CESifo WORKING PAPERS

9215
2021

July 2021

Preferences over Time and under Uncertainty: Theoretical Foundations

Ali al-Nowaihi, Sanjit Dhami

CESifo

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
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9215*

Preferences over Time and under Uncertainty: Theoretical Foundations

Abstract

We formulate a general theory of preferences over outcome-time-probability triplets and decompose uncertainty into risk and hazard. We define the delay, defer, shift and certainty functions that can be uniquely elicited from behaviour. These individually determine stationarity, the common difference effect and its converse; constant, decreasing and increasing impatience; additivity, subadditivity and super additivity; probability independence, the certainty effect and its converse. We propose a general discounted utility model which encompasses the main empirically supported discounted utility models. We show that our axioms on preferences are satisfied in our general discounted utility model. Finally, we discuss the various explanations of the common difference effect.

JEL-Codes: D150, D910, D810.

Keywords: time preferences, preferences under uncertainty, discounted utility models, common difference effect, impatience, additivity, certainty effect, probability weighting function, survival function.

Ali al-Nowaihi
Department of Economics, School of Business
University of Leicester
Brookfield, 266 London Road
United Kingdom - Leicester LE2 1RQ
aa10@le.ac.uk

*Sanjit Dhami**
Department of Economics, School of Business
University of Leicester
Brookfield, 266 London Road
United Kingdom - Leicester LE2 1RQ
sd106@le.ac.uk

*corresponding author

July 15, 2021
Declarations of interest: none.

1 Introduction

Typically decisions are taken in the present but their consequences occur in the future and are often uncertain. In this paper, we develop a general framework that encompasses both time and uncertainty. Furthermore, we decompose uncertainty into risk and hazard.

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. A significant body of evidence indicates that the EDU model is not able to account for the evidence from intertemporal choices in a satisfactory manner.[1] Several alternative models have been proposed.[2] The refutations of the EDU model, commonly referred to as *anomalies*, are not simply mistakes (Frederick et al., 2002, section 4.3).

Explanations of the anomalies of the EDU model have taken three main routes:

1. *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 (2006a, 2008a, 2008b, 2009, 2018).

2. *Imposing suitable restrictions on the discount function.* Arguably, the most serious anomaly, and certainly the most discussed, is the rejection of the *stationarity* property, which is termed the *common difference effect* (under certainty, the EDU model predicts stationarity). Examples of models that can take account of the common difference effect include hyperbolic discounting (Ainslie 1992; Phelps and Pollak, 1968; Laibson, 1997) and some attribute based models (Manzini and Mariotti, 2006).

3. *Interactions between risk and time preferences.* One of the most promising recent directions in the literature takes simultaneous account of time and risk preferences. This has provided (i) an alternative explanation of the common difference effect and its violations (Halevy, 2008; Chakraborty et al., 2020), and (ii) the magnitude effect (Baucells and Heukamp, 2012). It has also led to explanations of other phenomena such as the *common ratio effect* and issues of *time consistency* (Halevy, 2015).

Much attention has focussed on explaining the common difference effect that we outline next.

[1]See, for example, Thaler (1981), Read (2001), Dohmen et al. (2017), Read et al. (2013), Ericson et al. (2015), Scholten et al. (2016), Echenique, et al. (2019).

[2]For instance, by Loewenstein and Prelec (1992); Frederick et al. (2002); Rubinstein (2003); Read and Scholten (2006); Ok and Masatlioglu (2007); Manzini and Mariotti (2008). For a recent survey, see Dhami (2019b).

Example 1 *(Common difference effect; Thaler, 1981): Suppose that a decision maker is indifferent today between one apple received today and two apples received tomorrow. Then, stationarity implies that this decision maker must also be indifferent today between one apple received in* 50 *days' time and two apples received in* 51 *days' time. Thus, the preferences are unaltered if, ceteris-paribus, we move both outcome-time pairs by the same fixed time distance (*50 *days in this example). A decision maker violates stationarity if, say, today he is indifferent between one apple today and two apples tomorrow, but (also today) prefers two apples in* 51 *days' time to one apple in* 50 *days' time. Such a violation of stationarity is known as the common difference effect.*

There are two competing explanations of the common difference effect.

(i) The discount rate over a time interval of constant length decreases as that interval is moved towards the future. This is a violation of constant impatience, and it suggests *strictly decreasing impatience*, which leads to the well-known phenomenon of *present-biased preferences*.[3]

(ii) Under certainty, 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 (*strict subadditivity*) or relatively lower discounting (*strict superadditivity*). 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] When there is no uncertainty, the Read-Scholten (RS) *interval discount function* (Example 8, subsection 7.1, below) explains the common difference effect either through decreasing impatience, subadditivity, or both (Read, 2001; Read and Roelofsma, 2003; Scholten and Read, 2006; Scholten et al., 2016). Thus, it is the most general discount function available, yet surprisingly 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. Possibly the most important of these is that delayed payments may be considered risky (Halevy, 2008; Andreoni and Sprenger, 2012, section 2.3; Dhami, 2019b, sections 2.3, 3.3, 4.1.1); concave utility functions may impart caution that is mistakenly attributed to discount rates (Andersen et al., 2010; Andreoni and Sprenger, 2012; Dhami 2019b, section 4.1.2); and time preferences may not be stationary

[3]Experiments on animals and humans reveal a hyperbolically declining pattern of discount rates, giving rise to *hyperbolic discounting* (Ainslie, 1975, 1992; Thaler, 1981). For a survey of the evidence in support of hyperbolic discounting and possible confounds, see Dhami (2019b).

[4]There is also important evidence of subadditivity in other domains of decision making. For instance, subadditivity in probabilities and in events has been well documented in the domain of uncertainty; see Tversky and Koehler (1994) and the references therein. Various forms of subadditivity might have a common neural basis (Alvarado et al. 2007; Stanford et al., 2007).

(Halevy, 2015; Dhami, 2019b, section 4.2). Dohmen et al. (2017) find that after controlling for all these potential confounds, and using large representative samples, temporal behavior exhibits subadditive discounting. This calls for greater theoretical explorations of the implications of subadditivity in economics. Other anomalies of the EDU model and their explanations have been identified; for a survey see Dhami (2019b).

We describe the main contributions of our paper in subsections, 1.1 to 1.4 below.

1.1 Contribution 1: Axiomatic foundations

We provide an axiomatic foundation for decision making over time and under uncertainty (sections 2 and 3). We assume that at each moment in time, $r \geq 0$, a decision maker has a complete and transitive preference relation, $\preceq_r$, over future dated bundles of goods.

Specifically, given $r \geq 0$, $s \geq r$, $t \geq r$, $p, q \in [0, 1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, then $(\mathbf{x}, s, p) \preceq_r (\mathbf{y}, t, q)$ says that at time r the bundle of goods $\mathbf{y}$ offered for delivery at time t with probability q is strictly preferred to, or indifferent to, the bundle of goods $\mathbf{x}$ also offered at time r but for delivery at time s with probability p.

Such systems are, of course, not new. We are closest to Halevy (2015). Halevy (2015) assumes that we can compound forward, as well as discount backward. This rules out strict subadditive discount functions (see Remark 1). However, strict subadditivity has good empirical support (Read, 2001, Scholten and Read, 2006 and Dohmen et al., 2017). By contrast, we only assume discounting backwards. So we can accommodate strictly subadditive discount functions.[5]

1.2 Contribution 2: Special properties of preferences

In section 4, we consider *stationarity*, *constant impatience*, *additivity*, and *probability independence*[6]

Among other conclusions, we show that any decision maker who exhibits additivity and constant impatience must also exhibit stationarity (Theorem 1). We also show that any decision maker who exhibits (1) additivity and strictly declining impatience, or (2) strict subadditivity and constant impatience, or (3) strict subadditivity and strictly declining impatience, must also exhibit the common difference effect (Theorem 2).

Some of these conclusions were already known for specific functional forms, such as for the hyperbolic discount function (Loewenstein and Prelec, 1992), and for the Read-Scholten discount function (Scholten and Read, 2006). One of our contributions is to show that these conclusions hold for quite general time-preference relations, $\preceq_r$.

[5]Other differences with Halevy (2015) are as follows. Halevy (2015) considers *stationarity*, *time invariance* and *time consistency*; and proves that any two of them imply the third. We consider stationarity but not time invariance nor time consistency. On the other hand, we consider the *common difference effect* and its converse; *constant impatience* and its violations; and *additivity* and its violations. Halevy (2015) does not consider any of these.

[6]Probability independence, Definition 11, requires that preferences over two outcome time pairs, both discounted back to a common date, do not change if we change the probability with which the outcomes are received. A violation of this property leads to the *certainty effect* and its converse.

1.3 Contribution 3: Properties of preferences in terms of observable functions

In section 4, the properties of stationarity, constant impatience, additivity, probability independence, and their violations, are all defined in terms of the underlying preference relations, $\preceq_r$, $r \geq 0$. In practice, these are not directly observable (although some of their consequences are). The question then is: How can we empirically determine whether the behavior of a decision maker exhibits any of these properties? To answer this question, in section 5, we define four functions: The *delay function*, D, the *defer function*, Δ, the *shift function*, S, and the *certainty function*, C. These functions determine whether the behavior of a decision maker exhibits any of the properties considered in section 4.

The delay function, D, determines stationarity, the common difference effect and its converse (Theorem 3). The defer function, Δ, determines constant, decreasing and increasing impatience (Theorem 4). The shift function, S, determines additivity, subadditivity and superadditivity (Theorem 5). The certainty function, C, determines probability independence, the certainty effect and its converse (Theorem 6).

Each of these four functions can be uniquely elicited from behavior. Hence, they can provide parameter-free tests of whether any of the properties of section 4 hold.

1.4 Contribution 4: A general discounted utility model

In section 6 we propose a *general discounted utility model* that satisfies our axioms introduced in section 3 (subsection 6.3, Theorem 7). These axioms are: (1) *order* (completeness and transitivity), (2) *existence of time-neutral outcomes*, (3) *existence of present values*, (4) *consistency of gains and losses*, (5) *time monotonicity*, (6) *time sensitivity*, and (7) *probability sensitivity*.

In section 7, we consider time discount functions that take the following form:

$$\delta_r(r+t,p) = w(p\Pi(t))\,\delta_r^0(r+t), \tag{1.1}$$

where $r, t \in \mathbb{R}_+$, $p \in [0,1]$, w is a *probability weighting function*, Π is a *survival function* and δ^0 is a *riskless time discount function*, i.e., a discount function in the ordinary sense (such as the exponential, the quasi-hyperbolic, and the generalized hyperbolic: Examples 3-8 of subsection 7.1). The particular specification (1.1) allows us to add further results to those of sections 2-6. Under no uncertainty (no risk and no hazard), we have $p = 1, \Pi = 1$. It is a property of the probability weighting function that $w(1) = 1$, so in this case we have $\delta_r(r+t,p) = \delta_r^0(r+t)$, which is the case traditionally considered in time preferences.

We propose a generalization (Example 3) of the Read-Scholten interval discount function (Example 8), that we call the *generalized Read-Scholten discount function* (GRS). GRS also encompasses the exponential, quasi-hyperbolic, and generalized hyperbolic discount functions.

Chakraborty et al. (2020) consider the relation between behavior under risk and behavior over time, using the special case (in our notation) of (1.1) with $p = 1$, the constant survival function[7] $\Pi(t) = e^{-rt}$, where $r \in (0,1)$ is the constant hazard rate[8], the exponential discount function $\delta_r^0(r+t) = \delta^t$, $\delta \in (0,1)$, and a general probability weighting function, w. Because their focus is different, their conclusions are not special cases of our conclusions[9]. Furthermore, they consider consumption *streams*, which we do not in this paper.

1.5 Notation

We shall denote the set of non-negative real numbers by $\mathbb{R}_+ = [0,\infty)$ and the set of m-dimensional real vectors by $\mathbb{R}^m$. We shall use r, s, t, σ, τ, $\omega \in [0,\infty)$ to denote moments in time and $p, q \in [0,1]$ to denote probabilities. We shall use $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in \mathbb{R}^m$ to denote outcomes. We may also use subscripts or superscripts on these symbols as the need arises. To facilitate the readability of formulas with nested brackets, we shall often use the following hierarchy: $\{\langle[()]\rangle\}$.

2 Preferences

We assume that at each moment in time, $r \geq 0$, a decision maker has a preference relation, $\preceq_r$, over future dated bundles of goods.

Specifically, given $r, s, t \in [0,\infty)$, $p, q \in [0,1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, then

$$(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q) \tag{2.1}$$

says that at time r the bundle of goods $\mathbf{y}$ offered for delivery at time $r+t$ with probability q is strictly preferred to, or indifferent to, the bundle of goods $\mathbf{x}$ offered for delivery at time $r+s$ with probability p. As usual, we shall use the term "preferred to" to mean "strictly preferred to, or indifferent to".

Example 2 : *A simple example of a preference relation $\preceq_r$ is the exponentially discounted utility (EDU) model*

$$U_r(\mathbf{x}, r+t, p) = w(p\Pi(t))\, u(\mathbf{x})\, e^{-\beta t}, \beta > 0. \tag{2.2}$$

We explain the terms in (2.2), below.

At time r the bundle of goods $\mathbf{x}$ is promised for delivery at time $r+t$ with probability p, $u(\mathbf{x})$ is the prospect theory utility function with reference point $\mathbf{0}$ and $u(\mathbf{0}) = 0$ (Kahneman and Tversky, 1979). Even if $p = 1$, the decision maker may not actually receive $\mathbf{x}$. There are many reasons for this. For example, the provider of $\mathbf{x}$ may renege,

[7]Thus, they consider uncertainty arising from hazard only (we consider both risk, p, and hazard, Π). They use $p(t)$ for the survival function.

[8]This is unnecessarily restrictive. Their results go through with $r \in [0,\infty)$. They use $(1-r)^t$ for the probability of survival up to time t. It should be e^{-rt}.

[9]Their concepts are slightly different from the ones we use in this paper. They implicitly assume a reference point of zero, where the utility function is zero. They consider only gains (we consider gains and losses).

go bankrupt, be unable to locate the decision maker, or the decision maker may die before receipt of $\mathbf{x}$ (Halevy, 2008).

More generally, let $\Pi(t)$ be the probability with which the decision maker will receive the lottery $(\mathbf{x}, p)$ at time $r+t$ (Π is the survival function). Thus, the probability of actually receiving the outcome $\mathbf{x}$ at time $r+t$ is $p\Pi(t)$, and the probability of not receiving $\mathbf{x}$ is $1-p\Pi(t)$. These probabilities are then transformed using a *probability weighting function*, w (see subsection 7.2). Using a prospect theory evaluation, the resulting utility to the decision maker is then $U_r(\mathbf{x}, r+t, p) = \{[1-w(p\Pi(t))]\, u(\mathbf{0}) + w(p\Pi(t))\, u(\mathbf{x})\}\, e^{-\beta t}$. Since $u(\mathbf{0}) = 0$, we get $U_r(\mathbf{x}, r+t, p) = w(p\Pi(t))\, u(\mathbf{x})\, e^{-\beta t}$, which is (2.2). Then $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ if, and only if, $U_r(\mathbf{x}, r+s, p) \leq U_r(\mathbf{y}, r+t, q)$.

From (2.2), we can see two reasons for the restriction to $s \geq 0$ and $t \geq 0$, in (2.1). First, the survival function, $\Pi(t)$, need not be defined for $t < 0$ (see Example 10, subsection 7.3, below). Even if $\Pi(t)$ were defined for $t < 0$, the value of $p\Pi(t)$ may be greater than 1, in which case $w(p\Pi(t))$ need not be defined (see Example 9, subsection 7.2). In the case of certainty we have $p = 1$, $\Pi \equiv 1$ and, hence, $w(p\Pi(t)) = w(1) = 1$ (and, hence, $1 - w(p\Pi(t)) = 0$). In this case, (2.1) can be defined for $t < 0$. However, this need not be the case for more general discount functions, for example, the *interval discount function* of Read (2001) and Scholten and Read (2006) (see Example 8, subsection 7.1, below).

We are particularly interested in the following experimental situation. The experimenter chooses $p, q, r, s, t, \mathbf{x}, \mathbf{y}$. These are the same for all experimental subjects (recall (2.1)). However, subjects may face different hazards, have different probability weighting functions, and different discount rates (recall (2.2)); all possibly unknown to the experimenter.

We now return to the more general development of this section and the next.

Definition 1 *: We adopt the following standard terminology.*
(a) "$(\mathbf{x}, r+s, p) \sim_r (\mathbf{y}, r+t, q)$" stands for "$(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ and $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{x}, r+s, p)$".
(b) "$(\mathbf{x}, r+s, p) \prec_r (\mathbf{y}, r+t, q)$" stands for "$(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ but not $(\mathbf{x}, r+s, p) \sim_r (\mathbf{y}, r+t, q)$".
(c) "$(\mathbf{y}, r+t, q) \succeq_r (\mathbf{x}, r+s, p)$" stands for "$(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$".
(d) "$(\mathbf{y}, r+t, q) \succ_r (\mathbf{x}, r+s, p)$" stands for "$(\mathbf{x}, r+s, p) \prec_r (\mathbf{y}, r+t, q)$".

In the spirit of reference dependent models (e.g., prospect theory), we adopt a reference point relative to which we measure gains and losses. In Definition 2 below, we formalize the idea of a *time-neutral outcome* that serves as an appropriate reference point. In order to motivate this concept, consider Example 2. From (2.2), we see that $U_r(\mathbf{0}, r+t, p) = w(p\Pi(t))\, u(\mathbf{0})\, e^{-\beta t} = 0$, since $u(\mathbf{0}) = 0$. Then, the decision maker is indifferent between receiving $\mathbf{0}$ at time $r+s$ or receiving $\mathbf{0}$ at time $r+t$, because they both give the same utility, $U_r(\mathbf{0}, r+s, p) = U_r(\mathbf{0}, r+t, p)$ $(= 0)$, for all $s \geq 0$ and all $t \geq 0$.

Definition 2 *(time-neutrality): We say that the outcome* $\mathbf{w} \in \mathbb{R}^m$ *is time-neutral for the preference relation* $\preceq_r$ *and the probability* $p \in (0,1]$ *if, for all* $t \in [0,\infty)$, $(\mathbf{w}, r, p) \sim_r (\mathbf{w}, r+t, p)$.

From Definition 2, for the case of Example 2, $\mathbf{0} \in \mathbb{R}^m$ is a time-neutral outcome for the preferences given in that example. But $\mathbf{0}$ need not be the time-neutral outcome for all preference relations $\preceq_r$ and for all probabilities $p \in (0,1]$. Hence, we denote the time-neutral outcome more generally by $\mathbf{w} \in \mathbb{R}^m$.

We now define, for any outcome-time-probability triplet, its present value under the preference relation $\preceq_r$.

Definition 3 *(Present value): Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, $r, t \in [0,\infty)$, $p \in (0,1]$. *Suppose that* $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$, *then we call* $\mathbf{x}$ *a present value of* $\mathbf{y}$ *(more precisely,* $\mathbf{x}$ *is a present value at time* r *of* $\mathbf{y}$ *offered at time* r *for delivery at time* $r+t$*, both offered with probability* $p \in (0,1]$*, under the preference relation* $\preceq_r$*).*

In Definition 4 below, we define a gain or a loss relative to a time-neutral outcome. Consider Example 2. Suppose $u(\mathbf{x}) > 0$. Then (using the properties of w and Π, see subsections 7.2 and 7.3),
$U_r(\mathbf{x}, r+t, p) = w(p\Pi(t))\, u(\mathbf{x})\, e^{-\beta t} > 0 = w(p\Pi(t))\, u(\mathbf{0})\, e^{-\beta t} = U_r(\mathbf{0}, r+t, p)$ and, hence, $(\mathbf{0}, r+t, p) \prec_r (\mathbf{x}, r+t, p)$. For the preferences in Example 2, we term $\mathbf{x}$ offered at time $r+t$ as a *gain relative to* $\mathbf{0}$. Similarly, if $u(\mathbf{x}) < 0$, then $(\mathbf{x}, r+t, p) \prec_r (\mathbf{0}, r+t, p)$, i.e., for the preferences in Example 2, $\mathbf{x}$ offered at time $r+t$ is a *loss relative to* $\mathbf{0}$. We now formally state these ideas.

Definition 4 *(Gains and losses): Let* $r, t \in [0,\infty)$, $p \in (0,1]$, $\mathbf{w}, \mathbf{x} \in \mathbb{R}^m$, *where* $\mathbf{w}$ *is time-neutral for the preference relation* $\preceq_r$ *and the probability* p. *We say:*
(i) $\mathbf{x}$ *offered at time* $r+t$ *is a gain relative to* $\mathbf{w}$ *and according to* $\preceq_r$ *if* $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r+t, p)$.
(ii) $\mathbf{x}$ *offered at time* $r+t$ *is a loss relative to* $\mathbf{w}$ *and according to* $\preceq_r$ *if* $(\mathbf{x}, r+t, p) \prec_r (\mathbf{w}, r, p)$.

Throughout the paper, it is convenient to state our assumptions and conclusions separately for the cases of gains and losses, respective, $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r, p)$ and $(\mathbf{w}, r, p) \succ_r (\mathbf{z}, r, p)$, where $r \in [0,\infty)$, $\mathbf{w}, \mathbf{z} \in \mathbb{R}^m$, $p \in (0,1]$, and $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and probability p. A critical feature of prospect theory, strongly backed by the evidence, is that human behavior for gains and losses is very different (Kahneman and Tversky, 2000; Dhami, 2019a).

3 Axioms on preferences

In this section we introduce seven axioms that we assume hold for all the preference relations, $\preceq_r$, that we will consider.

Axiom 1 *(Order): Given $r \in [0, \infty)$, we assume that $\preceq_r$ satisfies:*
(a) Completeness: Given $s, t \in [0, \infty)$, $p, q \in [0, 1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, either $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ or $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{x}, r+s, p)$.
(b) Transitivity: Given $s, t, t' \in [0, \infty)$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, if $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ and $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{z}, r+t', q')$ then $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{z}, r+t', q')$.

Lemma 1 *:*
(a) Both $\prec_r$ and $\sim_r$ are transitive, both $\preceq_r$ and $\sim_r$ are reflexive[10], $\prec_r$ is complete[11] and $\sim_r$ is symmetric[12].
(b) Let $r, s, t, s', t' \in [0, \infty)$, $p, p', q, q' \in [0, 1]$ and $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in \mathbb{R}^m$. Suppose that $(\mathbf{x}, r+s, p) \sim_r (\mathbf{x}', r+s', p')$ and $(\mathbf{y}, r+t, q) \sim_r (\mathbf{y}', r+t', q')$. Then

$$(i)\ (\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q) \Leftrightarrow (\mathbf{x}', r+s', p') \preceq_r (\mathbf{y}', r+t', q'),$$
$$(ii)\ (\mathbf{x}, r+s, p) \prec_r (\mathbf{y}, r+t, q) \Leftrightarrow (\mathbf{x}', r+s', p') \prec_r (\mathbf{y}', r+t', q').$$

Axiom 2 *(Existence of time-neutral outcomes): Let $p \in (0, 1]$, $r \in [0, \infty)$, and $\preceq_r$ a preference relation, then there exists a time-neutral outcome, $\mathbf{w} \in \mathbb{R}^m$, i.e., for all $t \in [0, \infty)$, $(\mathbf{w}, r, p) \sim_r (\mathbf{w}, r+t, p)$.*

Axiom 3 *(Existence of present values): Let $r, t \in [0, \infty)$, $p \in (0, 1]$. We assume that, for each $\mathbf{y} \in \mathbb{R}^m$, there is an $\mathbf{x} \in \mathbb{R}^m$ (the present value of $\mathbf{y}$) such that $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$.*

Axiom 4 *(Consistency of gains and losses): Let $\mathbf{w}, \mathbf{x} \in \mathbb{R}^m$, $r \in [0, \infty)$, $p \in (0, 1]$, where $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and probability p.*
(i) Gains: If $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r+t_0, p)$ for some $t_0 \geq 0$,
then $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r+t, p)$ for all $t \geq 0$.
(ii) Losses: If $(\mathbf{x}, r+t_0, p) \prec_r (\mathbf{w}, r, p)$ for some $t_0 \geq 0$,
then $(\mathbf{x}, r+t, p) \prec_r (\mathbf{w}, r, p)$ for all $t \geq 0$.

Axiom 4 implies that, relative to the time-neutral outcome, a gain delayed, or brought forward, is still a gain; and a loss delayed, or brought forward, is still a loss. Our next axiom states that ceteris-paribus, the decision maker prefers to expedite a gain and postpone a loss.

Axiom 5 *(Time monotonicity): Let $r \in [0, \infty)$, $\mathbf{w}, \mathbf{z} \in \mathbb{R}^m$, $p \in (0, 1]$, where $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and probability p.*
(i) Gains: Let $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r, p)$, then, for all $\omega \geq 0$, $\sigma \geq 0$ and $\tau > 0$,
$(\mathbf{z}, r+\omega+\sigma+\tau, p) \prec_{r+\omega} (\mathbf{z}, r+\omega+\sigma, p)$.
(ii) Losses: Let $(\mathbf{w}, r, p) \succ_r (\mathbf{z}, r, p)$, then, for all $\omega \geq 0$, $\sigma \geq 0$ and $\tau > 0$,
$(\mathbf{z}, r+\omega+\sigma+\tau, p) \succ_{r+\omega} (\mathbf{z}, r+\omega+\sigma, p)$.

[10] $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{x}, r+s, p)$ and $(\mathbf{x}, r+s, p) \sim_r (\mathbf{x}, r+s, p)$ for all $r, s \in \mathbb{R}_+$, $p \in [0, 1]$ and $\mathbf{x} \in \mathbb{R}^m$.

[11] For all $r, s, t \in \mathbb{R}_+$, $p, q \in [0, 1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, either $(\mathbf{x}, r+s, p) \prec_r (\mathbf{y}, r+t, q)$, $(\mathbf{y}, r+t, q) \prec_r (\mathbf{x}, r+s, p)$ or $(\mathbf{x}, r+s, p) \sim_r (\mathbf{y}, r+t, q)$.

[12] $(\mathbf{x}, r+s, p) \sim_r (\mathbf{y}, r+t, q) \Rightarrow (\mathbf{y}, r+t, q) \sim_r (\mathbf{x}, r+s, p)$ for all $r, s, t \in \mathbb{R}_+$, $p, q \in [0, 1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$.

To see the intuition behind Axiom 5, take $\sigma = \omega = 0$. Part (i) then says that for gains, $(\mathbf{z}, r + \tau, p) \prec_r (\mathbf{z}, r, p)$ for all $\tau > 0$, i.e., delaying a gain is undesirable. Part (ii) says that for losses, $(\mathbf{z}, r + \tau, p) \succ_r (\mathbf{z}, r, p)$ for all $\tau > 0$, i.e., delaying a loss is desirable.

The more complex notation in Axiom 5 allows us to compare delays at various times (hence, σ is allowed to be positive), and for preference relations $\preceq_r$ and $\preceq_{r+\omega}$ at different moments of time (hence, ω is allowed to be positive).

Axiom 6 *(Time sensitivity): Let* $r, s, t \in [0, \infty)$, $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$. *Let* $\mathbf{w} \in \mathbb{R}^m$ *be time-neutral for the preference relation* $\preceq_r$ *and the probability* $p \in (0, 1]$.
(i) Gains: Assume $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p) \preceq_r (\mathbf{z}, r + s, p)$,
then, for some $T \geq 0$, $(\mathbf{y}, r + t, p) \sim_r (\mathbf{z}, r + s + T, p)$;
if $(\mathbf{y}, r, p) \prec_r (\mathbf{z}, r + s, p)$, *then,* $T > 0$.
(ii) Losses: Assume $(\mathbf{w}, r, p) \succ_r (\mathbf{y}, r, p) \succeq_r (\mathbf{z}, r + s, p)$,
then, for some $T \geq 0$, $(\mathbf{y}, r + t, p) \sim_r (\mathbf{z}, r + s + T, p)$;
if $(\mathbf{y}, r, p) \succ_r (\mathbf{z}, r + s, p)$, *then,* $T > 0$.

To see the intuition behind Axiom 6, consider the special case $t = 0$. Part (i) then says that if $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p) \preceq_r (\mathbf{z}, r + s, p)$, then $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s + T, p)$, for some $T \geq 0$; and if $(\mathbf{y}, r, p) \prec_r (\mathbf{z}, r + s, p)$, then $T > 0$. Thus, if $\mathbf{y}$ and $\mathbf{z}$ are both gains, with $\mathbf{z}$ preferred to $\mathbf{y}$, then by postponing the receipt of $\mathbf{z}$ sufficiently long, the receipt of $\mathbf{z}$ can be made indifferent to the receipt of $\mathbf{y}$. Similarly, part (ii) says that if $\mathbf{y}$ and $\mathbf{z}$ are both losses, with $\mathbf{y}$ preferred to $\mathbf{z}$, then by postponing the receipt of $\mathbf{z}$ sufficiently long, the receipt of $\mathbf{z}$ can be made indifferent to the receipt of $\mathbf{y}$.

The slightly more complex notation in Axiom 6 simplifies and clarifies the derivations later.

To motivate Axiom 7, below, consider the following question. Suppose that, under the preference relation $\preceq_r$, the receipt of $\mathbf{y}$ at time r with positive probability, q, is indifferent to the receipt of $\mathbf{z}$ at time $r + s$, also with probability, q. Now consider a possibly different positive probability, p. By postponing or bringing forward the receipt of $\mathbf{z}$, can we maintain the indifference between $\mathbf{y}$ and $\mathbf{z}$? Under the conditions of Axiom 7, the answer is yes.

Axiom 7 *(Probability sensitivity): Let* $r \in [0, \infty)$. *Consider the following situation:*
(a) $\mathbf{w} \in \mathbb{R}^m$ *is time-neutral for the preference relation* $\preceq_r$ *and for probabilities* $p, q \in (0, 1]$,
(b) $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ *are either both gains or both losses, for* $\preceq_r, \mathbf{w}, p, q$,
(c) $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r + s, q)$, *where* $s \in [0, \infty)$.
Then there exists a $T \in [0, \infty)$ *such that* $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + T, p)$.

Note that Axiom 5 relates the preference relation $\preceq_r$, at a time r, to the preference relation $\preceq_{r+t}$, at another time, $r + t$. On the other hand, the other axioms give the properties of a preference relation, $\preceq_r$, at a particular moment in time, r. Axiom 7 involves two probabilities, while the other Axioms involve a single probability.

4 Special properties of preferences

We assume that Axioms 1-7 hold for all the preference relations, $\preceq_r$, that we will consider. In this section we study the following further properties which have strong empirical support: (1) the *common difference effect*, (2) *decreasing impatience*, (3) *subadditivity*, (4) the *certainty effect*. However, we first define *stationarity*, *constant impatience*, *additivity* and *probability independence*, of which properties (1)-(4) are violations. For the special case of certainty, $p = 1$, $\Pi \equiv 1$, our definitions reduce to the standard ones. For completeness, we also define and study the converses of properties (1)-(4).

4.1 Stationarity, the common difference effect and its converse

Definition 5 *(Stationarity): Let $p \in (0,1]$, $r \in [0,\infty)$. The preference relation $\preceq_r$ is stationary for probability p, if for all $s,t \in [0,\infty)$, for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{w}$ is time-neutral for $\preceq_r$ and p (Definition 2), and $\mathbf{y}$ is either a gain or a loss, i.e., $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p)$ (Definition 4), then*

$$(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \sim_r (\mathbf{z},r+s+t,p)\,.$$

Note that if $\mathbf{y}$ is time-neutral for $\preceq_r$ and p, then so will $\mathbf{z}$. Hence, the statement "$(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \sim_r (\mathbf{z},r+s+t,p)$" always holds when $\mathbf{y}$ is time-neutral. Thus, Definition 5, as written above, is actually equivalent to the same definition but with the restriction "$\mathbf{y}$ is either a gain or a loss" removed. We have adopted the above form because it simplifies the proofs (otherwise, we would have to consider the case "$\mathbf{y}$ is time-neutral" separately).

Definition 5 allows for the case where $\preceq_r$ is stationarity for some levels of probability, p, and some values or r, but not for others; and, similarly, for the common difference effect, below.

Definition 6 *(Common difference effect and its converse): Let $p \in (0,1]$, $r \in [0,\infty)$. The preference relation $\preceq_r$ exhibits:*
(a) The common difference effect for probability p, if for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, $s > 0$, $t > 0$, where $\mathbf{w}$ is time-neutral,

$$\text{(i)}\ \ (\mathbf{w},r,p) \prec_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \prec_r (\mathbf{z},r+s+t,p)\,,$$
$$\text{(ii)}\ \ (\mathbf{w},r,p) \succ_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \succ_r (\mathbf{z},r+s+t,p)\,.$$

(b) The converse common difference effect for probability p, if for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, $s > 0$, $t > 0$, where $\mathbf{w}$ is time-neutral,

$$\text{(i)}\ \ (\mathbf{w},r,p) \prec_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \succ_r (\mathbf{z},r+s+t,p)\,,$$
$$\text{(ii)}\ \ (\mathbf{w},r,p) \succ_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \prec_r (\mathbf{z},r+s+t,p)\,.$$

Thaler's apples example (Example 1) is a popular illustration of the common difference effect under certainty and it is typically stated for the case of gains. However, one can also state it for losses as follows. A decision maker is indifferent between losing 1 apple today and losing 2 apples tomorrow. However, the same decision maker, today, prefers to lose 1 apple in 50 days to losing 2 apples in 51 days. [13]

4.2 Constant, decreasing and increasing impatience

Definition 7 *(Constant impatience): Let $p \in (0,1]$, $r \in [0,\infty)$. The preference relation $\preceq_r$ exhibits constant impatience for probability p if, for all $s,t \in [0,\infty)$, for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{w}$ is time-neutral for $\preceq_r$ and p (Definition 2), and $\mathbf{y}$ is either a gain or a loss, i.e., $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p)$ (4), then*

$$(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t,p).$$

Note that, for given $r \geq 0$, constant impatience is a property of the set of preference relations $\{\preceq_{r+t}: t \geq 0\}$. Also note that if $\mathbf{y}$ is time-neutral then $(\mathbf{y},r,p) \sim_r (\mathbf{y},r+s,p)$ will hold for all $s \in [0,\infty)$. Hence if the restriction "$\mathbf{y}$ is either a gain or a loss" is removed from Definition 7, then we would get "$(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{y},r+s+t,p)$ for all $s \in [0,\infty)$", i.e., under constant impatience, if $\mathbf{y}$ were time-neutral for $\preceq_r$, then it would also be time neutral for all $\preceq_{r+t}$, $t \geq 0$. But we have no reason to assume that this is generally the case.

Definition 8 *(Decreasing and increasing impatience): Let $p \in (0,1]$, $r \in [0,\infty)$. The preference relation $\preceq_r$ exhibits:*
(a) Strictly decreasing impatience for probability p if, for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, $s > 0, t > 0$, where $\mathbf{w}$ is time-neutral,

(i) $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \prec_{r+t} (\mathbf{z},r+s+t,p)$,

(ii) $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \succ_{r+t} (\mathbf{z},r+s+t,p)$.

(b) Strictly increasing impatience for probability p if, for all $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, $s > 0, t > 0$, where $\mathbf{w}$ is time-neutral,

(i) $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \succ_{r+t} (\mathbf{z},r+s+t,p)$,

(ii) $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \prec_{r+t} (\mathbf{z},r+s+t,p)$.

We use the term "decreasing (increasing) impatience" to mean "constant or strictly decreasing (increasing) impatience".[14]

[13] The reader can check that this pattern of preferences makes perfect sense, for instance, by using a model of (β,δ) preferences with $\beta = 0.5$ and $\delta = 1$. As we shall show later, the same preferences can also be explained by invoking subadditivity. We discuss the utility representation of these preferences in detail later in the paper.

[14] In the literature, "decreasing impatience" is used to denote what we have called "strict decreasing impatience". This change of terminology enables us to state some of our conclusions in a more compact form. Similar comments apply to Definitions 10, 21, 22, 25, 27, below.

4.3 Additivity, subadditivity and superadditivity

Definition 9 *(Additivity): Let* $p \in (0,1]$, $r \in [0,\infty)$. *The preference relation* $\preceq_r$ *is additive for probability* p *if, for all* $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $s, t \in [0,\infty)$, *where* $\mathbf{w}$ *is time-neutral for* $\preceq_r$ *and* p *(Definition 2), and* $\mathbf{x}$ *is either a gain or a loss, i.e.,* $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r, p)$ *or* $(\mathbf{w}, r, p) \succ_r (\mathbf{x}, r, p)$ *(Definition 4), then*

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \;\; \textit{and} \;\; (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p) \\ \Rightarrow (\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t, p)\,.$$

In other words, the preference relation $\preceq_r$ is additive if discounting a gain (or a loss) from time $r+s+t$ back to time r in one step is equivalent to discounting from time $r+s+t$ back to time $r+t$ and then discounting from time $r+t$ back to time r.[15] Departures from additivity take two forms: Less discounting over the longer interval (subadditivity) and more discounting over the longer interval (superadditivity). This motivates the following definitions.

Definition 10 *(Subadditivity and superadditivity): Let* $p \in (0,1]$, $r \in [0,\infty)$. *The preference relation* $\preceq_r$ *is:*
(a) Strictly subadditive for probability p *if, for all* $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $s > 0, t > 0$, *where* $\mathbf{w}$ *is time-neutral,*

$$\textit{(i)} \;\; (\mathbf{w}, r, p) \prec_r (\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \;\; \textit{and} \;\; (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p) \\ \Rightarrow (\mathbf{x}, r, p) \prec_r (\mathbf{z}, r+s+t, p)\,,$$
$$\textit{(ii)} \;\; (\mathbf{w}, r, p) \succ_r (\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \;\; \textit{and} \;\; (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p) \\ \Rightarrow (\mathbf{x}, r, p) \succ_r (\mathbf{z}, r+s+t, p)\,.$$

(b) Strictly superadditive for probability p *if, for all* $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $s > 0, t > 0$, *where* $\mathbf{w}$ *is time-neutral,*

$$\textit{(i)} \;\; (\mathbf{w}, r, p) \prec_r (\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \;\; \textit{and} \;\; (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p) \\ \Rightarrow (\mathbf{x}, r, p) \succ_r (\mathbf{z}, r+s+t, p)\,,$$
$$\textit{(ii)} \;\; (\mathbf{w}, r, p) \succ_r (\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \;\; \textit{and} \;\; (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p) \\ \Rightarrow (\mathbf{x}, r, p) \prec_r (\mathbf{z}, r+s+t, p)\,.$$

We use the term "sub- (super-) additive" to mean "additive or strictly sub- (super-) additive".

[15] Recall that we gave reasons for excluding the case "**y** is time-neutral" from the definitions of stationarity and constant impatience (Definitions 5, 7). However, the reason for excluding "**x** is time-neutral" from the definition of Additivity (Definition 9) is more subtle. If we allowed the latter then Theorem 5a, subsection 5.3, would no more hold in its present simple form.

4.4 Connections between stationarity, constant impatience, additivity, and their violations

Theorem 1 *: Let $p \in (0,1]$, $r \in [0,\infty)$. Assume that the preference relation $\preceq_r$ exhibits constant impatience for probability p (Definition 7) and is additive for probability p (Definition 9). Then $\preceq_r$ is stationary for probability p (Definition 5).*

Theorem 2 *: Let $p \in (0,1]$, $r \in [0,\infty)$.*
(a) Assume that the preference relation $\preceq_r$ exhibits decreasing impatience for probability p (Definition 8) and is subadditive for probability p (Definition 10); where, at least, one of these is strict. Then the preference relation $\preceq_r$ exhibits the common difference effect for probability p (Definition 6a).
(b) Assume that the preference relation $\preceq_r$ exhibits increasing impatience for probability p (Definition 8) and is superadditive for probability p (Definition 10); where, at least, one of these is strict. Then the preference relation $\preceq_r$ exhibits the converse common difference effect for probability p (Definition 6b).

Discussion of the results: Theorem 1 shows that constant impatience and additivity are jointly sufficient for stationarity. The common difference effect (a violation of stationarity) is possibly the most studied anomaly of the exponential discounted utility model. It is commonly thought to arise from decreasing impatience (as in models of hyperbolic discounting). However, Theorem 2a shows that it can also arise on account of subadditivity. Below we discuss discount functions that can take account of the common difference effect through either decreasing impatience or subadditivity (Example 8, subsection 7.1).

4.5 Probability independence, the certainty effect and its converse

Definition 11 *(Probability independence, the certainty effect and its converse): Consider the following four conditions: (I) $r, s \in [0,\infty)$; (II) $\mathbf{w} \in \mathbb{R}^m$ is time-neutral for the preference relation $\preceq_r$ and the probabilities $p, q \in (0,1]$; (III) $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ are either both gains or both losses, for $\preceq_r, \mathbf{w}, p, q$; (IV) $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$. Then the preference relation $\preceq_r$ exhibits:*
(a) Probability independence if, whenever (I)-(IV) hold, then

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s, p) .$$

(b) The certainty effect if, whenever (I)-(IV) hold, and $p > q$, then

$$\text{(i)}\ \ (\mathbf{w}, r, q) \prec_r (\mathbf{y}, r, q) \Rightarrow (\mathbf{y}, r, p) \succ_r (\mathbf{z}, r+s, p) ,$$
$$\text{(ii)}\ \ (\mathbf{w}, r, q) \succ_r (\mathbf{y}, r, q) \Rightarrow (\mathbf{y}, r, p) \prec_r (\mathbf{z}, r+s, q) .$$

(c) The converse certainty effect if, whenever (I)-(IV) hold, and $q > p$, then

$$\text{(i)}\ \ (\mathbf{w}, r, q) \prec_r (\mathbf{y}, r, q) \Rightarrow (\mathbf{y}, r, p) \prec_r (\mathbf{z}, r+s, p) ,$$
$$\text{(ii)}\ \ (\mathbf{w}, r, q) \succ_r (\mathbf{y}, r, q) \Rightarrow (\mathbf{y}, r, p) \succ_r (\mathbf{z}, r+s, p) .$$

It has been observed that the common difference effect strengthens with an increase in the probability of a gain (Keren and Roelofsma, 1995; Weber and Chapman, 2005). This has been attributed by Halevy (2008) to the certainty effect.

If a decision maker exhibits the certainty effect (Definition 11b), then that decision maker will appear more impatient as the probability of a gain increases. In particular, if that decision maker exhibits the common difference effect (Definition 6a) at all positive probabilities, then the common difference effect will strengthen with the increase in the probability of a gain.

5 Properties of preferences in terms of observable functions

In section 4, the properties stationarity, impatience, additivity, probability independence, and their violations, were all defined in terms of the underlying preference relations, $\preceq_r$, $r \geq 0$. In practice, these are not directly observable (although some of their consequences are). The question then is: How can we empirically determine whether the behavior of a decision maker exhibits any of these properties?

To answer this question, we define four functions: The *delay function*, D, the *defer function*, Δ, the *shift function*, S and the *certainty function*, C. Each of these four functions can be uniquely elicited from observed behavior. These functions determine whether the behavior of a decision maker exhibits any of the properties considered in section 4.

In particular, we show the following. (1) The delay function, D, determines stationarity, the common difference effect and its converse. (2) The defer function, Δ, determines constant, decreasing and increasing impatience. (3) The shift function, S, determines additivity, subadditivity and superadditivity. (4) The certainty function, C, determines probability independence, the certainty effect and its converse.

5.1 Delay function, D

Consider a fixed moment in time, $r \geq 0$, and a preference relation, $\preceq_r$, at time r. Let $\mathbf{w}$ be time-neutral (Definition 2). Let $\mathbf{y}$ offered at time r be either a gain or a loss (Definition 4). Suppose that at time r a decision maker reveals the following indifference, where $p > 0$,

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) \,. \tag{5.1}$$

Suppose that the receipt of $\mathbf{y}$ is *delayed* to time $r + t$. We ask, at what $T \geq 0$ will $\mathbf{z}$ offered at time $r + s + T$ be indifferent to $\mathbf{y}$ offered at time $r + t$, i.e., for what T does the following hold?

$$(\mathbf{y}, r + t, p) \sim_r (\mathbf{z}, r + s + T, p) \,. \tag{5.2}$$

Note that the same preference relation, $\preceq_r$, occurs in (5.1) and (5.2).

Let us conjecture that T depends on r, s, t through a functional relation, say, $T = D(r, s, t)$. We get, from (5.1) and (5.2), that $D(r, s, t)$ must satisfy

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) \Rightarrow (\mathbf{y}, r + t, p) \sim_r (\mathbf{z}, r + s + D(r, s, t), p).$$

The above discussion motivates the following definition.

Definition 12 *(Delay function): Let $p \in (0, 1]$, $r \in [0, \infty)$. Consider the preference relation $\preceq_r$. Let $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ where $\mathbf{w}$ is time-neutral and $\mathbf{y}$ is either a gain or a loss. Suppose that the function, $D : \mathbb{R}^3_+ \to [0, \infty)$, has the property that for all $s, t \in [0, \infty)$,*

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) \Rightarrow (\mathbf{y}, r + t, p) \sim_r (\mathbf{z}, r + s + D(r, s, t), p).$$

Then we call D a delay function (corresponding to the preference relation $\preceq_r$ and the probability p).

Note that, in general, the delay function will be a function of $\mathbf{w}, \mathbf{y}, \mathbf{z}$, p as well as r, s, t. However, and to simplify notation, we will explicitly indicate the dependence of D on r, s, t only. This will cause no problems, since $\mathbf{w}, \mathbf{y}, \mathbf{z}$, p will be held fixed in any particular context. Similar remarks will also apply to the defer function, Δ, the shift function S and the certainty function, C, that we will define in subsections 5.2 - 5.4, below.

Lemma 2 *(Existence and uniqueness of a delay function):*
(a) A delay function, $D(r, s, t)$, exists.
(b) $D(r, s, t)$ is unique.

From Definition 12, the delay function is easily elicited. One possible method is as follows. First, fix $\mathbf{y}, r, s$. Next, elicit a $\mathbf{z}$ for which the decision maker, at time r, expresses indifference between $\mathbf{y}$ offered at time r and $\mathbf{z}$ offered at time $r+s$ (both with probability $p > 0$). Finally, fix t and elicit the value T for which the decision maker, again at time r, expresses indifference between $\mathbf{y}$ offered at time $r + t$ and $\mathbf{z}$ offered at time $r + s + T$ (again, both with probability $p > 0$). Then $D(r, s, t) = T$. This argument, and in the light of Lemma 2, has established the following result.

Result 1 *: The delay function, D, can be uniquely elicited from behavior.*

Theorem 3 *: Let D be the delay function corresponding to the preference relation $\preceq_r$. Then $\preceq_r$:*
(a) Is stationary if, and only if, $D(r, s, t) = t$, for all $s \geq 0$ and $t \geq 0$.
(b) Exhibits the common difference effect if, and only if, $D(r, s, t) > t$, for all $s > 0$ and $t > 0$.
(c) Exhibits the converse common difference effect if, and only if, $D(r, s, t) < t$, for all $s > 0$ and $t > 0$.

5.2 Defer function, Δ

Consider a preference relation, $\preceq_r$, at time r. Let $\mathbf{w} \in \mathbb{R}^m$ be time-neutral (Definition 2). Let $\mathbf{y}$ be either a gain or a loss relative to $\mathbf{w}$ (Definition 4). As in subsection 5.1, suppose that at time r a decision maker reveals the following indifference, where $p > 0$,

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) . \tag{5.3}$$

Note that (5.3) is identical to (5.1) of subsection 5.1.

Now, consider the preference relation, $\sim_{r+t}$, at time $r + t$. As in subsection 5.1, suppose that the receipt of $\mathbf{y}$ is *delayed* to time $r + t$. However, we now ask, *according to the preference relation* $\sim_{r+t}$ *at time* $r+t$ (rather than $\sim_r$ *at time* r), at what T will $\mathbf{z}$ offered at time $r + T + t$ be indifferent to $\mathbf{y}$ offered at time $r + t$ (both with probability $p > 0$), i.e., for what T does the following hold?

$$(\mathbf{y}, r + t, p) \sim_{r+t} (\mathbf{z}, r + t + T, p) . \tag{5.4}$$

Comparing (5.4), of this subsection, with (5.2) of subsection 5.1, we see that the evaluation of the indifference is *deferred* to the preference relation $\sim_{r+t}$ (and, on the right-hand-side of (5.4) we have $r + T + t$ rather than $r + s + T$ of (5.2)).

Let us conjecture that T depends on r, s, t through a functional relation, say, $T = \Delta(r, s, t)$. We get, from (5.3) and (5.4), that $\Delta(r, s, t)$ must satisfy

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) \Rightarrow (\mathbf{y}, r + t, p) \sim_{r+t} (\mathbf{z}, r + \Delta(r, s, t) + t, p) .$$

This motivates the following definition.

Definition 13 *(Defer function): Let* $p \in (0, 1]$, $r \in [0, \infty)$. *Consider the preference relation* $\preceq_r$. *Let* $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ *where* $\mathbf{w}$ *is time-neutral for the preference relation* $\preceq_r$ *and* $\mathbf{y}$ *is either a gain or a loss. Suppose that the function,* $\Delta : \mathbb{R}^3 \to [0, \infty)$, *has the property that for all* $s, t \in [0, \infty)$,

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + s, p) \Rightarrow (\mathbf{y}, r + t, p) \sim_{r+t} (\mathbf{z}, r + \Delta(r, s, t) + t, p) .$$

Then we call Δ *a defer function (corresponding to the preference relation* $\preceq_r$ *and the probability* p*).*

Lemma 3 *(Uniqueness of a defer function): A defer function,* $\Delta(r, s, t)$, *if it exists, is unique.*

From Definition 13, a defer function, if it exists, is easily elicited. One possible method is as follows. First, fix $\mathbf{y}, r, s$. Next, elicit a $\mathbf{z}$ for which the decision maker, at time r, expresses indifference between $\mathbf{y}$ offered at time r and $\mathbf{z}$ offered at time $r + s$ (both with probability $p > 0$). Finally, fix t and elicit the value T for which the decision maker, at time $r + t$, expresses indifference between $\mathbf{y}$ offered at time $r + t$ and $\mathbf{z}$ offered at time $r + T + t$ (both with probability $p > 0$). Then $\Delta(r, s, t) = T$. This argument, and in the light of Lemma 3, has established the following result.

Result 2 *: The defer function, Δ, if it exists, can be uniquely elicited from behavior.*

Theorem 4 *(Existence and properties of a defer function):*
(a) Suppose $\preceq_r$ exhibits constant impatience. Then a defer function, $\Delta(r,s,t)$, exists and $\Delta(r,s,t) = s$, for all $s \geq 0$ and $t \geq 0$. Conversely, if a defer function, $\Delta(r,s,t)$, exists and satisfies $\Delta(r,s,t) = s$, for all $s \geq 0$ and $t \geq 0$, then $\preceq_r$ exhibits constant impatience.
(b) Suppose $\preceq_r$ exhibits strictly decreasing impatience. Then a defer function, $\Delta(r,s,t)$, exists and $\Delta(r,s,t) > s$, for all $s > 0$ and $t > 0$. Conversely, if a defer function, $\Delta(r,s,t)$, exists and satisfies $\Delta(r,s,t) > s$, for all $s > 0$ and $t > 0$, then $\preceq_r$ exhibits strictly decreasing impatience.
(c) Suppose $\preceq_r$ exhibits strictly increasing impatience. Then a defer function, $\Delta(r,s,t)$, exists and $\Delta(r,s,t) < s$, for all $s > 0$ and $t > 0$. Conversely, if a defer function, $\Delta(r,s,t)$, exists and satisfies $\Delta(r,s,t) < s$, for all $s > 0$ and $t > 0$, then $\preceq_r$ exhibits strictly increasing impatience.

5.3 Shift function, S

Consider the preference relations $\preceq_r$ at time $r \geq 0$. Let $\mathbf{w} \in \mathbb{R}^m$ be time-neutral for $\preceq_r$ (Definition 2). Let $\mathbf{x} \in \mathbb{R}^m$ be either a gain or a loss relative to $\mathbf{w}$ and according to $\preceq_r$ (Definition 4). Suppose that at time r a decision maker reveals the following indifference, where $p > 0$,

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p), \tag{5.5}$$

and also suppose that at time $r+t$ the decision maker reveals the following indifference.

$$(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+t+s, p). \tag{5.6}$$

We ask: at what time $r+T$ will the receipt of $\mathbf{z}$ be indifferent to the receipt of x at time r (both with probability $p > 0$)? That is

$$(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+T, p). \tag{5.7}$$

In other words, at what time, $r+T$, will discounting from time $r+T$ back to time r, in one step, be equivalent to first discounting from time $r+t+s$ back to time $r+t$, then discounting from time $r+t$ back to time r?

Comparing (5.6) with (5.7), we see that the receipt of $\mathbf{z}$ in (5.7) is *shifted*, from time $r+t+s$ to time $r+T$. Note that the receipt of $\mathbf{z}$ in (5.7), compared with that in (5.6), may be delayed or brought forward.

Let us conjecture that T depends on r, s, t through a functional relation, say, $T = S(r,s,t)$. We get, from (5.5) to (5.7), that $S(r,s,t)$ must satisfy

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \text{ and } (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)$$
$$\text{imply } (\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+S(r,s,t), p).$$

This motivates the following definition.

Definition 14 *(Shift function): Let* $p \in (0,1]$, $r \in [0,\infty)$. *Consider the preference relation* $\preceq_r$. *Let* $\mathbf{w},\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$ *where* $\mathbf{w}$ *is time-neutral and* $\mathbf{x}$ *is either a gain or a loss according to* $\preceq_r$. *Suppose that the function* $S : \mathbb{R}_+^3 \to [0,\infty)$ *has the property that for all* $s,t \in [0,\infty)$,

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p) \ \text{ and } \ (\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)$$
$$\text{imply } \ (\mathbf{x}, r, p) \sim_r (\mathbf{z}, r + S(r,s,t), p).$$

Then we call S *a shift function (corresponding to the preference relation* $\preceq_r$ *and the probability* p*).*

Lemma 4 *(Uniqueness of a shift function): A shift function,* $S(r,s,t)$*, if it exists, is unique.*

From Definition 14, the shift function, if it exists, is easily elicited. One possible method is as follows. First, fix $\mathbf{x}, r, t$. Next, elicit a $\mathbf{y}$ for which the decision maker, at time r, expresses indifference between $\mathbf{x}$ offered at time r and $\mathbf{y}$ offered at time $r+t$ (both with probability $p > 0$). Next, fix s and elicit a $\mathbf{z}$ for which the decision maker, at time $r+t$, expresses indifference between $\mathbf{y}$ offered at time $r+t$ and $\mathbf{z}$ offered at time $r+s+t$ (again, both with probability $p > 0$). Finally, elicit the value T for which the decision maker, at time r, expresses indifference between $\mathbf{x}$ offered at time r and $\mathbf{z}$ offered at time $r+T$ (both with probability $p > 0$). Then $S(r,s,t) = T$. This argument, and in the light of Lemma 4, has established the following result.

Result 3 *: The shift function,* S*, if it exists, can be uniquely elicited from behavior.*

Theorem 5 *(Existence and properties of a shift function):*
(a) Suppose $\preceq_r$ *is additive. Then a shift function,* $S(r,s,t)$*, exists and* $S(r,s,t) = s+t$*, for all* $s \geq 0$ *and* $t \geq 0$*. Conversely, if a shift function,* $S(r,s,t)$*, exists and satisfies* $S(r,s,t) = s+t$*, for all* $s \geq 0$ *and* $t \geq 0$*, then* $\preceq_r$ *is additive.*
(b) Suppose $\preceq_r$ *is strictly subadditive. Then a shift function,* $S(r,s,t)$*, exists and* $S(r,s,t) > s+t$*, for all* $s > 0$ *and* $t > 0$*. Conversely, if a shift function,* $S(r,s,t)$*, exists and satisfies* $S(r,s,t) > s+t$*, for all* $s > 0$ *and* $t > 0$*, then* $\preceq_r$ *is strictly subadditive.*
(c) Suppose $\preceq_r$ *is strictly superadditive. Then a shift function,* $S(r,s,t)$*, exists and* $S(r,s,t) < s+t$*, for all* $s > 0$ *and* $t > 0$*. Conversely, if a shift function,* $S(r,s,t)$*, exists and satisfies* $S(r,s,t) < s+t$*, for all* $s > 0$ *and* $t > 0$*, then* $\preceq_r$ *is strictly superadditive.*

5.4 Certainty function, C

Definition 15 *(Certainty function): Let* $r \in [0,\infty)$*. Consider the following situation:*
(i) $\mathbf{w} \in \mathbb{R}^m$ *is time-neutral for the preference relation* $\preceq_r$ *and both probabilities* $p, q \in (0,1]$,

(ii) $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ *are either both gains or both losses, for* $\preceq_r, \mathbf{w}, p, q$.
Suppose that the function, $C : [0, \infty) \to [0, \infty)$, *has the property that for all* $s \in [0, \infty)$,

$$(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r + s, q) \Rightarrow (\mathbf{y}, r, p) \sim_r (\mathbf{z}, r + C(s), p),$$

then we call C *a certainty function (corresponding to the preference relation* $\preceq_r$*, the probabilities* p, q *and the outcomes* $\mathbf{w}, \mathbf{y}, \mathbf{z}$*).*

Lemma 5 *(Existence and uniqueness of a certainty function):*
(a) A certainty function, C*, exists.*
(b) C *is unique.*

From Definition 15, the certainty function is easily elicited. One possible method is as follows. First, fix $p, q \in (0, 1]$, $r, s \in [0, \infty)$, $\mathbf{y} \in \mathbb{R}^m$. Next, elicit a $\mathbf{z}$ for which the decision maker, at time r, expresses indifference between $\mathbf{y}$ offered for delivery at time r and $\mathbf{z}$ offered at time $r + s$ (both with probability $q > 0$). Finally, elicit the value T for which the decision maker, again at time r, expresses indifference between $\mathbf{y}$ offered for delivery at time r and $\mathbf{z}$ offered at time $r + T$ (now both with probability $p > 0$). Then $C(s) = T$. This argument, and in the light of Lemma 5, has established the following result.

Result 4 *: The certainty function,* C*, can be uniquely elicited from behavior.*

Theorem 6 *: Let (i)* C *be the certainty function corresponding to the preference relation* $\preceq_r$*, (ii)* $\mathbf{w}$ *be time-neutral for* $\preceq_r, p, q$*, and (iii)* $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ *are either both gains or both losses, for* $\preceq_r, \mathbf{w}, p, q$*. Then* $\preceq_r$ *exhibits:*
(a) Probability independence if, and only if, $C(s) = s$*, for all* $s \geq 0$*,* $p, q \in (0, 1]$*,* $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$*, such that* $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r + s, q)$*.*
(b) The certainty effect if, and only if, $C(s) < s$*, for all* $s > 0$*,* $p, q \in (0, 1]$*,* $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$*, where* $p > q$*.*
(c) The converse certainty effect if, and only if, $C(s) > s$*, for all* $s > 0$*,* $p, q \in (0, 1]$*,* $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$*, where* $p > q$*.*

6 A general discounted-utility model

In this section we formulate a general discounted-utility model. Our basic ingredients are a *utility function*, u, and a *time discount function*, δ. We describe these in subsections 6.1 and 6.2, below. Then we derive the preference relation implied by these and show that the axioms of section 3 are satisfied (subsection 6.3).

6.1 Utility function, u

We adopt the prospect theory utility function with the reference point taken to be $\mathbf{0}$. This gives a particularly tractable form to our conclusions. Furthermore, this facilitates the explanation of *gain-loss asymmetry* and *delay-speedup asymmetry* (which, however, are not the focus of this paper).[16]

We stress that none of the conclusions in this paper depend on how the reference point for outcomes is determined. However, usually a reasonable choice of the reference outcome 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, 2019a; subsection 2.4.4).

Definition 16 *(Utility function, u): An instantaneous, time-invariant, utility function (or simply utility) is a continuous function $u : \mathbb{R}^m \to \mathbb{R}$ such that $u(x_1, x_2, ..., x_m)$ is strictly increasing in each x_i, $i = 1, 2, ..., m$, and $u(\mathbf{0}) = 0$.*

6.2 Discount function, δ

We define a general time discount function. This is motivated by the *interval discount function* of Read (2001) and Scholten and Read (2006), but goes beyond these contributions in that we also incorporate uncertainty. We shall denote the interval discount function by $\delta_r(t, p)$. It captures how the utility of an outcome $\mathbf{x} \in \mathbb{R}^m$ promised at time $r \in [0, \infty)$ for delivery at time $t \in [r, \infty)$ with probability $p \in [0, 1]$, is discounted. For example, for $p = 1$, the exponential time discount function can be written as $\delta_r(t, 1) = e^{-\beta(t-r)}$, $0 \leq r \leq t$; or as $\delta_r(r + t, 1) = e^{-\beta t}$, $r, t \in [0, \infty)$; where $\beta > 0$.

Definition 17 *(Discount functions): Let*

$$\nabla = \left\{(r, t, p) \in \mathbb{R}^3 : 0 \leq r \leq t, p \in [0, 1]\right\}. \tag{6.1}$$

A time discount function is a mapping, $\delta : \nabla \to (0, 1]$, satisfying:
(a) $\delta_r(t, 0) = 0$, $\delta_r(r, 1) = 1$, $p > 0 \Rightarrow \delta_r(t, p) > 0$.
(b) $\delta_r(t, p)$ is strictly increasing in p.
(c) For $p > 0$, $\delta_r(t, p)$ is strictly increasing in r but strictly decreasing in t.
(d) $\lim_{t \to \infty} \delta_r(t, p) = 0$.

Sometimes we find it convenient to work with the form $\delta_r(t, p)$, $0 \leq r \leq t$, but at other times we find it convenient to work with the form $\delta_r(r + t, p)$, $r, t \in [0, \infty)$.

Lemma 6 *: Let δ be a time discount function (Definition 17) such that $\delta_r(t, p)$ is continuous in t. Let $r, s, \tau \in [0, \infty)$, $p \in (0, 1]$. Then $\tau \mapsto \delta_r(r + s + \tau, p)$ maps $[0, \infty)$ onto $(0, \delta_r(r + s, p)]$.*

[16] Our choice of a prospect theory formulation is based on two considerations. First, prospect theory arguably provides the most accurate account of the evidence from situations of risk, uncertainty, and ambiguity (Kahneman and Tversky, 2000; Wakker, 2010; Dhami, 2019a; Ruggeri et al., 2020.). Second, expected utility and rank dependent utility can be recovered as special cases.

6.3 From utility to preferences

Given the utility function, u (subsection 6.1) and the discount function, δ (subsection 6.2), we can define an intertemporal utility function, U. From the latter, we define the preference relation $\preceq_r$. Finally, we show that $\preceq_r$ satisfies Axioms 1-7 of section 3.

Definition 18 *(Intertemporal discounted utility): Let u be a utility function (Definition 16) and δ a discount function (Definition 17). We define the intertemporal discounted utility function of the decision maker by*

$$U_r(\mathbf{x}, r+t, p) = \delta_r(r+t, p)\, u(\mathbf{x}), \tag{6.2}$$

where $\mathbf{x} \in \mathbb{R}^m$ is an outcome promised at time r for delivery at time $r+t$ with probability p, $p \in [0,1]$, $r, t \in [0, \infty)$.

Definition 19 *(Preferences): Let U be an intertemporal utility function of the decision maker (Definition 18). We define the preference relation, $\preceq_r$, by*

$$(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q) \Leftrightarrow U_r(\mathbf{x}, r+s, p) \leq U_r(\mathbf{y}, r+t, q).$$

We call $\preceq_r$ the preference relation induced by U_r and U_r a utility function that represents $\preceq_r$.

Lemma 7 *(Time-neutral outcomes, gains and losses): Let $\preceq_r$ be the preference relation induced by U_r (Definitions 18, 19). Then, for probability $p \in (0,1]$,*
(a) $\mathbf{w}$ is time-neutral (Definition 2) if, and only if, $u(\mathbf{w}) = 0$.
(b) $(\mathbf{x}, r+t, p)$ is a gain if, and only if, $u(\mathbf{x}) > 0$.
(c) $(\mathbf{x}, r+t, p)$ is a loss if, and only if, $u(\mathbf{x}) < 0$.

From Lemma 7, we see that if an outcome is time-neutral for a probability $p \in (0,1]$, then it is time-neutral for all probabilities $p \in (0,1]$. Similarly for gains and losses. We now state the most important conclusion of this section.

Theorem 7 : *Axioms 1-5 (section 3) hold for our discounted utility model (Definition 18). If the time discount function (Definition 17) is continuous in time, then Axioms 6 (time sensitivity) and 7 (probability sensitivity) also hold for our discounted utility model.*

In the light of Theorem 7, care must be taken that Axioms 6 (time sensitivity) and 7 (probability sensitivity), or the conclusions that depend on them, are not applied when the discount functions are not continuous in time, t.

6.4 Delay, defer, shift and certainty functions for the discounted-utility model

Lemma 8 : *Let $p, q \in (0, 1]$ and $r, s, t \in [0, \infty)$. Let δ_r be the discount function for our discounted-utility model (Definitions 16-19). Let D, Δ, S, C be, respectively, the delay, defer, shift and certainty functions (section 5) for that model. Then D, Δ, S, C satisfy:*
(a) $\delta_r(r+s,p)\,\delta_r(r+t,p) = \delta_r(r,p)\,\delta_r(r+s+D(r,s,t),p)$.
(b) $\delta_r(r+s,p)\,\delta_{r+t}(r+t,p) = \delta_r(r,p)\,\delta_{r+t}(r+\Delta(r,s,t)+t,p)$.
(c) $\delta_r(r+t,p)\,\delta_{r+t}(r+t+s,p) = \delta_{r+t}(r+t,p)\,\delta_r(r+S(r,s,t),p)$.
(d) $\delta_r(r,q)\,\delta_r(r+s,p) = \delta_r(r,p)\,\delta_r(r+C(s),q)$.

7 Separable time discount functions

Here, we consider time discount functions (Definition 17) that take the following form:

$$\delta_r(r+t,p) = w(p\Pi(t))\,\delta_r^0(r+t), \tag{7.1}$$

where $r, t \in \mathbb{R}_+$, $p \in [0, 1]$, w is a probability weighting function (subsection 7.2, below), Π is a survival function (subsection 7.3, below) and δ^0 is a riskless time discount function (subsection 7.1, below).

An interpretation of (7.1) is as follows. Suppose that at time r the experimenter promises to deliver the outcome $\mathbf{x}$ at time $r+t$ with probability p. There are two reasons why this promised reward, $\mathbf{x}$, may not actually materialize in the future. First, the reward is inherently risky because of the probability p of delivery. Second, there are various other reasons the subject may not actually get to play the lottery, $(\mathbf{x}, p)$, say, on account of death, or other risky factors that are not reflected in p (Halevy, 2008). Let $\Pi(t)$ be the probability that the subject will actually play the lottery $(\mathbf{x}, p)$. Thus, the joint probability that the subject will receive the outcome $\mathbf{x}$ is $p\Pi(t)$. This is then weighted by the probability weighting function, w. We could also consider ambiguity in the delivery of the reward by allowing the probability weighting function, w, to be source-dependent; for further discussion, references, and the evidence, see Wakker (2010), and Dhami (2019a, Section 4.4.2).

7.1 Riskless time discount function, δ^0

Definition 20 *(Riskless time discount functions): Let*

$$\nabla_0 = \left\{(r,t) \in \mathbb{R}^2 : 0 \le r \le t\right\}. \tag{7.2}$$

A riskless time discount function is a mapping, $\delta^0 : \nabla_0 \to [0, 1]$, satisfying:
(a) For each $r \in [0, \infty)$, $\delta_r^0(t)$ is a strictly decreasing function of $t \in [r, \infty)$ into $(0, 1]$ with $\delta_r^0(r) = 1$.
(b) For each $t \in [r, \infty)$, $\delta_r^0(t)$ is a strictly increasing function of $r \in [0, t]$ into $(0, 1]$.
(c) $\lim_{t\to\infty} \delta_r(t) = 0$.

Let δ^0 be a riskless time discount function (Definition 20). Let $r \in [0, \infty)$. Assume that $\delta_r^0(t)$ is continuous in $t \in [r, \infty)$. Let $s, \tau \in [0, \infty)$. Then $\tau \mapsto \delta_r(r + s + \tau)$ maps $[0, \infty)$ onto $(0, \delta_r(r + s)]$. The proof is similar to that of Lemma 6.

In section 4 we defined stationarity, constant impatience, and additivity, and their violations, for a preference relation $\preceq_r$. Below, we redefine these concepts but for a riskless time discount function. Under certainty, these two sets of concepts coincide as we shall see in Corollary 1, below. For instance, for the riskless time discount function, Corollary 1 will show that constant impatience of preferences, $\preceq_r$, in section 4 is equivalent to constant impatience of the riskless time discount function (Definition 21 below); additivity of preferences, $\preceq_r$, is equivalent to additivity of the riskless time discount function (Definition 22 below); and stationarity of preferences, $\preceq_r$, is equivalent to stationarity of the riskless time discount function (Definition 23 below). These equivalences also hold for violations of constant impatience, additivity, and stationarity. However, this equivalence between preferences and discount functions does not coincide under uncertainty (Theorem 10 below).

Definition 21 *: Let δ^0 be a riskless time discount function (Definition 20). Let $r \in [0, \infty)$. Then δ^0 exhibits :*
(a) constant impatience for r if $\delta_r^0(r + s) = \delta_{r+t}^0(r + s + t)$ for all $s \geq 0$ and $t \geq 0$,
(b) strictly decreasing impatience for r if $\delta_r^0(r + s) < \delta_{r+t}^0(r + s + t)$ for all $s > 0$ and $t > 0$,
(c) strictly increasing impatience for r if $\delta_r^0(r + s) > \delta_{r+t}^0(r + s + t)$ for all $s > 0$ and $t > 0$.
We use the term "decreasing (increasing) impatience" to mean "constant or strictly decreasing (increasing) impatience".

Definition 22 *: Let δ^0 be a riskless time discount function (Definition 20). Let $r \in [0, \infty)$. Then δ^0 is:*
(a) additive for r if $\delta_r^0(r + t)\, \delta_{r+t}^0(r + s + t) = \delta_r^0(r + s + t)$ for all $s \geq 0$ and $t \geq 0$,
(b) strictly subadditive for r if $\delta_r^0(r + t)\, \delta_{r+t}^0(r + s + t) < \delta_r^0(r + s + t)$ for all $s > 0$ and $t > 0$,
(c) strictly superadditive for r if $\delta_r^0(r + t)\, \delta_{r+t}^0(r + s + t) > \delta_r^0(r + s + t)$ for all $s > 0$ and $t > 0$.
We use the term "sub- (super-) additive" to mean "additive or strictly sub- (super-) additive".

Definition 23 *: Let δ^0 be a riskless time discount function (Definition 20). Let $r \in [0, \infty)$. Then δ^0:*
(a) is stationary for r if $\delta_r^0(r + s)\, \delta_r^0(r + t) = \delta_r^0(r + s + t)$, for all $s \geq 0$ and $t \geq 0$
(b) exhibits the common difference effect for r if $\delta_r^0(r + s)\, \delta_r^0(r + t) < \delta_r^0(r + s + t)$, for all $s > 0$ and $t > 0$,
(c) exhibits the converse common difference effect for r if $\delta_r^0(r + s)\, \delta_r^0(r + t) > \delta_r^0(r + s + t)$, for all $s > 0$ and $t > 0$.

We now give examples of several important time discount functions.

Example 3 *GRS: We propose a generalization of the Read-Scholten time discount function (Example 8, below) which we call the generalized Read-Scholten time discount function (GRS). We define it as follows.*
Let $Q:[0,\infty)\to[0,\infty)$ and $\phi:[0,\infty)\to\mathbb{R}$ be strictly increasing. Let $r,t\in[0,\infty)$. Then GRS is given by

$$\delta_r^0(r+t)=e^{-Q(\phi(r+t)-\phi(r))}.$$

Note that, in Example 3, $\phi(t)$ may be negative. However, since ϕ is strictly increasing, we have $\phi(r+t)-\phi(r)\geq 0$. Hence, $Q(\phi(r+t)-\phi(r))$ is well defined, and non-negative. Thus, $e^{-Q(\phi(r+t)-\phi(r))}\leq 1$, as required by a time discount function. Allowing $\phi(r)$ to be negative, enables us to incorporate the quasi hyperbolic time discount function (Example 5, below) as a special case of GRS.

If the functions Q and ϕ of Example 3 are onto $[0,\infty)$, then they are invertible and continuous, and their inverses are also continuous. The GRS riskless time discount function is then continuous in t.

The next five examples (Examples 4-8) are all special cases of GRS (Example 3). The claims we make about their properties can all be directly verified from Definitions 21-23.

Example 4 *(Exponential time discount function, EDF): The exponential time discount function (Samuelson, 1937), is given by*

$$\delta_r^0(r+t)=e^{-\beta t},$$

where $r,t\in[0,\infty)$, $\beta>0$. It is the special case of GRS (Example 3) with $\phi(s)=s$ and $Q(x)=\beta x$. For each $r\geq 0$, EDF is additive, stationary, and exhibits constant impatience.

Example 5 *(Quasi-hyperbolic time discount function, PPL): The quasi-hyperbolic time discount function (Phelps and Pollack, 1968; Laibson, 1997) is popular in applied work. While PPL is normally presented in discrete time, it is convenient to present the continuous time analogue.*
Let $r,t\in[0,\infty)$, $\alpha>0$, $\beta>0$. Then PPL is given by

$$\delta_r^0(r+t)=\begin{cases}1 & \text{if } r=t=0\\ e^{-\alpha-\beta t} & \text{if } r=0,\ t>0\\ e^{-\beta t} & \text{if } r>0\end{cases}.$$

It is the special case of GRS (Example 3) with $Q(x)=\beta x$, and $\phi(s)=s$ for $s>0$ but $\phi(0)=-\frac{\alpha}{\beta}$.
PPL is identical to EDF for $r>0$. Hence, for $r>0$, it is additive, stationary and exhibits constant impatience. However, for $r=0$, it is discontinuous in t, making a downward jump between $t=0$ and $t>0$; and it is this feature that distinguishes it from EDF. Due to this downward jump, for $r=0$ PPL (although still additive) exhibits strictly decreasing impatience and the common difference effect.

Example 6 *(Generalized hyperbolic time discount function, LP): The generalized hyperbolic time discount function (Loewenstein and Prelec, 1992; al-Nowaihi and Dhami, 2006a) is given by*

$$\delta_0^0(t) = (1 + \alpha t)^{-\frac{\beta}{\alpha}},$$

where $\alpha > 0$, $\beta > 0$. Example 7 shows how it can be obtained from the GRS discount function. LP is additive, exhibits strictly decreasing impatience and the common difference effect.[17]

Example 7 *(Generalized Loewenstein-Prelec time discount function, GLP): Let $0 \leq r \leq t$, $\alpha > 0$, $\beta > 0$. Then we define the generalized Loewenstein-Prelec time discount function (GLP) by*[18]

$$\delta_r^0(t) = \left(\frac{1 + \alpha t}{1 + \alpha r}\right)^{-\frac{\beta}{\alpha}},$$

setting $r = 0$ gives LP (Example 6). GLP can be obtained from GRS by setting $\phi(s) = \ln(1 + \alpha s)$ and $Q(x) = \frac{\beta}{\alpha} x$ in Example 3. For each $r \geq 0$, GLP is additive, exhibits strictly decreasing impatience and the common difference effect.

Note that GLP approaches EDF as $\alpha \to 0$.

Example 8 *(Read-Scholten interval discount function, RS): The Read-Scholten (RS) interval discount function explains the common difference effect either through decreasing impatience, subadditivity, or both (Read, 2001; Scholten and Read, 2006; Scholten et al., 2016). Thus, it is the most general discount function available. It is defined as follows. Let $0 \leq r \leq t$, $\alpha > 0$, $\beta > 0$, $\rho > 0$, $\tau > 0$, then RS is given by*[19]

$$\delta_r^0(t) = [1 + \alpha (t^\tau - r^\tau)^\rho]^{-\frac{\beta}{\alpha}}.$$

RS can be obtained from GRS by setting $\phi(s) = s^\tau$ and $Q(x) = \frac{\beta}{\alpha} \ln(1 + \alpha x^\rho)$ in Example 3.

(ai) If $\tau = 1$, then RS exhibits constant impatience for every $r \geq 0$.
(aii) If $0 < \tau < 1$, then RS exhibits strictly decreasing impatience for every $r \geq 0$.
(aiii) If $\tau > 1$, then RS exhibits strictly increasing impatience for every $r \geq 0$.
(b) If $0 < \tau \leq 1$ and $0 < \rho \leq 1$, then RS exhibits the common difference effect for every $r \geq 0$.
(c) Let $\tau > 0$ and $0 < \rho \leq 1$, then RS is strictly subadditivity for every $r \geq 0$.
(d) If $\rho > 1$, then RS can be neither subadditive nor additive.

[17] For an axiomatization of LP, see Loewenstein and Prelec (1992) and al-Nowaihi and Dhami (2006a, 2008a).

[18] Equivalently, GLP can be represented by $\delta_r^0(r+t) = \left[\frac{1+\alpha(r+t)}{1+\alpha r}\right]^{-\frac{\beta}{\alpha}}$, where $r, t \in [0, \infty)$.

[19] Equivalently, RS can be represented by $\delta_r^0(r+t) = \left\{1 + \alpha\left[(r+t)^\tau - r^\tau\right]^\rho\right\}^{-\frac{\beta}{\alpha}}$, where $r, t \in [0, \infty)$.

We can now see the interpretation of the parameters τ and ρ in the RS discount function.[20] τ controls impatience, independently of the values of the other parameters α, β and ρ; $0 < \tau < 1$, gives decreasing 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 the RS discount function can be neither subadditive nor additive because, for $\rho > 1$ we get $S(r, s, t) \lesseqqgtr s + t$, depending on the particular values of r, s and t.

In general, neither GLP nor RS is a special case of the other. However, for $r = 0$ (and only for $r = 0$), RS reduces to GLP when $\rho = \tau = 1$.

Read and Scholten, (2006), Scholten and Read (2010) and Read et al. (2016) presented a critique of the psychological basis for time discount models. They developed an *attribute model* that is based on firmer psychological foundations. al-Nowaihi and Dhami (2008b, subsection 5.1, pp. 38-40)[21] proposed the GRS discount function (Example 3, above) and argued that the Read and Scholten (2006) tradeoff model is equivalent to a discounted utility model with the GRS time discount function. Hence, the arguments in Read and Scholten, (2006), Scholten and Read (2010) and Read et al. (2016) in support of their tradeoff model also lend further support to their own time discount function, the RS time discount function (Example 8, above) and its generalization, GRS (Example 3, above).

7.2 Probability weighting function, w

Under expected utility we have linear probability weighting. This was contradicted by the Allais paradox in common ratio and common consequence forms in the 1950s and, since then, by a well developed body of empirical evidence (Kahneman and Tversky, 2000; Dhami, 2019a). This has led to significant developments in non-expected utility, such as rank dependent utility and prospect theory, that rely on non-linear probability weighting.

Definition 24 *(Probability weighting function): By a probability weighting function we mean a strictly increasing function* $w(p) : [0,1] \overset{onto}{\to} [0,1]$.

Lemma 9 *: A probability weighting function has the following properties:*
(a) $w(0) = 0$, $w(1) = 1$, $p \in (0,1) \Rightarrow w(p) \in (0,1)$.
(b) w *has a unique inverse,* w^{-1}, *and* w^{-1} *is also a strictly increasing function from* $[0,1]$ *onto* $[0,1]$.
(c) w *and* w^{-1} *are continuous.*

Definition 25 *: We say that the probability weighting function* w *(Definition 24) is:*
(a) Additive if $w(p)\,w(q) = w(pq)$ *for all* $p, q \in [0,1]$.
(b) Strictly subadditive if $w(p)\,w(q) < w(pq)$ *for all* $p, q \in (0,1)$.

[20] 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).

[21] Alternatively, see al-Nowaihi and Dhami (2018), subsection 7.1, pp. 24-27.

(c) Strictly superadditive if $w(p)w(q) > w(pq)$ for all $p, q \in (0,1)$.
We use the term "sub- (super-) additive" to mean "additive or strictly sub- (super-) additive".

Example 9 *(Prelec, 1998)[22]: The Prelec probability weighting function is given by*

$$w(0) = 0,\ w(1) = 1, \tag{7.3}$$

$$w(p) = e^{-\beta(-\ln p)^{\alpha}},\ p \in (0,1),\ \alpha > 0,\ \beta > 0. \tag{7.4}$$

Lemma 10 *: The Prelec probability weighting function (Example 9) is:*
(a) Additive for $\alpha = 1$, in which case $w(p) = p^{\beta}$.
(b) Strictly subadditive for $\alpha < 1$.
(c) Strictly superadditive for $\alpha > 1$.

For $\alpha = \beta = 1$, the Prelec probability weighting function reduces to the identity transformation, $w(p) = p$, as under expected utility. For $\alpha < 1$, the Prelec function overweights low probabilities but underweights high probabilities. The reverse holds for $\alpha > 1$.

7.3 Survival function, Π

We now introduce basic elements from *survival analysis* that allow us to take account of uncertainty in receiving a future prize (hazard) that arises over and above the probabilistic nature of the outcomes.

Definition 26 *: A survival function[23] is a map $\Pi : [0,\infty) \to (0,1]$ satisfying either:*
(a) $\Pi(t) = 1$ for all $t \geq 0$, or,
(b) $\Pi(0) = 1$ and $s < t \Rightarrow \Pi(s) > \Pi(t)$.
When Π is differentiable, the function $h(t) = -\Pi'(t)/\Pi(t)$ is known as the hazard function.

We wish to relate our analysis to the literature on time discounting which, in the main, assumes no hazard. This is why we have introduced case (a) into Definition 26. For the same reason, we do not assume that $\Pi(t) \to 0$, as $t \to \infty$ (which is standard in survival analysis).

Definition 27 *: The survival function Π (Definition 26) exhibits:*
(a) Constant hazard if $\Pi(s)\Pi(t) = \Pi(s+t)$ for all $s \geq 0$, $t \geq 0$,
(b) Strictly decreasing hazard if $\Pi(s)\Pi(t) < \Pi(s+t)$ for all $s > 0$, $t > 0$,
(c) Strictly increasing hazard if $\Pi(s)\Pi(t) > \Pi(s+t)$ for all $s > 0$, $t > 0$.
We use the term "decreasing (increasing) hazard" to mean "constant or strictly decreasing (increasing) hazard".

[22] For axiomatic derivations of the Prelec probability weighting function, see Prelec (1998), Luce (2001) and al-Nowaihi and Dhami (2006b).

[23] In the literature on survival analysis, the symbol S is usually used for the survival function. However, we have already used S for the shift function.

Example 10 : *The Weibull survival function is given by*

$$\Pi(t;\pi,k) = e^{-\pi t^k},\ \pi \in [0,\infty),\ k \in (0,\infty),\ t \geq 0, \tag{7.5}$$

hence, the hazard function is

$$h(t;\pi,k) = k\pi t^{k-1}. \tag{7.6}$$

Lemma 11 : *Let $\pi \in [0,\infty)$, $k \in (0,\infty)$. The Weibull survival function Π (Example 10) exhibits*
(a) Constant hazard if $k = 1$.
(b) Strictly decreasing hazard if $k < 1$ and $\pi > 0$.
(c) Strictly increasing hazard if $k > 1$ and $\pi > 0$.

If $\pi = 0$ in (7.5) and (7.6) then we have no hazard, which is the case usually considered in the literature on time discounting.

Strictly decreasing hazard (Definition 27b) is compatible with decreasing probability of survival ($\Pi'(t) < 0$) and even inevitable death ($\Pi(t) \to 0$, as $t \to \infty$). Many populations do exhibit decreasing hazard. In humans, as one gets older, one acquires greater immunity to disease, enjoys greater real income, benefits from progress in technology and in medicine, and learns to better handle dangerous situations (at least, until the onset of great old age).

Lemma 12, below, is the special case of Lemma 8, above, when the time discount function takes the separable form (7.1).

Lemma 12 : *Let $p, q \in (0,1]$ and $r, s, t \in [0,\infty)$. Let δ^0 be a riskless discount function for our discounted-utility model (Definitions 21-25). Let D, Δ, S, C be, respectively, the delay, defer, shift and certainty functions (section 5) for that model. Then D, Δ, S, C satisfy:*

$$\text{(a)}\ \frac{\delta_r^0(r+s)\,\delta_r^0(r+t)}{\delta_r^0(r+s+D(r,s,t))} = \frac{w(p)\,w(p\Pi(s+D(r,s,t)))}{w(p\Pi(s))\,w(p\Pi(t))},$$

$$\text{(b)}\ \frac{\delta_r^0(r+s)}{\delta_{r+t}^0(r+t+\Delta(r,s,t))} = \frac{w(p\Pi(\Delta(r,s,t)))}{w(p\Pi(s))},$$

$$\text{(c)}\ \frac{\delta_r^0(r+t)\,\delta_{r+t}^0(r+t+s)}{\delta_r^0(r+S(r,s,t))} = \frac{w(p)\,w(p\Pi(S(r,s,t)))}{w(p\Pi(t))\,w(p\Pi(s))},$$

$$\text{(d)}\ \frac{\delta_r^0(r+s)}{\delta_r^0(r+C(s))} = \frac{w(p)\,w(q\Pi(C(s)))}{w(q)\,w(p\Pi(s))}.$$

The next two theorems, Theorems 8 and 9, lead to the conclusion that the observation of the certainty effect, Definition 11b, implies the existence of both, hazard and non-additive probability weighting.

Theorem 8 *: Let $r, t \in [0, \infty)$. Let δ^0 be the riskless time discount function (Definition 20), w the probability weighting functions (Definition 24) and Π the survival function (Definition 26) for the separable time discount function $\delta_r(r+t, p) = w(p\Pi(t))\,\delta_r^0(r+t)$. Let $\preceq_r$ be the induced preference relation (Definition 19). Then $\preceq_r$ exhibits:*
(a) Probability independence if, and only if,

$$w(p)\,w(q\Pi(s)) = w(q)\,w(p\Pi(s)),$$

for all $s \geq 0$ and all $p, q \in (0, 1]$.
(b) The certainty effect if, and only if,

$$w(p)\,w(q\Pi(s)) > w(q)\,w(p\Pi(s)),$$

for all $s > 0$ and all $p, q \in (0, 1]$, where $p > q$.
(c) The converse certainty effect if, and only if,

$$w(p)\,w(q\Pi(s)) < w(q)\,w(p\Pi(s)),$$

for all $s > 0$ and all $p, q \in (0, 1]$, where $p > q$.

Theorem 9 *: Let $p, q \in (0, 1]$ and $r, s, t \in [0, \infty)$. Let $\delta_r(r+t, p) = w(p\Pi(t))\,\delta_r^0(r+t)$ be the separable time discount function for a discounted-utility model. Then:*
(a) If there is no hazard (i.e., $\Pi(t) = 1$, for all $t \geq 0$), then probability independence holds (Definition 11a).
(b) If probability weighting is additive (Definition 25a), then probability independence holds.
(c) In particular, the observation of the certainty effect, Definition 11b (or the converse certainty effect), implies the existence of both, hazard and non-additive probability weighting.

In section 5, above, we saw how stationarity, constant impatience and additivity, and their violations, can be characterized in terms of the delay, defer, shift and certainty functions, D, Δ, S and C. For a discounted utility model, with a separable time discount function (7.1), Theorem 10, below, gives an equivalent characterization, but in terms of the riskless time discount function, δ^0, the probability weighting function, w, and the survival function, Π.

Theorem 10 *: Let δ^0 be the riskless time discount function (Definition 20), w the probability weighting function (Definition 24) and Π the survival function (Definition 26) for the separable time discount function $\delta_r(r+t, p) = w(p\Pi(t))\,\delta_r^0(r+t)$. Let $\preceq_r$ be the induced preference relation (Definition 19). Then $\preceq_r$:*
(ai) Exhibits constant impatience for probability $p \in (0, 1]$ if, and only if,

$$\delta_r^0(r+s) = \delta_{r+t}^0(r+s+t),$$

for all $s \geq 0$ *and* $t \geq 0$.

(aii) Exhibits strictly decreasing impatience for probability $p \in (0,1]$ *if, and only if,*

$$\delta_r^0 (r+s) < \delta_{r+t}^0 (r+s+t),$$

for all $s > 0$ *and* $t > 0$.

(aiii) Exhibits strictly increasing impatience for probability $p \in (0,1]$ *if, and only if,*

$$\delta_r^0 (r+s) > \delta_{r+t}^0 (r+s+t),$$

for all $s > 0$ *and* $t > 0$.

(bi) Is additive for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+t) \, \delta_{r+t}^0 (r+s+t)}{\delta_r^0 (r+s+t)} = \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s \geq 0$ *and* $t \geq 0$.

(bii) Is strictly subadditive for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+t) \, \delta_{r+t}^0 (r+s+t)}{\delta_r^0 (r+s+t)} < \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s > 0$ *and* $t > 0$.

(biii) Is strictly superadditive for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+t) \, \delta_{r+t}^0 (r+s+t)}{\delta_r^0 (r+s+t)} > \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s > 0$ *and* $t > 0$.

(ci) Is stationary for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+s) \, \delta_r^0 (r+t)}{\delta_r^0 (r+s+t)} = \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s \geq 0$ *and* $t \geq 0$.

(cii) Exhibits the common difference effect for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+s) \, \delta_r^0 (r+t)}{\delta_r^0 (r+s+t)} < \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s > 0$ *and* $t > 0$.

(ciii) Exhibits the converse common difference effect for probability $p \in (0,1]$ *if, and only if,*

$$\frac{\delta_r^0 (r+s) \, \delta_r^0 (r+t)}{\delta_r^0 (r+s+t)} > \frac{w(p) \, w(p\Pi(s+t))}{w(p\Pi(s)) \, w(p\Pi(t))},$$

for all $s > 0$ *and* $t > 0$.

From Theorem 10a, we see that the riskless time discount function, δ^0, determines constant, decreasing and increasing impatience. Also from Theorem 10a, note that if constant impatience (respectively, decreasing impatience and increasing impatience) holds for some probability $p \in (0,1]$, then it holds for all probabilities $p \in (0,1]$.

On the other hand, from Theorem 10b, we see that additivity, subadditivity and superadditivity are jointly determined by the riskless time discount function, δ^0, the probability weighting function, w, and the survival function, Π. Similarly, from Theorem 10c, we see that stationarity, the common difference effect, and its converse, are jointly determined by the riskless time discount function, δ^0, the probability weighting function, w, and the survival function, Π.

Remark 1 *: Consider the special case of certainty, i.e., $p = 1$, $\Pi \equiv 1$ (hence, $\frac{w(p)w(p\Pi(s+t))}{w(p\Pi(s))w(p\Pi(t))} = 1$). Theorem 10bii then gives "A decision maker exhibits strict subadditivity if, and only if, $\delta_r^0(r+t)\,\delta_{r+t}^0(r+s+t) < \delta_r^0(r+s+t)$". This can be rewritten in the following equivalent form "A decision maker exhibits strict subadditivity if, and only if, $\delta_r^0(s)\,\delta_s^0(t) < \delta_r^0(t)$, for all s,t such that $r < s < t$".[24] Now, suppose the restriction $r < s < t$ is removed. Theorem 10bii would then read: "The decision maker exhibits strict subadditivity if, and only if, $\delta_r^0(s)\,\delta_s^0(t) < \delta_r^0(t)$, for all s,t". In particular, for $r = t$, we would get "For a decision maker who exhibits strict subadditivity, $\delta_r^0(s)\,\delta_s^0(r) < \delta_r^0(r) = 1$, for all s". The interpretation of this last statement is that, for $r < s$, discounting a magnitude of 1 from time s back to time r, then compounding the same magnitude forward to time s, would result in a magnitude less than 1. However, this compound operation should leave the magnitude of 1 unchanged (because it is equivalent to "do nothing"), i.e., we would get $1 < 1$. Hence, the restriction $r \leq s$ in the expression $\delta_r^0(s)$ is needed.*

We now show, in Corollary 1 below, that there is an equivalence between the properties of the riskless discount function (Definitions 21, 22, 23) on the one hand, and the underlying properties of the preference relation $\preceq_r$ in Definitions 5 to 10.

Corollary 1 *: Consider the special case of certainty ($p = 1$, $\Pi \equiv 1$). Let the riskless time discount function be δ^0. Let $r \geq 0$. Then the induced preference relationship, $\preceq_r$, exhibits (a) stationarity/the common difference effect/the converse common difference effect (b) constant/decreasing/increasing impatience (c) additivity/subadditivity/superadditivity, according to δ^0 exhibiting any of these properties.*

8 Explaining the common difference effect

From Theorem 2a, section 4, we saw that the common difference effect (Definition 6a) can be explained by the decision maker exhibiting subadditivity and decreasing impatience provided that at least one of these is strict. If preferences are given by a separable time discount function (7.1), then more can be said.

[24] Start with $\delta_r^0(r+t)\,\delta_{r+t}^0(r+s+t) < \delta_r^0(r+s+t)$. Now replace r,s,t with x,y,z, respectively, to get $\delta_x^0(x+z)\,\delta_{x+z}^0(x+y+z) < \delta_x^0(x+y+z)$. Next replace x by r; $x+z$ by s; and $x+y+z$ by t. This gives: $\delta_r^0(s)\,\delta_s^0(t) < \delta_r^0(t)$.

First, consider the case of certainty (no risk, $p = 1$; no hazard, $\Pi \equiv 1$). From Theorem 2a and Corollary 1, if the riskless time discount function δ^0 exhibits subadditivity and decreasing impatience, with at least one of these strict, then the induced preference relation $\preceq_r$ exhibits the common difference effect for every $r \geq 0$. Theorem 11, below, extends this analysis to the case of uncertainty.

Theorem 11 *: Let δ^0 be the riskless time discount function (Definition 20), w the probability weighting functions (Definition 24) and Π the survival function (Definition 26) for the separable time discount function $\delta_r(r+t,p) = w(p\Pi(t))\,\delta_r^0(r+t)$, $p \in (0,1]$. Suppose that δ^0 is subadditive and exhibits decreasing impatience. Then, for each $r \geq 0$, the induced preference relation $\preceq_r$ exhibits the common difference effect for any of the cases (a) to (c) below.*
(a) $p = 1$, w is strictly subadditive, Π is strictly decreasing and exhibits decreasing hazard.
(b) $p = 1$, w is subadditive and Π exhibits strictly decreasing hazard.
(c) w is additive and Π exhibits strictly decreasing hazard.

Discussion of Theorem 11: Under certainty (no risk or hazard) the exponential discount function, Example 4, exhibits stationarity. Hence, the observation of the common difference effect used to be regarded as a refutation of exponential discounting. This led to the development of several time discount functions that can explain the common difference effect, under certainty, in terms of either strict decreasing impatience, strict subadditivity or both. Instances of such time discount functions are given by Examples 5, 7, and 8b.

However, Halevy (2008) pointed out that even when an experimenter offers a reward with probability 1, there is still some chance that the subject will not receive the reward (caused by, for example, death). Experimenters try to reduce, as far as possible, the uncertainties around an experiment. Paradoxically, this increases overweighting of small probabilities. For example, for the Prelec probability weighting function with $\alpha < 1$, we have $\lim\limits_{p \to 0} \frac{w(p)}{p} = \infty$ (Example 9). This could sharpen the rejection of exponential time discounting, even when true.

Halevy (2008), using the Weibull survival function under constant hazard (Definition 27, Example 10 and Lemma 11), weighted by a probability weighting function, produced the common difference effect, even when time discounting is exponential. This is a subcase of our case (a).

Furthermore, the observation of the common difference effect is also consistent with exponential time discounting, even when there is no probability weighting, hence, no certainty effect[25], provided hazard is strictly decreasing. This is a subcase of our case (c).

[25] This is contrary to the view of Halevy, 2008, where it is stated that the presence of the certainty effect is essential for the observation of the common difference effect under exponential discounting. This is because Halevy (2008) assumed constant hazard.

9 Summary

In this paper we formulated a general theory of preferences over outcome-time-probability triplets (sections 2-4).

We defined the *delay*, *defer*, *shift* and *certainty* functions (section 5). Each of these functions can be uniquely elicited from behavior and determine various aspects of preferences over outcome-time-probability triplets. The delay function determines stationarity, the common difference effect and its converse. The defer function determines constant, decreasing and increasing impatience. The shift function determines additivity, subadditivity and superadditivity. The certainty function determines probability independence, the certainty effect and its converse. Thus, these functions can provide parameter-free tests of these properties.

In section 6 we proposed a general discounted utility model that satisfies the axioms of section 3 and which encompasses the main empirically supported discounted utility models. However, it goes beyond these in two respects: (1) it allows discounting back to an arbitrary point in time, r (not just 0), (2) it uses a discount function that is probability dependent (as well as, of course, time dependent). In section 7 we studied the special case where the discount function is a product of a riskless time discount function (a discount function in the ordinary sense) and a term that captures risk, hazard, and probability weighting.

Finally, in section 8 we discussed the various explanations of the common difference effect. We believe that this extends and adds to the known conclusions in this area.

In this paper we kept close to the usual experimental setup, which compares smaller-sooner outcomes (SS) to later-larger outcomes (LL). However, we went beyond this setup by including uncertainty. An obviously interesting extension would be to general lotteries (not just binary lotteries we used in this paper). Another interesting extension would be to streams of outcomes (this is carried out in al-Nowaihi and Dhami, 2008b, but for the case of certainty only).

Acknowledgements

We are grateful to Junaid Arshad, Andrew Colman, Clive Fraser, Itzhak Gilboa, Herbert Gintis, Yingyu Guo, Adriana Iorga, George Loewenstein, Briony Pulford, Larry Samuelson, and Marc Scholten, for valuable comments. All errors are our own.

10 Proofs

We start by establishing a lemma.

Lemma 13 *(The intermediate value theorem): Let C be a connected subset of $\mathbb{R}^m$. Let $F: C \to \mathbb{R}$ be continuous. Let $\mathbf{x}, \mathbf{z} \in C$. Let $y \in \mathbb{R}$ such that $F(\mathbf{x}) \leq y \leq F(\mathbf{z})$. Then, there exists a $\mathbf{y} \in C$ such that $F(\mathbf{y}) = y$.*

Proof of Lemma 13

Since C is connected, there exists a continuous curve, $f : [0,1] \to C$ such that $f(0) = \mathbf{x}$ and $f(1) = \mathbf{z}$. Since $f : [0,1] \to C$ and $F : C \to \mathbb{R}$ are continuous, it follows that $F \circ f : [0,1] \to \mathbb{R}$ is also continuous. We also have $F \circ f(0) = F(f(0)) = F(\mathbf{x})$ and $F \circ f(1) = F(f(1)) = F(\mathbf{z})$. Hence, $F \circ f(0) \leq y \leq F \circ f(1)$. Hence, by the intermediate value theorem for one dimension, there exists an $s \in [0,1]$ such that $F \circ f(s) = y$. Let $\mathbf{y} = f(s)$. Then $F(\mathbf{y}) = F(f(s)) = F \circ f(s) = y$. ■

Proof of Lemma 1

(a) The proof follows from Definition 1 and Axiom 1.

(b) Since $(\mathbf{x}, r+s, p) \sim_r (\mathbf{x}', r+s', p')$ and $(\mathbf{y}, r+t, q) \sim_r (\mathbf{y}', r+t', q')$, it follows that $(\mathbf{x}', r+s', p') \preceq_r (\mathbf{x}, r+s, p)$ and $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{y}', r+t', q')$ (Definition 1a).

(i) Suppose $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$. We thus have $(\mathbf{x}', r+s', p') \preceq_r (\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q) \preceq_r (\mathbf{y}', r+t', q')$. Hence, by transitivity (Axiom 1b), $(\mathbf{x}', r+s', p') \preceq_r (\mathbf{y}', r+t', q')$. The proof of the converse implication is similar.

(ii) Suppose $(\mathbf{x}, r+s, p) \prec_r (\mathbf{y}, r+t, q)$. From Definition 1b, $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$. From part (i) we get $(\mathbf{x}', r+s', p') \preceq_r (\mathbf{y}', r+t', q')$. If $(\mathbf{x}', r+s', p') \sim_r (\mathbf{y}', r+t', q')$, then we would have $(\mathbf{x}, r+s, p) \sim_r (\mathbf{x}', r+s', p') \sim_r (\mathbf{y}', r+t', q') \sim_r (\mathbf{y}, r+t, q)$. By transitivity (part a), $(\mathbf{x}, r+s, p) \sim_r (\mathbf{y}, r+t, q)$; which is not the case (Definition 1b). Hence, $(\mathbf{x}', r+s', p') \prec_r (\mathbf{y}', r+t', q')$ (Definition 1b). The proof of the converse implication is similar. ■

Proof of Theorem 1

Let $p \in (0,1]$. Suppose $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $r \geq 0$, $s \geq 0$, $t \geq 0$, where $\mathbf{w}$ is time-neutral for p and $\preceq_r$ (Definition 2) and $\mathbf{y}$ is either a loss or a gain relative to $\mathbf{w}$ and according to $\preceq_r$ (Definition 4). Assume $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s, p)$. By constant impatience (Definition 7), $(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)$. From Axiom 3 (existence of present values) it follows that there exists an $\mathbf{x} \in \mathbb{R}^m$ such that $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$. Since $\mathbf{y}$ is either a loss or a gain relative to $\mathbf{w}$ and according to $\preceq_r$, the same holds for $\mathbf{x}$. Hence, we have: $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$ and $(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)$. Using additivity (Definition 9) we get $(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t, p)$. Recalling that $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$, we get $(\mathbf{y}, r+t, p) \sim_r (\mathbf{z}, r+s+t, p)$. Thus, $\preceq_r$ is stationary (Definition 5). ■

Proof of Theorem 2

(a) Let $p \in (0,1]$. Suppose $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, $r \geq 0$, $s > 0$, $t > 0$. Assume

$$(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s, p). \tag{10.1}$$

From Axiom 3 (existence of present values) it follows that there exists an $\mathbf{x} \in \mathbb{R}^m$ such that

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p). \tag{10.2}$$

We concentrate on the proof for the case of gains (Definition 4i), then indicate how the proof can be modified for the case of losses.

Let $\mathbf{w}$ be time-neutral, given $\preceq_r$. For gains we have:

$$(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p). \tag{10.3}$$

From (10.3), Axiom 4i (Consistency of gains) gives

$$(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r+t, p)\,. \tag{10.4}$$

From (10.2) and (10.4),

$$(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r, p)\,. \tag{10.5}$$

From (10.1) and (10.3), we get

$$(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r+s, p)\,. \tag{10.6}$$

From (10.6) and Axiom 4i (Consistency of gains), we get

$$(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r, p)\,. \tag{10.7}$$

Part 1: Additivity and strictly decreasing impatience

From (10.1) and (10.3), strict decreasing impatience for gains (Definition 8ai) gives

$$(\mathbf{y}, r+t, p) \prec_{r+t} (\mathbf{z}, r+s+t, p)\,. \tag{10.8}$$

From (10.4) and (10.8), Axiom 6i (time sensitivity for gains) gives

$$(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t+T, p)\,, \text{ for some } T>0. \tag{10.9}$$

From (10.2) and (10.9), additivity (Definition 9) gives

$$(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t+T, p)\,. \tag{10.10}$$

From (10.2) and (10.10), we get

$$(\mathbf{y}, r+t, p) \sim_r (\mathbf{z}, r+s+t+T, p)\,. \tag{10.11}$$

From (10.7) and Axiom 5i (time monotonicity for gains), with $\sigma = s+t$, $\omega = 0$ and $\tau = T$, we get

$$(\mathbf{z}, r+s+t+T, p) \prec_r (\mathbf{z}, r+s+t, p)\,. \tag{10.12}$$

From (10.11) and (10.12), we get

$$(\mathbf{y}, r+t, p) \prec_r (\mathbf{z}, r+s+t, p)\,. \tag{10.13}$$

Hence, the common difference effect holds for gains (Definition 6ai).

The proof for losses, $(\mathbf{w}, r, p) \succ_r (\mathbf{y}, r, p)$, is similar but uses Axioms 4ii, 5ii and 6ii and Definitions 6aii and 8aii.

Part 2: Strict subadditivity and constant impatience

From (10.1), constant impatience (Definition 7) gives

$$(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)\,. \tag{10.14}$$

Applying strict subadditivity for gains (Definition 10ai) to (10.2) and (10.14), we get

$$(\mathbf{x},r,p)\prec_r(\mathbf{z},r+s+t,p)\,. \tag{10.15}$$

From (10.2) and (10.15), we get $(\mathbf{y},r+t,p)\prec_r(\mathbf{z},r+s+t,p)$. Hence, the common difference effect holds for gains (Definition 6ai).

The proof for losses, $(\mathbf{w},r,p)\succ_r(\mathbf{y},r,p)$, is similar, except that we use Definitions 6aii and 10aii.

Part 3: Strict subadditivity and strictly decreasing impatience

The proof is similar to that of Part 1. However, instead of using additivity (Definition 9) to derive (10.10), we use strict subadditivity (Definition 10ai) to derive:

$$(\mathbf{x},r,p)\prec_r(\mathbf{z},r+s+t+T,p)\,. \tag{10.16}$$

The proof for losses, $(\mathbf{w},r,p)\succ_r(\mathbf{y},r,p)$, is similar but uses Axioms 4ii, 5ii and 6ii and Definitions 6aii, 8aii and 10aii.

(b) The proof is similar to that for part a but uses Definitions 6b, 8b and 10b. ■

Proof of Lemma 2

Let $p\in(0,1]$.

(a) Let $\mathbf{y}$ be a gain, i.e., $(\mathbf{w},r,p)\prec_r(\mathbf{y},r,p)$. Assume $(\mathbf{y},r,p)\sim_r(\mathbf{z},r+s,p)$. Axiom 6i (time sensitivity for gains) gives $(\mathbf{y},r+t,p)\sim_r(\mathbf{z},r+s+T,p)$, for some $T\geq 0$. Set $D(r,s,t)=T$ to get $(\mathbf{y},r,p)\sim_r(\mathbf{z},r+s,p)\Rightarrow(\mathbf{y},r+t,p)\sim_r(\mathbf{z},r+s+D(r,s,t),p)$. The proof for losses is similar, except that we use Axiom 6ii.

(b) Let D_1 and D_2 be two delay functions. Assume $(\mathbf{y},r,p)\sim_r(\mathbf{z},r+s,p)$. Then, from the definition of delay functions (Definition 12),
$(\mathbf{y},r+t,p)\sim_r(\mathbf{z},r+s+D_1(r,s,t),p)$ and
$(\mathbf{y},r+t,p)\sim_r(\mathbf{z},r+s+D_2(r,s,t),p)$. Hence,
$(\mathbf{z},r+s+D_1(r,s,t),p)\sim_r(\mathbf{z},r+s+D_2(r,s,t),p)$. Suppose $D_1(r,s,t)\neq D_2(r,s,t)$. Without loss of generality, assume that $D_1(r,s,t)<D_2(r,s,t)$. Suppose that $\mathbf{y}$ is a gain, i.e., $(\mathbf{w},r,p)\prec_r(\mathbf{y},r,p)$. Hence, also, $(\mathbf{w},r,p)\prec_r(\mathbf{z},r+s,p)$. Axiom 4i (Consistency of gains) then gives $(\mathbf{w},r,p)\prec_r(\mathbf{z},r,p)$. Hence, Axiom 5i (time monotonicity for gains), with $\omega=0$, $\sigma=s+D_1(r,s,t)$ and $\tau=D_2(r,s,t)-D_1(r,s,t)>0$, gives $(\mathbf{z},r+s+D_2(r,s,t),p)\prec_r(\mathbf{z},r+s+D_1(r,s,t),p)$, which is not the case. Hence, $D_1(r,s,t)=D_2(r,s,t)$.

Suppose that $\mathbf{y}$ is a loss. A similar argument to the one above, but using Axioms 4ii and 5ii shows that $(\mathbf{z},r+s+D_2(r,s,t),p)\succ_r(\mathbf{z},r+s+D_1(r,s,t),p)$, which is not the case. Hence, again, $D_1(r,s,t)=D_2(r,s,t)$. ■

Proof of Theorem 3

Let $p\in(0,1]$. Let D be the delay function corresponding to the preference relation $\preceq_r$. Let $r,s,t\in[0,\infty)$ and $\mathbf{w},\mathbf{y},\mathbf{z}\in\mathbb{R}^m$, where $\mathbf{w}$ is time-neutral and $\mathbf{y}$ is either a gain or a loss, i.e., $(\mathbf{w},r,p)\prec_r(\mathbf{y},r,p)$ or $(\mathbf{w},r,p)\succ_r(\mathbf{y},r,p)$. Assume that

$$(\mathbf{y},r,p)\sim_r(\mathbf{z},r+s,p)\,. \tag{10.17}$$

Hence, from the definition of a delay function (Definition 12) we get

$$(\mathbf{y}, r+t, p) \sim_r (\mathbf{z}, r+s+D(r,s,t), p). \tag{10.18}$$

(a) If $D(r,s,t) = t$ then, from (10.17), (10.18) and Definition 5, we see that $\preceq_r$ is stationary.

Conversely, if $\preceq_r$ is stationary, so that $(\mathbf{y}, r+t, p) \sim_r (\mathbf{z}, r+s+t, p)$, then, from the uniqueness of D (Lemma 2b), we get that $D(r,s,t) = t$.

(b) First, consider gains, i.e., $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p)$. Hence, also, $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r+s, p)$ (Axiom 4i, Consistency of gains). From (10.17), we get $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r+s, p)$. Axiom 4i (Consistency of gains) then gives $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r, p)$. Suppose $D(r,s,t) > t$. Axiom 5i (time monotonicity for gains), with $\omega = 0$, $\sigma = s+t$ and $\tau = D(r,s,t) - t > 0$, then gives $(\mathbf{z}, r+s+D(r,s,t), p) \prec_r (\mathbf{z}, r+s+t, p)$. Hence, from (10.18), $(\mathbf{y}, r+t, p) \prec_r (\mathbf{z}, r+s+t, p)$. Thus, the common difference effect holds for gains (Definition 6ai).

Conversely, suppose the common difference effect holds for gains (Definition 6ai). Then
$(\mathbf{y}, r+t, p) \prec_r (\mathbf{z}, r+s+t, p)$. Using (10.18) we then get
$(\mathbf{z}, r+s+D(r,s,t), p) \prec_r (\mathbf{z}, r+s+t, p)$. If $D(r,s,t) = t$,
then $(\mathbf{z}, r+s+t, p) \prec_r (\mathbf{z}, r+s+t, p)$, which cannot be. Suppose $D(r,s,t) < t$. Axiom 5i (time monotonicity for gains), with $\omega = 0$, $\sigma = s+D(r,s,t)$ and $\tau = t - D(r,s,t) > 0$, then gives $(\mathbf{z}, r+s+t, p) \prec_r (\mathbf{z}, r+s+D(r,s,t), p)$, which is not the case. Hence, $D(r,s,t) > t$.

The proof for losses is similar, except that we use Axiom 5ii and Definition 6aii.

(c) The proof is similar to that of part (b), except that the inequalities $>$ and $\prec_r$ are reversed. ■

Proof of Lemma 3

Let $p \in (0,1]$, $r, s, t \in [0, \infty)$.

Let Δ_1 and Δ_2 be two defer functions. Assume $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s, p)$. Then, from the definition of defer functions (Definition 13), $(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+\Delta_1(r,s,t)+t, p)$ and $(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+\Delta_2(r,s,t)+t, p)$.
Hence, $(\mathbf{z}, r+\Delta_1(r,s,t)+t, p) \sim_{r+t} (\mathbf{z}, r+\Delta_2(r,s,t)+t, p)$. Suppose $\Delta_1(r,s,t) \neq \Delta_2(r,s,t)$. Without loss of generality, assume that $\Delta_1(r,s,t) < \Delta_2(r,s,t)$.

Suppose that $\mathbf{y}$ is a gain, i.e., $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p)$. Hence, also, $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r+s, p)$. Axiom 4i (Consistency of gains) then gives $(\mathbf{w}, r, p) \prec_r (\mathbf{z}, r, p)$. Hence, Axiom 5i (time monotonicity for gains), with $\omega = t$, $\sigma = \Delta_1(r,s,t) \geq 0$ and $\tau = \Delta_2(r,s,t) - \Delta_1(r,s,t) > 0$, gives $(\mathbf{z}, r+\Delta_2(r,s,t)+t, p) \prec_{r+t} (\mathbf{z}, r+\Delta_1(r,s,t)+t, p)$, which is not the case. Hence, $\Delta_1(r,s,t) = \Delta_2(r,s,t)$.

Suppose that $\mathbf{y}$ is a loss, i.e., $(\mathbf{w}, r, p) \succ_r (\mathbf{y}, r, p)$. A similar argument to the one above, but using Axiom 5ii, shows that $(\mathbf{z}, r+\Delta_2(r,s,t)+t, p) \succ_{r+t} (\mathbf{z}, r+\Delta_1(r,s,t)+t, p)$, which is not the case. Hence, again, $\Delta_1(r,s,t) = \Delta_2(r,s,t)$. ■

Proof of Theorem 4

(a) Suppose $\preceq_r$ exhibits constant impatience (Definition 7).

Let $p \in (0,1]$, $r,s,t \in [0,\infty)$. Let $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$ where $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and $\mathbf{y}$ is either a gain or a loss (Definitions 2 and 4). Suppose that $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$ By constant impatience, we get $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t,p)$. Set $\Delta(r,s,t) = s$.

Conversely, suppose a defer function, $\Delta(r,s,t)$, exists and satisfies $\Delta(r,s,t) = s$, for all $s \geq 0$ and $t \geq 0$. Let $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$. From the definition of a defer function (Definition 13), we get $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+\Delta(r,s,t)+t,p)$ and, hence, $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t,p)$. Hence, $\preceq_r$ exhibits constant impatience.

(b) Suppose $\preceq_r$ exhibits strictly decreasing impatience (Definition 8a).

Let $p \in (0,1]$, $r \in [0,\infty)$, $s,t \in (0,\infty)$. Let $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$ where $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and $\mathbf{y}$ is a gain (Definitions 2 and 4i). Suppose that $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$ By strictly decreasing impatience for gains (Definition 8ai), we get $(\mathbf{y},r+t,p) \prec_{r+t} (\mathbf{z},r+s+t,p)$. By time sensitivity for gains (Axiom 6i), we get $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t+T,p)$ for some $T > 0$. Set $\Delta(r,s,t) = s+T > s$. Then $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p) \Rightarrow (\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+\Delta(r,s,t)+t,p)$ and $\Delta(r,s,t) > s$. The case if $\mathbf{y}$ is a loss is similar, except that we use Definitions 4ii and 8aii and Axiom 6ii.

Conversely, suppose a defer function, $\Delta(r,s,t)$, exists and satisfies $\Delta(r,s,t) > s$, for all $s > 0$ and $t > 0$. Let $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$. From the definition of a defer function (Definition 13), we get $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+\Delta(r,s,t)+t,p)$. Suppose $\mathbf{y}$ is a gain (Definition 4i). Hence, $\mathbf{z}$ is also a gain. Since $\Delta(r,s,t) > s$ we get, from time monotonicity for gains (Axiom 5i), $(\mathbf{y},r+t,p) \prec_{r+t} (\mathbf{z},r+s+t,p)$. Hence, $\preceq_r$ exhibits strictly decreasing impatience for gains (Definition 8ai). The case if $\mathbf{y}$ is a loss is similar, except that we use Definitions 4ii and 8aii and Axiom 5ii.

(c) The proof for part c is similar to that of part b except that we use Definition 8b.

■

Proof of Lemma 4

Let $p \in (0,1]$, $r,s,t \in [0,\infty)$. Consider the preference relation $\preceq_r$. Let $\mathbf{w},\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$ where $\mathbf{x}$ is either a gain or a loss according to $\preceq_r$, i.e., $(\mathbf{w},r,p) \prec_r (\mathbf{x},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{x},r,p)$. Suppose that

$$(\mathbf{x},r,p) \sim_r (\mathbf{y},r+t,p), \tag{10.19}$$

$$(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t,p). \tag{10.20}$$

Let S_1 and S_2 be two shift functions. From (10.19), (10.20) and the definition of shift functions (Definition 14), we get $(\mathbf{x},r,p) \sim_r (\mathbf{z},r+S_1(r,s,t),p)$ and $(\mathbf{x},r,p) \sim_r (\mathbf{z},r+S_2(r,s,t),p)$. Hence, $(\mathbf{z},r+S_1(r,s,t),p) \sim_r (\mathbf{z},r+S_2(r,s,t),p)$. Suppose $S_1(r,s,t) \neq S_2(r,s,t)$. Without loss of generality, assume that $S_1(r,s,t) < S_2(r,s,t)$.

Suppose that $\mathbf{x}$ is a gain, i.e., $(\mathbf{w},r,p) \prec_r (\mathbf{x},r,p)$. Hence, also, $(\mathbf{w},r,p) \prec_r (\mathbf{z},r+S_2(r,s,t),p)$. Axiom 4i (Consistency of gains) then gives $(\mathbf{w},r,p) \prec_r (\mathbf{z},r,p)$. Hence, Axiom 5i (time monotonicity for gains), with $\omega = 0$, $\sigma = S_1(r,s,t)$ and $\tau = S_2(r,s,t) - S_1(r,s,t) > 0$, gives $(\mathbf{z},r+S_2(r,s,t),p) \prec_r (\mathbf{z},r+S_1(r,s,t),p)$, which is not the case. Hence, $S_1(r,s,t) = S_2(r,s,t)$.

Suppose that $\mathbf{x}$ is a loss. A similar argument to the one above, but using Axioms 4ii and 5ii, shows that $(\mathbf{z}, r+S_2(r,s,t), p) \succ_r (\mathbf{z}, r+S_1(r,s,t), p)$, which is not the case. Hence, again, $S_1(r,s,t) = S_2(r,s,t)$. ■

Proof of Theorem 5

Let $p \in (0,1]$, $r.s.t \in [0,\infty)$. Consider the preference relation $\preceq_r$. Let $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ where $\mathbf{x}$ is either a gain or a loss according to $\preceq_r$, i.e., $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r, p)$ or $(\mathbf{w}, r, p) \succ_r (\mathbf{x}, r, p)$. Suppose that

$$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p), \tag{10.21}$$

$$(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p). \tag{10.22}$$

(a) Suppose $\preceq_r$ is additive (Definition 9). From (10.21), (10.22) we then get $(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t, p)$. Set $S(r,s,t) = s+t$.

Conversely, suppose a shift function exists and satisfies $S(r,s,t) = s+t$. From (10.21), (10.22) and the definition of a shift function (Definition 14) we get $(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t, p)$. Therefore, $\preceq_r$ is additive.

(b) Suppose $\preceq_r$ is strictly subadditive for gains (Definition 10ai). From (10.21), (10.22) we then get $(\mathbf{x}, r, p) \prec_r (\mathbf{z}, r+s+t, p)$. Axiom 6i (time sensitivity for gains) then gives $(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+s+t+T, p)$ for some $T > 0$. Set $S(r,s,t) = s+t+T > s+t$.

Conversely, suppose a shift function, $S(r,s,t)$, exists and satisfies $S(r,s,t) > s+t$. Axiom 5i (time monotonicity for gains), with $\omega = 0$, $\tau = S(r,s,t) - s - t > 0$ and $\sigma = s+t$, then gives $(\mathbf{z}, r+S(r,s,t), p) \prec_r (\mathbf{z}, r+s+t, p)$. From (10.21), (10.22) and the definition of a shift function (Definition 14) we get $(\mathbf{x}, r, p) \sim_r (\mathbf{z}, r+S(r,s,t), p) \prec_r (\mathbf{z}, r+s+t, p)$. Hence we have
$(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$, $(\mathbf{y}, r+t, p) \sim_{r+t} (\mathbf{z}, r+s+t, p)$ and $(\mathbf{x}, r, p) \prec_r (\mathbf{z}, r+s+t, p)$. Hence, $\preceq_r$ is strictly subadditive for gains (Definition 10ai).

The proof for losses is similar, except that we use Axioms 4ii and 5ii and Definition 10aii.

(c) The proof is similar to that of part b, except that the inequalities $>$ and $\prec_{r+t}$ are reversed. ■

Proof of Lemma 5

Let $r \in [0,\infty)$, $\mathbf{w} \in \mathbb{R}^m$ is time-neutral for the preference relation $\preceq_r$ and both probabilities $p, q \in (0,1]$, $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ are either both gains or both losses, for $\preceq_r, \mathbf{w}, p, q$. Let $s \in [0,\infty)$, $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$.

(a) By Axiom 7 (probability sensitivity), there exists a $T \in [0,\infty)$ such that $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+T, p)$. Set $C(s) = T$.

(b) Consider the case of gains. Assume that $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+T_1, p)$ and $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+T_2, p)$, where $T_1 \geq 0$ and $T_2 \geq 0$. Hence, $(\mathbf{z}, r+T_1, p) \sim_r (\mathbf{z}, r+T_2, p)$. Without loss of generality, assume that $T_1 \leq T_2$. Let $\omega = 0$, $\sigma = T_1$ and $\tau = T_2 - T_1 \geq 0$. If $\tau > 0$ then, from Axiom 5i (time monotonicity for gains), we would get $(\mathbf{z}, r+\sigma+\tau, p) \prec_r (\mathbf{z}, r+\sigma, p)$, i.e., $(\mathbf{z}, r+T_2, p) \prec_r (\mathbf{z}, r+T_1, p)$, which is not the case. Hence, $\tau = 0$ and, hence, $T_1 = T_2$. Hence, $C(s)$ is unique.

The case of losses is similar. ■

Proof of Theorem 6

(a) Let $r \in [0,\infty)$. Suppose that $\preceq_r$ exhibits probability independence. Let $p,q \in (0,1]$, $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{w}$ is time-neutral for $\preceq_r, p, q$ and $\mathbf{y},\mathbf{z}$ are either both gains or both losses, for $\preceq_r, \mathbf{w}, p, q$. Let $s \geq 0$ and assume that $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$. From Definitions 11a and 15 we get that $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s, p)$ and $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+C(s), p)$. From the uniqueness of the certainty function (Lemma 5b), we get that $C(s) = s$.

Conversely, suppose that probability independence (Definition 11a) does not hold. Hence, for some $s \geq 0$, $p,q \in (0,1]$, $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{w}$ is time-neutral for $\preceq_r, p, q$ and $\mathbf{y},\mathbf{z}$ are either both gains or both losses, for $\preceq_r, \mathbf{w}, p, q$, we have $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$ and either $(\mathbf{y}, r, p) \prec_r (\mathbf{z}, r+s, p)$ or $(\mathbf{y}, r, p) \succ_r (\mathbf{z}, r+s, p)$. From Definition 15 we get $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+C(s), p)$. Hence, $(\mathbf{z}, r+C(s), p) \prec_r (\mathbf{z}, r+s, p)$ or $(\mathbf{z}, r+C(s), p) \succ_r (\mathbf{z}, r+s, p)$. If $C(s) = s$, then we would get $(\mathbf{z}, r+s, p) \prec_r (\mathbf{z}, r+s, p)$ or $(\mathbf{z}, r+s, p) \succ_r (\mathbf{z}, r+s, p)$. But neither can be the case. Hence, $C(s) \neq s$.

(b) Let $r \geq 0$, $p,q \in (0,1]$, $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $p > q$, $\mathbf{w}$ is time-neutral for $\preceq_r, p, q$ and $\mathbf{y},\mathbf{z}$ are either both gains or both losses with respect to $\preceq_r, p, q, \mathbf{w}$. Let $s > 0$. Suppose $C(s) < s$. Start with gains. Suppose $(\mathbf{w}, r, q) \prec_r (\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$. Definition 15 then gives $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+C(s), p)$. Since $C(s) < s$, we get $C(s) + \tau = s$, $\tau > 0$. Hence, $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+s-\tau, p)$. But, by Axiom 5i (with $\omega = 0$ and $\sigma = s - \tau$), $(\mathbf{z}, r+s, p) \prec_r (\mathbf{z}, r+s-\tau, p)$. Hence, $(\mathbf{z}, r+s, p) \prec_r (\mathbf{z}, r+C(s), p)$ and, hence, $(\mathbf{z}, r+s, p) \prec_r (\mathbf{y}, r, p)$. Thus, the certainty effect holds (Definition 11bi). The case of losses is similar, except that we use Axiom 5ii and Definition 11bii.

Conversely, suppose that the certainty effect holds.

Let $r \geq 0$, $p,q \in (0,1]$, $\mathbf{w},\mathbf{y},\mathbf{z} \in \mathbb{R}^m$, where $p > q$, $\mathbf{w}$ is time-neutral for $\preceq_r, p, q$ and $\mathbf{y},\mathbf{z}$ are both either gains or both losses with respect to $\preceq_r, p, q, \mathbf{w}$. First, consider gains. Let $s > 0$ and assume that $(\mathbf{w}, r, q) \prec_r (\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+s, q)$. Definition 11bi then gives $(\mathbf{y}, r, p) \succ_r (\mathbf{z}, r+s, p)$ and Definition 15 gives $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+C(s), p)$. Hence, $(\mathbf{z}, r+C(s), p) \succ_r (\mathbf{z}, r+s, p)$. Suppose that $C(s) \geq s$. If $C(s) = s$, then $(\mathbf{z}, r+C(s), p) \sim_r (\mathbf{z}, r+s, p)$, which is not the case. Hence, $C(s) > s$ and, hence, $C(s) = s + \tau$ for some $\tau > 0$. It follows that $(\mathbf{z}, r+s+\tau, p) \succ_r (\mathbf{z}, r+s, p)$, which contradicts Axiom 5i (with $\omega = 0$ and $\sigma = s$). Hence, $C(s) < s$. The case of losses is similar, except that we use Axiom 5ii.

(c) The proof is similar to that of part (b), except that we use Definition 11c. ■

Proof of Lemma 6

Let $p \in (0,1]$, $s \geq 0$, and assume that δ is continuous in t. Let $\gamma \in (0, \delta_r(r+s,p)]$. Since $\lim_{t\to\infty} \delta_r(r+s+t,p) = 0$ (follows from Definition 17d), we get $\delta_r(r+s+t_1,p) < \gamma$ for some $t_1 \in [0,\infty)$. Hence, $\delta_r(r+s+t_1,p) < \gamma \leq \delta_r(r+s,p)$. Since δ is continuous in t it follows, from the intermediate value theorem, that $\delta_r(r+s+\tau,p) = \gamma$, for some $\tau \in [0,\infty)$. Hence, $\tau \mapsto \delta_r(r+s+\tau,p)$ maps $[0,\infty)$ onto $(0, \delta_r(r+s,p)]$. ■

Proof of Lemma 7

Let $p > 0$. Then $\delta_r(r+t,p) > 0$ (Definition 17a).

(a) Suppose that $u(\mathbf{w}) = 0$. Then $U_r(\mathbf{w}, r+t, p) = \delta_r(r+t, p)\, u(\mathbf{w}) = 0$, for all $t \geq 0$, i.e., $(\mathbf{w}, r+t, p) \sim_r (\mathbf{w}, r, p)$, for all $t \geq 0$. Hence, $\mathbf{w}$ is time-neutral.

Conversely, suppose $u(\mathbf{w}) > 0$. Then $U_r(\mathbf{w}, r+t, p) = \delta_r(r+t, p)\, u(\mathbf{w})$ is strictly decreasing in t (Definition 17c), contradicting the time-neutrality of $\mathbf{w}$ (Definition 2). Analogously, we can argue that $u(\mathbf{w}) < 0$ cannot be the case. Hence, $u(\mathbf{w}) = 0$.

(b) Suppose $(\mathbf{x}, r+t, p)$ is a gain. Hence $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r+t, p)$, i.e., $\delta_r(r, p)\, u(\mathbf{w}) < \delta_r(r+t, p)\, u(\mathbf{x})$. From part (a), it follows that $0 < \delta_r(r+t, p)\, u(\mathbf{x})$ and, hence, $u(\mathbf{x}) > 0$. Conversely, if $u(\mathbf{x}) > 0$ then $\delta_r(r+t, p)\, u(\mathbf{x}) > 0 = \delta_r(r, p)\, u(\mathbf{w})$, using part (a). Hence, $(\mathbf{w}, r, p) \prec_r (\mathbf{x}, r+t, p)$, i.e., $(\mathbf{x}, r+t, p)$ is a gain.

The proof in case (c) is similar to case (b). ■

Proof of Theorem 7

Axiom 1 (Order) holds

Let $r \in [0, \infty)$. Let $\preceq_r$ be the preference relation induced by U_r (Definition 19). Given $s, t \in [0, \infty)$, $p, q \in [0, 1]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, we have either $U_r(\mathbf{x}, r+s, p) \leq U_r(\mathbf{y}, r+t, q)$ or $U_r(\mathbf{y}, r+t, q) \leq U_r(\mathbf{x}, r+s, p)$. Hence, either $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ or $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{x}, r+s, p)$. Thus, completeness holds.

Given $s, t, t' \in [0, \infty)$ and $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$, assume $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{y}, r+t, q)$ and $(\mathbf{y}, r+t, q) \preceq_r (\mathbf{z}, r+t', q')$. Hence, $U_r(\mathbf{x}, r+s, p) \leq U_r(\mathbf{y}, r+t, q)$ and $U_r(\mathbf{y}, r+t, q) \leq U_r(\mathbf{z}, r+t', q')$. Consequently, $U_r(\mathbf{x}, r+s, p) \leq U_r(\mathbf{z}, r+t', q')$. Hence, $(\mathbf{x}, r+s, p) \preceq_r (\mathbf{z}, r+t', q')$. Thus, transitivity holds.

Axiom 2 (Existence of time-neutral outcomes) holds

By Definition 16, $u(\mathbf{0}) = 0$. Hence, by Lemma 7a, $\mathbf{0}$ is time-neutral.

Axiom 3 (Existence of present values) holds

Let $r \in [0, \infty)$. Let $\preceq_r$ be the preference relation induced by U_r (Definition 19). Let $\mathbf{y} \in \mathbb{R}^m_+$, $t \in [0, \infty)$, $p \in (0, 1]$. We want to show that there exists an $\mathbf{x} \in \mathbb{R}^m_+$ such that $(\mathbf{x}, r, p) \sim_r (\mathbf{y}, r+t, p)$, i.e., $\delta_r(r, p)\, u(\mathbf{x}) = \delta_r(r+t, p)\, u(\mathbf{y})$.

Let $\mathbf{w} \in \mathbb{R}^m_+$ be time-neutral (Definition 2). Then, $u(\mathbf{w}) = 0$ (Lemma 7a). Since $p \in (0, 1]$, we get $\delta_r(r, p), \delta_r(r+t, p) \in (0, 1]$ and $0 < \frac{\delta_r(r+t,p)}{\delta_r(r,p)} \leq 1$ (Definition 17).

Suppose $u(\mathbf{y}) \geq 0$. Hence,

$$u(\mathbf{w}) = 0 \leq \frac{\delta_r(r+t, p)}{\delta_r(r, p)} u(\mathbf{y}) \leq u(\mathbf{y}). \tag{10.23}$$

From (10.23), the continuity of u, and the intermediate value theorem (Lemma 13), it follows that there exists an $\mathbf{x} \in \mathbb{R}^m_+$ such that $u(\mathbf{x}) = \frac{\delta_r(r+t,p)}{\delta_r(r,p)} u(\mathbf{y})$. Thus, we have shown that there exists an $\mathbf{x} \in \mathbb{R}^m_+$ such that $\delta_r(r, p)\, u(\mathbf{x}) = \delta_r(r+t, p)\, u(\mathbf{y})$.

The case $u(\mathbf{y}) \leq 0$ is similar. ■

Axiom 4 (Consistency of gains and losses) holds

This is an immediate consequence of Lemma 7.

Axiom 5 (Time monotonicity) holds

Let $r \in [0, \infty)$. Let $\preceq_r$ be the preference relation induced by U_r (Definition 19). Let $\mathbf{w}, \mathbf{z} \in \mathbb{R}^m$, $p \in (0, 1]$, where $\mathbf{w}$ is time-neutral for the preference relation $\preceq_r$ and probability p.

(i) Gains. From Lemma 7b, we get $u(\mathbf{z}) > 0$. Let $\omega \geq 0$, $\sigma \geq 0$ and $\tau > 0$. Then, $U_{r+\omega}(\mathbf{z}, r+\omega+\sigma+\tau, p) = \delta_{r+\omega}(r+\omega+\sigma+\tau, p)\, u(\mathbf{z})$ and $U_{r+\omega}(\mathbf{z}, r+\omega+\sigma, p) = \delta_{r+\omega}(r+\omega+\sigma, p)\, u(\mathbf{z})$. Since and $\delta_{r+\omega}(t,p) > 0$ is strictly decreasing in t (Definition 17c) and $u(\mathbf{z}) > 0$, we get $U_{r+\omega}(\mathbf{z}, r+\omega+\sigma+\tau, p) < U_{r+\omega}(\mathbf{z}, r+\omega+\sigma, p)$. Hence, $(\mathbf{z}, r+\omega+\sigma+\tau, p) \prec_{r+\omega} (\mathbf{z}, r+\omega+\sigma, p)$. Hence, Axiom 5i (time monotonicity for gains) holds.

(ii) Losses. A similar argument shows that Axiom 5ii (time monotonicity for losses) also holds.

Axiom 6 (Time sensitivity) holds

Let $r, s, t \in [0, \infty)$, $\mathbf{y}, \mathbf{z} \in \mathbb{R}^m$. Let $\mathbf{w} \in \mathbb{R}^m$ be time-neutral for the preference relation $\preceq_r$ and the probability $p \in (0, 1]$.

(i) Gains: Assume $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p) \preceq_r (\mathbf{z}, r+s, p)$.

From Axiom 5i (time monotonicity for gains) with $\sigma = \omega = 0$, $\tau = t$ and $\mathbf{y}$ instead of $\mathbf{z}$, we get $(\mathbf{y}, r+t, p) \preceq_r (\mathbf{y}, r, p)$. Hence, $(\mathbf{y}, r+t, p) \preceq_r (\mathbf{z}, r+s, p)$. Using Definitions 18 and 19, we then get $\delta_r(r+t, p)\, u(\mathbf{y}) \leq \delta_r(r+s, p)\, u(\mathbf{z})$. From Lemma 7b, we get $u(\mathbf{y}) > 0$ and, hence, also $u(\mathbf{z}) > 0$. It follows that $0 < \delta_r(r+t, p) \frac{u(\mathbf{y})}{u(\mathbf{z})} \leq \delta_r(r+s, p)$. From Lemma 6, it follows that $\delta_r(r+t, p) \frac{u(\mathbf{y})}{u(\mathbf{z})} = \delta_r(r+s+T, p)$, for some $T \geq 0$. Hence, $\delta_r(r+t, p)\, u(\mathbf{y}) = \delta_r(r+s+T, p)\, u(\mathbf{z})$, for some $T \geq 0$. If $(\mathbf{y}, r, p) \prec_r (\mathbf{z}, r+s, p)$ then, also, $(\mathbf{y}, r+t, p) \prec_r (\mathbf{z}, r+s, p)$ and, hence, $\delta_r(r+t, p)\, u(\mathbf{y}) < \delta_r(r+s, p)\, u(\mathbf{z})$ It follows that, in this case, $T > 0$.

(ii) The proof for losses is similar, except that we use Axiom 5ii (time monotonicity for losses) and Lemma 7c. ■

Axiom 7 (Probability sensitivity) holds

Let $r \in [0, \infty)$. Let $\mathbf{w} \in \mathbb{R}^m$ be time-neutral for $\preceq_r, p, q \in (0, 1]$ (Definition 2). Hence, $u(\mathbf{w}) = 0$ (Lemma 7a). Let $\mathbf{y}, \mathbf{z}$ be either both gains or both losses relative to $\preceq_r, p, q, \preceq_r$(Definition 4). Hence, either $u(\mathbf{y}) > 0$ and $u(\mathbf{z}) > 0$ or $u(\mathbf{y}) < 0$ and $u(\mathbf{z}) < 0$ (Lemma 7b,c).

Suppose $(\mathbf{y}, r, q) \sim_r (\mathbf{z}, r+t, q)$, $t \geq 0$. Hence, $U_r(\mathbf{y}, r, q) = U_r(\mathbf{z}, r+t, q)$. Hence, $\delta_r(r, q)\, u(\mathbf{y}) = \delta_r(r+t, q)\, u(\mathbf{z})$. Consider the case when $\mathbf{y}$ and $\mathbf{z}$ are both gains. Then, $u(\mathbf{y}) > 0$ and $u(\mathbf{z}) > 0$. Definition 17c gives $0 < \delta_r(r+t, q)\, u(\mathbf{z}) \leq \delta_r(r, q)\, u(\mathbf{z})$. Hence, $0 < \delta_r(r, q)\, u(\mathbf{y}) \leq \delta_r(r, q)\, u(\mathbf{z})$. Therefore, $u(\mathbf{y}) \leq u(\mathbf{z})$. Hence, $0 < \delta_r(r, p)\, u(\mathbf{y}) \leq \delta_r(r, p)\, u(\mathbf{z})$. Therefore, $0 < \delta_r(r, p) \frac{u(\mathbf{y})}{u(\mathbf{z})} \leq \delta_r(r, p)$. From Lemma 6 (with $s = 0$), it follows that, for some $T \geq 0$, $\delta_r(r, p) \frac{u(\mathbf{y})}{u(\mathbf{z})} = \delta_r(r+T, p)$. Hence, for some $T \geq 0$, $\delta_r(r, p)\, u(\mathbf{y}) = \delta_r(r+T, p)\, u(\mathbf{z})$, i.e., for some $T \geq 0$, $(\mathbf{y}, r, p) \sim_r (\mathbf{z}, r+T, p)$. The case when $\mathbf{y}$ and $\mathbf{z}$ are losses is similar. ■

Proof of Lemma 8

Let $p, q \in (0, 1]$. Let $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$ and $r, s, t \in [0, \infty)$. Let $\preceq_r$ and $\sim_{r+t}$ be the preference relations induced by U_r and U_{r+t}, respectively, (Definitions 18 and 19), and let $\mathbf{w}$ be time-neutral for the preference relation $\preceq_r$ and the probability p (Definition 2).

(a) Let $(\mathbf{w}, r, p) \prec_r (\mathbf{y}, r, p)$ or $(\mathbf{w}, r, p) \succ_r (\mathbf{y}, r, p)$. Hence, $u(\mathbf{y}) \neq 0$ (Lemma

7). Let $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$, i.e., $\delta_r(r,p)\,u(\mathbf{y}) = \delta_r(r+s,p)\,u(\mathbf{z})$. Hence, also, $u(\mathbf{z}) \neq 0$. From the definition of a delay function (Definition 12) we have $(\mathbf{y},r+t,p) \sim_r (\mathbf{z},r+s+D(r,s,t),p)$, i.e., $\delta_r(r+t,p)\,u(\mathbf{y}) = \delta_r(r+s+D(r,s,t),p)\,u(\mathbf{z})$. Hence, $\delta_r(r+t,p)\,\delta_r(r+s,p)\,u(\mathbf{z}) = \delta_r(r,p)\,\delta_r(r+s+D(r,s,t),p)\,u(\mathbf{z})$. Since $u(\mathbf{z}) \neq 0$, we get $\delta_r(r+t,p)\,\delta_r(r+s,p) = \delta_r(r,p)\,\delta_r(r+s+D(r,s,t),p)$.

(b) Let $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p)$. Hence, $u(\mathbf{y}) \neq 0$ (Lemma 7). Let $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$, i.e., $\delta_r(r,p)\,u(\mathbf{y}) = \delta_r(r+s,p)\,u(\mathbf{z})$. Hence, also, $u(\mathbf{z}) \neq 0$. From the definition of a defer function (Definition 13) we have $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+\Delta(r,s,t)+t,p)$, i.e., $\delta_{r+t}(r+t,p)\,u(\mathbf{y}) = \delta_{r+t}(r+\Delta(r,s,t)+t,p)\,u(\mathbf{z})$. Hence, $\delta_r(r+s,p)\,\delta_{r+t}(r+t,p)\,u(\mathbf{z}) = \delta_r(r,p)\,\delta_{r+t}(r+\Delta(r,s,t)+t,p)\,u(\mathbf{z})$. Since $u(\mathbf{z}) \neq 0$, we get $\delta_r(r+s,p)\,\delta_{r+t}(r+t,p) = \delta_r(r,p)\,\delta_{r+t}(r+\Delta(r,s,t)+t,p)$.

(c) Let $(\mathbf{w},r,p) \prec_r (\mathbf{x},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{x},r,p)$. Hence, $u(\mathbf{x}) \neq 0$ (Lemma 7). Let $(\mathbf{x},r,p) \sim_r (\mathbf{y},r+t,p)$ and $(\mathbf{y},r+t,p) \sim_{r+t} (\mathbf{z},r+s+t,p)$, i.e., $\delta_r(r,p)\,u(\mathbf{x}) = \delta_r(r+t,p)\,u(\mathbf{y})$ and $\delta_{r+t}(r+t,p)\,u(\mathbf{y}) = \delta_{r+t}(r+s+t,p)\,u(\mathbf{z})$. Hence, also, $u(\mathbf{y}) \neq 0$, $u(\mathbf{z}) \neq 0$ From the definition of a shift function (Definition 14), we get $(\mathbf{x},r,p) \sim_r (\mathbf{z},r+S(r,s,t),p)$, i.e., $\delta_r(r,p)\,u(\mathbf{x}) = \delta_r(r+S(r,s,t),p)\,u(\mathbf{z})$. Hence, $\delta_r(r+t,p)\,u(\mathbf{y}) = \delta_r(r+S(r,s,t),p)\,u(\mathbf{z})$. Therefore,
$\delta_r(r+t,p)\,\delta_{r+t}(r+t+s,p)\,u(\mathbf{z}) = \delta_{r+t}(r+t,p)\,\delta_r(r+S(r,s,t),p)\,u(\mathbf{z})$. Since $u(\mathbf{z}) \neq 0$, we get $\delta_r(r+t,p)\,\delta_{r+t}(r+t+s,p) = \delta_{r+t}(r+t,p)\,\delta_r(r+S(r,s,t),p)$.

(d) Let $(\mathbf{w},r,p) \prec_r (\mathbf{y},r,p)$ or $(\mathbf{w},r,p) \succ_r (\mathbf{y},r,p)$. Hence, $u(\mathbf{y}) \neq 0$ (Lemma 7). Let $(\mathbf{y},r,p) \sim_r (\mathbf{z},r+s,p)$, i.e., $\delta_r(r,p)\,u(\mathbf{y}) = \delta_r(r+s,p)\,u(\mathbf{z})$. Hence, also, $u(\mathbf{z}) \neq 0$. From the definition of a certainty function (Definition 15) we have $(\mathbf{y},r,q) \sim_r (\mathbf{z},r+C(s),q)$, i.e., $\delta_r(r,q)\,u(\mathbf{y}) = \delta_r(r+C(s),q)\,u(\mathbf{z})$ and, hence,
$\delta_r(r,q)\,\delta_r(r,p)\,u(\mathbf{y}) = \delta_r(r,p)\,\delta_r(r+C(s),q)\,u(\mathbf{z})$. It follows that
$\delta_r(r,q)\,\delta_r(r+s,p)\,u(\mathbf{z}) = \delta_r(r,p)\,\delta_r(r+C(s),q)\,u(\mathbf{z})$. Since $u(\mathbf{z}) \neq 0$, we get $\delta_r(r,q)\,\delta_r(r+s,p) = \delta_r(r,p)\,\delta_r(r+C(s),q)$. ■

Proof of Lemma 9

These properties follow immediately from Definition 24. ■

Proof of Lemma 10

(a) Let $p,q \in [0,1]$. From (7.4) $w(pq) = e^{-\beta[-\ln(pq)]} = e^{\beta\ln(pq)} = \left[e^{\ln(pq)}\right]^{\beta} = (pq)^{\beta} = p^{\beta}q^{\beta} = w(p)\,w(q)$.

(b) Let $p,q \in (0,1)$. From (7.4) $w(p)\,w(q) = e^{-\beta(-\ln p)^{\alpha}}e^{-\beta(-\ln q)^{\alpha}} = e^{-\beta[(-\ln p)^{\alpha}+(-\ln q)^{\alpha}]}$ and $w(pq) = e^{-\beta[-\ln(pq)]^{\alpha}} = e^{-\beta(-\ln p-\ln q)^{\alpha}}$. Hence, $w(p)\,w(q) < w(pq)$ if, and only if, $(-\ln p)^{\alpha} + (-\ln q)^{\alpha} > (-\ln p - \ln q)^{\alpha}$. Setting $x = -\ln p$ and $y = -\ln q$, the latter is equivalent to $x^{\alpha} + y^{\alpha} > (x+y)^{\alpha}$; which is valid for $x > 0$, $y > 0$, $0 < \alpha < 1$ (which all hold in part b).

(c) Similar to part (b) except that we use the mathematical identity $x^{\alpha}+y^{\alpha} < (x+y)^{\alpha}$ for $x > 0$, $y > 0$, $\alpha > 1$. ■

Proof of Lemma 11

Follows from (7.5), Example 10 and the mathematical fact that, for $s > 0$ and $t > 0$, $s^k + t^k \gtreqless (s+t)^k \Leftrightarrow k \lesseqgtr 1$. ■

Proof of Lemma 12

Follows by substituting from (7.1) into Lemma 8. ■

Proof of Theorem 8

From Lemma 12d, for $r, s \geq 0, p, q \in (0,1]$, we have

$$\frac{w(p)\,w(q\Pi(C(s)))}{w(q)\,w(p\Pi(s))} = \frac{\delta_r^0(r+s)}{\delta_r^0(r+C(s))}, \tag{10.24}$$

or, equivalently,

$$\begin{aligned} & \delta_r^0(r+C(s))\,w(p)\,w(q\Pi(C(s))) \\ = \; & \delta_r^0(r+s)\,w(q)\,w(p\Pi(s)). \end{aligned} \tag{10.25}$$

(a) Suppose probability independence holds (Definition 11a). From Theorem 6a, we get $C(s) = s, s \geq 0, p, q \in (0,1]$. Hence, from (10.24), we get $\frac{w(p)w(q\Pi(s))}{w(q)w(p\Pi(s))} = \frac{\delta_r^0(r+s)}{\delta_r^0(r+s)} = 1$ and, hence, $w(p)\,w(q\Pi(s)) = w(q)\,w(p\Pi(s))$.

Suppose $w(p)\,w(q\Pi(s)) = w(q)\,w(p\Pi(s))$, $s \geq 0$, $p, q \in (0,1]$. From (10.25), we get

$$\begin{aligned} & \delta_r^0(r+C(s))\,w(q\Pi(C(s))) \\ = \; & \delta_r^0(r+s)\,w(q\Pi(s)), \end{aligned} \tag{10.26}$$

which is satisfied by $C(s) = s$. The left hand side of (10.26) is strictly decreasing in C, while the right hand side is independent of C, hence, $C(s) = s$ is the unique solution to (10.26). Hence, $C(s) = s$.

(b) Suppose the certainty effect holds. From Theorem 6b, we get $C(s) > s$, $s > 0$, $p, q \in (0,1]$, $p > q$. Since the left hand since of (10.25) is strictly decreasing in C, we get $w(p)\,w(q\Pi(s)) > w(q)\,w(p\Pi(s))$. Conversely, if $w(p)\,w(q\Pi(s)) > w(q)\,w(p\Pi(s))$ then we get, from (10.25),

$$\begin{aligned} & \delta_r^0(r+C(s))\,w(q\Pi(C(s))) \\ < \; & \delta_r^0(r+s)\,w(q\Pi(s)). \end{aligned} \tag{10.27}$$

Since both sides of (10.27) are equal for $C = s$, and since the left hand side of (10.27) is strictly decreasing in C, while the right hand side is independent of C, we get that $C(s) > s$.

The proof for part (c) is similar to that of part (b), but with the inequalities reversed. ■

Proof of Theorem 9

From Theorem 8a, we know that probability independence (Definition 11a) holds if, and only if,

$$w(p)\,w(q\Pi(s)) = w(q)\,w(p\Pi(s)), \tag{10.28}$$

for all $s \geq 0$ and $p, q \in (0,1]$.

(a) Suppose no hazard, i.e., $\Pi \equiv 1$. We then get $w(p)\,w(q\Pi(s)) = w(p)\,w(q) =$

$w(q)\,w(p) = w(q)\,w(p\Pi(s))$.
Hence, (10.28) holds and, hence, probability independence holds.
(b) Suppose probability weighting is additive (Definition 25a). We then get
$w(p)\,w(q\Pi(s)) = w(p)\,w(q)\,w(\Pi(s)) = w(q)\,w(p)\,w(\Pi(s)) = w(q)\,w(\Pi(s)\,p)$. Hence, again, (10.28) holds and, hence, probability independence holds.
(c) This is an immediate consequence of (a) and (b). ■

Proof of Theorem 10

These conclusions follow from Theorems 3, 4, 5, Lemma 12a-c and Definitions 20, 24, 26. ■

Proof of Corollary 1

Set $p = 1$ and $\Pi \equiv 1$ in Theorem 10, then apply Definitions 21-23. ■

Proof of Theorem 11

Let

$$LHS = \frac{\delta_r^0(r+s)\,\delta_r^0(r+t)}{\delta_r^0(r+s+t)}, \tag{10.29}$$

$$RHS = \frac{w(p)\,w(p\Pi(s+t))}{w(p\Pi(s))\,w(p\Pi(t))},\ p \in (0,1], \tag{10.30}$$

then, from Theorem 10c, the common difference effect for probability $p \in (0,1]$ (Definition 6a) holds if, and only if,

$$LHS < RHS, \text{ for all } s > 0,\ t > 0. \tag{10.31}$$

From (10.29) and Definitions 21, 22, a straightforward calculation shows that

$$LHS \leq 1 \text{ for all } s > 0,\ t > 0. \tag{10.32}$$

On the other hand, from (10.30) and Definitions 25, 27, we get

$$RHS > 1 \text{ for all } s > 0,\ t > 0, \tag{10.33}$$

in each of cases (a)-(c). Hence, from (10.31), (10.32) and (10.33), it follows that the common difference effect holds for any of the cases (a)-(c). ■

References

[1] Ainslie, G. (1992). *Picoeconomics*. Cambridge: Cambridge University Press.

[2] Alvarado J.C., Vaughan J.W., Stanford T.R., and Stein B.E. (2007). Multisensory versus unisensory integration: Contrasting modes in the superior calculus. *Journal of Neurophysiology* 97: 3193–205.

[3] 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.

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

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

[6] 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.

[7] al-Nowaihi, A., Dhami, S. (2009). A value function that explains the magnitude and sign effects. *Economics Letters* 105, 224-229.

[8] al-Nowaihi, A., Dhami, S. (2018). Foundations for Intertemporal Choice. CESifo Working Paper No. 6913.

[9] Andersen, S., Harrison, G., Morten, I. L. and Rutstroem, E. (2010) Discounting behavior: A reconsideration. *European Economic Review.* 71, 15-33.

[10] Andreoni, J., and Sprenger, C. (2012) Estimating Time Preferences from Convex Budgets. *American Economic Review.* 102(7): 3333-3356.

[11] Baucells, M. and Heukamp, F.H. (2012). Probability and time trade-off. *Management Science* 58(4):831–842.

[12] Chakraborty, A., Halevy, Y., Saito, K. (2020) The relation between behavior under risk and over time. *American Economic Review: Insights* 2(1): 1-16.

[13] Dhami, S. (2019a) *Foundations of behavioral economic analysis. Volume I Behavioral Decision Theory.* Oxford University Press: Oxford.

[14] Dhami, S. (2019b) *Foundations of behavioral economic analysis: Volume III Behavioral time discounting.* Oxford University Press: Oxford.

[15] Dohmen, T., Falk, A., Huffman, D., Sunde, U. (2017) The Robustness and Pervasiveness of Sub-Additivity in Intertemporal Choice. Mimeo. University of Bonn.

[16] Echenique, F., Imai, T., Saito, K. (2019). Testable implications of models of intertemporal choice: Exponential discounting and its generalization. Forthcoming in *American Economic Journal.*

[17] Ericson, K., White, J., Laibson, D., Cohen, J. (2015). Money earlier or later? Simple heuristics explain intertemporal choices better than delay discounting. *Psychological Science*, 26 (6), 826-833, doi: 10.1177/0956797615572232.

[18] 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.

[19] Halevy, Y. (2008). Strotz Meets Allais: Diminishing Impatience and the Certainty Effect. *American Economic Review*. 98(3): 1145-1162.

[20] Halevy, Y. (2015) Time consistency: Stationarity and time invariance. *Econometrica*. 83(1): 335-352.

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

[22] Kahneman D., Tversky A. (2000). *Choices, Values and Frames*, Cambridge University Press, Cambridge.

[23] Keren, G., and Roelofsma, P. (1995). Immediacy and Certainty in Intertemporal Choice. *Organizational Behavior and Human Decision Making*, 63(3): 287–97.

[24] Köszegi, Botond and Matthew Rabin. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics* 121 (4): 1133-1165.

[25] Köszegi, Botond and Matthew Rabin. (2007). Reference-dependent risk attitudes. *The American Economic Review* 97 (4): 1047-1073.

[26] Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics* 112, 443-477.

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

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

[29] 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.

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

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

[32] Prelec, D. (1998). The probability weighting function. *Econometrica*. 60: 497-528.

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

[34] 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.

[35] Read, D., and Roelofsma, P. H. M. P. (2003). Subadditive versus hyperbolic discounting: A comparison of choice and matching. Organizational Behavior and Human Decision Processes 91(2): 140-153.

[36] Read, D. and Scholten, M. (2006). Beyond Discounting: The Trade-off Model of Intertemporal Choice. Working Paper LSE-OR 06.88. London: London School of Economics and Political Science.

[37] Rubinstein, A. (2003). Is it economics and psychology? The case of hyperbolic discounting, *International Economic Review*, 44 (4), 1207-1216.

[38] Ruggeri, K., Alí, S., Berge, M.L. et al. (2020) Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour* 4, 622–633.

[39] Samuelson, P. (1937). A note on measurement of utility. *Review of Economic Studies*, 4, 155-161.

[40] Scholten, M., and Read, D. (2006). Discounting by intervals: A generalized model of intertemporal choice, *Management Science*, 52 (9), 1424-1436.

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

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

[43] Stanford T.R., Stein B.E. (2007). Superadditivity in multisensory integration: Putting the computation in context. NeuroReport18:787–92.

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

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

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

[47] Tversky, A., and Koehler, D.J. (1994). Support theory: a nonextensional representation of subjective probability. *Psychological Review*. 101(4): 547–567.

[48] Wakker, P. P. (2010). *Prospect Theory for Risk and Ambiguity*. Cambridge University Press.

[49] Weber, B. J., and Chapman, G. B. (2005). The Combined Effects of Risk and Time on Choice: Does Uncertainty Eliminate the Immediacy Effect? Does Delay Eliminate the Certainty Effect? *Organizational Behavior and Human Decision Processes*, 96(2): 104–18.