support for RT theory and, in particular, their own RS-discount function.

The second development we consider is Halevy (2007), H for short, who shows that the common difference effect is compatible with exponential discounting, provided subjects are non-expected utility maximizers and exhibit the *certainty effect*. The certainty effect was first proposed as an explanation of the Allais paradox: subjects are much more sensitive to a change from certainty to uncertainty than they are to changes in the middle range of probabilities. The third is the theory of vague time preferences of Manzini and Mariotti (2006), MM for short. Again, they can explain the common difference effect without departing from exponential discounting.

We believe that the importance of H and MM far transcends their ability to explain the common difference effect. On the other hand, and because in their present formulations they do not include any reference dependence, they are unable to explain *gain-loss asymmetry*, *delay-speedup asymmetry*, *subadditivity* and (apparent) *intransitivity*. By contrast, RT theory can explain all the anomalies. Nevertheless, we believe that it is desirable, and easy, to extend RT theory to incorporate uncertainty, as in H, and multiple criteria, as in MM. While we do not formally prove this assertion, we make this point, below, in the context of simple examples.

The fourth recent development we discuss here is the theory of intransitive preferences and relative discounting of Ok and Masatlioglu (2007), OM for short. This is the most radical of all the theories considered so far. From the outset it neither assumes transitivity nor additivity and, hence, is compatible with these two phenomena. However, in its present formulation, it cannot account for either *gain-loss asymmetry* or *delay-speedup asymmetry*. Furthermore, the lack of transitivity will make it hard to work with this theory, as the authors themselves explain. On the other hand, these problems can all be resolved in the special case of a transitive preference relation. But then their model becomes additive, in which case OM reduces to a standard discounting model.

Finally, all five theories (SR, H, MM, OM and RT) can explain the magnitude effect, when combined with the SIE value function.

6.1 The tradeoff model of intertemporal choice

Read and Scholten's critique of discounting models, including their own, led them to develop their *tradeoff model* of intertemporal choice (Read and Scholten, 2006; Scholten and Read, 2010). This model is an important advance in attribute based models of intertemporal choice that can explain many anomalies of the EDU model. We argue that the tradeoff model of Scholten and Read (2010) and Read and Scholten (2006) can be incorporated within RT-theory. If this is accepted, then their tradeoff model lends further support to the RT-theory and, in particular, their own discount function (17) and its generalization (18), above.

We proceed by first recasting their model 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. They develop their model through three successive versions. We concentrate on their fourth and final version, page 15.

Let $r \geq 0$ be the reference point for time.[21] The tradeoff 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 the reference 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, y received at time t is equivalent to x received at time s. These relationship are established using three functions, a *value function*, u, a *tradeoff function* Q and a *delay-perception function*, ϕ.

We assume that $Q : [0, \infty) \overset{\text{onto}}{\to} [0, \infty)$ is strictly increasing, $\phi : [0, \infty) \overset{\text{onto}}{\to} [0, \infty)$ is

[21] In Scholten and Read (2006b), $r = 0$. To ease the burden of notation, we shall suppress reference to the reference point for wealth, w_0. Thus, in what follows, we write $\prec_r$ and $\sim_r$ when we should have written $\prec_{r,w_0}$ and $\sim_{r,w_0}$, respectively.

strictly increasing (the same as in (18), above).[22]

Let $x > 0$, $y > 0$ and $s \geq r \geq 0$, $t \geq r$. The decision criteria in this model is given by:

$$(x,s) \sim_r (y,t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] = u(y) - u(x), \tag{31}$$

$$(x,s) \prec_r (y,t) \Leftrightarrow Q[\phi(t) - \phi(r)] - Q[\phi(s) - \phi(r)] < u(y) - u(x). \tag{32}$$

To understand these inequalities, suppose that $x < y$ and $s < t$. The decision maker then has a choice between a *smaller-sooner* (SS) reward, (x,s), and a *larger later* (LL) reward, (y,t). The two *attributes* are outcome and time. 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 inequalities in (31), (32). Alternatively this may be termed as the *disadvantage of the LL reward*. The *advantage of the LL reward* (alternatively the *disadvantage of the SS reward*) is that it offers a higher outcome; this is indicated by the RHS of the inequalities in (31), (32). Thus, in (32), the LL reward is strictly preferred to the SS reward if the advantage of the SS reward is smaller than the advantage of the LL reward.

We now state the analogue of the decision criteria in (31), (32) when the outcomes are losses: $x < 0$, $y < 0$ and (as before) $s \geq r \geq 0$, $t \geq r$. In this case:

$$(x,s) \sim_r (y,t) \Leftrightarrow Q(\phi(t) - \phi(r)) - Q(\phi(s) - \phi(r)) = u(x) - u(y), \tag{33}$$

$$(x,s) \prec_r (y,t) \Leftrightarrow Q(\phi(t) - \phi(r)) - Q(\phi(s) - \phi(r)) > u(x) - u(y). \tag{34}$$

For losses, let $y < x < 0$. Then, the RHS of the inequality in (34), $u(x) - u(y)$, is the advantage of the SS reward, (x,s). Since both rewards are losses, the LHS of (34) becomes the advantage of the LL reward, (y,t).

For completeness, we also need (again, $s \geq r \geq 0$, $t \geq r$) to specify the following properties:

$$(0,s) \sim_r (0,t), \tag{35}$$

$$x < 0 \Rightarrow (x,s) \prec_r (0,t), \tag{36}$$

$$y > 0 \Rightarrow (0,s) \prec_r (y,t), \tag{37}$$

$$x < 0, y > 0 \Rightarrow (x,s) \prec_r (y,t). \tag{38}$$

From (35), the time at which a zero outcome is received is irrelevant. From (36), a reward of zero is always preferred to a negative reward, irrespective of the time. From (37), a positive reward is always preferred to a zero reward, irrespective of the time. From (38), a positive outcome is always preferred to a negative one, irrespective of the time.

[22] 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 w are in line with what they intend (see their equations (2) and (5) for the earlier, and simpler, versions of their model).

To get the tradeoff model of Read and Scholten, set $r = s$ in the above equations.[23] However, RT theory allows for the reference time $r \geq s$.

To define a discount function, D, that expresses these preferences, let

$$v(x) = e^{u(x)}, \text{ for } x > 0, \quad (39)$$

$$v(x) = -e^{-u(x)}, \text{ for } x < 0. \quad (40)$$

Then all the above relations, (31) to (38), can be summarized by the following. For all x, y and all r, s, t such that $s \geq r \geq 0$, $t \geq r$:

$$(x, s) \sim_r (y, t) \Leftrightarrow v(x) e^{-Q[\phi(s)-\phi(r)]} = v(y) e^{-Q[\phi(t)-\phi(r)]}, \quad (41)$$

$$(x, s) \prec_r (y, t) \Leftrightarrow v(x) e^{-Q(\phi(s)-\phi(r))} < v(y) e^{-Q[\phi(t)-\phi(r)]}. \quad (42)$$

(41) and (42) suggest we take our discount function to be the generalized RS function (18), which is a generalization of the discount function (17) of Scholten and Read (2006). Thus, RT-theory can incorporate the tradeoff model.

6.2 The certainty effect

A test of a theory (T) is always a test of T plus auxiliary assumptions (A). Thus, a refutation of T&A may be a refutation of A rather than T. However, since A is often left implicit, a refutation of T&A may be misconstrued as a refutation of T rather than A. A case in point may be T = ‘exponential discounting’ and A = ‘uncertainty is not relevant’.

Despite the experimenters’ best efforts to eliminate uncertainty, there will always be a residual uncertainty in the minds of the experimental subjects that they may not receive their promised payoffs. If subjects were expected utility (EU) maximizers, then risk would not matter (Example 7, below). However, if subjects overweight low probabilities and underweight high probabilities, as in many non-EU theories, then risk matters (Example 8, below). Moreover, the *lower* the residual risk the *greater* will be its effect (Example 9, below).[24] Thus, Halevy (2007) argues that the common difference effect may, in fact, be a refutation of EU rather than exponential discounting. We now show how Halevy’s ideas can be incorporated within the framework of RT theory. We do not give any general results but argue through examples.

The above points are illustrated by the following three examples. They all involve a choice between receiving \$1000 now or \$1100 next year and, simultaneously, a choice

[23] They explicitly state only (31) and (33) (with $r = s$). However, we believe that our other equations are in line with their framework.

[24] So, experimenters, by doing their best to reduce residual risk with the aim of getting a sharper refutation of exponential discounting may, actually, achieve the opposite.

between receiving these two sums 10 and 11 years from now, respectively. We use the SIE utility function (2), for the parameter values in (21) so that

$$v(1000) = 1000.5 \text{ and } v(1100) = 1100.5. \tag{43}$$

Let the discount function be $D(s,t)$ and the *probability weighting function* be $w(p)$, where p is the probability perceived by experimental subjects that the payoff will actually be paid one year from now. Probability weighting functions are a fundamental feature of almost all non-expected utility theories, in particular, rank dependent utility, prospect theory, cumulative prospect theory etc.

Definition 5 *(Probability weighting function): By a probability weighting function we mean a strictly increasing function* $w : [0,1] \overset{onto}{\longrightarrow} [0,1]$.

To fix ideas, readers may, for the moment, consider $w(p)$ as the subjective weight placed by decision makers on the objective probability, p. Expected utility theory corresponds to the special case $w(p) = p$.

We assume independence across years so that the probability of receiving the payoff t years from now is p^t. Let $(x,s) \prec (y,t)$ mean (y,t) is strictly preferred to (x,s). We take the current level of wealth, w_0, and present time, $r = 0$, to be the reference points for wealth and time, respectively (and, to simplify notation, we have dropped the subscripts, w_0, r, from $\prec_{w_0,r}$). We thus have:

$$\text{No common difference effect: } (1100,1) \prec (1000,0) \Rightarrow (1100,11) \prec (1000,10) \tag{44}$$

A decision maker who follows non-expected utility (say, rank dependent utility or prospect theory or cumulative prospect theory) evaluates the preferences expressed in (44) as follows.

$$(1100,1) \prec (1000,0) \Leftrightarrow v(1100)\,D(0,1)\,w(p) < v(1000), \tag{45}$$

$$(1100,11) \prec (1000,10) \Leftrightarrow v(1100)\,D(0,11)\,w\left(p^{11}\right) < v(1000)\,D(0,10)\,w\left(p^{10}\right). \tag{46}$$

Example 7 *(Exponential discounting with expected utility): Assume exponential discounting, so* $D(s,t) = e^{-\beta(t-s)}$, $\beta > 0$, *and expected utility, so* $w(p) = p$. *Substitute in (45), (46) to get:*

$$(1100,1) \prec (1000,0) \Leftrightarrow v(1100)\,e^{-1}p < v(1000), \tag{47}$$

$$(1100,11) \prec (1000,10) \Leftrightarrow v(1100)\,e^{-11\beta}p^{11} < v(1000)\,e^{-10\beta}p^{10}, \tag{48}$$

(48) is equivalent to:

$$(1100,11) \prec (1000,10) \Leftrightarrow v(1100)\,e^{-1}p < v(1000). \tag{49}$$

From (47) and (49) we see that $(1100, 1) \prec (1000, 0) \Leftrightarrow (1100, 11) \prec (1000, 10)$. *Thus, exponential discounting together with expected utility*[25] *imply no common difference effect.*
Conclusion from Example 7: The observation of a common difference effect is a rejection of the joint hypothesis of exponential discounting and expected utility. Thus, it would imply the rejection of one or the other (or both) but not, necessarily, exponential discounting.

Example 8 *(Exponential discounting with non-expected utility): We take cumulative prospect theory (Tversky and Kahneman, 1992) as our model of non-expected utility. Take* $D(r, t) = e^{-\beta(t-r)}$, $\beta = 0.04$, *and* $w(p) = e^{-(-\ln p)^{\alpha}}$ *(the Prelec probability weighting function*[26]*) with* $\alpha = 0.5$. *Assume that there is 'near certainty', so,* $p = 0.98$. *Substitute in (45), (46), using (43), to get:*

$$(1100, 1) \prec (1000, 0) \Leftrightarrow 1100.5e^{-0.04}e^{-(-\ln 0.98)^{0.5}} < 1000.5, \tag{50}$$

$$(1100, 11) \prec (1000, 10) \Leftrightarrow 1100.5e^{-0.44}e^{-\left(-\ln(0.98)^{11}\right)^{0.5}} < 1000.5e^{-0.4}e^{-\left(-\ln(0.98)^{10}\right)^{0.5}}. \tag{51}$$

A calculation shows that the inequality in (50) holds while the corresponding inequality in (51) does not hold. Hence, $(1100, 1) \prec (1000, 0)$ *but* $(1100, 11) \succ (1000, 10)$.
Conclusion from Example 8: Exponential discounting may be consistent with an observation of a common difference effect, if subjects do not behave according to expected utility.

Example 9 *(It is the certainty effect that is doing the work): Consider Example 8, except that now there is considerable uncertainty,* $p = 0.4$, *rather than near certainty,* $p = 0.98$. *Substitute in (45), (46), using (43), to get:*

$$(1100, 1) \prec (1000, 0) \Leftrightarrow 1100.5e^{-0.04}e^{-(-\ln 0.4)^{0.5}} < 1000.5, \tag{52}$$

$$(1100, 11) \prec (1000, 10) \Leftrightarrow 1100.5e^{-0.44}e^{-\left(-\ln(0.4)^{11}\right)^{0.5}} < 1000.5e^{-0.4}e^{-\left(-\ln(0.4)^{10}\right)^{0.5}}. \tag{53}$$

A calculation shows that the inequalities in (52) and (53) both hold. Hence, $(1100, 1) \prec (1000, 0)$ *and* $(1100, 11) \prec (1000, 10)$.
Conclusion from Example 9: The common difference effect is due to the certainty effect in particular, rather than uncertainty as such.

Example 9 suggests that if the common difference effect is due to the certainty effect alone, rather than a combination of the certainty effect and non-exponential discounting, then the phenomenon should disappear for probabilities around 0.4.

[25] Or no risk so that $p = 1$ and, hence, $w(p) = w(1) = 1$.
[26] See Prelec (1998), Luce (2001) and al-Nowaihi and Dhami (2006b).

6.3 Vague time preferences

Manzini and Mariotti (2006) develop a theory of vague time preferences and discuss the psychological foundations for such an approach. In this theory, the choice between, say, receiving \$1000 now and \$1100 next year is 'clearer' than the choice between these two sums received 10 and 11 years from now, respectively. MM propose three criteria to choose between (x, s) and (y, t). The *primary criterion* is to choose whichever has the highest present utility value. If the two present values are not 'significantly' different, then the *secondary criteria* requires the subject to choose the option with the highest monetary value. If the secondary criterion fails, then the subject behaves according to the third criterion: 'choose the outcome that is delivered sooner'. If all three criteria fail, then the subject is indifferent. Thus, MM achieve a complete but intransitive ordering. In particular, indifference here is not an equivalence relationship. Suppose that two present values are significantly different if their difference is greater than $\mu > 0$.

We now restate the model of vague time preferences in terms of RT theory using the preference relation $\prec_{w_0,r}$. $(x, s) \prec_{w_0,r} (y, t)$ if, and only if, one of the following holds in the order given below:

1. $v(y) D(r,t) - v(x) D(r,s) > \mu$, or
2. $|v(y) D(r,t) - v(x) D(r,s)| \leq \mu$, and $x < y$, or
3. $|v(y) D(r,t) - v(x) D(r,s)| \leq \mu$, $x = y$ and $t < s$.

Obviously, if x and y are vectors, then extra criteria can be added.[27] Present utility values whose difference is less than μ are regarded as not significantly different. This could be because, for example, the decision maker is not sure of the appropriate utility function or discount function to use. Therefore, the decision maker does not want the decision to depend too critically on the choice of these functions. On the other hand, the decision maker might be absolutely sure that more is better than less and sooner is better than later. Example 10, below, shows how this theory can explain the common difference effect.

Example 10 : *Consider the choice between receiving \$1000 now and \$1100 next year and the choice between these two sums received 10 and 11 years from now, respectively. We use the SIE value function (2), for the same parameter values as in (21) so that (43) holds. We use the exponential discount function (14) with* $\beta = 0.1$ *and the reference time* $r = 0$*, so* $D(0,t) = e^{-0.1t}$. *We take* $\mu = 3$*, so that present utility values whose difference is less than 3 are regarded as not significantly different. Using these values, we get* $v(1000) - v(1100) e^{-0.1} = 1000.5 - 1100.5e^{-0.1} = 4.7264 > 3$. *Hence, the*

[27] More generally, σ is a 'vagueness function', in which case '$|v(y) D(r,t) - v(x) D(r,s)| \leq \mu$' is replaced by '$v(y) D(r,t) - v(x) D(r,s) \leq \sigma(x; r, s)$ and $v(x) D(r,s) - v(y) D(r,t) \leq \sigma(y; r, t)$'

primary criterion holds and the decision maker prefers \$1000 now to \$1100 next year. Next, consider the two choices in 10 and 11 years time: $|v(1100)e^{-1.1} - v(1000)e^{-1}| = |1100.5e^{-1.1} - 1000.5e^{-1}| = 1.7388 < 3$*. Hence, the primary criterion fails, and the decision maker considers the second criterion. Since* $1000 < 1100$*, the second criterion holds. The decision maker prefers \$1100 received 11 years from now to \$1000 received 10 years from now. Thus, we have an illustration of the common difference effect.*

The experimental results of Roelofsma and Read (2000) supported 'sooner is better than larger' against 'larger is better than sooner'. However, if the order of the secondary criteria is reversed, so that sooner is better than larger (in agreement with the experimental results of Roelofsma and Read, 2000), then \$1000 received 10 years from now would be better than \$1100 received 11 years from now, and we would not get a common difference effect.

However, whether MM's explanation of the common difference effect is acceptable or not, to us the main contribution of their paper lies in the use of primary and secondary criteria. This appears to us to be a more accurate description of actual decision making compared to the assumption of a single criterion.

6.4 Intransitive preferences and relative discounting

Ok and Masatlioglu (2007) (henceforth OM) develop a theory of intransitive time preferences. At time 0, a decision maker has a binary relationship, $\preceq$, on the set $\chi = X \times [0, \infty)$, where X is a non-empty set. Let $x, y \in X$ and $s, t \in [0, \infty)$, then $(x, s) \preceq (y, t)$ is to be interpreted in the following manner. If the individual is asked at time 0 to commit to a choice between (x, s) and (y, t) he states that 'y received at time t is (weakly) preferred to x received at time s'. For this reason, $\preceq$ may be described as representing the *commitment preferences* of the individual.

Let $\sim$ and $\prec$ be the symmetric and asymmetric parts of $\preceq$, respectively. For each $t \in [0, \infty)$, $\preceq_t$ is the t-th time projection of $\preceq$ onto X, i.e., $x \preceq_t y$, if, and only if, $(x, t) \preceq (y, t)$. In particular, $\preceq_0$ is the projection of $\preceq$ onto X at time 0 (and, similarly, for $\sim_t$ and $\prec_t$). Thus, $\preceq_0$ describes the undated preferences of the individual in the present, i.e., time 0.

If X is a metric space, then further structure can be imposed on $\preceq$.

Definition 6 *(Time Preference; OM, p.218): Let* X *be a metric space, then* $\preceq$ *is a time preference on* χ *if*
(i) $\preceq$ *is complete and continuous,*
(ii) $\preceq_0$ *is complete and transitive,*
(iii) $\preceq_t = \preceq_0$ *for each* $t \in [0, \infty)$.

In Definition 6, note that transitivity is imposed on $\preceq_0$ (and, hence, also on $\preceq_t$) but transitivity is not imposed on $\preceq$. In other words, static preferences ($\preceq_t$, $\preceq_0$) are always transitive but intransitivity may arise with the passage of time. Hence, neither $\prec$ nor $\sim$ are, necessarily, transitive. In particular, $\sim$ is *not*, in general, an equivalence relationship.

Example 11 *: Suppose that $x, y, z \in X$. Then Definition 6(ii) implies that if $(x, 0) \preceq_0 (y, 0)$ and $(y, 0) \preceq_0 (z, 0)$ then, by transitivity, we get $(x, 0) \preceq_0 (z, 0)$. Further, from Definition 6(iii) we get that for any $t \in [0, \infty)$ the t-th time projection of $\preceq$ onto X satisfies: if $(x, t) \preceq_t (y, t)$ and $(y, t) \preceq_t (z, t)$ then, $(x, t) \preceq_t (z, t)$. However, from Definition 6(i) we may have for $s, t, u \in [0, \infty)$: $(x, s) \preceq (y, t)$ and $(y, t) \preceq (z, u)$ but $(z, u) \preceq (x, s)$, which violates transitivity.*

OM introduce a set of six properties, P1-P6. We only give a heuristic discussion of these properties because we only need the theorems that OM derive from them. *Time sensitivity* (P1) requires that any outcome can be made less desirable than another outcome by postponing it for a sufficiently long period. *Outcome sensitivity* (P2) requires that any delay can be compensated by a high enough prize. *Monotonicity* (P3) requires that $\preceq$ be increasing in outcomes and decreasing in time. Suppose that $(x, 0) \sim (y, 1)$ and $(z, 0) \sim (w, 1)$, then the *delay premium* in these respective cases is $y - x$ and $w - z$. *Separability* (P4) requires that if the delay premium for delaying a reward x at any time $t \in [0, \infty)$ for a time interval of length, l, is $y - x$, then over the same interval, $w - z$ is the delay premium for delaying a reward of w at time t. *Path independence* (P5) is less persuasive as compared to the other properties. It requires that for any outcomes $x, y, w, z \in X$ and times $t_1, t_2, t_3 \in [0, \infty)$ if $(x, t_1) \sim (y, t_2) \sim (w, t_3)$ then $(x, t_2) \sim (z, t_3)$ where $(z, t_1) \sim (w, t_2)$. Finally *monotonicity in outcomes* (P6) requires that someone who uses the preference $\preceq_0$ prefers a higher to a lower amount.

The next proposition gives the implications of these properties. Let $\mathbb{R}$ be the set of real numbers, $\mathbb{R}_+$ the set of non-negative reals and $\mathbb{R}_{++}$ the set of positive reals. Recall that a homeomorphism is a mapping that is 1-1, onto, continuous and its inverse is also continuous.

Proposition 8 *(OM, Theorem 1): Let X be an open interval in $\mathbb{R}$ and $\preceq$ a binary relation on χ. $\preceq$ is a time preference on χ that satisfies properties (P1)-(P6) if, and only, there exists an increasing homeomorphism $U : X \to \mathbb{R}_{++}$ and a continuous map $D : \mathbb{R}_+^2 \to \mathbb{R}_{++}$ such that, for all $x, y \in X$ and $s, t \in [0, \infty)$,*

$$(x, s) \preceq (y, t) \;\; \textit{iff} \; U(x) \leq U(y) D(s, t), \tag{54}$$

while (i) for given s, $D(s, t)$ is decreasing in t with $D(s, \infty) = 0$, and (ii) $D(t, s) = 1/D(s, t)$.

Suppose $s \leq t$. Then (54) says that y received at time t is (weakly) preferred to x received at time s if, and only if, the (undated) utility of x is less or equal to the (undated) utility of y discounted from time t back to time s by the discount factor $D(s,t)$. In this case, part (i) of Proposition 8 implies the following. Fix the time, s, at which x is received. Let the time, t, at which y is received, recede into the future. Then the value of the utility of y, discounted back to time s, decreases. In the limit, as the receipt of y is indefinitely postponed, the value of its utility, discounted back to time s, approaches zero. Part (ii) of Proposition 8 says that compounding forward, from time s to time t, is the inverse of discounting backwards from time t to time s.

For each $r \in [0,\infty)$, let $\preceq^r$ be the restriction of $\preceq$ to $X \times [r,\infty)$, i.e., to times $t \geq r$. Thus, for $r \leq s$ and $r \leq t$, $(x,s) \preceq^r (y,t)$ if, and only if, $(x,s) \preceq (y,t)$. We can now point to the main differences between RT and OM.

First, note that U in Proposition 8 can take only positive values while v in (4)-(5) takes both positive and negative values.[28] To bypass this problem, we consider only the domain of strictly positive gains. Let w_0 be the reference point for wealth. Take $X = \{w - w_0 : w > w_0\} = (0,\infty)$ and let $\preceq$ satisfy the conditions of Proposition 8. Let (U,D) be the representation of $\preceq$ guaranteed by Proposition 8, where U is the utility function and D is the interval discount function.

Recall that, for each $r \in [0,\infty)$, $\preceq_{w_0,r}$ is a complete transitive order on $(-\infty,\infty)\times[r,\infty)$ and, hence, also on $X \times [r,\infty)$. The second point we wish to make is that, in general, $\preceq_{w_0,r}$, unlike $\preceq^r$, is *not* the restriction to $X \times [r,\infty)$ of some complete binary relationship on $X \times [0,\infty)$. Thus OM and RT are not compatible and neither is a special case of the other.

Third, $\preceq_{w_0,r}$ is *transitive* while, in general, $\preceq^r$ is *not transitive*. To elaborate this point, consider (x,r), (y,s) and (z,t), where $x,y,z \in X$ and $s,t \in [r,\infty)$. Suppose $(x,r) \preceq_{w_0,r} (y,s)$ and $(y,s) \preceq_{w_0,r} (z,t)$. Since $\preceq_{w_0,r}$ is transitive, we can conclude that $(x,r) \preceq_{w_0,r} (z,t)$. Now, suppose that $(x,r) \preceq^r (y,s)$ and $(y,s) \preceq^r (z,t)$. Since $\preceq^r$ is not, in general, transitive, we *cannot* conclude that $(x,r) \preceq^r (z,t)$.[29] More generally, given a compact subset $C \subset X \times [r,\infty)$, there is no guarantee in OM that it has a maximum under $\preceq^r$ (i.e., an $m \in C$ such that $c \preceq^r m$ for all $c \in C$). This, obviously, will cause great difficulty for any economic theory formulated in the OM framework. On the other hand, in RT theory, and if D is continuous, C will always have a maximum under $\preceq_{w_0,r}$.

Fourth, and finally, these problems with OM can all be resolved in the special case where $\preceq$ is transitive. But then $\preceq$ would also be additive. In this case, OM would reduce

[28] Hence, in its present formulation, OM cannot explain *gain-loss asymmetry*. However, it can explain the *magnitude effect* using the SIE utility function.

[29] As OM clearly explain, it is for this reason that we should think of $D(s,t)$ in their theory as the *relative* discount function between times s and t.

to the standard discounting model.

7 Conclusions

We generalize the framework of Loewenstein and Prelec (1992) and introduce a *reference time* in addition to a *reference outcome* level. We show that there are profound implications for time preference if the discount function is non-additive. We explain some observed intransitivities in time preferences as due to shifts in the reference point for time under non-additive discount functions. We also show that some of the recent advances in models of time preference can be incorporated within our framework. The framework that we propose is tractable, more general, and suitable for use in applications.

8 Appendix: Proofs

Proof of Proposition 3: Let $r \in [0,\infty)$ and $t \in [r,\infty)$. Let $\{t_n\}_{n=1}^{\infty}$ be a sequence in $[r,\infty)$ converging to t. We want to show that $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. It is sufficient to show that any monotone subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. In particular, let $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ be a decreasing subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$. Since $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ is bounded below by $D(r,t)$, it must converge to, say, q, where $D(r,t) \leq q \leq D(r,t_{n_i})$, for all i. Since D is onto, there is a $p \in [r,\infty)$ such that $D(r,p) = q$. Moreover, $t_{n_i} \leq p \leq t$, for each i. Suppose $D(r,t) < q$. Then $t_{n_i} < p$, for each i. Hence also $t_{n_i} < t$, for each i. But this cannot be, since $\{t_{n_i}\}_{i=1}^{\infty}$, being a subsequence of the convergent sequence $\{t_n\}_{n=1}^{\infty}$, must also converge to the same limit, t. Hence, $D(r,t) = q$. Hence, $\{D(r,t_{n_i})\}_{i=1}^{\infty}$ converges to $D(r,t)$. Similarly, we can show that any increasing subsequence of $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. Hence, $\{D(r,t_n)\}_{n=1}^{\infty}$ converges to $D(r,t)$. Hence, $D(r,t)$ is continuous in t. ■.

Proof of Proposition 4 (Time sensitivity): Let $D(r,t)$ be a continuous discount function and $r \geq 0$. Suppose $0 < x \leq y$. From (4) and (5), it follows that $0 < v(x) \leq v(y)$ and, hence, $0 < \frac{v(x)}{v(y)} \leq 1$. Since, by Definition 2(c), $D(r,t) : [r,\infty) \overset{\text{onto}}{\to} (0,1]$, it follows that $\frac{v(x)}{v(y)} = D(r,t)$ for some $t \in [r,\infty)$. A similar argument applies if $y < x < 0$. ■.

Proof of Proposition 5 (Existence of present values): Let $r \leq t$ and $y \geq 0$. Then, $0 < D(r,t) \leq 1$. Hence, $0 = v(0) \leq v(y) D(r,t) \leq v(y)$. Since v is continuous, (8), and strictly increasing, (4), it follows that $v(y) D(r,t) = v(x)$ for some $x \in [0,y]$. Similarly, if $y \leq 0$, then $v(y) D(r,t) = v(x)$ for some $x \in [y,0]$. ■.

To facilitate the proof of Proposition 6, below, we first establish Lemmas 9 and 10.

Lemma 9 *: Let $x \geq 0$ and $y \geq 0$. Then:*
(a) $\rho \geq 1 \Rightarrow x^{\rho} + y^{\rho} \leq (x+y)^{\rho}$.

(b) $0 < \rho \leq 1 \Rightarrow x^{\rho} + y^{\rho} \geq (x+y)^{\rho}$.

Proof of Lemma 9: Clearly, the results hold for $x = 0$. Suppose $x > 0$. Let $z = \frac{y}{x}$ and $f(z) = (1+z)^{\rho} - 1 - z^{\rho}$. Then $f(z) = 0$ and $f'(z) = \rho\left[(1+z)^{\rho-1} - z^{\rho-1}\right]$, for $z > 0$. Suppose $\rho \geq 1$. Then $f'(z) \geq 0$. Since f is continuous, it follows that $f(z) \geq 0$ for $z \geq 0$. Part (a) follows from this. Now suppose $0 < \rho \leq 1$. Then $f'(z) \leq 0$. Since f is continuous, it follows that $f(z) \leq 0$ for $z \geq 0$. Part (b) follows from this. ■.

Lemma 10 *: Let $\tau > 0$, $0 \leq s < t$ and $r > 0$. Let $f(r) = (t+r)^{\tau} - (s+r)^{\tau} - (t^{\tau} - s^{\tau})$. Then:*
(a) $\tau > 1 \Rightarrow f(r) > 0$.
(b) $0 < \tau < 1 \Rightarrow f(r) < 0$.

Proof of Lemma 10: Clearly, $f(0) = 0$. Also, $f'(r) = \tau\left[(t+r)^{\tau-1} - (s+r)^{\tau-1}\right]$. If $\tau > 1$, then $f'(r) > 0$ for $r > 0$. Since f is continuous, it follows that $f(r) > 0$ for $r > 0$. This establishes part (a). If $0 < \tau < 1$, then $f'(r) < 0$ for $r > 0$. Since f is continuous, it follows that $f(r) < 0$ for $r > 0$. This establishes part (b). ■.

Proof of Proposition 6: (a) Suppose $0 < \rho \leq 1$. Let $0 \leq r < s < t$. From (17), we get:

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

and, hence,

$$\begin{aligned}
D(r,s)\,D(s,t) &= \left\{1 + \alpha\left[\left(s^{\tau} - r^{\tau}\right)^{\rho} + \left(t^{\tau} - s^{\tau}\right)^{\rho}\right] + \alpha^2\left(s^{\tau} - r^{\tau}\right)^{\rho}\left(t^{\tau} - s^{\tau}\right)^{\rho}\right\}^{-\frac{\beta}{\alpha}}, \\
&< \left\{1 + \alpha\left[\left(s^{\tau} - r^{\tau}\right)^{\rho} + \left(t^{\tau} - s^{\tau}\right)^{\rho}\right]\right\}^{-\frac{\beta}{\alpha}}, \text{ since } \alpha^2\left(s^{\tau} - r^{\tau}\right)^{\rho}\left(t^{\tau} - s^{\tau}\right)^{\rho} > 0 \text{ and } -\frac{\beta}{\alpha} < 0, \\
&\leq \left[1 + \alpha\left(s^{\tau} - r^{\tau} + t^{\tau} - s^{\tau}\right)^{\rho}\right]^{-\frac{\beta}{\alpha}}, \text{ by Lemma 9b and since } -\frac{\beta}{\alpha} < 0, \\
&= \left[1 + \alpha\left(t^{\tau} - r^{\tau}\right)^{\rho}\right]^{-\frac{\beta}{\alpha}}, \\
&= D(r,t).
\end{aligned}$$

(b) It is sufficient to give an example. Let $\alpha = \tau = 1$ and $\rho = 2$. Hence, $D(0,1)\,D(1,2) = 4^{-\beta} > 5^{-\beta} = D(0,2)$. Hence, for $\alpha = \tau = 1$ and $\rho = 2$, D cannot be additive or subadditive. However, for the same parameter values, we have $D(0,10)\,D(10,20) = 10201^{-\beta} < 401^{-\beta} = D(0,20)$. Hence, D cannot be superadditive either.[30]

[30] 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 $\rho > 1$.

(c) Let $0 \le s < t$. (ii) is obvious from inspecting (17). Let $r > 0$. (iii) For $\tau > 1$, Lemma 10(a) gives $D(s+r, t+r) = \{1 + \alpha[(t+r)^{\tau} - (s+r)^{\tau}]^{\rho}\}^{-\frac{\beta}{\alpha}} < \{1 + \alpha[t^{\tau} - s^{\tau}]^{\rho}\}^{-\frac{\beta}{\alpha}} = D(s,t)$. (i) For $0 < \tau < 1$, Lemma 10(b) gives $D(s+r, t+r) = \{1 + \alpha[(t+r)^{\tau} - (s+r)^{\tau}]^{\rho}\}^{-\frac{\beta}{\alpha}}$ $> \{1 + \alpha[t^{\tau} - s^{\tau}]^{\rho}\}^{-\frac{\beta}{\alpha}} = D(s,t)$. ■.

Proof of Proposition 7:

$$((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \preceq_{w0,r} ((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)),$$
$$\Leftrightarrow V_r((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \le V_r((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)),$$
$$\Leftrightarrow D(0,r) V_r((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \le D(0,r) V_r((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)),$$
$$\Leftrightarrow D(0,r) \Sigma_{i=1}^{m} v(x_i) D(r, s_i) \le D(0,r) \Sigma_{i=1}^{n} v(y_i) D(r, t_i),$$
$$\Leftrightarrow \Sigma_{i=1}^{m} v(x_i) D(0,r) D(r, s_i) \le \Sigma_{i=1}^{n} v(y_i) D(0,r) D(r, t_i),$$
$$\Leftrightarrow \Sigma_{i=1}^{m} v(x_i) D(0, s_i) \le \Sigma_{i=1}^{n} v(y_i) D(0, t_i), \text{ by additivity,}$$
$$\Leftrightarrow V_0((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \le V_0((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)),$$
$$\Leftrightarrow ((x_1, s_1), (x_2, s_2), ..., (x_m, s_m)) \preceq_{w0,0} ((y_1, t_1), (y_2, t_2), ..., (y_n, t_n)). \blacksquare.$$

References

al-Nowaihi, A., Dhami, S. 2006a. 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., 2006b. A simple derivation of Prelec's probability weighting function. *The Journal of Mathematical Psychology* 50 (6), 521-524.

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.

Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. *Journal of Economic Literature* 40, 351-401.

Halevy, Y., 2007. Strotz meets Allais: Diminishing impatience and the certainty effect. Forthcoming in *American Economic Review*.

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

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

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

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

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

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

Luce R. D. 2001. Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology* 45, 167-179.

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.

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.

Prelec D. 1998. The probability weighting function. *Econometrica* 60, 497-528.

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

Read, D., and Scholten, M., 2006. Beyond Discounting: The Tradeoff Model of Intertemporal. Choice. Working Paper LSE-OR 06.88. London: London School of Economics and Political Science

Roelofsma, P. and Read, D. 2000. Intransitive intertemporal choice. *Journal of Behavioral Decision Making* 13, 161-177.

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. 2010. The Psychology of Intertemporal Tradeoffs, Psychological Review, 117: 925–944.

Scholten, M. and Read, D. 2006.Discounting by Intervals: A generalized model of intertemporal choice. *Management Science*, 52, 1424-1436.

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